path of the vehicle if sums S'i and S"i are both large and significantly larger than sum S"',, since the sums indicate that the patch is associated with a line or lines that are vertical and/or horizontal and in the direction of the roadway.
The value of the weighting function ai is assigned to the patch based on the degree to which the patch
W, is deemed to be a projection of the roadway (case (a) above).

It will be appreciated that, if, for a patch W'*, the sum S"i is relatively large, indicating that the patch is associated with a line that is horizontal and in the direction of the roadway, but sums S'i and S..., are relatively small, indicating that the patch is not associated with a line that is vertical or horizontal and perpendicular to the direction of the roadway, it generally cannot be determined from the set of sums generated for the patch whether the patch is a projection of the roadway (case (a) above) or a projection of an object to the side of the path of the vehicle (case (c) above). However, since the patch is not associated with a line that is vertical, it will generally not be deemed to be a projection of an obstacle (case (b) above). In that case, an assessment as to whether the patch is a projection of the roadway (case (a) above) or a projection of an object to the side of the path of the vehicle (case (c) above) by referring to patches adjacent thereto.

As noted above, the value of the gradient strength Pi for a patch reflects the degree to which the patch the contains a texture, and thus will more likely to contain useful information for use in determining ego motion of the vehicle. The gradient strength Pi corresponds to
Image available on "Original document"

For relatively uniform patches, the value of the
Image available on "Original document"

will be relatively low for all motions, in which case the value ofp ; will be relatively low. On the other hand, for patches with texture, the value of the SSD will be relatively high for most motions, in which case the value of p, will be relatively high.

With this background, operations performed by the ego-motion estimation system processor 14 will be describe in connection with the flow chart depicted in FIG. 2. In connection with FIG.

2, it is assumed that the ego-motion estimation system processor 14 already has image T, which it may have used in connection with determining the translational and rotational motion up to the location at which image recorded. With reference to FIG. 2, after the ego-motion estimation system processor 14 has received image T' (step 100), it will rectify the image according to information provided during the camera 13 calibration operation (described below) to provide that the optical axis is parallel to the plane defined by the roadway (step 101). In addition, the egomotion estimation system processor 14 will generate an initial guess as to the translational and rotational motion, using the previous motion estimate and, perhaps information from other sensors if available (step 102). For example, the ego-motion estimation system processor 14 may make use of information from the vehicle 10's speedometer, as well as information as to the time period between the time at which image T was recorded and the time at which image T'was recorded, in generating the initial guess. Generally, it will be appreciated that the time period will be fixed, and will preferably the same for each

IPR2025-00944
Tesla EX1002 Page 947

successive pair of images T and T'. After the ego-motion estimation system processor 14 has generated the initial guess, it will use the initial guess to warp image T'toward image T, thereby to generate a warped image (step 103).

After the ego-motion estimation system processor 14 has generated the warped image (step 103), it will divide the image # and the warped image #' into patches and, for each pair of corresponding patches in the two images T and', generate the weighting value Xi. In that operation, the ego-motion estimation system processor 14 will select a patch in the image T (step 104) and generate values for P2 (step 105), P, (equation 20) (step 106) and Xj (equation 22) (step 107) as described above. In addition, the ego-motion estimation system processor 14 can generate the value for Pi (equation 23) and ai (step 108). After the ego-motion estimation system processor 14 has generated performed steps 105 through 108 for the selected patch, it will determine whether all of the patches in image T have been processed (step 109) and if not, return to step 104 to select another patch and perform steps 105 through 109 in connection therewith.

The ego-motion estimation system processor 14 will perform steps 104 through 109 in connection with each patch in the image T.. After the ego-motion estimation system processor 14 has performed steps 104 through 109 in connection with all of the patches in the image, it will sequence from step 109 to step 110 to search for the motion m that maximizes the value provided by equation (19) (step 110). That motion m will comprise values for translation tz and rotation wx, wy parameters that will constitute the estimate of the motion of the vehicle 10 as between the point in time at which image T was recorded and the point in time at which image T'is recorded. The ego-motion estimation system processor 14 can perform operations described above in connection with each successive pair of images T and T'to estimate the motion of the vehicle 10.

In performing steps 106 (to generate the values for P_i) and 110 (to determine the motion m that maximizes the value provided by equation (19)), the ego-motion estimation system processor
14 can perform a gradient descent that is limited to a selected cube-shaped region around the initial guess.

In determining the initial guess (step 102) for each new image T', the ego-motion estimation system processor 14 can use the estimate of the motion generated for the previously-received image.

In addition, the size of the region # can be adjusted adaptively.

A second methodology for determining the ego-motion of a vehicle 10 will be described in connection with FIG. 3. As with the methodology described above in connection with FIG. 2, it is assumed that the images have been rectified so that the image planes are perpendicular to the plane represented by the roadway on which the vehicle 10 is traveling, and their horizontal ("x") axes are parallel to the plane represented by the roadway. By way of background, in that case, the equation of a plane AX+ BY+ CZ = 1 (24) (reference equation (3)) reduces to
BY = 1 (25), in which case
image available on "Original document"

is the height of the optical axis of the camera 13 (or, more specifically, the Z axis) above the road. Since, for a point p (x, y) in an image that is a projection of a

Tesla EX1002 Page 948 IPR2025-00944

point P (X, Y, Z) y in three-dimensional space, y = f-, equation (25) becomes
1 -= boy (26)
The brightness constraint is ulX+vly+l=0 (27) for each point, where, at each point (x, y) in the image, Ix and Iy are the horizontal and vertical components of the spatial gradient of the luminance and It is the time gradient of the luminance. In addition, the equations for the components (u, v) of the flow vector (reference equation (1)) can be written
Image available on "Original document"

where [w] x is a skew-symmetric matrix, in which case
Image available on "Original document"

where
Image available on "Original document"

For motion constrained to a plane, equation (29) reduces to (ax + by + c) STt + ST[w} xp + It = 0
Since the images are rectified, equation (26) will hold. In addition, since rotational motion is constrained so as to occur only around the vertical ("Y") axis, wx=wz=0. Accordingly, equation (31) will correspond to
Image available on "Original document"

Expanding equation 32, and using equation (30),
Image available on "Original document"

Generally, the portion uwy of the vertical component of the flow vector that arises due to rotation (wy) of the vehicle 10 will be larger than the portion ux that arises due to translation (tx) along the "X"axis, and the portions will have a very different form. The portion ux that arises due to translation along the"X"axis will depend on the distance that the vehicle 10 moves between the times the images are recorded, and will be larger for points towards the bottom of the image than for points further up. In that case, expressions for uwy and ux are
Image available on "Original document"

Since, for small rotations around the vertical"Y,"axis, the angle of rotation wy is approximately t wy ~ x (36), tz
Image available on "Original document"

It will be appreciated that values for ty, the component of the translation t in the vertical direction, and wx and wz, the X and Z components of rotation w, will be zero. Accordingly, after the egomotion estimation system processor 14 receives a new image T', it will determine the values for t, and t3, the components of the translation t in the forward (along the Z axis) and side (along the X axis) directions, and wy, the component of rotation around the vertical (Y) axis. In that operation, the ego-motion estimation system processor 14 will generate an initial estimate as to the motion (step 150, FIG. 3) and use that initial estimate to generate a warped image sus (step 151). The egomotion estimation system processor 14 can use information from a number of

IPR2025-00944
Tesla EX1002 Page 949

sources in connection with generating the initial estimate (step 150), including information from, for example, the vehicle 10's speedometer. Thereafter, the ego-motion estimation system processor 14 divides the image T and Y'into patches (step 152) and determines which patches are likely to comprise images of the roadway (step 153). In determining which patches are likely to comprise images of the roadway (reference step 153), the ego-motion estimation system processor 14 can generate an SSD (equation 16) for corresponding patches in the images T and V"and deten-nine that patches in the warped image F'that comprise images of the roadway will be those patches with a relatively high SSD value. The ego-motion estimation system processor 14 then uses the patches identified in step 153 to minimize a cost function of the form image available on "Original document"

where W, and Wz are weighting matrices that essentially describe confidence in the priors or smoothness values (step 154). Equation (38) can be formalized in the form of a Kalman filter, and the value of'p"can be selected to be one or two depending on whether the L, or L2 norm is to be used.

As noted above, for methodologies described above in connection with both FIGS. 2 and 3, the ego-motion estimation system processor 14 will initially rectify the images as received from the camera 13. In the above description, the images I and I'are images as rectified by the ego-motion estimation syst%n processor 14. To rectify the image3 as received from the camera 13, the camera 13 will need to be calibrated during a calibration operation prior to use in connection with recording images for use in estimating vehicle 10 motion as described above. Before describing operations to be performed during calibration, it would be helpful to consider the effects of incorrect calibration.

If, for example, a vehicle is driving along a straight road, and if the optical axis of the camera is aligned with the direction of motion, the flow field of successively recorded images will be an expansion field with the focus of expansion located at the center of the respective images, that is, at the origin (x, y) = (0, 0) of the image's rectilinear coordinate system. On the other hand, if the camera is mounted on the vehicle with a small rotation around the vertical ("Y") axis in three-dimensional space, then the focus of expansion will be displaced along the image's horizontal ("x") axis. In that case, the motion model defined by equation (11) will not account for the flow field, but will be well approximated by a forward translation and a rotational velocity wy around the vertical ("Y") axis.

Accordingly, errors in the orientation of the camera around the vertical ("Y") axis in three dimensional space will create a bias in the rotation estimate, in which case a curved path would be estimated instead of a straight path. Similarly, errors in the camera's orientation around the horizontal ("X") axis in three dimensional space will cause a bias in the pitch estimate. Based on these observations, a calibration operation can be performed by having the camera record a sequence of images while the vehicle is being driven down a straight roadway. The vehicle's ego-motion is estimated as described above in connection with FIGS. 2 or 3, and calibration parameters are estimated that would cause the ego-motion to integrate into a straight path.

The invention provides a number of advantages. In particular, the invention provides an arrangement for determining ego-motion of a vehicle 10 on a roadway from a series of images recorded by a camera 13 mounted on the vehicle 10, at least a portion of the images comprising projections of the roadway, and without requiring mechanical

Tesla EX1002 Page 950 IPR2025-00944

sensors which are normally not provided with a vehicle 10 and that would, if provided, increase the cost and maintenance expenses thereof.

It will be appreciated that a system in accordance with the invention can be constructed in whole or in part from special purpose hardware or a general purpose computer system, or any combination thereof, any portion of which may be controlled by a suitable program. Any program may in whole or in part comprise part of or be stored on the system in a conventional manner, or it may in whole or in part be provided in to the system over a network or other mechanism for transferring information in a conventional manner. In addition, it will be appreciated that the system may be operated and/or otherwise controlled by means of information provided by an operator using operator input elements (not shown) which may be connected directly to the system or which may transfer the information to the system over a network or other mechanism for transferring information in a conventional manner.

The foregoing description has been limited to a specific embodiment ofthis invention. It will be apparent, however, that various variations and modifications may be made to the invention, with the attainment of some or all of the advantages of the invention. It is the object of the appended claims to cover these and such other variations and modifications as come within the true spirit and scope of the invention.

What is claimed as new and desired to be secured by Letters Patent of the United States is :

Espacenet

Claims: EP1257971 (A2) — 2002-11-20

SYSTEM AND METHOD FOR ESTIMATING EGO-MOTION OF A MOVING VEHICLE USING SUCCESSIVE IMAGES RECORDED ALONG THE VEHICLE'S PATH OF MOTION

Claims not available for EP1257971 (A2)
Claims of corresponding document: WO0139120 (A2)

A high quality text as facsimile in your desired language may be available amongst the following family members:

CA2392652 (A1) WO0139120 (A2)

- Original claims
- Claims tree

The EPO does not accept any responsibility for the accuracy of data and information originating from other authorities than the EPO; in particular, the EPO does not guarantee that they are complete, up-to-date or fit for specific purposes.

CLAIMS 1. An ego-motion determination system for generating an estimate as to the ego-motion of a vehicle moving along a roadway comprising :
A. an image information receiver configured to receive image information relating to a series of at least two images recorded as the vehicle moves along a roadway ; and
B. a processor configured to process the image information received by the image receiver to generate an ego-motion estimate of the vehicle.

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Veröffentlichungsnummer:
(11) Publication number: **EP 1 257 971 A0**
(11) Numéro de publication:

Internationale Anmeldung veröffentlicht durch die Weltorganisation für geistiges Eigentum unter der Nummer:

WO 01/039120 (art. 158 des EPÜ).

International application published by the World Intellectual Property Organisation under number:

WO 01/039120 (art. 158 of the EPC).

Demande internationale publiée par l'Organisation Mondiale de la Propriété sous le numéro:

WO 01/039120 (art. 158 de la CBE).

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 622 072 B1

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention of the grant of the patent:
07.07.2010 Bulletin 2010/27

(51) Int Cl.:
G06N 3/00 *(2006.01)*

(21) Application number: **04291912.6**

(22) Date of filing: **27.07.2004**

(54) **An automated action-selection system and method and application thereof for training prediction machines and for driving the development of self-developing devices**

Ein automatisiertes System zur Aktionsanwahl, sowie das Verfahren und dessen Anwendung, um Prognosemaschinen auszubilden und die Entwicklung sich selbst entwickelnder Geräte zu unterstützen

Système de sélection d'actions, méthode et application associées, pour former des machines de prévision et pour conduire le développement de dispositifs capables de s'auto-développer

(84) Designated Contracting States:
DE FR GB

(43) Date of publication of application:
01.02.2006 Bulletin 2006/05

(73) Proprietor: **Sony France S.A.**
92110 Clichy La Garenne (FR)

(72) Inventors:
- **Kaplan, Frédéric**
  75005 Paris (FR)
- **Oudeyer, Pierre-Yves**
  75013 Paris (FR)

(74) Representative: **Thévenet, Jean-Bruno et al**
Cabinet Beau de Loménie
158, rue de l'Université
75340 Paris Cédex 07 (FR)

(56) References cited:
- **F. KAPLAN, P. OUDEYER: "Motivational principles for visual know-how development" PROCEEDINGS OF THE THIRD INTERNATIONAL WORKSHOP ON EPIGENETIC ROBOTICS, 5 August 2003 (2003-08-05), XP002329049 BOSTON - USA**
- **F. KAPLAN, P. OUDEYER: "Maximizing Learning Progress: An Internal Reward System for Development" EMBODIED ARTIFICIAL INTELLIGENCE: INTERNATIONAL SEMINAR, 11 July 2003 (2003-07-11), XP002329050 DAGSTUHL CASTLE - GERMANY**
- **SCHMIDHUBER J ED - INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS: "Curious model-building control systems" PROCEEDINGS OF THE INTERNATIONAL JOINT CONFERENCE ON NEURAL NETWORKS. SINGAPORE, NOV. 18 - 21, 1991, NEW YORK, IEEE, US, vol. VOL. 2, 18 November 1991 (1991-11-18), pages 1458-1463, XP010054430 ISBN: 0-7803-0227-3**
- **P. OUDEYER, F. KAPLAN: "Intelligent Adaptive Curiosity: a source of self-development" PROCEEDINGS OF THE FOURTH INTERNATIONAL WORKSHOP ON EPIGENETIC ROBOTICS, 27 August 2004 (2004-08-27), XP002329051 GENOA, ITALY**

Note: Within nine months of the publication of the mention of the grant of the European patent in the European Patent Bulletin, any person may give notice to the European Patent Office of opposition to that patent, in accordance with the Implementing Regulations. Notice of opposition shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

EP 1 622 072 B1

Printed by Jouve, 75001 PARIS (FR)

Description

Field of the Invention

[0001] The present invention relates to the optimization of series of actions or trials which are performed in order to learn. More particularly, the invention is directed to an automated method and system for selecting which trials or actions should be tried next in order to achieve efficient learning.

[0002] The invention also relates to application of this technique in methods for training prediction machines, and to control the progress of active learning, or to organize the behaviour of, a self-developing device or system (e.g. a robot). The invention further relates to systems for training prediction machines using this technique and to systems applying this technique for controlling active learning or behaviour-organization in a self-developing device or system. The invention yet further relates to prediction machines trained using this technique and to self-developing devices applying this technique.

Background of the Invention

[0003] There is a wide variety of applications in which it is necessary or useful to perform a series of actions on or affecting a setup or system S in order to acquire knowledge about the behaviour of the system S. The setup S can be a natural system or an artificial (man-made) system - use of the expression "setup" does not imply that the system has been actively organized or configured (although that could be the case). Often the desired "knowledge" consists in learning some relationship that holds between a set of the system's parameters. In general, it is desirable to minimize the number of actions that must be carried out.

[0004] In general, any setup or system S can be described in terms of its own current state and the state of an environmental context in which the system finds itself. The system's current state can be described in terms of values taken by a number of different parameters characterizing the system ("system-state parameters"). In a similar way, the system's environmental context can be described in terms of values taken by a number of different parameters characterizing that context ("context parameters"). In many cases it is not possible to determine clearly whether a given parameter is a system-state or context parameter. Indeed this distinction is immaterial as far as the operation of the present invention is concerned. According both kinds of parameter will be referred to herein as system-state/context parameters.

[0005] Actions can be taken which affect the system S. For example, it may be possible to control a parameter which affects the system S, so that this parameter takes a selected value. The controlled (or "tuned") parameter may be a system-state/context parameter but this is not obligatory. For a particular system *S* it is possible to select a set of parameters (or "action variables") whose values will be tuned in actions that are intended to elicit information about the behaviour of the system *S*.

[0006] The present invention concerns the selection of actions, affecting a system *S*, which are to be taken deliberately with a view to observing how the system reacts or behaves. In the present document, the expression "action" is not intended to imply a necessity for physical motion or the operation of an actuator; on the contrary, the expression "action" relates to the deliberate setting of a set of one or more action variables to respective particular values. In many cases the tuned variables will relate to a physical quantity (e.g. the amount of water to supply to a crop, the level of a voltage to apply to an electrode, etc.), but this is not a requirement of the present invention.

[0007] It is helpful to explain what is meant in the present document by the expression "system-state/context/action space".

[0008] Consider an application in which a user wishes to discover the relationships inherent in a setup *S* in which circuit boards are being photographed for quality control purposes. The circuit boards can have substrates made of different materials and, hence, can have different brightness values. The user wishes to discover how the exposure time and ambient illumination affect the contrast between the circuit boards and the conductive paths they carry, as observed in the photographic images. More particularly, he wishes to find out what conditions produce high-contrast images.

[0009] It is assumed that the photographic equipment includes a digital camera mounted over a conveyor which carries the circuit boards into a fixed position relative to the camera. An image-processing unit determines automatically the contrast between the circuit boards and the conductive traces they carry, by processing the image data generated by the digital camera. For simplicity, it is assumed that the image processing unit can accurately determine what within the image corresponds to the circuit board and what corresponds to the conductive traces.

[0010] This particular system *S* can be described in terms of the brightness of the circuit board substrate, ***BR***, the ambient illumination, ***AI***, and the exposure period, ***EP***, at the time when the photographic images are generated. Only the exposure-period parameter ***EP*** can be set to different values under the user's control (or automatically), the other two parameters are considered to be outside the user's control. Thus the circuit board brightness, ***BR***, and ambient illumination, ***AI***, can be considered to be system-state/context parameters describing the system, and the exposure period ***EP*** can be considered to be an action variable whose value can be varied or "tuned".

IPR2025-00944
Tesla EX1002 Page 955

[0011] It can be considered that these parameters ***AI***, ***BR*** and ***EP*** define a multi-dimensional space which, in this simple case, is a three-dimensional space as illustrated in Fig.1. In a more realistic example the multi-dimensional space is likely to have considerably more dimensions than three, however this would be difficult, if not impossible, to represent in a drawing.

[0012] The above-described multi-dimensional space defined by the system-state/context parameter(s) and action variable(s) of a system is referred to in the present document as "system-state/context/action space".

[0013] At any given moment, the above-mentioned example system can be described in terms of a vector defining the values taken by the circuit board brightness, ***BR***, ambient illumination, ***AI***, and the exposure duration, ***EP***. Fig.1 shows an example of a vector, ***A***, corresponding to one particular combination of values of ***BR***, ***AI*** and ***EP***.

[0014] The cuboid shown in Fig. 1 corresponds to the system-state/context/action space for the contrast-measurement system *S* described above. It will be noted that Fig.1 shows limits of 0 and 1 on the values of each system-state/context parameter and action variable. This represents the system-state/context/action space in the case where each of the system-state/context parameter values and action variable values is normalised so as to range from 0.0 to 1.0.

Description of the Prior Art

[0015] Various proposals have already been made in the field of statistics, and in the field of developmental robotics, with regard to how a series of trials or actions can be scheduled so as to optimize learning.

[0016] In the field of statistics, "optimal experiment design" seeks to determine how it is possible to minimize the number of examples that it is necessary to consider in order to achieve a given level of performance in generalization.

[0017] In the field of developmental robotics, one of the main goals is to produce robots which can develop. In this context "development" consists in a progressive increase in the complexity of the activities in which the robot can engage, with an associated increase in the robot's capabilities. As the robot develops, it can be said that the robot is engaged in "learning". In order to develop, the robot must explore its sensory-motor space; in other words, it must discover the consequences of performing given actions. For example, this could mean determining the visual changes which result from setting the speed of the robot's wheels to a particular value in a given environment (or from setting the joints in the robot's neck in a particular configuration, etc.). The way in which the robot explores the sensory-motor state-space is by performing a series of actions in a given context (or environment) and noting the perceived results. This set of interactions enables the robot to learn.

[0018] It is desirable to maximize the efficiency of robot learning, in other words to reduce the number of actions that the robot must perform in order to improve its capabilities to a particular extent. Randomly exploring the sensory-motor state-space would be immensely inefficient; to gain knowledge about its sensory-motor mapping, the robot needs to explore the sensory-motor state-space methodically.

[0019] In the field of developmental robotics, it has been found that the efficiency of robot learning can be improved by ensuring that the robot first tackles tasks that are relatively simple but then moves on to tackling tasks of progressively increasing difficulty. In general, humans control the situations or tasks encountered by the robot at a given time, ensuring that there is a progression towards tasks of increasing complexity. Thus the learning process can be described as "passive learning" or "passive development".

[0020] Currently, there is considerable interest in producing a mechanism to enable a robot (or other man-made system) to develop autonomously. In this case a learning process can be the engine of development. More particularly, it is desired to produce a robot (or other man-made system) which, when it encounters a complex, continuous, dynamic environment, is capable of determining, without pre-programmed knowledge, which tasks or situations in this environment have a complexity which is suited for efficient learning at a given stage in its development. A robot/man-made system of this type would learn easy parts of its sensory-motor mapping first of all and then progressively shift its focus of attention to situations of increasing complexity.

[0021] In the field of machine learning, various "active learning" proposals have been made as to how to choose actions which will provide a maximal gain in knowledge or "know-how". These proposals differ with regard to the manner in which "knowledge" and gains in knowledge are quantified. They differ, also, with regard to the way in which it is decided which action is liable to result in the maximal gain in knowledge. Some of the main known proposals are discussed below.

"Choose the action which gives the least predictable result"

[0022] One approach was proposed by J. Denzler and C. Brown in "Information Theoretic Sensor Data Selection for Active Object Recognition and State Estimation" (see IEEE Transactions on Pattern Analysis and Machine Intelligence, 24, vol. 2, pp. 145-157, (2001)). This paper concerns a robotic system which selects its next action based indirectly on the output of a prediction module which produces predictions of the sensory consequences of a given candidate action. For example, the prediction module might predict that a particular predicted variable will take a value 2.0. A decision-making module evaluates the degree of certainty of the predictions produced by the prediction module (e.g. "the predicted

IPR2025-00944
Tesla EX1002 Page 956

value 2.0 will be correct 65% of the time) and selects, as the next action to perform, the candidate action for which the prediction module's prediction has the lowest degree of certainty.

[0023] In some applications this method can be efficient. However, in an environment which contains a random element, this approach can lead the robot to concentrate on taking actions which explore that random element (because it is maximally unpredictable). However, in general there is nothing useful to be learned by exploring a random element of this type. Thus, development of the robot will be side-tracked.

[0024] Moreover, techniques in which the certainties of predictions are evaluated are computationally expensive and cannot readily be implemented for continuous system-state/context/action spaces.

"Choose the action which is expected to result in the greatest prediction error"

[0025] Another approach was proposed by S. Thrun in "Exploration in Active Learning", in Handbooks of Brain Science and Neural Networks, ed. M. Arbib, MIT Press (1995). This proposal concerns another system which makes use of a prediction module which produces predictions of the consequences of a given candidate action. Once again, there is a module which evaluates the performance of the prediction module, this time by evaluating the prediction error that is expected to result from implementing a given candidate action. This decision-making module selects as the next action to be performed that candidate action for which it expects the greatest prediction error to arise.

[0026] This method can be implemented very efficiently and run in high-dimensional spaces in real-time on small robots. Moreover, it can be useful in controlled environments (e.g. when an artificial device is learning to navigate in a static environment). However, it is counter-productive in an uncontrolled non-deterministic environment because, once again, it can lead to a focus of attention on a random element in the environment.

"Choose the action which will maximally reinforce the certainty of the predictions which are currently uncertain"

[0027] A further technique was proposed by N. Roy and A. McCallum in "Towards Optimal Active Learning through Sampling Estimation of Error Reduction", in Proceedings of the 18th International Conference on Machine Learning (ICML-2001). This proposal arises in the statistical field of "optimal experiment design" and takes a probabilistic approach.

[0028] According to this technique, a robot contains a prediction module, arranged to predict the outcome of certain candidate actions. There is a further module which predicts the degree of certainty of these predictions. This further module is updated in a dynamic manner so that the reliability of the various predictions can become more certain. In this system, the candidate action which is selected for performance is the action which will lead to a maximum reinforcement of the certainty of the predictions for which the robot is currently unsure.

[0029] Unlike the above-described Denzler and Brown system, according to this technique the evaluation of a candidate action includes a step of predicting how the robot's knowledge will be changed, and how this will affect the certainties of its predictions. It will be seen, then, that in this system there is an attempt to evaluate the gain in knowledge that will arise from choice of a particular candidate action.

[0030] In an environment which contains a random element, this Roy and McCallum approach can avoid the situation where the robot is trapped in exploration of that random element. However, the robot needs to consider a certain minimum number of examples before it will determine that it is focusing on a random element (a part of the sensory-motor state-space or "system-state/context/action space" where its actions have no effect on the certainties of its predictions). Thus, time and computational resources are wasted on evaluating and performing actions which can lead to no useful gain in knowledge. This wasted time may be minimized in a system where the noise is independently normally distributed. However, that will not be the case in a non-controlled, realistic environment.

[0031] Furthermore, this method requires use of considerable computational resources and no efficient implementation exists for the case where the sensory-motor state-space (or, more generally, system-state/context/action space) is continuous.

"Choose the action which (you predict) will maximally decrease prediction error"

[0032] An earlier active learning proposal by the present inventors is described in copending European patent application no. 03 291 359.2. According to this technique, a self-developing device contains a prediction module designed to predict the outcome of candidate actions, and a meta-prediction module which predicts the prediction errors that are associated with the different predictions made by the prediction module. The self-developing device updates the meta prediction module, and monitors its current level of prediction errors. The candidate action that is selected for performance is that action which, the meta prediction module predicts, will lead to the largest decrease in prediction error (possibly several time steps in the future).

[0033] The above-described technique has proved useful in certain situations, and successfully prevents a robot from focussing its attention on a random element in the environment. However, in some cases this technique can lead the

IPR2025-00944
Tesla EX1002 Page 957

robot to enter a loop in which it alternately focuses its attention on a first situation associated with a high prediction error rate, and a second situation associated with a low prediction error rate. For example, in the case of a robot moving in a space enclosed by walls, the robot would spend time alternately bumping randomly against walls (where the error rate of predicting what would happen next is high) and then stopping in front of the wall, looking at it (a situation in which it is simple to predict what will happen next).

[0034] The inventors' earlier paper "motivational principles for visual know-how development" (published in the Proceedings of the Third International Workshop on Epigenetic Robotics, 5 August 2003, Boston, USA) describes a self-developing robot which measures various parameters of its own state and its environment and attempts to determine, autonomously, which action to take next. This self-developing robot calculates a reward function associated with different possible actions and selects an action associated with a high value of the reward function. The reward function is based on the values of one or more particular task-independent motivational variables. One motivational variable is a "predictability" variable that is based on the accuracy with which the self-developing device can predict the outcome of a particular function. Calculation of the predictability motivational variable involves predicting what values will be taken by certain variables after a given action and then comparing the prediction with the actual values.

Summary of the Invention

[0035] In most systems, the behaviour of the system varies depending upon "where the system is" in system-state/context/action space. In other words, for different regions in the system-state/context/action space, different relationships exist between the various system-state/context parameters and/or action variables. Each different region in system-state/context/action space corresponds to a different type of situation in the system, ***S***.

[0036] Fig.2 illustrates the case where the system-state/context/action space of Fig.1 can be divided into three regions in which respective different relationships hold between the ambient illumination, substrate brightness and exposure period. These three regions are labelled R_1, R_{21} and R_{22}. The vector ***A***, corresponding to one particular combination of values of the system-state/context parameters and action variable, has an end-point in region R_1.

[0037] As mentioned above, when investigating the behaviour of a setup ***S*** and trying to select which action affecting the setup should be performed next, it can be helpful to consider what learning progress can be expected to result from the performance of different candidate actions. The present inventors have realised that the best results are not obtained when learning progress is assessed based on analysis of the immediately-preceding series of actions that have been performed. Instead, it is advantageous to consider the previous actions of the same kind as the present one that have been performed when the setup was in a similar situation to the present one.

[0038] Clearly, it is necessary to have some way of determining which actions are "similar" to the presently-considered one and which situations are "similar" to the current situation. According to the present invention, action/situation pairs are judged to be similar to each other if they occupy the same region in system-state/context/action space. In other words, a situation ***A*** corresponding to system-state/context parameter and action variable values $SCA_A(t)$ and a situation *B* corresponding to system-state/context parameter and action variable values $SCA_B(t)$ will be considered to be "similar" if $SCA_A(t)$ and $SCA_B(t)$ define vectors having end-points in a common region of the sensory-motor state-space.

[0039] The above concept is illustrated in a highly-simplified form in Fig.3 which shows three vectors, in the system-state/context/action space of Fig.1 having three regions defined as in Fig.2. In the system ***S***, photographs can be taken in the situations ***A***, ***B*** and ***C***, corresponding to combinations of values of the system-state/context parameters and the action variable that are represented by $SCA_A(t)$, $SCA_B(t)$ and $SCA_C(t)$, respectively. According to the present invention, the situations ***A*** and ***B*** are considered to be similar to each other because the respective vectors representing the combinations of values of the system-state/context parameters and action variable when the system is in these situations ***A*** and ***B*** have end-points in a common region (region R_1 in Fig.3).

[0040] On the other hand, situation ***C*** is not similar to either of situations ***A*** or B because the vector representing the corresponding combination of values of the system-state/context parameters and action variable when the system is in this situation ***C*** has an end-point in region R_{21} of Fig.3, which is different from the region R_1 containing the end-points of the vectors corresponding to situations ***A*** and ***B***.

[0041] According to the present invention, it is considered to be advantageous to use different functions for assessing learning progress for situations in respective different regions of system-state/context/action space.

[0042] According to a first aspect thereof, the present invention provides an automated action-selection method and system for selecting which action affecting a setup ***S*** should be performed next when seeking to acquire knowledge efficiently about the behaviour of the setup ***S***. By using the action-selection technique of the present invention, the number of actions that are required in order to gain knowledge about the behaviour of setup ***S*** can be reduced. Moreover, efficient learning can be achieved even in a complex, continuous, dynamic environment. Furthermore, the learning process does not become trapped in a loop between highly-predictable and highly-unpredictable situations, or become side-tracked into exploring a random element in the setup ***S***.

[0043] Preferred embodiments according to a first aspect of the present invention provide an automated action-se-

IPR2025-00944
Tesla EX1002 Page 958

lection system and method which is capable of determining automatically which parts of system-state/context/action space should be explored next so as to match the complexity of the explored situation to the level of learning ("development") at the time.

[0044] The action-selection technique according to preferred embodiments of the first aspect of the present invention involves use of a prediction machine. More particularly, the prediction machine attempts to predict a respective value for each of a set of one or more variables, ***VAR***, that is dependent on or affected by the behaviour of setup ***S***. In order to select which action should be performed next, a set of possible next actions ("candidate actions") is evaluated. For each candidate action, the prediction machine predicts what respective value(s) it expects the predetermined set of variables ***VAR*** to take after this candidate action has been applied to the setup under investigation, taking into account the current state and context of that setup. A meta prediction machine then assesses how accurate it expects those predictions to be, notably by predicting what will be the prediction error.

[0045] One of the candidate actions is selected for performance so as to affect the setup ***S***. After a selected action has been performed, feedback is received regarding the effect of this action on the setup's state and context parameters and on the values of the set of predicted variables (unless these values can already be derived from the observed values of the system-state/context parameters). This feedback is used to adapt the function applied by the prediction machine, whereby to improve the accuracy of its predictions. This feedback is also used by the meta prediction machine so as to improve the accuracy of its assessments in the future.

[0046] According to the present invention, an assessment is made of the gain in knowledge that is expected to result from performance of a given candidate action. In certain preferred embodiments according to the first aspect of the invention, this expected gain in knowledge is evaluated by determining what reduction in prediction error is expected to result from performance of this candidate action.

[0047] A particularity of the present invention is that the expected gain in knowledge is not evaluated by comparing the level of knowledge immediately before the selected candidate action has been performed with the expected level of knowledge immediately after performance of this action. On the contrary, the expected gain in knowledge is evaluated by comparing the expected level of knowledge after the selected candidate action with the level of knowledge that existed on one or more previous occasions when the setup had values of system-state/context parameters and action variables that are similar to those applicable at the time when the selected candidate action will be performed.

[0048] In other words, according to the present invention learning progress is assessed based on study of some parameter which indicates the system's level of knowledge, and the evolution of this "knowledge-indicator" parameter is evaluated "locally" (that is, in the region of system-state/context/action space which contains the current values of the system-state/context parameter(s) and action variable(s)).

[0049] Because the expected change in the value of the knowledge-indicator parameter is evaluated based on actual values of this knowledge-indicator parameter observed in a comparable situation to the present one, and not simply based on the immediately preceding values of the knowledge-indicator parameter, the technique of the present invention enables efficient learning to take place whilst avoiding becoming trapped in a loop between highly-predictable and highly-unpredictable situations. Moreover, the technique of the present invention manages to avoid focussing attention on situations which contain a random element (and which, hence, cannot give rise to the learning of general rules).

[0050] In preferred embodiments of the invention, a candidate action is considered to produce a large knowledge gain if performance of this action is expected to cause a large decrease in prediction error. In other words, a prediction is made related to performance of a candidate action, and the expected prediction error is computed. The expected prediction error for this candidate action is compared with actual prediction error values observed when this kind of action was performed previously in similar circumstances to the present ones (that is, when the values of system-state/context parameter(s) and of the action variable(s) were in the same region of system-state/context/action space as they are now).

[0051] Once a candidate action is selected for performance, and has been performed, the data defining this action, the system-state/context parameters when this action was performed and the consequences of performance of this action (in terms of the observed values of the set of predicted variables, ***VAR***) can be considered to constitute a "training example".

[0052] Each training example can be represented using a multi-dimensional vector having as its different components the respective values of each system-state/context parameter at the time when this action was performed, the respective values of each action variable defining this action, and the respective values of each of the set of variables ***VAR*** that were observed as a result of performance of this action. According to this representation, each training example can be considered to be an association of two vectors, the first one defining the position of this training example in system-state/context/action space and the second defining the location of this training example in "outcome space" (that is, a space whose dimensions correspond to the actual values of the different predicted variables ***VAR***).

[0053] According to the preferred embodiments of the action-selection technique according to the first aspect of the invention, it is preferable to divide system-state/context/action space into regions in a progressive manner: starting by dividing the whole space up into a single region (or small number of regions) and then sub-dividing that region (or those

IPR2025-00944
Tesla EX1002 Page 959

regions) later on.

[0054] As actions are selected and performed, training examples are accumulated. The training example data is stored in association with the region in system-state/context/action space which contains the end-point of the vector defining this training example. When a decision is taken to divide a region of system-state/context/action space into two or more new regions, the training examples that previously were associated with the original region become assigned to one or other of the new regions. Thus, at a given moment each training example is associated with only one region.

[0055] According to the preferred embodiments of the action-selection technique according to the first aspect of the invention, a given region of system-state/context/action space will be sub-divided into two (or more) smaller regions when a first criterion is met. According to one example of this first criterion, sub-division of a region, ***R***, will take place when it is determined that the latest training example is the Nth training example having a combination of system-state/context parameters and action variables which falls within this region R (with N being a predetermined number).

[0056] Preferably, when a region, ***R***, is sub-divided into two or more smaller new regions, the boundaries separating the new regions from each other will be set based on a second criterion. According to one example of this second criterion, the region boundaries are set so as to minimize the variance in "outcome space" of the locations of the training examples which are assigned to the same new region of system-state/context/action space (whilst keeping a balance between the numbers of training examples located in the new regions of system-state/context/action space).

[0057] It will be seen that as system-state/context/action space is sub-divided in this progressive manner a tree structure builds up: the initial region can be considered to be the root, the sub-divisions of this initial region constitute the branches, sub-divisions of those sub-divisions constitute twigs, etc. A region at the end of a given chain of sub-divisions can be considered to be a "leaf".

[0058] When an action is selected and performed, the action-selection system according to the preferred embodiment of the first aspect of the invention receives feedback enabling it to determine the actual prediction error that occurred. The actual prediction error data will be stored in relation to a particular region of system-state/context/action space, namely that region corresponding to the combination of values of system-state/context parameters and action variable(s) applicable at the time when the prediction was made. The size of this region will depend upon whether this combination of values is located in a part of system-state/context/action space that has been highly sub-divided or barely divided at all.

[0059] In the action-selection technique according to the first aspect of the invention, it has been found to be advantageous if, every so often, there is a random selection of the next action. In other words, on a majority of occasions the next action will be selected so as to maximise the expected knowledge gain (evaluated locally) but, on a minority of occasions, there will be a random selection of one of the candidate actions.

[0060] According to a second aspect, the present invention provides a system and method for training a prediction machine. The prediction machine is trained by causing it to participate in the selection of a series of actions using the action-selection technique according to the first aspect of the invention.

[0061] In the prediction-machine training method according to preferred embodiments of the second aspect of the present invention, for each of a set of candidate actions which can be taken to affect a setup ***S***, the prediction machine attempts to predict what would be the value of a set of variables, ***VAR***, if that candidate action were to be performed (the respective values of the set of variables ***VAR*** being dependent on or affected by the behaviour of the setup ***S*** and the setup ***S*** being describable in terms of a number of system-state/context parameters). A meta prediction module evaluates the expected accuracy of the prediction (expected prediction error), and a knowledge-gain assessment module assesses what reduction in prediction error would result from performance of the candidate action. The expected knowledge-gain is assessed by comparing an indicator of prediction error expected to occur if this action is performed with a comparable indicator of the actual prediction error that existed on one or more previous occasions when the setup had a combination of values of system-state/context parameters and action variables that is similar to that applicable at the time of the current experiment.

[0062] According to the prediction-machine training technique of the second aspect of the invention, system-state/context/action space is divided up into regions in a dynamic manner. The prediction, meta prediction and knowledge-gain assessment are made separately for each region and, as the series of actions progresses (the number of training examples increases), they become better and better adapted to suit the properties of their assigned region. For each candidate action, the expected knowledge gain is assessed using a prediction function, meta prediction module and knowledge-gain assessment module assigned to the region in system-state/context/action space corresponding to that candidate action and the current input values of the system-state/context parameters.

[0063] An action-selection module selects one of the candidate actions for performance. On a majority of occasions the selected candidate action is the one which is expected to produce the greatest reduction in prediction error (as assessed "locally" in system-state/context/action space). The selected action is performed, feedback is received enabling the actual prediction error to be determined, and the operation of the prediction machine modules assigned to the relevant region in system-state/context/action space is modified taking into account the received feedback.

[0064] Once the prediction machine has been trained, in a sequence of actions selected using the action-selection technique of the present invention, it can be considered to be programmed with the relationships inherent in the setup

IPR2025-00944
Tesla EX1002 Page 960

S. Such a trained prediction machine has commercial value in its own right, in view of the fact that it can be used to predict how the setup S will behave for a given combination of values for the state and context parameters, and under the effect of different actions. Assuming that the setup S has parameters which are related to one another in a non-random fashion, this prediction will tend to be accurate because the prediction machine has already learned the relationships that hold between the system parameters.

[0065] Thus, the second aspect of the present invention also provides a prediction machine trained according to the above-described prediction-machine training technique.

[0066] The action-selection technique and prediction-machine training technique of the first and second aspects of the present invention can be implemented using purpose-designed hardware or using suitably-programmed general-purpose computers. In either case, it is necessary to configure the purpose-designed machine/programmed computer so that it knows what are the system-state/context parameters applicable to the setup S that is to be investigated, what actions can be taken affecting the setup S, notably what are the action variables which can be tuned (and what are the possible tuning values or ranges), and what variables(s) (***VAR***) should be predicted.

[0067] The necessary configuring can be performed by the manufacturer before supplying the purpose-designed machine, computer program or programmed general-purpose computer to the end-user. Alternatively, the purpose-designed machine, computer program, or programmed computer can be supplied to the end-user in a general state, allowing the user to identify what are the system-state/context parameters, action variables and predicted variable(s) (***VAR***) applicable to the setup he wishes to investigate. In the latter case, it may be possible for the user to re-initialise the purpose-designed machine/programmed computer so that he can input new system-state/context parameters, action variables and predicted variables in order to investigate a different setup.

[0068] After the initial configuration phase, a user, *U*, wishing to learn about the behaviour of a setup *S*, or to train the prediction machine to learn the behaviour of setup *S*, can input to the purpose-designed machine/programmed computer the current values of the system-state/context parameters. The purpose-designed machine/programmed computer will then indicate which action should be taken now. In suitable cases, the purpose-designed machine/programmed computer may automatically cause the selected action to be performed. Otherwise, the purpose-designed machine/programmed computer may simply indicate to the user which action should be taken next.

[0069] Once the selected action has been performed, the user inputs to the purpose-designed machine/programmed computer updated values of the system-state/context parameters and (if need be) the actual values of the set of predicted variables, and the purpose-designed machine/programmed computer will then determine what action should be performed next (what values should be assigned to the one or more action variables). Thus, the action-selection technique and prediction-machine training technique of the invention determine the sequence of actions that are performed affecting the setup S. This will enable the user or prediction machine to acquire knowledge about the setup S in a structured and efficient manner, reducing the number of actions that are required.

[0070] The action-selection technique and prediction-machine training technique according to the first and second aspects of the present invention are not limited to the case where a user (human or robot) inputs the current values of the system-state/context parameters and provides feedback on the consequences of performance of each action. In appropriate cases, the inputting of system-state/context parameter values and providing of feedback will be performed automatically, for example by sensors coupled to the purpose-designed hardware/programmed computer implementing the technique according to the invention.

[0071] According to a third aspect of the present invention, the above-described action-selection technique can be applied as a mechanism for driving the development of, or managing learning by, a robot or other self-developing device or system. More particularly, the third aspect of the present invention uses the action-selection technique of the first aspect of the invention to decide a sequence of actions to be performed by a self-developing device or system. The self-developing device or system performs the selected actions and provides feedback enabling the consequences of the action to be determined.

[0072] Preferred embodiments of the third aspect of the invention provide a decision-making module for a self-developing device or system. This decision-making module determines which action should be performed next by the self-developing device or system in order to promote efficient learning by that device or system and/or to achieve a desired organization of the behaviour thereof. The decision-making module is capable of determining for itself which parts of its system-state/context/action space should be explored next so as to match the complexity of the explored situation to the self-developing device or system's level of development at the time. The decision-making module can be integrated into the self-developing device or system itself, or can be physically separate from, but in communication with, that device or system.

[0073] The third aspect of the present invention enables active learning of complicated non-deterministic mappings in an efficient manner. Application of this approach by a robot (or other self-developing device or system) leads to organization of the robot's behaviour such that the robot selectively chooses the focus of its action (and "attention") in a manner leading to autonomous scaling of the complexity of the learning situations it encounters. The ability to organize the behaviour of a self-developing device or system in this way is an important achievement.

IPR2025-00944
Tesla EX1002 Page 961

[0074] The method according to the third aspect of the present invention constitutes a drive motivating the selection of actions by a self-developing device or system. This drive can be termed "Intelligent Adaptive Curiosity". This drive can be termed "intelligent" because the action-selection method ensures that the self-developing device avoids situations that are overly predictable or overly unpredictable. This drive can be termed "adaptive" because the situations which will be attractive to the self-developing device change over time. This drive can be designated as "curiosity" because it maximizes learning progress, thereby pushing the self-developing device towards exploration of novel situations (where things can be learnt).

[0075] In many cases this "Intelligent Adaptive Curiosity" drive will be one of several drives motivating the behaviour of a self-developing device or system. For example, in a robot there may be a number of drives relating to maintenance of the robot's viability: for example, temperature maintenance, energy maintenance, maintenance of physical integrity, etc.

[0076] There are a wide variety of techniques allowing these different, and often competing, drives to be integrated with each other: for example, simple weighting of the different drives; allocation of reward functions to each drive and determination of the overall reward associated with a proposed action; etc. The technical literature in this field can be consulted for further examples. In a context where the method according to the third aspect of the present invention constitutes one of a number of drives motivating the behaviour of a self-developing device, the invention can be applied regardless of which technique is selected for integrating the different drives.

[0077] Further features and advantages of the various aspects of the present invention will become clear from the following description of preferred embodiments thereof, given by way of example, illustrated by the accompanying drawings, in which:

Fig.1 is a graph illustrating a multi-dimensional "system-state/context/action" space applicable to a simple setup or system, S;
Fig.2 is a graph illustrating different regions in the system-state/context/action space of Fig.1;
Fig.3 is a graph illustrating different vectors in the system-state/context/action space of Fig.1;
Fig.4 is a block diagram indicating schematically the main components of a preferred embodiment of action-selection system according to a first aspect of the present invention;
Fig.5 is a block diagram indicating schematically the main components of a preferred embodiment of prediction-machine training system according to a second aspect of the present invention;
Fig.6 is a schematic block diagram showing the main elements in one example of a fully-automated system applying the first (or second) aspect of the present invention in the field of biology;
Fig.7 is a schematic block diagram showing the main elements in one example of a fully-automated system applying the first (or second) aspect of the present invention in the field of finance;
Fig.8 is a schematic representation of an experimental set-up, involving a simple robot, used to test an embodiment according to the third aspect of the present invention;
Fig.9 is a block diagram indicating schematically the main components of a preferred embodiment of training system according to a third aspect of the present invention, for a self-developing device or system;
Fig.10 is a graph showing how prediction error evolves when actions are selected according to a preferred embodiment of the action-selection method of the present invention, by the simple robot in the experimental set-up of Fig.8;
Fig. 11 is a graph indicating the percentage of time the simple robot of Fig.8 spends in different situations when its actions are selected according to the preferred embodiment of the action-selection method of the present invention;
Fig. 12 is a graph showing how prediction error evolves when, according to a first comparative example, actions are selected in a random manner by the simple robot in the experimental set-up of Fig.8;
Fig. 13 is a graph showing how prediction error evolves when, according to a second comparative example, actions are selected according to a "maximizing" algorithm by the simple robot in the experimental set-up of Fig.8;
Fig. 14 is a graph indicating the percentage of time the simple robot of Fig.8 spends in different situations when actions are selected according to the "maximizing" algorithm;
Fig.15 is a graph showing how prediction error evolves when, according to a third comparative example, actions are selected according to a "minimizing" algorithm by the simple robot in the experimental set-up of Fig.8;
Fig.16 is a graph indicating the percentage of time the simple robot of Fig.8 spends in different situations when actions are selected according to the "minimizing" algorithm; and
Fig.17 is a graph comparing how learning progresses when the actions of the simple robot of Fig.8 are selected using the action-selection algorithm according to the preferred embodiments of the present invention and using algorithms according to the comparative examples.

FIRST ASPECT OF THE INVENTION

[0078] Fig.4 is a block diagram which illustrates schematically the main components of an action-selection system 1 according to a preferred embodiment of the first aspect of the present invention, which is used to investigate the behaviour

IPR2025-00944
Tesla EX1002 Page 962

of a setup or system S. It is to be understood that, in practice, this action-selection system will often be implemented in software; the different "modules" shown in Fig.4 are identified merely to aid understanding of the various functions that are performed in the overall system. Moreover, the distribution of functions between the various modules shown in Fig. 4 could be changed and/or these functions could be performed using a lesser or greater number of modules than that shown in Fig.4.

[0079] As shown in Fig.4, the action-selection system 1 of the preferred embodiment of the first aspect of the invention includes a region-identifier module 2, a knowledge-gain assessment module 3, an action-selection module 4 and a module 5 for calculating actual prediction error. The action-selection system 1 selects which action should be performed at each of a series of time steps, these time steps being separated by a predetermined interval, τ.

[0080] The region-identifier module 2 receives signals SC(t) indicating what are the values of the system-state/context parameters of the setup or system ***S*** at a first time step t. This data can be obtained from the setup ***S*** in a number of different ways, as indicated above. In general, the values of the system-state/context parameters will either be input by a user or received directly from sensing devices coupled to the action-selection system 1.

[0081] The region-identifier module 2 also receives signals A(t) indicating what is the set of candidate actions that could be performed next. Typically, the kth candidate action, $A_k(t)$, assigns respective values v_1, v_2, ..., v_n to each of a set of n action variables AV_1, AV_2, ..., AV_n. This can be expressed, as follows:

$$A_k(t) = \{AV_1(t) = v_1; AV_2(t) = v_2; \ldots; AV_n(t) = v_n\}.$$

[0082] It will be noted that each candidate action assigns values to the action variables for the current time step ***t***. In fact, the action variables would not take the assigned values v_1 etc. until a time $t+\delta t$, because it takes the action-selection system 1 a finite time δt to choose a candidate action and to arrange for this candidate action to be performed. However, this will not perturb the good functioning of the action-selection system 1 provided that the system-state/context parameters do not change significantly from SC(t) during the period ***t*** to $t+\delta t$. In many applications this will inevitably be the case because the system's state and context will not change until the selected action is performed. In other applications, notably real-time applications, the time δt required to compute and select the action (typically a few milliseconds), and perform the action, is so short that the setup's state and context will not have changed during that brief time.

[0083] The set of candidate actions can be input to the action-selection system 1 of the present invention, for example by a human user or an external device or system. Alternatively, the action-selection system 1 may itself include a candidate action generator module (not shown). In either case, given the setup's current state and context, it may be impossible to assign certain values to the action variables. In order to avoid redundant processing, it is preferable for the set of candidate actions to be generated taking into account the current values of the system-state/context parameters, ***SC(t)***.

[0084] In certain preferred embodiments of the invention, the action variables and system-state/context parameters are normalized within the action-selection system 1 of the invention, so as to take values running from 0.0 to 1.0. This simplifies processing within the system 1. The actual inputs to and outputs from the system can be converted between the normalised values and the corresponding "real-world" value using well-known techniques. For example, the "real-world" range of a voltage variable, ***VV***, could be from 0 volts to 15 volts, such that the assignment of a value to this voltage variable in the following terms - "set ***VV*** to 0.2" - means "set voltage variable, ***VV***, to 3 volts".

[0085] The region-identifier module 2 has a dual role. Firstly, it serves to divide up system-state/context/action space into a number of regions, as will be explained in greater detail below. Secondly, for each candidate action, $A_k(t)$, the region-identifier module 2 identifies the region of system-state/context/action space that contains the combination of this candidate action $A_k(t)$ and the current values ***SC(t)*** of the system-state/context.

[0086] The action-selection unit 1 according to the invention is adapted to assess knowledge gain separately for the different regions of system-state/context/action space. This is achieved by arranging the knowledge-gain assessment module 3 to include a number of expert units, ***EXP,*** one for each region of system-state/context/action space, and by virtue of the fact that the region-identifier module 2 is capable of identifying which region of system-state/context/action space contains the combination of each candidate action $A(t)$ and the current values $SC(t)$ of the system-state/context parameters.

[0087] Fig.4 shows three expert units, $EXPR_{R1}$, EXP_{R21} and EXP_{R22}, in the knowledge-gain assessment module 3. However, the actual number of expert units defined in the knowledge-gain assessment module 3 at a particular time varies in a dynamic manner as the region identifier 2 sub-divides system-state/context/action space into new regions.

[0088] According to the preferred embodiment of the invention, each expert unit includes its own prediction module 12, meta prediction module 13, knowledge-gain assessment module 14 and data memory 15. When the region identifier 2 determines that the candidate action would put the setup ***S*** into a situation that is in a region R_{21} of system-state/context/action space, the expert unit EXP_{R21} associated with region R_{21} is used to predict the outcome of performing

IPR2025-00944
Tesla EX1002 Page 963

this candidate action, to assess the expected knowledge gain that will result, and to output expected-knowledge-gain signals to the action selection module 4.

[0089] The operation of an expert unit EXP_{R21} associated with a region R_{21} of system-state/context/action space will now be considered in more detail. A counting variable q can be used to designate the successive occasions on which the setup S is in a situation corresponding to region R_{21} of system-state/context/action space. These occasions may be spaced apart in time in an irregular fashion.

[0090] When the region identifier 2 identifies region R_{21} as the region of system-state/context/action space that contains the situation corresponding to a given candidate action $A_k(t)$, taking into account the current values $SC(t)$ of the system-state/context parameters, the expert unit EXP_{R21} associated with region R_{21} is activated. For this candidate action $A_k(t)$, the prediction module 12 of expert unit EXP_{R21}, predicts the value of each a set of one or more variables VAR at the next time step (t+1). The nature of these variables VAR depends upon the system being investigated by the action-selection system of the invention. Typically, the prediction module 12 could attempt to predict what would be the value of one or more system-state/context parameters, or some variable(s) derivable from those parameters.

[0091] The predictions made by the prediction module 12 of expert unit EXP_{R21} can be labelled $VAR_{R21}(t+1)$, and they are based on the system's current system-state/context parameter values $SC(t)$, as well as on the given candidate action $A_k(t)$. In other words, $P_{R21}(A_k(t),SC(t)) = VAR_{R21}(t+1)$. More generally, for a candidate action $A(t)$ corresponding to a situation in region, R, of system-state/context/action space, the prediction can be expressed by the equation:

$$P_R(A(t),SC(t)) = VAR_R(t+1) \qquad (1)$$

[0092] A large number of different prediction techniques are known, for example: the nearest neighbour algorithm, methods using multi-layer perceptron neural networks, radial basis function neural networks, support vector machines, Bayesian machines, etc.. The present invention is not particularly limited with regard to which method is used to perform the prediction process. However, in preferred embodiments of the invention the prediction function applied by the prediction modules 12 of the expert units is the nearest neighbour algorithm.

[0093] Moreover, it is advantageous for the prediction implemented by the prediction module 12 of a given expert unit, for example EXP_{R21}, to be tailored to the characteristics of its region of system-state/context/action space, e.g. R_{21}. More particularly, over time the expert unit EXP_{R21} will adaptively modify the prediction function applied by its prediction module unit 12 so as to improve the prediction accuracy. (This will be the case for all of the above-listed examples of prediction techniques). The appropriate adaptation is implemented when the expert unit receives training example data for training examples located within the region of system-state/context/action space for which it is responsible. As each training example is assigned only to one region, this adaptation will be specific to the region R_{21}. Thus, even if the prediction modules 12 of all of the expert units initially apply the same prediction function (and initial use of the same function is not obligatory), after a time these functions will be different.

[0094] Each expert unit, EXP, also includes a meta prediction module 13 which estimates what will be the prediction error for each prediction made by the prediction module 12 of the same expert unit. For example, considering the operation of an expert unit EXP_R, it should be explained that, for each candidate action $A(t)$ corresponding to a situation in region R of system-state/context/action space, the meta prediction module 13 of this expert unit EXP_R computes a value indicating the respective errors $Ep_R(t)$ expected to be associated with the predictions $VAR_R(t+1)$ made by the prediction module 12 of this same expert unit, EXP_R, as follows:

$$MP_R(A(t),SC(t)) = Ep_R(t) \qquad (2)$$

[0095] Once again, various techniques (for example the nearest neighbour algorithm, methods using multi-layer perceptron neural networks, radial basis function neural networks, support vector machines, Bayesian machines, etc) are known for performing meta prediction and the present invention is not particularly limited with regard to which method is used to perform the meta prediction process. However, in preferred embodiments of the invention the meta prediction function applied by the meta prediction module 13 of expert unit EXP_R is a simple function according to which it is assumed that the prediction error for the currently-considered candidate action will be the same as the last actual prediction error observed for a training example in the region R for which this expert unit EXP_R is responsible.

[0096] Once again, the meta prediction function applied by the meta prediction module 13 of expert unit EXP_R can be tailored to the characteristics of the region R, such that the meta prediction function evolves over time whereby to increase in accuracy. Thus, even if the meta prediction modules 13 of all of the expert units initially apply the same prediction function (and, once again, this is not obligatory), after a time these functions will usually be different.

IPR2025-00944
Tesla EX1002 Page 964

[0097] Each expert unit, ***EXP***, also includes a knowledge-gain assessment module 14 which examines the value of the prediction error that is expected to arise for a given candidate action, $A_k(t)$, corresponding to a situation in the region, ***R***, covered by this expert unit. The knowledge-gain assessment module 14 compares the expected prediction error with actual prediction errors that have been observed on one or more earlier occasions when the setup *S* had a combination of system-state/context parameters and action variables located in this same region R. This actual prediction error data is stored in a memory 15 associated with this expert unit. The operation of a preferred embodiment of the knowledge-gain assessment module 14 will be described in more detail later on.

[0098] The knowledge-gain assessment module 14 outputs, to the action selector module 4, data indicating the knowledge gain (decrease in prediction error) expected to arise if this particular candidate action is performed, together with data indicating the candidate action for which this expected knowledge-gain has been assessed.

[0099] On a majority of occasions, the action-selection module 4 selects, as the next action to be performed on or in the setup, *S*, that candidate action which is expected to lead to the greatest knowledge gain (which, in this case, is the greatest expected decrease in prediction error).

[0100] For a given set of candidate actions, it is fairly likely that knowledge-gain assessment modules 14 of more than one expert unit will output expected knowledge-gain information to the action selection module 4. In a variant of the present embodiment, the knowledge-gain assessment module 3 may itself include a unit which calculates which candidate action is expected to lead to the greatest knowledge gain. In this variant, the knowledge-gain assessment module 3 outputs to the action selection module 4 an indication of which candidate action is expected to yield the greatest knowledge gain (learning progress).

[0101] The action selection module 4 outputs a signal indicating what the selected candidate action is. This signal causes the selected action to be performed (either automatically or with intervention, for example, by a human user). The action selection module 4 also indicates to a calculator 5 which candidate action was selected for performance. The calculator 5 also receives a signal indicating the region, ***R***, of system-state/context/action space corresponding to the selected candidate action. This enables the calculator to retrieve the prediction $\boldsymbol{VAR_R(t+1)}$ from the appropriate prediction module within the knowledge-gain assessment module 3.

[0102] After an action is selected by the action-selection module 4 and performed at time t (more precisely, time $t+\delta t$), the calculator 5 determines what is the actual value, $\boldsymbol{VAR_{Ract}(t+1)}$, of the predicted parameter(s) at the time $t+1$. It may be that the respective values of the predicted variables in the set ***VAR*** can be calculated from the values of the system-state/context parameters. In such a case, the calculator 5 receives as inputs the values of the system-state/context parameters at the time $t+1$. Alternatively, it may be necessary or convenient to input to the calculator 5 the actual values of the predicted variables in the set ***VAR*** that are observed at time $t+1$ as the outcome of performing the selected candidate action.

[0103] Based on the inputs it receives, the calculator 5 calculates what is the actual prediction error, $\boldsymbol{E_{Ract}(t)}$, that is, what is the difference between the set of predictions $\boldsymbol{VAR_R(t+1)}$ made by the prediction module 12 for this candidate action and the actual outcome $\boldsymbol{VAR_{Ract}(t+1)}$ that has been observed.

[0104] For each predicted variable in the set ***VAR***, the actual prediction error, $\boldsymbol{e_{Ract}(t)}$, can be determined, as follows:

$$e_{Ract}(t), = [\text{absolute value } \{VAR_R(t+1) - VAR_{Ract}(t+1)\}]^k \qquad (3)$$

where k is an integer greater than or equal to 1 (typically, equal to 2). The overall prediction error $\boldsymbol{E_{Ract}(t)}$ can be calculated by combining the actual prediction errors obtained for each predicted variable (for example, by summing $|VAR_R(t+1) - VAR_{Ract}(t+1)|^2$).

[0105] The calculator 5 supplies this overall actual prediction error data $\boldsymbol{E_{Ract}(t)}$ to the memory 15 associated with expert unit $\boldsymbol{EXP_R}$, so that it can be taken into account by this expert unit in later calculations.

[0106] Now it is helpful to replace the variable t by a corresponding value of the counting variable q (recalling that q is the variable counting the number of times the system has encountered a situation in region R of system-state/action/context space, in other words q counts the number of training examples observed in region ***R***). If q^t indicates the value of counting variable q at time t, then the actual prediction error observed at time t can be designated $E_{Ract}(q^t)$.

[0107] At time step t the system is likely to have already been in situations in region R of system-state/context/action space on a number of occasions. Thus, there will be a series of actual prediction error values, $E_{Ract}(1)$, $\boldsymbol{E_{Ract}(2)}$, $\boldsymbol{E_{Ract}(3)}$, ..., $\boldsymbol{E_{Ract}(q^t-1)}$, stored in the memory 15 associated with expert unit $\boldsymbol{EXP_R}$. It can be impractical to base calculations on the entire series of actual prediction error values observed for the region ***R***, so it is preferred to consider only the $N+1$ last observed values for this region. Moreover, in order to reduce the effects of noise, the average is determined of the current value and the N previous values of actual prediction error, as follows:

IPR2025-00944
Tesla EX1002 Page 965

$$Em_{Ract}(q^t) = \frac{1}{N+1}\{E_{Ract}(q^t) + E_{Ract}(q^t\text{-}1) + E_{Ract}(q^t\text{-}2) + \ldots + E_{Ract}(q^t\text{-}N)\} \quad (4)$$

It will be seen that the successive terms in the above equation, reading from right to left, correspond to successive occasions on which the system encountered situations in region ***R*** of system-state/context/action space. In general, these successive occasions will be spaced apart in time in an irregular manner.

[0108] Now, the knowledge-gain assessment module 14 of the expert unit EXP_R can define knowledge-gain (or "learning progress") achieved at time t, on the q^t-*th* occasion that the system 1 encounters a state in region ***R***, as follows:

$$LP_R(q^t) = Em_{Ract}(q^t) - Em_{Ract}(q^t\text{-}N) \quad (5)$$

[0109] Now, at the time when the knowledge-gain module 14 estimates the learning progress expected to be achieved because of performance of a given candidate action at time t, it does not yet know the actual prediction error that will result from performance of this action, so it cannot calculate the average value $Em_{Ract}(_qt)$. However, the knowledge-gain module 14 does have available to it the estimate $Ep_R(t)$ produced by the meta prediction module 13 of this same expert unit EXP_R. Accordingly, the knowledge-gain assessment module 14 uses this estimate $Ep_R(t)$ instead of $E_{Ract}(q^t)$ when calculating a value for $Em_{Ract}(q^t)$, and assesses expected learning progress by comparing the thus-calculated value of $Em_{Ract}(q^t)$ with $Em_{Ract}[(q^t\text{-}N)]$.

[0110] In the preferred embodiments of the first aspect of the present invention, the action selection module 4 is arranged so that on a majority of occasions it selects the candidate action for which the expected learning progress is the greatest. However, it is beneficial if, on some occasions, the action selection module 4 makes a random choice of the next action. This allows the action-selection system 1 to discover new regions in system-state/context/action space that are interesting from the point of view of yielding information as to the behaviour of the setup or system S.

[0111] In certain successful experiments (described below in relation to Figs. 10, 11 and 17), the action selection module 4 was arranged so that there was a probability of 0.85 that it would select the candidate action expected to yield the greatest learning progress, and a probability of 0.15 that it would select a random action. The present invention is not particularly limited with regard to the value of this latter probability. As long as the probability of selecting an action at random is greater than zero, the system 1 will discover new interesting regions of system-state/context/action space. However, if the value of this probability is equal to 1, then the system 1 will just be selecting actions at random. So, in the preferred embodiments of the invention, the probability, ***p***, that the action selection module 4 selects an action at random is $0<p<1$.

[0112] In some applications, the action selected by the action selector 4 will be automatically performed, under control of the action-selection system 1. In other cases, the action-selection system 1 may merely output signals indicating what are the desired values of the respective one or more action variables. In the latter case, the selected action is performed either by a human user or by another machine or system. The outcome of the action is fed back to the action-selection system 1, either in terms of the resulting values of the system-state/context parameters at some later time (e.g. *t*+1), or directly in terms of the actual values of the set of predicted variables ***VAR*** as observed at the later time (e.g. *t*+1). The action-selection system 1 associates this outcome data with data defining the performed action as well as with data defining the system description at the time when the action was performed. The resultant ensemble of data constitutes one training example.

[0113] As mentioned above, the training example data comprises:

- the performed action (that is, the values of the set of action variables defining the selected action performed at time t);
- the description of the setup/system S at time t (that is, the values of the set of system-state/context parameters at time t); and
- the outcome of the performed action (that is, the actual values of the set of predicted variables ***VAR*** at a predetermined later time, e.g. time *t*+1).

Incidentally, it is not essential for the feedback to be received by the system 1 at the time step subsequent to performance of the action. A number of time steps may intervene.

[0114] The training example data is supplied to the region identifier 2, as will be explained in greater detail below.

[0115] So far nothing has been said with regard to the way in which system-state/context/action space is divided up into the various regions. It is advantageous if this division of system-state/context/action space is performed by the

IPR2025-00944
Tesla EX1002 Page 966

region-identifier module 2 itself, progressively refining regions into smaller and smaller areas as it learns more about the different situations that can present themselves. According to the preferred embodiment of the first aspect of the present invention, this can be achieved as follows:

[0116] The action-selection unit 1 is arranged to perform memory-based learning. In other words, the action-selection system 1 is adapted to "remember" (store data on) each of the training examples encountered since the system 1 was last initialised, that is to log training example data for each time step. This training example data is made available to the region identifier module 2, and is associated with the particular region of system-state/context/action state which contains the end-point of the vector defined by this training example's values of system-state/context parameters and action variables.

[0117] Initially the region identifier 2 has no information as to which situations are similar to each other; the whole of system-state/context/action space is considered to be a single region and there is a corresponding single expert unit in the knowledge-gain assessment module 3. As the action-selection system 1 operates it will build up a set of training examples corresponding to values (system-state/context parameter values, actual values of the predicted variables, etc.) encountered at each time step. The example data is allocated to the single existing expert unit which handles predictions for the single existing region of system-state/context/action space.

[0118] As data is gathered on more and more training examples, the region identifier 2 applies a criterion ***C1*** in order to decide whether or not to spit the initial single region of system-state/context/action space into two or more new regions. If the initial region is split into new regions, the original expert unit is split into a plurality of expert units - one for each of the new regions - (or new expert units are created), and the training example data that has already been collected is re-allocated to the expert unit associated with the relevant new region (that is, the new region which contains the combination of system-state/context parameter values and action variable values applicable at the time of this training example).

[0119] According to preferred embodiments of the first aspect of the invention, the criterion ***C1*** used for deciding whether or not to divide up a region is a simple judgement based on the number of examples encountered in the original region. Notably, when the number of training examples encountered for a given region of system-state/context/action space has exceeded a predetermined number, ***NS***, a decision is taken to partition the initial region into a plurality of new regions (for example, two new regions). In the experiments discussed below with reference to Figs. 10, 11 and 17, the predetermined number ***NS*** was 250.

[0120] At the time of splitting a region, a criterion ***C2*** determines how the original region is split, that is, what positions should be chosen for the boundaries of the new regions. In preferred embodiments according to the first aspect of the invention, the following criterion ***C2*** is used: when a region ***R*** is split into new regions (for example, two new regions $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$) the location of the boundary is set so that the example data that had already been collected for region R is split in a balanced fashion between the new regions and with a minimal variance. That is, after the split each new region will have inherited a number of training examples which is similar to the number of training examples inherited by the other new regions, and the variance of the example data will be minimized. This variance can be minimized within each new region of system-state/context/action space, or within "outcome space" (that is, within a space defined by the actual observed values of the set of predicted variables, ***VAR***), or within a higher-dimensional system-state/context/action/outcome space defined by the values of the sets of system-state/context parameters, action variables and predicted variables. This condition can be expressed, as follows:

$$\text{minimize } [\{\text{card } (\boldsymbol{R}1) \text{ x variance } (\boldsymbol{R}1)\} + \{\text{card } (\boldsymbol{R}2) \text{ x variance } (\boldsymbol{R}2)\}]$$

[0121] According to the preferred embodiments of the first aspect of the present invention, when a region of system-state/context/action space is split in this way, the process of determining the location of the boundaries consists in first deciding which dimension (which parameter or variable, or combination thereof) to cut and then deciding the location of the (or each) cut point. The boundaries can be defined in simple terms, e.g. the first new region will contain examples in which a variable $\boldsymbol{V}_1$ takes a value 0 to 0.5, the second new region will contain examples in which variable $\boldsymbol{V}_1$ takes a value 0.51 to 1.0. The region identifier 2 maintains a list of cut dimensions and cut values.

[0122] The above-described region-splitting process is recursively iterated. When the criterion ***C1*** is met for any of the regions recognised by the region identifier 2, that region in its turn is split into new regions. After a while, a tree structure develops, with the initial region (and its expert unit) forming the root, the first level of division of the root region corresponds to a branch which in its turn may be divided into further branches, ending in regions (and expert units) which can be considered to represent "leaves" of the tree structure. This tree structure corresponds to a segmentation of the input space, with each expert unit being a specialist for its own part of the system-state/context/action space.

[0123] The action-selection unit 1 continues to operate and, at each time step, the region identifier 2 determines what "leaf" region contains the combination of a given candidate action and the current values of system-state/context parameters. This determination is performed by comparing the candidate action variable values and the current values of

IPR2025-00944
Tesla EX1002 Page 967

the system-state/context parameters with the cut data defining the various regions of the tree structure. The expert unit applicable for this leaf region then makes its predictions. When the expected-knowledge gain has been assessed for each of the candidate actions, one candidate action is selected for performance. When the actual result is available, data indicative of the actual values of the set of predicted variables (***VAR***) is stored, together with other data for this training example, in the memory of the applicable leaf expert unit.

[0124] When the action-selection system 1 is used to select a series of actions, after some time it will build up a decision-tree architecture which reflects substantially all of the relationships that exist between the system-state/context parameters and action variables of the system, *S*. There will be a levelling-off in the prediction errors observed when predictions are made by the various prediction modules 12 in the expert units. This evolution of the prediction error values can be used as a sign that the learning process has succeeded in acquiring the desired knowledge about the setup S. Accordingly, the series of actions can be stopped. The action-selection system of the invention can include a control module (not shown) adapted to monitor the evolution of the prediction error values in the different expert units whereby to trigger a halt in the operation of the action-selection system 1 when a stop criterion is satisfied (e.g. when the percentage change in actual prediction error from one training example to the next has fallen below a threshold value for every "leaf" region defined by the region identifier 2).

[0125] By using the action-selection mechanism described above, the preferred embodiment of the first aspect of the present invention avoids transition effects: that is, when switching from an unpredictable situation to a highly predictable situation the assessment of expected knowledge gain will be based upon the previous actions in the unpredictable situation not merely upon the previous actions in time.

[0126] Moreover, by progressively refining the definitions of the regions it recognises in system-state/context/action space, the action-selection unit of the preferred embodiments of the present invention ensures that the learning process will first master the easy parts of the mapping to be learned, then progressively move on to the more difficult parts of the mapping. Not all parts of system-state/context/action space will be explored in detail, only those parts which are determined to be interesting.

[0127] By combining region-identification based on a "decision tree" architecture with predictions based on the nearest neighbour algorithm the action-selection unit of the preferred embodiment of the invention is an extremely efficient learning approach. Learning is maximally fast and prediction is accurate. In particular, the prediction is extremely fast compared with plain nearest-neighbour prediction: for example, with one million examples and ***NS***=250, a plain nearest neighbours algorithm would involve one million computations, whereas the technique of the preferred embodiment of the first aspect of the present invention would involve only 250 computations (i.e. it is 4000 times faster). Moreover the technique according to the preferred embodiment of the first aspect of the invention gives more accurate results than plain nearest-neighbour prediction because the region-splitting reduces the over-fitting which characterizes the nearest-neighbours prediction algorithm.

THE SECOND ASPECT OF THE INVENTION

[0128] When the action-selection system 1 of the first aspect of the present invention has been run so as to select a series of actions, and the prediction errors in the various prediction modules have all reached plateau values, it can be considered that the action-selection system 1 has become an expert on the system or setup S. More particularly, the prediction modules of the various expert units will tend to produce fairly accurate predictions of the effect of different candidate actions in different setup conditions (that is in for different values of the system-state/context parameters).

[0129] In these circumstances, the combination of the region identifier 2 and the prediction modules 12 in the knowledge-gain assessment module 3 can be considered to constitute a prediction machine that has been trained to make accurate predictions with regard to the behaviour of the setup S. More particularly, the combination of the region identifier 2 and the prediction modules 12 can predict the respective values of the set of one or more variables, ***VAR***, for different combinations of values of the action variables and system-state/context parameters.

[0130] Thus, according to a second aspect thereof, the present invention provides a method of training a prediction machine by causing it to form part of an action-selection system according to the first aspect of the invention and causing that system to select a series of actions until prediction errors level off in the prediction modules. The second aspect of the invention also provides a prediction-machine training system implementing this method.

[0131] Fig.5 shows, schematically, the main components of a preferred embodiment of prediction-machine training system 11 according to a preferred embodiment of this second aspect of the invention. It will be noted that this prediction-machine training system 11 has a similar construction and functioning to that of the action-selection system 1 that is illustrated schematically in Fig.4. Accordingly, the same reference numerals are used in Fig.5 to designate elements having a similar function to those in the action-selection system 1 of Fig.4.

[0132] Once again, it is to be understood that this prediction-machine training system 11 would probably be implemented in software and the various "modules" shown in Fig.5 are provided for illustrative purposes only. The various functions performed in the prediction-machine training system 11 may be split between modules in different ways from

IPR2025-00944
Tesla EX1002 Page 968

that shown in Fig.5 and/or a greater or lesser number of modules can be used to perform these functions.

[0133] The prediction machine training system 11 shown in Fig.5 includes a region identifier 2, knowledge-gain assessment module 3, action selector 4 and prediction error calculator 5 comparable to those used in the experiment-selection system 1 of Fig.4. As indicated above, it is the combination of the region identifier 2 and the prediction modules 12 of the knowledge-gain assessment module 3 which constitutes the prediction machine to be trained by this training system 11.

[0134] The prediction machine training system 11 operates in two modes and has an operation mode setting module 16 which controls the different components of the system 11 dependent upon the mode in which the system is currently operating. In the first mode of operation the prediction machine is being trained and, in the second mode of operation, the prediction machine is already trained and is making predictions.

[0135] When the prediction-machine training system 11 is in the first operating mode, the region identifier 2, knowledge-gain assessment module 3 (and its expert units), action selector 4 and calculator 5 operate substantially as described above in relation to Fig.4. In other words, a sequence of actions affecting a setup S is selected and performed, and training example data is accumulated regarding each performed action, the setup's description (state and context) at the time the action was performed and the outcome of the action. This training example data is used by the region identifier 2 to progressively build-up a decision-tree architecture based on dividing system-state/context/action space into finer and finer regions, corresponding to the different types of situation inherent in the setup *S*.

[0136] The training example data for each region is also used to adapt the prediction function applied by the prediction module 12 of the expert unit assigned to handle that region in the knowledge-gain assessment module 3. In other words, as the sequence of actions is selected and performed, the functions applied by the different prediction modules 12 in the knowledge-gain assessment module 3 become more and more closely adapted to the properties/behaviour of a respective region in system-state/context/action space that corresponds to a certain type of situation in the setup S.

[0137] The prediction modules 12 can be considered to be trained when the actual prediction errors made by the prediction modules 12 stop falling (that is, they reach plateau values). This can be detected by a module (not shown) that monitors the evolution of the actual prediction errors for each prediction module. When this situation is detected, the operation mode setter 16 switches over from the first operating mode to the second operating mode of the prediction-machine training system 11.

[0138] In the second mode, the operation mode setter 16 sends signals to the knowledge-gain assessment module 3 and action selector 4 so as to disable certain of their functions. More specifically, when the prediction-machine training system 11 is in the second mode of operation, the user inputs data identifying a possible action that could be performed, that is, specifying values for the set of action variables. The user also inputs data describing the setup's state and context, more particularly data specifying respective values for the set of one or more system-state/context parameters.

[0139] The region identifier 2 identifies which region of system-state/context/action space contains the end-point of the vector defined by this set of action variable and system-state/context parameter values, and activates the prediction module 12 of the expert unit which handles this region. The activated prediction module outputs its prediction of the values of the set of one or more predicted variables, VAR, to a prediction-outputting interface 18, which supplies them to the user.

[0140] In some applications, the user may use the trained prediction machine to make a prediction of the outcome of a particular action, perform that action, and feed back to the system 11 the outcome of the performed action. In such a case, the fed-back data can be used to generate a further training example which can be used to further adapt the prediction function applied by the appropriate prediction module.

EXAMPLES OF APPLICATIONS OF THE FIRST AND SECOND ASPECTS OF THE PRESENT INVENTION IN DIFFERENT FIELDS

[0141] The action-selection/prediction-machine-training technique of the present invention can be applied in a wide variety of fields. A few examples of applications in different fields will be outlined below. In each of the examples below it is assumed that an action-selection system (prediction-machine training system) as described above in relation to Fig. 4 (Fig.5) is used.

Example 1 - Biological Field

[0142] In this particular example, a human user wishes to find out how a certain organ of an organism reacts when injected with a mixture of chemicals. The mixture of chemicals can be defined by specifying the concentration of each chemical ($C_1, C_2, C_3, ..., C_n$) and the reaction of the studied organ can be measured by determining what is the concentration (CC_{30}) of cells of a certain type in that organ thirty minutes after injection of the mixture of chemicals.

[0143] Each trial or action involves injection of chemicals into the organ and a consequent reaction which could entail damage to or destruction of that organ. Time and costs are also involved in performing each trial. Clearly it is desirable

IPR2025-00944
Tesla EX1002 Page 969

to be able to minimize the number of trials that are performed.

[0144] The user can ensure that knowledge regarding the reaction of the organism to chemical mixtures can be acquired in an efficient manner, thereby reducing the number of trials that are required, by making use of an action-selection system according to a preferred embodiment of the present invention, such as that shown in Fig.4.

[0145] Preferably, the user employs a fully-automated assembly comprising the action-selection system of the present invention, together with a set of one or more automated robots which performs the selected actions and provides feedback to the action-selection system.

[0146] An example of such a fully-automated system is shown, schematically, in Fig.6. In the case illustrated in Fig. 6, an action-selection system according to the first aspect of the present invention (e.g. as shown in Fig.4) is used to select between different candidate actions that are possible at a given time. The selected action is output to an automatic laboratory which includes a set of experimenter robots. The selected action is performed automatically and feedback on the outcome is fed from the automatic laboratory back to the action-selection system.

[0147] More particularly, in this example the action-selection system 1 may be configured to consider the concentration values C_1 to C_n as action variables. In this example, the action-selection system will be configured such that its prediction module(s) seek to predict what CC_{30} will be, that is what will be the concentration of cells of the specified type in the organ thirty minutes after injection of a particular mixture of chemicals.

[0148] In this particular example, there are no system-state/context parameters to be taken into consideration. Thus, in this case the region identifier 2 and the actual prediction error calculator 5 do not have system-state/context parameter inputs from the setup S under investigation, the only input they receive from the setup S is the actual value of CC_{30} at a given time. However, if desired, the user could have configured the action-selection system 1 so that it recognises as system-state/context parameters (and monitors) the temperature, blood pressure etc. of the organism containing the organ under test.

[0149] The relevant prediction modules 12 of the action-selection system predict CC_{30} for each of a number of candidate actions, that is, for each of a number of different possible combinations of values of the action variables ***C1***, ***C2***, ..., ***Cn***. In the fully-automated assembly, the different sets of candidate actions that are considered may be generated automatically within the action-selection system 1 itself, or by a separate candidate-action generation module, or by the human user.

[0150] The candidate action to be taken is chosen (on a majority of occasions) based on the result of an evaluation by the action selector 4 as to which candidate action is expected to yield the greatest increase in knowledge compared to the last occasion on which the system had a combination of values of C_1 to C_n that was in the same region of system-state/context/action space as the combination specified in the candidate action. On a minority of occasions, a random action is selected. Output is given as to which mixture of chemicals has been selected. This output is supplied to the set of automated robots and these mix up the selected mixture of chemicals and inject it into the organ under investigation.

[0151] After thirty minutes, one of the automated robots measures the concentration (CC_{30}) of the specified cells in the organ. The automated robot supplies the measured concentration (CC_{30}) to the region identifier 2 and actual prediction error calculator 5 of the action-selection system 1 as feedback, enabling the system 1 to determine what was the actual error on the prediction made by the prediction module for this candidate action (this mixture of chemicals).

[0152] Initially the action-selection system 1 will recognise only a single region in system-state/context/action space. However, as more and more actions are performed, new expert units will be created and a decision-tree structure will develop sub-dividing the system-state/context/action space (which, in this case, reduces to "action space") into different regions. The structure of this decision tree will reflect the relationships that hold between the concentrations of the various chemicals injected into the organ and the reaction of the organ to those chemicals (to the extent that any such relationships exist).

[0153] When each relationship has been learned by a prediction module 12 of the action-selection system 1 according to the invention operating on the relevant region of system-state/context/action space, the prediction errors made by the respective prediction module 12 will level off at a plateau value. The action-selection system 1, or the user, can monitor the evolution of the prediction error values so as to determine when the prediction modules 12 have, in effect, learned all that they can. At this point the series of actions can be stopped.

[0154] At the end of the series of actions the action-selection system 1 of the invention has developed into a prediction machine that is trained to predict how the studied organ will react to different mixtures of chemicals. If the user is interested in knowing what will be the effect of administering a particular mixture of chemicals to the studied organ, he can simply input the appropriate combination of values of C_1 to C_n to the action-selection system 1, as if it were the only candidate action, and then sample the output from the appropriate prediction module - this will be a prediction of the resultant CC_{30}.

[0155] As an alternative to using the fully-automated assembly described above, the user could make use only of the action-selection system 1 of the invention and himself mix-up and administer the selected combination of chemicals and provide feedback to the action-selection system. The human user may also generate the various sets of candidate actions to be evaluated by the action-selection system 1 (although this can be done within the action-selection system 1 itself or by a separate candidate-action generation module).

IPR2025-00944
Tesla EX1002 Page 970

[0156] It will be seen that, in this example of an application in the field of biology, the first aspect of the invention has provided the user with a machine which selects the actions in a sequence which enable the user to learn the effect of injecting particular combinations of drugs into an organ. The first aspect of the invention also provides a fully-automated assembly in which the action-selection system 1 determines which actions should be performed and a set of one or more automated robots performs the selected actions and provides feedback on the outcomes.

[0157] Moreover, in this example of an application in the field of biology, if the action-selection system of the invention were to be replaced by a prediction-machine training system such as that shown in Fig.5, the second aspect of the invention would provide the user with a prediction-machine training system which, in an interactive process, trains a prediction machine so that it can predict the effect of injecting particular combinations of compounds into an organ. As a result of this interactive process, the user is also provided with a prediction machine that is trained to make predictions of what will be the concentration of cells of a particular type in an organ 30 minutes after injection of different mixtures of compounds. This trained prediction machine could, for example, take the form of a computer system or an expert system.

Example 2 - Agricultural Field

[0158] Next consider an example in which a biotechnology firm is designing a particular species of yeast intended for sale to farmers. Yeast can be cultivated under different conditions, for example, using a greater or lesser quantity of water, fertilising chemicals and pesticides. The quality and quantity of the adult yeast will vary depending on the settings of these parameters as well as based on the agricultural context (for example, the type of soil, amount of sun, etc.). Different farmers may have different priorities: one wishing to use few chemicals whilst still assuring sufficient yield, another wishing to maximise the yield of yeast even if large quantities of chemicals are required.

[0159] The biotechnology firm wishes to study the way in which the amounts of water, fertilising chemicals and pesticides used during cultivation of its yeast affect the final yield, in different agricultural contexts.

[0160] The quantity of water, fertilising chemicals and pesticide applied per week during cultivation of the yeast can be specified using parameters ***W***, ***F*** and ***P***, respectively, the soil quality can be specified using a number of parameters $\boldsymbol{SQ_1}$, $\boldsymbol{SQ_2}$, etc., and the yield can be specified using a parameter ***Y*** indicating the number of kilograms produced per hectare a given number of weeks after planting.

[0161] Each action or trial involves cultivating a crop of the biotechnology company's yeast using water, fertilising chemicals and pesticide. There are labour costs, costs of water/chemicals, and extended periods of time required for performing each trial. Thus, it is desirable for the biotechnology company to be able to minimize the number of trials that are performed.

[0162] In this example, the action-selection system according to a preferred embodiment of the invention, such as that of Fig.4, may be configured to recognise the parameters ***W***, ***F*** and ***P*** as action variables, and to input as system-state/context parameters the values SQ_i defining the nature of the soil (e.g. in terms of its nitrogen content, clay content, etc.).

[0163] In this example, the action-selection system 1 will be configured such that its prediction module(s) 12 seek to predict what will be the yield ***Y*** of yeast at the end of a specified time period from planting.

[0164] The action-selection system 1 according to the preferred embodiment of the invention will predict ***Y*** for each of a number of candidate actions, that is, for each of a number of different possible combinations of applied quantities of water, fertilising chemicals and pesticides (different combinations of values of action variables ***W***, ***F***, ***P***), taking into account the current values of the system-state/context parameters $\boldsymbol{SQ_i}$.

[0165] Once again, on a majority of occasions the candidate action that is selected by the action selector 4 for performance will be the one that is expected to produce the greatest increase in knowledge compared to the last occasion on which there was a combination of values for **W**, ***F, P*** and $\boldsymbol{SQ_i}$ that was in the same region of system-state/context/action space as the combination of current $\boldsymbol{SQ_i}$ values with the ***W***, ***F*** and ***P*** values specified in the candidate action. On a minority of occasions a random action will be selected. Output is given as to which combination of values of ***W***, ***F, P*** has been selected.

[0166] The biotechnology company receives the output from the action-selection system 1 and may then cultivate a crop of yeast under the specified conditions. After the requisite time period, the yield ***Y*** is determined. The measured yield ***Y*** is input to the action-selection system 1 as feedback, enabling the system's calculator 5 to determine what was the actual error on the prediction made by the relevant prediction module 12 for this candidate action (this set of cultivation conditions), taking into account the values $\boldsymbol{SQ_i}$ at the time of the prediction.

[0167] As in the first example, as the series of trials progresses the action-selection system 1 will develop expert units which sub-divide system-state/context/action space, developing a decision-tree structure which reflects the relationships inherent in the system.

[0168] At the end of the series of trials, the combination of the region identifier 2 and the prediction modules 12 of the various expert units of the action-selection system 1 has developed into a prediction machine that is trained to predict

IPR2025-00944
Tesla EX1002 Page 971

how the yield of the biotechnology company's yeast is affected by different cultivation conditions. This prediction machine can be used by the biotechnology company, or supplied to farmers along with the biotechnology company's yeast, so as to enable them to evaluate the yield likely to result from use of different combinations of cultivation conditions, given their soil conditions.

[0169] It will be seen that, in this example of an application in the field of agriculture, the first aspect of the invention has provided the biotechnology company with a machine which selects the actions in a sequence of actions which enables the company to learn the effect of different cultivation conditions on the yield of its new yeast.

[0170] Moreover, in this example of an application in the field of agriculture, the second aspect of the invention provides the biotechnology company with a prediction-machine training system which, in an interactive process, trains a prediction machine so that it can predict the effect of different cultivation conditions on the yield of yeast, taking into account different soil conditions. As a result of this interactive process, the biotechnology company/farmer is provided with a prediction machine (e.g. in the form of a computer system or expert system) that is trained to make predictions of the yield of yeast for different combinations of soil and cultivation conditions.

Example 3 - Field of Industrial Production/Manufacturing

[0171] This example concerns the case of a company manufacturing combustion engines or motors. Different compositions of fuel can be used to power these motors. However, the precise choice of composition will affect the functioning of the motor: some compositions may be cheap but cause the motor to deteriorate, cutting short its lifespan; certain other compositions may extend the lifespan of the engine but at the cost of increased energy consumption. The effect of the fuel composition on the engine's performance/lifespan may be particularly critical in the case of large engines used in factories and the like. It is desirable to know how variation of the fuel composition affects variables such as the lifespan of the engine, the fuel consumption, etc.

[0172] Clearly, tests involving determination of factors such as an engine's lifespan for different compositions of fuel can be extremely costly. Once again, it is desirable to be able to reduce the number of tests that must be performed in order to learn the effect of different fuel compositions on the engine and its performance.

[0173] The action-selection technique of the present invention can be used in order to determine the sequence of actions which should be performed so as to learn efficiently about the effect of different fuel compositions on the engine's performance. More particularly, below there is described the case where the engine manufacturer uses an action-selection system 1 as shown in Fig.4 to select actions in a sequence that will investigate the performance of his engine when supplied with different fuel compositions.

[0174] In this example, the percentage by weight of different components ***a, b,*** **...m,** in the fuel composition can be specified using parameters C_a, C_b, ... C_m, respectively. The amount of fuel consumed by the engine per hour of operation can be specified using a parameter ***FC***, the power output rating of the engine can be specified using a parameter ***P***, and the lifetime of the engine can be specified using a parameter, ***LT***.

[0175] In this example, the action-selection system 1 according to a preferred embodiment of the invention may be configured to recognise the parameters C_a, C_b, ... C_m as action variables. In this example, the action-selection system 1 will be configured such that its prediction module(s) 12 seek to predict what will be the useful lifespan of the engine ***(LT)***, the engine's fuel consumption (***FC***) and the power rating of the engine (***P***). In other words, in this example the set of predicted variables ***VAR*** includes three variables. This is another example in which no system-state/context parameters are defined for the action-selection system 1.

[0176] The action-selection system 1 according to the preferred embodiment of the invention will predict ***LT, FC*** and ***P*** for each of a number of candidate actions, that is, for each of a number of different possible fuel compositions.

[0177] Once again, on a majority of occasions the candidate action that the action selector 4 selects for performance, based on the output from the relevant knowledge-gain assessment modules 14, will be the one that is expected to produce the greatest increase in knowledge compared to the last occasion on which the system had a combination of values of C_a, C_b, ..., and C_m that are in the same region of system-state/context/action space as the combination specified in the candidate action. On a minority of occasions, a random action will be selected. Output is given as to which combination of values of C_a, C_b, ... C_m has been selected.

[0178] The engine-manufacturing company receives the output from the action-selection system 1, and may then operate an engine under the specified conditions and note the useful lifetime of the engine, the fuel consumption and the power output when so operated. The measured data ***LT, FC*** and ***P*** is input to the action-selection system as feedback (actual values of ***VAR***), enabling the calculator 5 of the action-selection system 1 to determine what was the actual error on the prediction made by the relevant prediction module 12 for this candidate action (this fuel composition).

[0179] As in the above-described first and second examples, as the series of tests progresses the action-selection system 1 will create new expert units (with respective prediction modules) which sub-divide system-state/context/action space according to a tree-like structure which reflects the relationships that exist between the different components of the fuel composition, the power rating of the engine, the fuel consumption of the engine and the lifetime of the engine.

IPR2025-00944
Tesla EX1002 Page 972

[0180] At the end of the series of tests, the combination of the region identifier 2 and the prediction modules 12 of the action-selection system 1 has developed into a prediction machine that is trained to predict how the lifetime of the engine, its fuel consumption and power output are affected by the composition of the fuel supplied to it. When selling an engine, the manufacturer can supply this trained prediction machine to the end-user, enabling that user to evaluate the engine's performance for different possible fuel compositions.

[0181] It will be seen that, in this example of an application in the field of industrial production, the first aspect of the invention has provided the engine manufacturer with a machine which selects the actions in a sequence which enables the company to learn the effect of different fuel compositions on the performance of its engine.

[0182] Moreover, in this example of an application in the field of industrial production, the second aspect of the invention provides the engine manufacturer with a prediction-machine training system which, in an interactive process, trains a prediction machine so that it can predict the effect of different fuel compositions on the engine's performance. As a result of this interactive process, the engine manufacturer/engine user is provided with a prediction machine (e.g. a computer system or expert system) that is trained to make predictions of the engine's fuel consumption, power output and lifetime when using different fuel compositions.

Example 4 - Financial Field

[0183] Consider an example in which a person wishes to be able to predict what will be the value of a particular financial object (e.g. a building or other asset; shares, a warrant, an option or other financial instrument; etc.) thirty days after purchase.

[0184] The value of the financial object will be affected by prevailing economic conditions; these can be considered to be the system's context and can be described using context parameters such as the end-of-day level of the local stock market index (***SMI***), the current interest rate **(*IR*)**, etc.

[0185] The value of the financial object will often also be affected by a number of parameters defining that object. This can be understood by considering an example such as a (call) share option. This share option gives its owner the right to buy a certain quantity of shares on a particular day at a particular price, known as the "strike price". (An option to sell shares would be known as a "put" option). The share option itself can be traded and so will have a price; this is affected by the time left until the date on which the option can be exercised as well as the difference between the strike price and the actual market price for the same shares. On any given day it is usually possible to buy different options on the same shares, notably options having different strike prices and different exercise dates. The actual share price at the time of purchase of the option, and the price of the share option, can be considered to define the "state" of this kind of financial object and could be described using respective state parameters CP (current price of shares on date of purchase of option) and OP (current price of option).

[0186] In order to find out how the value of a call share option 30 days after purchase varies with economic conditions and the "state" of the option itself, it would be necessary to buy a number of options having different strike prices and exercise dates, in circumstances where the economic conditions (including the underlying share price) vary. Clearly each such purchase involves costs and, potentially, financial loss. Thus, it is desirable to be able to discover the relationships inherent in this system in an efficient manner. Use of the action-selection system according to the present invention enables this to be achieved.

[0187] Preferably, the user employs a fully-automated arrangement in which the action-selection system of the invention is adapted to be able to access online financial databases, electronic marketplaces and/or electronic trading systems (e.g. by connecting itself to websites), so as to be able to gather for itself the current values of the system-state/context parameter values, to initiate share option purchases and to obtain feedback on the outcome of those purchases.

[0188] An example of such a fully-automated system is shown, schematically, in Fig.7. In the case illustrated in Fig. 7, an action-selection system according to the first aspect of the invention is capable of accessing Internet financial databases (for example the NASDAQ website) in order to obtain current values for various economic indicators and for parameters defining potential products that could be purchased. The action-selection system is also capable of making purchases in the international financial marketplace (e.g. by making use of an electronic trading system or online trading service) and of obtaining feedback from that marketplace, automatically, with regard to the outcome of those purchases.

[0189] In this particular example, the action-selection system 1 shown in Fig.4 may be used in the fully-automated assembly of Fig. 7 and may be configured to recognise the end-of-day level of the local stock market index (***SMI***), the current interest rate (***IR***), the current share price (***PP***) and current option price (***OP***) as system-state/context parameters. The share option exercise date (***ED***) or time until exercise (***TE***), as well as the strike price (***SP***) are action variables.

[0190] The region identifier 2 and actual prediction error calculator 5 receive as inputs the values of the end-of-day level of the local stock market index (***SMI***), the current interest rate (***IR***), the current share price *(**PP**)* and current option price (***OP***) at a given time, t, when a purchase is going to be made. The prediction module(s) 12 of the action-selection system are configured to predict the value of the share option 30 days after purchase (based on the assumption that the values of the various system-state/context parameters that are input at time ***t*** will still hold when the actual purchase

IPR2025-00944
Tesla EX1002 Page 973

is made). The latter assumption is reasonable, given that the action-selection system 1 can select an action, and output data identifying that action, within a few milliseconds, and economic variables such as share prices vary much more slowly than that.

[0191] In this example the action-selection system 1 of the invention may be configured to process action variables ED and SP, corresponding to tuning of the exercise date and strike price parameters.

[0192] Once the action-selection system 1 has been configured, it accesses financial databases or the like so as to input details of the current values of the stock market index, ***SMI***, the official central bank interest rate, ***IR***, and the current share price, ***PP***. The action-selection system 1 may itself include a candidate action generator so as to generate possible combinations of values for the exercise date and strike price, ***ED*** and ***SP***. Alternatively, the different sets of candidate actions may be generated by a separate module or by the user. The action-selection system 1 accesses finanacial databases or the like in order to determine the prices, ***OP***, of call share options having the combinations of exercise dates and strike prices specified in the candidate actions.

[0193] For each of these different potential call-share-option purchases, the action-selection system 1 of the invention seeks to predict what will be the value of the bought option in 30 days time, taking into account the input values for the economic indicators (***SMI*** and ***IR***), the current share price (***PP***), the current price of the option (***OP***), and the time until exercise of the option (***TE***) which the system can calculate for itself, provided that it knows the current date.

[0194] As an alternative to using a fully-automated arrangement, the user could make use of the action-selection system 1 only to select actions, and himself input the values of the various system-state/context parameters, make the selected purchases, and provide feedback as to the outcomes.

[0195] The action selector 4 of the action-selection system will select which action (purchase) should be tried next. In general, the selected action will be the one which is expected to produce the greatest increase in knowledge (assessed "locally" in system-state/context/action space), although random actions will be selected with probability ***p***. The share option having the selected combination of strike price and exercise date can then be bought automatically for example via an electronic trading system (or by the user). The value of the purchased share option after 30 days is noted and fed back to the action-selection system automatically, from internet financial databases for example, or manually by the user.

[0196] Initially, the action-selection system 1 of the invention only recognises one region in context/system-state/action space and so all predictions are made a common prediction module. However, as the purchases suggested by the system are made, and feedback is received on the actual value of the purchased share options 30 days after purchase, the system will begin to divide up the space into different regions, creating new expert units (and corresponding new prediction modules 12) as required.

[0197] After a certain number of purchases have been made, provided that there is some non-random element in the system, the action-selection mechanism 1 of the invention will have developed expert units having prediction modules 12 that are tailored to the different situations that can arise (different combinations of economic indicators, share price, option parameters, etc.). The series of actions (purchases) can then be stopped and the user has available to him a machine which can predict what will be the value in thirty days time of share options defined by different combinations of strike price and exercise date, taking into account the current economic context and share price. He can compare these predicted values with the respective current prices of the share options in order to decide which purchases are more likely to provide financial gain.

[0198] The trained action-selection system 1 could be sold to third parties, as a trained option-value prediction machine. If the third party acts on the suggestions made to it by the trained prediction machine, he can feedback the results of his purchases (i.e. actual option value after 30 days) so as to provide the prediction machine with further examples to process. In other words, even after the prediction machine has been trained it is still possible to generate further "training examples" and refine the operation of the prediction machine yet further.

[0199] Incidentally, in this particular application it is not essential for the automated system (or the user) to actually make the selected purchases (perform the selected actions). The price of the selected share-option can be determined 30 days after it would have been purchased, and this price fed back to the action-selection system 1. It is true that the actual purchase of the share-option could have had an effect on the market value quoted for options having the same specification. However, unless an extremely large purchase was intended (or the market in shares of this type is extremely small) the absence of an actual purchase will no have a significant effect on the market value for this option in 30 days time.

[0200] It will be seen that, in this example of an application in the financial field, the first aspect of the invention provides a user with a machine which selects the actions in a sequence which enables the user to learn the effect of making different purchases in different conditions. The first aspect of the invention also provides a fully-automated arrangement in which the action-selection system 1 obtains for itself the system-state/context parameter values it requires (for example, by connecting itself to on-line financial databases), determines which actions should be performed, optionally performs them (by connecting itself to electronic trading systems), and obtains feedback on the outcomes (for example, by once again connecting itself to online financial databases). In this fully-automated arrangement the action-selection system 1 could be considered to comprise a software robot (or "softbot") making the required connections to websites/trading

IPR2025-00944
Tesla EX1002 Page 974

systems.

[0201] Moreover, in this example of an application in the financial field, the second aspect of the invention provides the user with a prediction-machine training system which, in an interactive process, trains a prediction machine so that it can predict the effect of making different purchase in different conditions. As a result of this interactive process, the user is provided with a prediction machine (e.g. in the form of a computer system or expert system) that is trained to make predictions of the future value of a particular financial instrument given different combinations of conditions at the time of purchase.

THE THIRD ASPECT OF THE INVENTION

[0202] According to the third aspect thereof, the present invention provides a technique and system for managing learning by, or organizing behaviour of, a self-developing device or system, as well as a self-developing device or system making use of such a technique.

[0203] Some types of self-developing devices are robotic devices, or other devices which have actuators and sensors enabling them to interact with their environment. According to the third aspect of the invention, it is possible to manage learning by such a self-developing device, and/or organize its behaviour, by treating that self-developing device interacting with its environment as a system ***S***, and investigating its behaviour in a series of actions that are chosen in an autonomous fashion according to the action-selection technique of the first aspect of the invention. In general, the system-state/context parameters of such a system are the signals received by the self-developing device's sensors and the action variables will be the setting of the self-developing device's actuators. In this case, the system-state/context/action space equates to the self-developing device's sensory-motor state-space.

[0204] However, it is to be understood that the third aspect of the present invention is not limited to application with single self-developing devices which have sensors and actuators.

[0205] More particularly, the present invention is also applicable to systems which contain distinct separate components or entities working together, in which case the selected action may be the setting of respective values for action variables of one or more of the system components/entities.

[0206] Furthermore, the third aspect of the present invention is applicable to self-developing devices or systems which can be considered to be passive, that is, which do not have physical actuators and may not even have sensors. For example, the present invention can be applied to a self-developing computer system which receives or generates certain inputs, and produces outputs, thereby interacting with an environment. According to the third aspect of the invention, it is possible to manage the learning of such a self-developing system, or organize its behaviour, by treating that self-developing computer interacting with its environment as a setup or system S, and investigating the behaviour of this setup in a series of actions that are chosen in an autonomous fashion according to the action-selection technique of the first aspect of the invention. The system-state/context parameters in this case will generally be the inputs received or generated by the self-developing computer, and the set of action variables will be one or more variables whose value (s) are set based on one or more of the outputs from the self-developing computer.

[0207] For example, for a self-developing system constituted by a computer, the system-state/context parameters could include: data indicating the state of certain internal elements of the computer (e.g. status of flags, full/empty status of buffers, existence of a state of overflow, etc.), data input by a human operator, data indicating the state of a context or environment with which the computer system interacts, etc. This latter environment could be physical, for example in a case where the computer system controls a physical device or system, or it could be non-physical, for example in a case where the computer system is trying to teach itself what patterns there are in the data it is processing. In either case, management of learning by the computer system, or organization of its behaviour, can be achieved using the method according to the third aspect of the present invention.

[0208] Moreover, the third aspect of the present invention can be used to manage learning by, or organize the behaviour of, an expert system. For example, an expert system may be intended to interact with human users, by asking them a series of questions in order to make some determination (for example, as to the identity of the user). However, before using the expert system for that purpose, it may be useful for the expert system to perform a series of tests, in which it tries out different possible questions or sequences of questions in order to determine which sequence is most likely to enable it to make accurate determinations later on. Such a series of tests could be selected using a technique according to the third aspect of the invention, in order to organize the behaviour of the expert system/manage its learning during the test phase. In this example application, the action variables would be different possible questions that could be posed.

[0209] The third aspect of the invention will be described below with reference to a simple experimental set-up. As shown in Fig. 8, this experimental set-up involves a simple robot, RBT, moving in a square space that is enclosed by walls, WL. The robot is formed of a box mounted on a left wheel and a right wheel, and it can control the speed of each wheel separately. The robot also has two infrared sensors at the front, enabling it to sense the walls, a "toy sensor" to be discussed in greater detail below, and can produce sounds at a selected one of three frequencies (labelled ***f1***, ***f2*** or ***f3***). It can be considered that the robot has two wall sensors (the two infrared sensors), one toy sensor, and three

IPR2025-00944
Tesla EX1002 Page 975

actuators (the controller of left wheel speed, the controller of right wheel speed, and the controller of sound frequency). The robot has an internal clock and updates the readings from its sensors, as well as the settings of its actuators, once every 1000 ms.

[0210] In the example discussed here, the robot has a decision-making unit integrated into it which determines which action the robot should take at a given time step. This decision-making unit comprises an action-selection system according to the first aspect of the invention, for example the action-selection system 1 of Fig.4. In the present case, this action-selection system can be considered to be a module that manages the robot's learning process thereby obtaining a particular kind of organization of the robot's behaviour. It should be noted that the invention is not limited to the case where the action-selection system according to the first aspect of the invention (or the decision-making unit) is on-board the robot. This system and unit could be separate from the robot but in communication with it.

[0211] Furthermore, for simplicity it is assumed in the following discussion that the decision-making unit bases its decisions on which action the robot should take solely on the selections made by the action-selection system according to the first aspect of the invention - in other words, in this example there is only one drive motivating the robot's actions.

[0212] In the experimental set-up of Fig.8 there is an intelligent toy, T, in the same space as the robot. The toy moves according to the sounds produced by the robot. In particular, when the robot emits sounds at frequency ***f1***, the toy moves randomly around the space at a speed higher than the robot can move; consequently it seems to the robot as if the toy is appearing and disappearing at random positions. When the robot emits sounds at frequency ***f2***, the toy stops moving. When the robot emits sounds at frequency ***f3***, the toy jumps into the robot. The robot can detect the toy, here using a vision system, and calculate the toy's position relative to the robot, in terms of x and y co-ordinates. This vision-based system for detecting the toy and determining its position is designated the "toy sensor".

[0213] The above-described experimental set-up was simulated using the well-known "Webot" mobile robot simulation software provided by Cyberbotics Sarl of Lausanne, Switzerland (available at www.cyberbotics.com).

[0214] At each time step, the robot (or, strictly speaking the action-selection system thereof) is arranged to decide what action to take: that is, what settings should be given to the speed of the left wheel, the speed of the right wheel, and the frequency of the sound the robot produces. Thus, in this case the action variables of the system are the left-wheel speed, right-wheel speed and sound frequency produced by the robot.

[0215] Simulations were performed in which the robot attempted to predict what would be the distance to the toy one time step (i.e. 1000 ms) after the robot took an action in a particular context (i.e. for a particular set of sensor readings). After elapsing of the 1000 ms period, the robot determined the actual position of the toy relative to the robot's own position in terms of (x,y) co-ordinates, using the toy sensor. Based on the sensed co-ordinate values, the robot calculated the actual distance to the toy.

[0216] When the robot produces the sound at frequency ***f1***, the toy moves around randomly. In this case, it is impossible for the robot to determine a relationship which will accurately predict the distance between the robot and the toy. There is no correlation between the distance to the toy and the robot's sensor values and wheel speeds. Nothing worthwhile can be learned by exploring this situation. In other words, the robot will not learn efficiently if it spends a long time emitting sound at frequency f1.

[0217] When the robot produces the sound at frequency f3 the toy jumps inside the robot and so the distance between the robot and the toy is zero. It is fairly easy for the robot's action-selection system to learn this relationship between production of sound at frequency f3 and zero distance to the toy (even in the presence of noise, for example due to the sensors). This relationship has very low complexity.

[0218] When the robot produces the sound at frequency f2 the toy stops moving, but because of the robot's own movements it will still seem to the robot as if the toy *is* moving. Nevertheless, if the robot makes a series of observations while continuing to emit sound at frequency f2, its action-selection system should eventually learn that the distance to the toy is related to the robot's own movements.

[0219] The interval between selecting an action and observing its outcome was set at 1000 ms, which is a relatively long period of time. Accordingly there could be substantial changes in the toy's position relative to the robot. This made it more difficult for the robot's action-selection system to learn the above-mentioned relationships that exist between the motor speeds of the robot's wheels, the sound frequency produced by the robot and the distance to the toy.

[0220] More particularly, in this example the goal was for the robot's action-selection system to learn mappings to enable it to predict, accurately, at a time ***t*** what would be the distance to the toy in 1000 ms time (i.e. at ***t***+1) IF the robot takes a particular action A(t) in the robot's current sensory-motor context ***SC(t)***. The values of the system-state/context parameters and action variables in this case are the set of values measured by the robot's sensors (wall sensors and toy sensors) at the time ***t***, as well as the settings of the robot's actuators at the time ***t***.

[0221] The only real complexity in the mapping that the robot's action-selection system must learn resides in that part of the mapping which relates to production of sound at frequency f2. When sound at frequency f2 is produced, the distance to the toy varies, as a function of the sensory context SC(t) and the actions that are performed **A**(***t***), in a complicated but learnable manner.

[0222] If the robot is to learn efficiently it should spend a lot of time exploring situations in which sound is produced

IPR2025-00944
Tesla EX1002 Page 976

at frequency f2 (where there is a lot to learn), a little time in situations where sound is produced at frequency f3 (where there is a small amount to learn), and practically no time in situations where sound is produced at frequency f1 (where there is nothing that can be learned). The third aspect of the present invention provides a technique which enables the robot to autonomously organize its own behaviour so as to focus attention appropriately on the situations which will render its learning efficient.

[0223] Fig.9 shows, schematically, the main components of a preferred embodiment of action-selection system 21 that can be used to manage learning by, and/or organization of the behaviour of, the robot of Fig.8. It will be noted that this action-selection system 21 has a similar construction and functioning to that of the action-selection system 1 that is illustrated schematically in Fig.4. Accordingly, the same reference numerals are used in Fig.9 to designate elements having a similar function to those in the action-selection system 1 of Fig.4.

[0224] Once again, it is to be understood that this action-selection system 21 would probably be implemented in software and the various "modules" shown in Fig.9 are provided for illustrative purposes only. The various functions performed in the action-selection system 21 may be split between modules in different ways from that shown in Fig.9.

[0225] The action-selection system 21 of Fig.9 receives inputs from sensors and actuators, these inputs representing the robot's sensory-motor state. In the present example, the inputs will be from the robot's infrared wall sensors, from the toy sensor, and from the robot's actuators (signals indicating the values of right wheel speed, left wheel speed and sound frequency - these signals may be supplied from controllers controlling the wheels/sound generator, rather than from the wheels/sound generator themselves).

[0226] In the preferred embodiments according to the third aspect of the invention, the signals from each of the sensors and actuators are normalised so as to fall within an interval 0.0 to 1.0. In the case of the sound frequency produced by the robot, f1 is produced when the "frequency" variable is in the range 0 to 0.33, f2 is produced when the "frequency" variable is in the range 0.34 to 0.66, f3 is produced when the "frequency" variable is in the range 0.67 to 1.0.

[0227] The action-selection system 21 includes a candidate action generator 25 which generates a list of possible actions that the robot could perform next, given the current values of its system-state/context parameters $\boldsymbol{SC}(t)$.

[0228] For each candidate action, the prediction module 12 in the relevant expert unit of the knowledge-gain assessment module 3 makes a prediction, $\boldsymbol{d}(t+1)$, of the distance to the toy at the next time step (in 1000 ms time). For each prediction made by a prediction module 12, the meta prediction module 13 of the same expert unit produces an estimate, $\boldsymbol{Ep}(t+1)$, of what will be the error in the predicted distance to the toy. The knowledge-gain assessment module 14 of this same expert unit compares the expected prediction error for the candidate action with actual prediction error data that has been observed and stored in the memory 15 of this expert unit, and outputs to the candidate action selector 4 a signal indicating the knowledge gain expected to arise if this candidate action is performed.

[0229] The action selection module 4 receives the expected-knowledge-gain signals, selects which action should be performed next and outputs appropriate control signals to the robot's actuators. The action selection module 16 also indicates to the calculator 5 which candidate action has been selected.

[0230] The calculator 5 calculates what is the actual prediction error, $\boldsymbol{E}_{\boldsymbol{Ract}}(t+1)$, and stores this data in the memory 15 of the appropriate expert unit.

Experimental results

[0231] A number of simulations were performed to demonstrate the efficiency of the technique of the third aspect of the present invention for managing learning in a self-developing device. These simulations used the experimental set-up of Fig.8 and ran for a large number of time steps. In these simulations the robot's action-selection system had no prior "knowledge", in particular it was not programmed to know that there is a difference between setting the speed of the wheels and setting the sound frequency - these parameters were, in effect, variables of unidentified type as far as the robot was concerned. The robot's action-selection system was designed to treat as an action variable any variable that it could tune, and to treat as a system-state/context parameter any parameter for which it could determine a value.

[0232] Fig. 10 is a graph indicating how actual prediction error varied over time (as the number of performed actions increased) when the action-selection system 21 of the preferred embodiment of the invention was used to help the robot of Fig.8 to learn to predict the distance to the toy T. In this simulation, the action selection module 4 had a 0.85 probability of selecting the candidate action expected to produce the greatest knowledge gain and a 0.15 probability of selecting a random action.

[0233] In Fig.10, a separate line is drawn to represent the evolution of actual prediction error for each "leaf" expert unit. On splitting of an expert unit, each of the new "leaf" expert units inherits from the initial expert unit those training examples (and actual prediction error values) which relate to the sub-region of system-state/context/action space now assigned to this new "leaf" expert. Thus, the prediction error traces in Fig.10 overlap at the left-hand side of the graph and progressively split from each other at the times when the corresponding expert units split. Thus, a tree structure appears in Fig. 10 and this corresponds to the decision-tree structure developed by the region identifier 2.

[0234] It will be seen that three groups of expert units developed during the simulation. Tests confirmed that the

IPR2025-00944
Tesla EX1002 Page 977

different branches of the decision-tree structure did indeed correspond to different situations; the topmost branch in Fig. 10 corresponds to situations in which the robot produces sounds at f1, the middle branch corresponds to situations in which the robot produces sounds at f2, and the lowest branch corresponds to situations in which the robot produces sounds at f3.

[0235] As shown in Fig. 10, the decision-tree structure quickly developed so as to distinguish those situations in which sound was being produced at f3. After only 250 time steps the robot (or, rather, its action-selection system 21) was able to differentiate situations in which sound was being produced at f3 from situations where sounds were being produced at other frequencies. At this time, the initial (or "root") expert unit in the action-selection system 21 was split into a first "branch" expert unit handling situations where the "sound frequency" variable took values 0.67 to 1.0 (i.e. sounds at f3) and a second "branch" expert unit handling situations where the sound frequency variable took values 0.0 to 0.66 (i.e. sounds produced at f1 or f2).

[0236] The robot discovered that, initially, the situations with sounds at f3 were the most important sources of learning progress. Thus, the majority of the time actions were selected which kept the robot in situations where the robot produced sounds at frequency f3. However, the robot very quickly learned all there is to know about such situations, namely that the distance to the toy is substantially zero in these cases. Accordingly, the prediction error quickly fell to around zero and these situations were no longer a source of learning progress.

[0237] During the initial stage of the simulation, although the robot preferred to explore situations with sound at f3 (because these offered the greatest learning progress) it nevertheless performed some actions which involved production of sounds at f1 or f2 because of the random selection of actions by the action-selection module 4. After a time, the "branch" expert unit handling situations where the sound frequency variable took values 0.0 to 0.66 (i.e. sounds produced at f1 or f2) had accumulated sufficient examples to meet above-mentioned criterion C1. This occurred at around the 320th time step, at which time this branch expert unit split into an "upper branch" expert unit handling situations where the sound frequency variable took values 0.0 to 0.33 (i.e. sounds produced at f1) and a "central branch" expert unit handling situations where the sound frequency variable took values 0.34 to 0.66 (i.e. sounds produced at f2).

[0238] At about the 400th time step the robot (or, strictly speaking, the knowledge-gain assessment modules of the expert units in the action-selection system 21) accurately determined that situations with sounds at f2 would provide greater learning progress than situations with sounds produced at f3 or situations with sounds produced at f1. Accordingly the action-selection module 4 tended to select actions which would involve production of sounds at f2.

[0239] Because the action-selection module 4 occasionally chooses random actions, the robot sometimes performs actions involving production of sounds at f1. However, when sounds are produced at f3 the expert units cannot learn anything to help them improve the accuracy of their predictions of the distance to the toy. Accordingly, the expert units quickly learn that actions resulting in the production of sound at f3 can produce no knowledge gain. It will be seen in Fig. 10 that the branch corresponding to situations with sounds at f1 ceases to be explored.

[0240] On the other hand, as more and more training examples are accumulated for situations in which sounds are produced at f2, the expert units can begin to see relationships between the robot's actions and the distance to the toy, thus enabling their predictions to improve. Accordingly, the robot explores different situations in which sound is produced at f2, experimenting with varying the speed of its left and right wheels. It has been found that, when using the active-learning management algorithm of the preferred embodiment of the present invention, this exploration of motor speed settings is organized and the speed settings are themselves selected in a manner which optimizes learning progress.

[0241] For the above-described simulation, it is interesting to consider how much time the robot spends in situations where sounds are produced at f1, f2 and f3, respectively. In order to investigate this, the simulation was repeated and Fig. 11 indicates how the percentage of time the robot spent in performing actions of different kinds varied as time went on (after the robot had performed a greater and greater number of actions). In Fig.11, separate traces represent the percentage of time spent performing actions in which sound is produced at f1, f2 and f3, respectively.

[0242] Figure 11 shows that the robot initially concentrates its attention on situations in which sounds are produced at f3. After the robot has mastered these situations it turns its attention to situations in which sounds are produced at f2.

[0243] From a consideration of the results of these simulations it will be seen that, by managing the robot's learning using the preferred embodiment of the third aspect of the invention, the robot has been able, autonomously, to scale the complexity of the situations it investigates, starting with situations of relatively low complexity and moving on to those of greater complexity. It has also been able to avoid spending time exploring situations where there is nothing to be learned. Thus, the third aspect of the present invention enables the robot to engage in efficient learning. Moreover, use of this technique also enables the robot to organize its own behaviour without outside intervention. This is a highly-desirable achievement.

[0244] As mentioned above, the active-learning management technique of the third aspect of the invention can be termed an Intelligent Adaptive Curiosity drive.

IPR2025-00944
Tesla EX1002 Page 978

Comparative Examples

[0245] Figs. 12 to 17 are graphs representing the results of simulations that were performed for the purposes of comparing the intelligent adaptive curiosity drive of the third aspect of the present invention with other known techniques for managing active learning.

First Comparative Example

[0246] Fig.12 shows how prediction error evolved over time in a simulation comparable to that describe above, in which the robot used expert units handling different regions of sensory-motor state-space and progressively divided up that space into regions according to the same criteria C1 and C2. However, in the case of this first comparative example, there was no assessment of expected knowledge gain; instead, actions were selected entirely at random. The algorithm used in this first comparative example can be denoted the "random" algorithm (RAN).

[0247] Fig. 12 shows a typical example of results obtained for the RAN algorithm after 15000 random actions. Sensory-motor state-space has been divided into three main regions, respectively corresponding to situations in which sounds are produced at f1, f2 and f3. Early on, situations have been explored in which sounds are produced at f3 and the system has learned how to predict the distance to the toy in those situations. Moreover, over time the prediction error associated with situations in which sound is produced at f2 has decreased. However, the system continues to perform actions which involve the robot producing sounds at f3. It is completely unproductive to do so, because the robot can learn nothing in these situations. Thus, in contrast to the intelligent adaptive curiosity drive of the present invention, the RAN algorithm does not produce efficient learning.

Second Comparative Example

[0248] Fig. 13 shows how prediction error evolved over time in another simulation comparable to the first described above. Once again the robot used expert units handling different regions of sensory-motor state-space and progressively divided up that space into regions according to the criteria C1 and C2. However, in the case of this second comparative example, actions were selected according to a conventional approach that can be summarized "choose the action which you expect to maximize prediction error". This can be termed the "maximizing" algorithm (MAX). It will be seen that this algorithm seeks out novel situations (on the assumption that these situations offer the greatest potential for learning).

[0249] Fig. 13 shows a typical example of results obtained using the MAX algorithm. Once again, sensory-motor state-space has been divided into three main regions corresponding to situations with sounds produced at f1, f2 and f3, respectively and, at an early stage, the system learns how to predict the distance to the toy in situations where sounds are produced at f3. However, situations in which sounds are produced at f2 and f3 are not explored in detail; instead, the system concentrates on situations where sounds are produced at f1. This is made even clearer by considering Fig. 14, which shows the percentage of time the robot spent in the three different types of situations. After only 1000 time steps, the robot is already spending 35% of its time in situations in which sounds are produced at f1.

[0250] It can be seen that, when seeking to manage active development, the MAX algorithm is counter-productive: it leads the robot to explore situations in which nothing can be learned.

Third Comparative Example

[0251] Fig.15 shows how prediction error evolved over time in a fourth simulation. This third comparative example was similar to the second comparative example except that actions were selected according to a conventional approach that can be summarized "choose the action which you expect to minimize prediction error" (a "minimizing" algorithm (MIN)).

[0252] In fact, Figs. 15 and 16 are the same sort of graphs as Figs.10 and 11 but, this time, plotted for a simulation using the MIN algorithm. In this case the algorithm drives the robot to stay in situations in which sound is produced at f3, the most easily learnable and predictable situations. This is not as inefficient as the MAX algorithm, but it is still counter-productive, very quickly the robot will have learnt all there is to know about the simple situations but it will not progress to explore more complex situations.

Advantages of the Third Aspect of the Invention

[0253] The action-selection technique according to the third aspect of the present invention allows a robot (or other self-developing device or system) to scale the complexity of its learning situations by successively and actively focusing its activity on problems of progressively increasing difficulty. This is achieved through the efficient and operational discovery, within system-state/context/action space, of regions of varying complexity, then evaluation of the relative

IPR2025-00944
Tesla EX1002 Page 979

complexity of these regions (in terms of learning potential), and the exploitation of this information by the robot (or other self-developing device or system) to self-organize its behaviour.

[0254] Other techniques for managing learning do not lead to development of this kind of organized behaviour except, perhaps, in very particular kinds of environment: for example, in an environment in which perturbations of increasing complexity are continuously generated. Such environments can be produced by human intervention but, otherwise, they are believed to be extremely rare. By way of contrast, as long as there is some potential complexity in the robot/environment, the third aspect of the present invention can efficiently manage open-ended learning without complicated dynamics of this kind.

[0255] The efficiency of the active learning achievable using the third aspect of the present invention is greater than that obtained using the algorithms of the comparative examples. This was shown by a further set of simulations which involved computing, at different stages in the simulation, how much the robot has learned about the more interesting situations, that is, the situations in which sounds are produced at f2.

[0256] In this set of simulations, the robot used the active-learning management technique of the preferred embodiment of the invention, the RAN, MIN and MAX algorithms, respectively, to predict the distance to the toy at each of a large number of time steps. After every 150 time steps, each simulation was temporarily stopped and the robot was exposed to a test database to see how well the robot could predict the distance to the toy for a number of situations in which sounds were being produced at f2. When each simulation was restarted, the examples in the test database were not maintained in memory (so as to avoid biasing the test results).

[0257] Incidentally, the test database was generated as follows. The robot was operated such that it selected actions at random. For each situation in which sounds were produced at f2, a note was made of the robot's sensory-motor data at that time as well as the distance to the toy.

[0258] Fig. 17 shows the results of this set of simulations, more especially it shows, for each of the different algorithms, how the mean prediction errors produced on analysing the test database varied over time.

[0259] It will be seen from Fig.17 that, initially, the RAN algorithm gives the lowest prediction error. In other words, early on it is the random algorithm that lets the robot learn the most about the situations where sounds are produced at f2. However, after about 3000 actions (time steps), the active-learning management technique according to the third aspect of the present invention begins to produce markedly lower prediction errors for these situations with sound at f2. This corresponds to the fact that the robot has turned its attention away from situations having sound at f3 (which the robot has thoroughly explored) and towards the situations with sounds at f2. In fact, the robot is spending 85% of its time exploring situations with sounds at f2 and rapidly learns how to improve its predictions of the distance to the toy in such situations.

[0260] After about 3000 time steps (actions) the active-learning management technique according to the third aspect of the present invention is the algorithm which has enabled the robot to learn most about the situations in which sounds are produced at f2. The mean prediction error using this algorithm stabilises at a low value of around 0.090. The mean prediction error produced by the MAX and MIN algorithms, respectively, remains at a substantially higher level (around 0.11). The mean prediction error produced by the RAN algorithm does, eventually, decrease to stabilise at a low value. However, this low value is around 0.092 which is still higher than the mean prediction error produced by the algorithm according to the present invention. Moreover, this low value is attained by the RAN algorithm after around 11000 actions. In other words, both the RAN algorithm and the algorithm according to the present invention enable the robot to learn about situations in which sounds are produced at frequency f2, but this learning is about 2.6 times faster when using the algorithm according to the third aspect of the present invention.

[0261] Although the various aspects of the present invention have been described above in terms of embodiments which are presently preferred, it is to be understood that the invention is not limited by the particular details of the preferred embodiments. In particular, the skilled person will readily appreciate that various changes and developments can be made in the preferred embodiment, within the scope of the annexed claims.

[0262] For example, although the prediction modules in the preferred embodiments of the invention use a nearest-neighbour prediction algorithm, the invention is not limited to the case where this algorithm is used. Similarly, although the action-selection module 4 of the preferred embodiments of the invention chooses random actions with a probability of 0.15, the invention is not limited to use of this probability value.

[0263] Also, although the region identifier used in preferred embodiments of the invention splits a region ***R*** into two or more smaller regions when the number of training examples in region ***R*** has reached a value *NS*=250, this number is merely illustrative, not limitative. Moreover, as indicated above, other splitting criteria could be used which do not count the number of training examples in a region.

[0264] Moreover, as indicated above, the distribution of functions between modules in the systems of Figs. 4, 5 and 9 is purely illustrative. Other distributions of the described functions between the modules are possible and/or a greater or lesser number of modules could be used to perform the same functions - for example, one module could be used for defining regions in system-state/action/context space and another module could be used for identifying which region applies to a particular combination of system-state/context parameters and action variables. Moreover, it is not necessary

IPR2025-00944
Tesla EX1002 Page 980

to include distinct modules in the action-selection systems and training systems of the invention. The various functions of the preferred embodiments of the invention could be performed, for example, in a computer program, by one or more software modules having overlapping functions.

[0265] Also, the skilled person will readily appreciate that certain of the functions ascribed to the different modules in Figs.4, 5 and 9 can be replaced by other functions, or combinations of functions, having an equivalent result. For example, as shown in Fig.4, the action selector 4 sends the calculator 5 of actual prediction error a signal indicating which action was actually selected for performance. The calculator 5 then retrieves the prediction relating to the action that has been selected for performance. However, the system 1 could be arranged such that calculator 5 receives all predictions made by the prediction modules 12 for a given set of candidate actions and identifies which of these predictions to use when assessing actual prediction error by studying the signal received from the action selector.

[0266] Furthermore, although the above description refers to distinct expert units, each assigned to calculate the expected knowledge-gain for different regions of system-state/context/action space, and refers to old expert units being split, or to new expert units being "created", this is not limitative. More particularly, it is to be understood that the important factor is that the knowledge gain should be assessed in an appropriate manner for each respective region in system-state/context/action space, whether this is achieved using an expandable set of distinct expert units, or using a single device (or software module) configured to load the prediction function, meta-prediction function and stored training example data applicable to the relevant region in system-state/context/action space, or using some other means.

[0267] Moreover, systems embodying the present invention could include a specific module for collating and storing the training example data and supplying it, as needed, to the expert unit handling the appropriate region in system-state/context/action space.

[0268] Furthermore, as indicated above the intelligent adaptive curiosity drive provided by the present invention can be just one of a plurality of drives motivating the selection of actions by a self-developing device. In such a case, the action-selection system 21 of Fig.9 would not itself make the final selection of which action the robot would take. More especially, the action selection module 4 of Fig.9 could be omitted and the knowledge-gain values produced by the various knowledge-gain assessment modules 14 could be output to a central controller as reward values indicating how desirable the associated candidate actions are considered to be. The central controller would process these reward values together with others generated by other drives motivating the robot's behaviour in order to decide which action the robot should take next. The central controller would send a signal to the module 21 indicating which candidate action has been selected, thus enabling actual prediction error to be calculated and stored in the memory 15 of the appropriate expert unit ***EXP***.

[0269] Finally, it is to be understood that robotic devices, artificial intelligence systems, expert systems and the like are not limited to using only the third aspect of the invention. The first aspect of the present invention can be applied for selecting the actions to be taken by such devices and systems, even when these devices/systems do not form part of the setup S whose behaviour is being investigated. The second aspect of the present invention can be applied in order to train prediction machines included in these devices and systems, even when these devices/systems do not form part of the setup S whose behaviour is being learned.

Claims

1. An automated action-selection system (1) adapted to generate signals specifying values for a set of one or more action variables defining an action that can be taken whereby to affect a setup (***S***), the automated action-selection system (1) comprising:

   input means for receiving signals indicative of the value, at a time ***t***, of a set of system-state/context parameters (***SC(t)***) describing the state and/or context of the setup (***S***);
   a region definer (2) adapted to define a set of regions in a multi-dimensional system-state/context/action space, each dimension of the system-state/context/action space being defined by a respective different parameter or variable of the sets of system-state/context parameters and action variables;
   means for determining a set of candidate actions, each candidate action consisting of a possible set of values for the action variables;
   a region identifier (2) for identifying the region in system-state/context/action space containing the combination of a given candidate action with values of any system-state/context parameters at time ***t***;
   a prediction unit (12) adapted to predict the value of a set of one or more predicted variables (***VAR***) a predetermined interval after time ***t***, wherein a prediction function applied by the prediction unit (12) depends upon the region in system-state/context/action space containing the combination of this given candidate action with any system-state/context parameters at time ***t***;
   calculator means adapted to calculate, for selected candidate actions, a respective indicator of the expected

IPR2025-00944
Tesla EX1002 Page 981

error in the prediction made by the prediction unit (12) for said selected candidate action;
memory means for storing indicators of actual prediction errors made by the prediction unit (12) for respective candidate actions selected on one or more previous occasions;
assessment means (14) adapted to evaluate the expected improvement in the performance of the prediction unit (12) if a given candidate action is performed, wherein an assessment performed by the assessment means (14) depends upon the region (***R***) in system-state/context/action space containing the combination of this given candidate action with the values, at time ***t***, of any system-state/context parameters, and the assessment means is further adapted to evaluate said expected improvement by comparing an indicator of said expected error in prediction with an indicator of the actual prediction error that existed on one or more occasions, previous to time ***t***, when the setup (***S***) had a combination of system-state/context parameters and action variables located in the same region (***R***) of the system-state/context/action space;
means for generating a signal indicating the desirability of selecting a given candidate action for performance, said signal being dependent on the expected improvement in the performance of the prediction unit (12) evaluated by the assessment unit (14) for said given candidate action; and
an action selector (4) adapted to select one of the candidate actions for performance based on said signals indicating the desirability of selecting a given candidate action and further adapted to output a signal indicating the selected candidate action.

2. The automated action-selection system according to claim 1, and comprising an action selector (4) adapted to select an action for performance to affect the setup (***S***), the action selector (4) is adapted to have a probability of ***p***, where $0 < p < 1$, of selecting that one of the set of candidate actions that the assessment unit (14) evaluates to be expected to yield the greatest improvement in performance of the prediction unit (12); wherein the action-selection system (1) is adapted to output data defining the action selected by the action selector (4).

3. The automated action-selection system according to claim 1 or 2, wherein the region definer (2) is adapted to divide an existing region (***R***) into two or more new regions when a first criterion (***C1***) is met.

4. The automated action-selection system according to claim 3, and comprising counting means for counting the number of occasions on which a first and second condition are both met, the first condition being that an action is taken affecting the setup (***S***), and the second condition being that the combination of the action variable values defining said action and the values of system-state/context parameter values at said time said action is taken falls within the region (***R***);
wherein the first criterion (***C1***) is met when the counting means has counted up to a predetermined number (***NS***).

5. The automated action-selection system according to any one of claims 1 to 4, and comprising meta prediction means (13) adapted to evaluate the expected error in the values of said set of variables (***VAR***) predicted by the prediction means (12), a prediction function applied by the meta prediction means (13) depending upon the region in system-state/context/action space containing the combination of said given candidate action with the system-state/context parameters at time ***t***; wherein the assessment means (14) is adapted to evaluate the expected improvement in the performance of the prediction means (12) if a given candidate action is performed by evaluating the decrease in prediction error that is expected to result from performance of said given candidate action.

6. The automated action-selection system according to any of claims 1 to 5, and comprising means for receiving feedback regarding the actual values of the predicted variables (***VAR***) resulting from performance of a given action; the prediction means (12) is responsive to the feedback data whereby to adapt the prediction function that is applied for candidate actions in the same region of system-state/context/action space as the region containing the combination of the action which produced the feedback data and values of any system-state/context parameters at the time when said given action was performed.

7. The automated action-selection system according to claim 6, and comprising memory means for storing training example data; for a given performed action said training example data comprising action variable values, any system-state/context parameter values applicable at the time said given action was performed, and feedback data defining the actual values of the set of predicted variables (***VAR***).

8. The automated action-selection system according to claim 7, wherein the region definer (2) is adapted to divide a region (***R***) of system-state/context/action space into two or more new regions, when a first criterion (***C1***) is met, by defining boundaries of the new regions so that there is a balance between the numbers of training examples in each new region, and the variance of the training examples is minimized in system-state/context/action space, or in a

IPR2025-00944
Tesla EX1002 Page 982

space defined by the set of predicted variables, or in a multi-dimensional space combining system-state/context/action space and said space defined by the set of predicted variables.

9. The automated action-selection system according to any one of claims 2 to 8, wherein the action selector (4) is adapted to select either a random action or that one of the set of candidate actions that the assessment unit (14) evaluates to be expected to yield the greatest improvement in performance of the prediction unit (12).

10. The automated action-selection system according to any one of claims 1 to 9, wherein the prediction unit (12) is adapted to make predictions by applying nearest-neighbours algorithms.

11. An automated prediction-machine-training system comprising an automated action-selection system according to one of claims 1 to 10, wherein the combination of the region identifier (2) and the prediction unit (12) of the automated action-selection system constitute a trainable prediction machine; the automated prediction-machine-training system comprising means for receiving feedback regarding the actual values of the predicted variables (***VAR***) resulting from performance of a given action; and wherein the prediction means (12) is responsive to the feedback data whereby to adapt the prediction function that is applied for candidate actions in the same region of system-state/context/action space as the region containing the combination of the action which produced the feedback data and any system-state/context parameters applicable at the time when said given action was performed.

12. The automated prediction-machine-training system according to claim 11, and comprising:

monitoring means for monitoring the evolution of the respective prediction functions applied by the prediction unit (12) for the different regions of system-state/context/action space; and
an operation-mode setter (16) adapted to change over operation of the prediction-machine-training system (11) from a training mode to a prediction mode when the monitoring means determines that the rate of change of the prediction functions has fallen below a threshold level.

13. An automated action-selection system according to any one of claims 1 to 10, adapted to control the actions performed by a robot or other self-developing device or system, said actions affecting the setup (***S***).

14. An automated action-selection system (21) according to claim 13, and comprising means for supplying the signal indicating the desirability of selecting a given candidate action for performance to a decision-making unit which selects the actions to be performed by said robot or other self-developing device or system.

15. An automated action-selection system (21) according to claim 13 or 14, wherein the action-selection system (21) is separate from, but in communication with the robot or other self-developing device or system.

16. A self developing robot, or other self-developing device or system, comprising an automated action-selection system (21) according to claim 13 or 14.

17. An automated action-selection method adapted to generate signals specifying values for a set of one or more action variables defining an action that can be taken whereby to affect a setup (***S***), the automated action-selection method comprising the steps of:

deceiving signals indicative of the value, at a time *t*, of a set of system-state/context parameters (***SC(t)***) describing the state and/or context of the setup (***S***);
defining a set of regions in a multi-dimensional system-state/context/action space, each dimension of the system-state/context/action space being defined by a respective different parameter or variable of the sets of system-state/context parameters and action variables;
determining a set of candidate actions, each candidate action consisting of a possible set of values for the action variables;
identifying the region in system-state/context/action space containing the combination of a given candidate action with values of any system-state/context parameters at time ***t***;
predicting the value of a set of one or more predicted variables (***VAR***) a predetermined interval after time ***t***, wherein a prediction function applied in the predicting step depends upon the region in system-state/context/action space containing the combination of this given candidate action with any system-state/context parameters at time ***t***;
calculating, for selected candidate actions, a respective indicator of the expected error in the prediction made

IPR2025-00944
Tesla EX1002 Page 983

in the predicting step for said selected candidate action;
storing indicators of actual prediction errors made in the predicting step for respective candidate actions selected on one or more previous occasions;
evaluating the expected improvement in the performance of the predicting step if a given candidate action is performed, wherein an assessment performed in the evaluating step depends upon the region (***R***) in system-state/context/action space containing the combination of this given candidate action with the values, at time t, of any system-state/context parameters, and the evaluating step is further adapted to evaluate said expected improvement by comparing an indicator of said expected error in prediction with an indicator of the actual prediction error that existed on one or more occasions, previous to time ***t***, when the setup (***S***) had a combination of system-state/context parameters and action variables located in the same region (***R***) of the system-state/context/action space;
generating a signal indicating the desirability of selecting a given candidate action for performance, said signal being dependent on the expected improvement in the performance of the prediction unit (12) evaluated in the evaluating step for said given candidate action;
selecting one of the candidate actions for performance based on said signals indicating the desirability of selecting a given candidate; and
outputting a signal indicating the selected candidate action.

18. The automated action-selection method according to claim 17, and comprising the step of feeding back data indicative of the actual values of the predicted variables (***VAR***) resulting from performance of a given action, and data indicative of the applicable values of any system-state/context parameters at the time said given action was performed, wherein the predicting step is responsive to the feedback data whereby to adapt the prediction function that is applied for candidate actions in the same region of system-state/context/action space as the region containing the combination of the action which produced the feedback data and any system-state/context parameters applicable at the time when said given action was performed.

19. An automated action-selection method according to claim 17 or 18, and comprising the step of controlling the actions performed by a robot or other self-developing device or system, said actions affecting the setup (***S***).

20. An automated action-selection method according to claim 19, and comprising the step of supplying the signal indicating the desirability of selecting a given candidate action for performance to a decision-making unit which selects the actions to be performed by said robot or self-developing device or system.

21. A self-developing robot, or other self-developing device or system, trained by having been caused to participate in the performance of a series of actions, each of the actions in the series having been selected using the automated action-selection system of any one of claims 13 to 15 and/or using the automated action-selection method of claim 19 or 20, wherein the system-state/context parameters (***SC***(***t***)) resulting from performance, by said robot or other self-developing device, of a given action in said series constituted the system-state/context parameters (***SC***(***t***)) used by the automated action-selection system or method in selection of the next action in said series.

22. A prediction machine trained by having been caused to predict a series of sets of predicted variables (VAR) using the automated prediction-machine-training system of claim 11 or 12, data indicative of the actual values of the predicted variables (***VAR***) resulting from performance of a given action, as well as data indicative of the applicable values of any system-state/context parameters at the time said given action was performed, having been fed back to the prediction machine such that an adaptation, responsive to the feedback data, was made of the prediction function applied for candidate actions in the same region of system-state/context/action space as the region containing the combination of the action which produced the feedback data and any system-state/context parameters applicable at the time when said given action was performed.

23. A computer system adapted to predict the value of a set of one or more predicted variables (***VAR***), said computer system having been trained by having been caused to predict a series of sets of predicted variables (VAR) using the automated prediction-machine-training system of claim 11 or 12, data indicative of the actual values of the predicted variables (***VAR***) resulting from performance of a given action, as well as data indicative of the applicable values of any system-state/context parameters at the time said given action was performed, having been fed back to the computer system such that an adaptation, responsive to the feedback data, was made of the prediction function applied for candidate actions in the same region of system-state/context/action space as the region containing the combination of the action which produced the feedback data and any system-state/context parameters applicable at the time when said given action was performed.

IPR2025-00944
Tesla EX1002 Page 984

24. An expert system adapted to predict the value of a set of one or more predicted variables (***VAR***), said expert system constituting a prediction-machine having been trained by having been caused to predict a series of sets of predicted variables (VAR) using the automated prediction-machine-training system of claim 11 or 12, data indicative of the actual values of the predicted variables (***VAR***) resulting from performance of a given action, as well as data indicative of the applicable values of any system-state/context parameters at the time said given action was performed, having been fed back to the prediction machine such that an adaptation, responsive to the feedback data, was made of the prediction function applied for candidate actions in the same region of system-state/context/action space as the region containing the combination of the action which produced the feedback data and any system-state/context parameters applicable at the time when said given action was performed.

Patentansprüche

1. Automatisiertes Aktionsauswahlsystem (1), welches dazu ausgebildet ist, Signale zu erzeugen, die Werte für einen Satz aus einer oder mehreren Aktionsvariablen spezifizieren, welche eine Aktion definieren, die ausgeführt werden kann, wodurch eine Konfiguration (S) beeinflusst wird,
wobei das automatische Aktionsauswahlsystem (1) aufweist:

eine Eingabeeinrichtung zum Empfangen von Signalen, die zu einem Zeitpunkt t den Wert eines Satzes von Systemzustands-/-kontextparametern (SC(t)) anzeigen, die den Zustand und/oder den Kontext der Konfiguration (S) beschreiben,
einen Bereichsdefinierer (2), welcher dazu ausgebildet ist, einen Satz von Bereichen in einem multidimensionalen Systemzustands-/-kontext-/- aktionsraum zu definieren, wobei jede Dimension des Systemzustands-/-kontext-/-aktionsraums definiert wird durch einen jeweiligen unterschiedlichen Parameter oder eine Variable aus den Sätzen von Systemzustands-/- kontextparametern und Aktionsvariablen,
eine Einrichtung zum Bestimmen eines Satz von Kandidatenaktionen, wobei jede Kandidatenaktion aus einem möglichen Satz von Werten für die Aktionsvariablen besteht,
einen Bereichsidentifizierer (2) zum Identifizieren des Bereichs im Systemzustands-/-kontext-/-aktionsraum, welcher die Kombination aus einer gegebenen Kandidatenaktion mit Werten eines jeglichen Systemzustands-/-kontextparameters zu einem Zeitpunkt t aufweist,
eine Vorhersageeinheit (12), welche dazu ausgebildet ist, den Wert eines Satzes einer oder mehrerer vorhergesagter Variablen (VAR) eine vorbestimmte Zeitspanne nach der Zeit t vorherzusagen, wobei eine Vorhersagefunktion, die durch die Vorhersageeinheit (12) angewandt wird, abhängt von dem Bereich im Systemzustands-/-köntext-/-aktionsraum, welcher die Kombination dieser gegebenen Kandidatenaktion mit jeglichen Systemzustands-/-kontextparametern zum Zeitpunkt t aufweist,
einer Berechnereinrichtung, welche dazu ausgebildet ist, für ausgewählte Kandidatenaktionen einen jeweiligen Indikator für den erwarteten Fehler in der Vorhersage zu berechnen, die von der Vorhersageeinheit (12) für die ausgewählte Kandidatenaktion gemacht wird,
eine Speichereinrichtung zum Speichern von Indikatoren tatsächlicher Vorhersagefehler, die von der Vorhersageeinheit (12) für jeweilige Kandidatenaktionen gemacht werden, die zu einer oder mehreren vorangehenden Gelegenheiten ausgewählt wurden,
eine Prüfeinrichtung (14), welche dazu ausgebildet ist, die erwartete Verbesserung in der Performance der Vorhersageeinheit (12) zu bewerten, falls eine gegebene Kandidatenaktion ausgeführt wird, wobei eine Prüfung, die von der Prüfeinrichtung (14) ausgeführt wird, von dem Bereich (R) im Systemzustands-/-kontext-/-aktionsraum abhängt, welcher die Kombination dieser gegebenen Kandidatenaktion mit den Werten eines jeglichen Systemzustands-/-kontextparameters zum Zeitpunkt t enthält, und wobei die Prüfeinrichtung des Weiteren dazu ausgebildet ist, die erwartete Verbesserung zu bewerten durch Vergleichen eines Indikators des erwarteten Fehlers in der Vorhersage mit einem Indikator des tatsächlichen Vorhersagefehlers, der zu einer oder mehreren Gelegenheiten vor dem Zeitpunkt t existierte, wenn die Konfiguration (S) eine Kombination von Systemzustands-/- kontextparametern und Aktionsvariablen aufwies, die im selben Bereich (R) des Systemzustands-/-kontext-/-aktionsraums lokalisiert war,
eine Einrichtung zum Erzeugen eines Signals, welches die Erwünschtheit des Auswählens einer gegebenen Kandidatenaktion für die Performance anzeigt, wobei das Signal abhängig ist von der erwarteten Verbesserung in der Performance der Vorhersageeinheit (12), welche durch die Prüfeinheit (14) für die gegebene Kandidatenaktion bewertet wurde, und
eine Aktionsauswahleinrichtung (4), welche dazu ausgebildet ist, eine der Kandidatenaktionen für die Performance auf der Grundlage der Signale auszuwählen, welche die Erwünschtheit des Auswählens einer gegebenen Kandidatenaktion anzeigen, und welche des Weiteren dazu ausgebildet ist, ein Signal, welches die ausgewählte

IPR2025-00944
Tesla EX1002 Page 985

Kandidatenaktion anzeigt, auszugeben.

2. Automatisiertes Aktionsauswahlsystem nach Anspruch 1,
welches des Weiteren eine Aktionsauswahleinrichtung (4) aufweist, die dazu ausgebildet ist, eine Aktion für die Performance zum Beeinflussen der Konfiguration (S) auszuwählen, wobei die Aktionsauswahleinrichtung (4) dazu ausgebildet ist, eine Wahrscheinlichkeit p, wobei $0 < p < 1$ gilt, für das Auswählen derjenigen aus dem Satz von Kandidatenaktionen hat, welche die Prüfeinheit (14) auswertet, von der erwartet wird, dass sie die größte Verbesserung in der Performance der Vorhersageeinheit (12) erzielt, wobei das Aktionsauswahlsystem (1) dazu ausgebildet ist, diejenigen Daten auszugeben, die die von der Aktionsauswahleinrichtung (4) ausgewählte Aktion definieren.

3. Automatisiertes Aktionsauswahlsystem nach Anspruch 1 oder 2,
wobei der Bereichsdefinierer (2) dazu ausgebildet ist, einen existierenden Bereich (R) in zwei oder mehr neue Bereiche zu unterteilen, wenn ein erstes Kriterium (C1) erfüllt ist.

4. Automatisiertes Aktionsauswahlsystem nach Anspruch 3,
welche des Weiteren eine Zähleinrichtung aufweist zum Zählen der Anzahl von Gelegenheiten, bei welchen eine erste und eine zweite Bedingung beide erfüllt sind, wobei die erste Bedingung darin besteht, dass eine Aktion, die ausgeführt wird, die Konfiguration (S) beeinflusst und wobei die zweite Bedingung darin besteht, dass die Kombination der Aktionsvariablenwerte, welche die Aktion definieren, und der Systemzustands-/-kontextparameterwerte zu dem Zeitpunkt, bei welchem die Aktion ausgeführt wird, in den Bereich (R) fällt, und
wobei das erste Kriterium (C1) erfüllt ist, wenn die Zähleinrichtung bis zu einer vorbestimmten Zahl (NS) hochgezählt hat.

5. Automatisiertes Aktionsauswahlsystem nach einem der Ansprüche 1 bis 4,
wobei des Weiteren eine Metavorhersageeinrichtung (13) vorgesehen ist, welche dazu ausgebildet ist, den erwarteten Fehler in den Werten des Satzes von Variablen (VAR) zu bewerten, die durch die Vorhersageeinrichtung (12) vorhergesagt wurden, wobei eine Vorhersagefunktion, die von der Metavorhersageeinrichtung (12) angewandt wird, abhängig ist vom Bereich im Systemzustands-/-kontext-/-aktionsraum, welcher die Kombination der gegebenen Kandidatenaktion mit den Systemzustands-/-kontextparametern zum Zeitpunkt t enthält, wobei die Prüfeinheit (14) dazu ausgebildet ist, die erwartete Verbesserung in der Performance der Vorhersageeinrichtung (12) zu bewerten, wenn eine gegebene Kandidatenaktion ausgeführt wird, durch Auswerten des Abfalls im Vorhersagefehler, von dem erwartet wird, dass er aus der Performance der gegebenen Kandidatenaktion resultiert.

6. Automatisiertes Aktionsauswahlsystem nach einem der Ansprüche 1 bis 5,
wobei eine Einrichtung vorgesehen ist zum Empfangen einer Rückmeldung in Bezug auf die tatsächlichen Werte der vorhergesagter Variablen (VAR), die aus der Performance einer gegebenen Aktion resultieren,
wobei die Vorhersageeinrichtung (12) auf die Rückmeldedaten reagierend ist, um die Vorhersagefunktion anzupassen, die für Kandidatenaktionen im selben Bereich des Systemzustands-/-kontext-/-aktionsraums angewandt wird wie im Bereich, welcher die Kombination enthält aus der Aktion, welche die Rückmeldungsdaten erzeugte, und den Werten eines jeglichen Systemzustands-/-kontextparameters zu dem Zeitpunkt, bei dem die gegebene Aktion ausgeführt wurde.

7. Automatisiertes Aktionsauswahlsystem nach Anspruch 6,
welches eine Speichereinrichtung aufweist zum Speichern von Trainingsbeispieldaten,
wobei für eine gegebene ausgeführte Aktion die Trainingsbeispieldaten Aktionsvariablenwerte, jegliche Systemzustands-/-kontextparameterwerte, die zu dem Zeitpunkt anwendbar waren, zu dem die gegebene Aktion ausgeführt wurde, und Rückmeldungsdaten, welche die tatsächlichen Werte des Satzes vorhergesagter Variablen (VAR) definieren, aufweisen.

8. Automatisiertes Aktionsauswahlsystem nach Anspruch 7,
wobei der Bereichsdefinierer (2) dazu ausgebildet ist, einen Bereich (R) des Systemzustands-/-kontext-/-aktionsraums in zwei oder mehr neue Bereiche zu unterteilen, wenn ein erstes Kriterium (C1) erfüllt ist, durch Definieren von Grenzen der neuen Bereiche derart, dass ein Gleichgewicht vorliegt zwischen den Anzahlen der Trainingsbeispiele in jedem neuen Bereich und dass die Varianz der Trainingsbeispiele im Systemzustands-/-kontext-/- aktionsraum oder in einem Raum, der definiert wird durch den Satz vorhergesagter Variablen oder in einem multidimensionalen Raum, der den Systemzustands-/-kontext-/-aktionsraum und den durch den Satz vorhergesagter Variablen definierten Raum miteinander kombiniert, minimiert wird.

IPR2025-00944
Tesla EX1002 Page 986

9. Automatisiertes Aktionsauswahlsystem nach einem der Ansprüche 2 bis 8,
wobei die Aktionsauswahleinrichtung (4) dazu ausgebildet ist, entweder eine zufällige Aktion oder diejenige aus dem Satz von Kandidatenaktionen auszuwählen, welche die Prüfeinheit (14) auswertet, von der erwartet wird, dass sie die größte Verbesserung in der Performance der Vorhersageeinheit (12) liefert.

10. Automatisiertes Aktionsauswahlsystem nach einem der Ansprüche 1 bis 9,
wobei die Vorhersageeinheit (12) dazu ausgebildet ist, Vorhersagen durch Anwenden von Nächstnachbaralgorithmen zu machen.

11. Automatisiertes Vorhersagemaschinentrainingssystem mit einem automatisierten Aktionsauswahlsystem nach einem der Ansprüche 1 bis 10,
wobei die Kombination des Bereichsidentifizierers (2) und der Vorhersageeinheit (12) des automatisierten Aktionsauswahlsystems eine trainierbare Vorhersagemaschine bildet,
wobei das automatisierte Vorhersagemaschinentrainingssystem eine Einrichtung aufweist zum Empfangen einer Rückmeldung in Bezug auf die tatsächlichen Werte der vorhergesagten Variablen (VAR), welche aus der Performance einer gegebenen Aktion resultieren, und
wobei die Vorhersageeinrichtung (12) auf die Rückmeldedaten reagierend ist, um die Vorhersagefunktion anzupassen, die für Kandidatenaktionen im selben Bereich des Systemzustands-/-kontext-/-aktionsraums angewandt wird wie im Bereich, welcher die Kombination enthält aus der Aktion, welche die Rückmeldungsdaten erzeugte, und jeglichen Systemzustands-/- kontextparametern, die anwendbar waren zum Zeitpunkt, bei welchem die gegebene Aktion ausgeführt wurde.

12. Automatisiertes Vorhersagemaschinentrainingssystem nach Anspruch 11,
welches eine Überwachungseinrichtung aufweist zum Überwachen der Entwicklung der jeweiligen Vorhersagefunktionen, die von der Vorhersageeinheit (12) für die verschiedenen Bereichen des Systemzustands-/-kontext/-aktionsraums angewandt werden, und
mit einer Betriebsmoduseinstelleinrichtung (16), welche dazu ausgebildet ist, den Betrieb des Vorhersagemaschinentrainingssystems (11) zu wechseln von einem Trainingsmodus zu einem Vorhersagemodus, wenn die Überwachungseinrichtung ermittelt, dass die Änderungsrate der Vorhersagefunktionen unterhalb eines Schwellwerteniveaus gefallen ist.

13. Automatisiertes Aktionsauswahlsystem nach einem der Ansprüche 1 bis 10,
welches dazu ausgebildet ist, die Aktionen zu steuern, die von einem Roboter oder von einer anderen selbstentwickelnden Einrichtung oder System ausgeführt werden und welche die Konfiguration (S) beeinflussen.

14. Automatisiertes Aktionsauswahlsystem (21) nach Anspruch 13,
mit einer Einrichtung zum Zuführen des Signals, welches die Erwünschtheit des Auswählens einer gegebenen Kandidatenaktion für die Performance einer Entscheidungseinheit anzeigt, welche die von dem Roboter oder der anderen selbstentwickelnden Einrichtung oder System auszuführenden Aktionen auswählt.

15. Automatisiertes Aktionsauswahlsystem (21) nach Anspruch 13 oder 14,
wobei das Aktionsauswahlsystem (21) getrennt ist von dem aber in Kommunikation steht mit dem Roboter oder der anderen selbstentwickelnden Einrichtung oder System.

16. Selbstentwickelnder Roboter oder andere selbstentwickelnde Einrichtung oder System,
mit einem automatisierten Aktionsauswahlsystem (21) nach Anspruch 13 oder 14.

17. Automatisiertes Aktionsauswahlverfahren, welches dazu ausgebildet ist, Signale zu erzeugen, welche Werte spezifizieren für einen Satz einer oder mehrerer Aktionsvariablen, welche eine Aktion definieren, die ausgeführt werden kann, wodurch eine Konfiguration (S) beeinflusst wird,
wobei das automatisierte Aktionsauswahlverfahren die Schritte aufweist:

Empfangen von Signalen, die zu einem Zeitpunkt t anzeigend sind für den Wert eines Satzes von Systemzustands-/-kontextparametern (SC(t)), welche den Zustand und/oder den Kontext der Konfiguration (S) beschreiben,
Definieren eines Satzes von Bereichen in einem multidimensionalen Systemzustands-/-kontext-/-aktionsraum, wobei jede Dimension des Systemzustands-/-kontext-/-aktionsraums definiert wird durch einen jeweils unterschiedlichen Parameter oder eine Variable der Sätze von Systemzustands-/-kontextparametern und Aktions-

IPR2025-00944
Tesla EX1002 Page 987

variablen,
Bestimmen eines Satzes von Kandidatenaktionen, wobei jede Kandidatenaktion aus einem möglichen Satz von Werten für die Aktionsvariablen besteht,
Identifizieren des Bereichs im Sysiemzustands-/-kontext-/- aktionsraum, welcher die Kombination einer gegebenen Kandidatenaktion mit Werten eines jeglichen Systemzustands-/-kontextparameters zu einem Zeitpunkt t enthält,
Vorhersagen des Werts eines Satzes einer oder mehrerer vorhergesagter Variablen (VAR) eine vorbestimmte Zeitspanne nach der Zeit t, wobei eine Vorhersagefunktion, die im Vorhersageschritt angewandt wird, abhängt von dem Bereich im Systemzustands-/-kontext-/-aktionsraum, welcher die Kombination dieser gegebenen Kandidatenaktion mit einem jeglichen Systemzustands-/-kontextparameter zum Zeitpunkt t enthält,
für ausgewählte Kandidatenaktionen Berechnen eines jeweiligen Indikators des erwarteten Fehlers in der Vorhersage, der im Vorhersageschritt für die ausgewählte Kandidatenaktion gemacht wurde,
Speichern von Indikatoren tatsächlicher Vorhersagefehler, die im Vorhersageschritt für jeweilige Kandidatenaktionen, die bei einer oder mehreren vorangehenden Gelegenheiten ausgewählt wurden,
Bewerten der erwarteten Verbesserung in der Performance des Vorhersageschritts, falls eine gegebene Kandidatenaktion ausgeführt wird, wobei eine Überprüfung, die im Auswerteschritt ausgeführt wurde, abhängig ist von dem Bereich (R) im Systemzustands-/-kontext-/-aktionsraum, welcher die Kombination dieser gegebenen Kandidatenaktion mit den Werten eines jeglichen Systemzustands-/-kontextparameters zum Zeitpunkt t enthält, und wobei der Bewertungsschritt des Weiteren ausgebildet ist, um die erwartete Verbesserung zu bewerten durch Vergleichen eines Indikators des erwarteten Fehlers in der Vorhersage mit einem Indikator des tatsächlichen Vorhersagefehlers, der bei einer oder mehreren Gelegenheiten vor dem Zeitpunkt t existierte, wenn die Konfiguration (S) eine Kombination von Systemzustands-/-kontextparametern und Aktionsvariablen hatte, die im selben Bereich (R) des Systemzustands-/-kontext-/-aktionsraums lokalisiert war,
Erzeugen eines Signals, welches die Erwünschtheit des Auswählens einer gegebenen Kandidatenaktion für die Performance anzeigt, wobei das Signal abhängig ist von der erwarteten Verbesserung in der Performance der Vorhersageeinheit (12), die im Bewertungsschritt für die gegebene Kandidatenaktion bewertet wurde,
Auswählen einer der Kandidatenaktionen für die Performance auf der Grundlage der Signale, die die Erwünschtheit des Auswählens eines gegebenen Kandidaten anzeigen, und
Ausgeben eines Signals, welches die ausgewählte Kandidatenaktion anzeigt.

18. Automatisiertes Aktionsauswahlverfahren nach Anspruch 17,
mit einem Schritt des Rückmeldens von Daten, die anzeigend sind für die tatsächlichen Werte der vorhergesagten Variablen (VAR), welche aus der Performance einer gegebenen Aktion resultieren, und von Daten, die anzeigend sind für die anwendbaren Werte eines jeglichen Systemzustands-/- kontextparameters zum Zeitpunkt, bei welchem die gegebene Aktion ausgeführt wurde, wobei der Vorhersageschritt reagierend ist auf die Rückmeldedaten, wodurch die Vorhersagefunktion angepasst wird, welche angewandt wird für die Kandidatenaktion im selben Bereich des Systemzustands-/- kontext-/-aktionsraums wie dem Bereich, welcher die Kombination der Aktion, welche die Rückmeldungsdaten erzeugte, und jegliche Systemzustands/-kontextparamtern enthält, welche zu dem Zeitpunkt anwendbar waren, als die gegebene Aktion ausgeführt wurde.

19. Automatisiertes Aktionsauswahlverfahren nach Anspruch 17 oder 18,
mit dem Schritt des Steuerns der Aktionen, die ausgeführt werden durch einen Roboter oder eine andere selbstentwickelnde Einrichtung oder ein System, wobei die Aktionen die Konfiguration (S) beeinflussen.

20. Automatisiertes Aktionsauswahlverfahren nach Anspruch 19,
mit dem Schritt des Zuführens des Signals, welches die Erwünschtheit des Auswählens einer gegebenen Kandidatenaktion für die Performance an eine Entscheidungseinheit, welche die Aktionen auswählt, die durch den Roboter oder die selbstentwickelnde Einrichtung oder das System auszuführen sind.

21. Selbstentwickelnder Roboter oder eine andere selbstentwickelnde Einrichtung oder System,
welche trainiert wurden, indem bewirkt wurde, dass sie in der Performance einer Serie von Aktionen teilnehmen, wobei jede der Aktionen in der Serie ausgewählt wurde durch Verwenden des automatisierten Aktionsauswahlsystems nach einem der Ansprüche 13 bis 15 und/oder durch Verwenden des automatisierten Aktionsauswahlverfahrens nach Anspruch 19 oder 20,
wobei die Systemzustands-/-kontextparameter (SC(t)), die aus der Performance in Bezug auf eine gegebene Aktion aus der Serie durch den Roboter oder die andere selbstentwickelnde Einrichtung resultieren, die Systemzustands-/-kontextparameter (SC(t)) bilden, die vom automatisierten Aktionsauswahlsystem oder dem Verfahren bei der Auswahl der nächsten Aktion in der Serie verwendet werden.

IPR2025-00944
Tesla EX1002 Page 988

22. Vorhersagemaschine, die trainiert wurde, indem bewirkt wurde, dass sie eine Serie von Sätzen von vorhergesagten Variablen (VAR) unter Verwendung des automatisierten Vorhersagemaschinentrainingssystems nach Anspruch 11 oder 12 vorhersagt,
wobei Daten, die anzeigend sind für die tatsächlichen Werte der vorhergesagten Variablen (VAR), die aus der Performance einer gegebenen Aktion resultieren, und Daten, die anzeigend sind für die anwendbaren Werte eines jeweiligen Systemzustands-/-kontextparameters zu dem Zeitpunkt, bei welchem die gegebene Aktion ausgeführt wurde, an die Vorhersagemaschine derart zurück gemeldet wurden, dass eine Adaption in Antwort auf die Rückmeldedaten für die Vorhersagefunktion, welche für Kandidatenaktionen angewandt wurde im selben Bereich des Systemzustands-/-kontext-/- aktionsraums wie dem Bereich, der die Kombination der Aktion, welche die Rückmeldungsdaten erzeugte, und jegliche Systemzustands-/- kontextparameter aufweist, die zu dem Zeitpunkt anwendbar sind, bei dem die gegebene Aktion ausgeführt wurde, durchgeführt wurde.

23. Computersystem, welches dazu ausgebildet ist, den Wert eines Satzes aus einer oder mehreren vorhergesagten Variablen (VAR) vorherzusagen, wobei das Computersystem trainiert wurde, indem bewirkt wurde, dass sie eine Serie von Sätzen von vorhergesagten Variablen (VAR) unter Verwendung des automatisierten Vorhersagemaschinentrainingssystems nach Anspruch 11 oder 12 vorhersagt,
wobei Daten, die anzeigend sind für die tatsächlichen Werte der vorhergesagten Variablen (VAR), die aus der Performance einer gegebenen Aktion resultieren, und Daten, die anzeigend sind für die anwendbaren Werte eines jeglichen Systemzustands-/-kontextparameters zu dem Zeitpunkt, bei welchem die gegebene Aktion ausgeführt wurde, an das Computersystem derart zurück gemeldet wurden, dass eine Adaption in Antwort auf die Rückmeldedaten in Bezug auf die Vorhersagefunktion, welche für Kandidatenaktionen angewandt wurde im selben Bereich des Systemzustands-/- kontext-/-aktionsraunis wie dem Bereich, der die Kombination der Aktion, welche die Rückmeldungsdaten erzeugte, und jegliche Systemzustands-/- kontextparameter aufweist, die zu dem Zeitpunkt anwendbar sind, bei dem die gegebene Aktion ausgeführt wurde, durchgeführt wurde.

24. Expertensystem, die trainiert wurde, indem bewirkt wurde, dass sie eine Serie von Sätzen von vorhergesagten Variablen (VAR) unter Verwendung des automatisierten Vorhersagemaschinentrainingssystems nach Anspruch 11 oder 12 vorhersagt,
wobei Daten, die anzeigend sind für die tatsächlichen Werte der vorhergesagten Variablen (VAR), die aus der Performance einer gegebenen Aktion resultieren, und Daten, die anzeigend sind für die anwendbaren Werte eines jeweiligen Systemzustands-/-kontextparameters zum Zeitpunkt, bei welchem die gegebene Aktion ausgeführt wurde, an die Vorhersagemaschine derart zurück gemeldet wurden, dass eine Adaption in Antwort auf die Rückmeldedaten in Bezug auf die Vorhersagefunktion, welches für Kandidatenaktionen angewandt wurde im selben Bereich des Systemzustands-/- kontext-/-aktionsraums als den Bereich, der die Kombination der Aktion, welche die Rückmeldungsdaten erzeugt, und jeglichen Systemzustands-/- kontextparameters aufweist, der zu dem Zeitpunkt anwendbar ist, bei dem die gegebene Aktion ausgeführt wurde, durchgeführt wurde.

Revendications

1. Système de sélection d'action automatisée (1) conçu pour générer des signaux spécifiant des valeurs pour un ensemble d'une ou de plusieurs variables d'action définissant une action qui peut être prise pour affecter une configuration (S), le système de sélection d'action automatisée (1) comprenant :

un moyen d'entrée destiné à recevoir des signaux indicatifs de la valeur, à un temps t, d'un ensemble de paramètres d'état-de-système/contexte (SC(t)) décrivant l'état et/ou le contexte de la configuration (S) ;
un définisseur de région (2) conçu pour définir un ensemble de régions dans un espace d'état-de-système/contexte/action multidimensionnel, chaque dimension de l'espace d'état-de-système/contexte/action étant défini par une variable ou un paramètre différent respectif des ensembles des variables d'action et paramètres d'état-de-système/contexte;
un moyen destiné à déterminer un ensemble d'actions candidates, chaque action candidate se composant d'un ensemble possible de valeurs pour les variables d'action ;
un identificateur de région (2) destiné à identifier la région dans un espace d'état-de-système/contexte/action contenant la combinaison d'une action candidate donnée avec des valeurs de tout paramètre d'état-de-système/contexte au temps t ;
une unité de prévision (12) conçue pour prévoir la valeur d'un ensemble d'une ou de plusieurs variables prévues (VAR) à un intervalle prédéterminé après le temps t, dans lequel une fonction de prévision appliquée par l'unité de prévision (12) dépend de la région de l'espace d'état-de-système/contexte/action contenant la combinaison

IPR2025-00944
Tesla EX1002 Page 989

de cette action candidate donnée avec tout paramètre d'état-de-système/contexte au temps t ;
un moyen de calcul conçu pour calculer, pour les actions candidates sélectionnées, un indicateur respectif de l'erreur attendue lors de la prévision effectuée par l'unité de prévision (12) pour ladite action candidate sélectionnée ;
un moyen de mémoire destiné à stocker des indicateurs d'erreurs de prévision réelles réalisées par l'unité de prévision (12) pour les actions candidates respectives sélectionnées à une ou plusieurs occasions précédentes ;
un moyen d'évaluation (14) conçu pour évaluer l'amélioration attendue sur les performances de l'unité de prévision (12) si une action candidate donnée est réalisée, dans lequel une évaluation effectuée par le moyen d'évaluation (14) dépend de la région (R) dans l'espace d'état-de-système/contexte/action contenant la combinaison de cette action candidate donnée avec les valeurs, au temps t, de toute paramètre d'état-de-système/contexte, et le moyen d'évaluation est en outre conçu pour évaluer ladite amélioration attendue en comparant un indicateur de ladite erreur de prévision attendue avec un indicateur de l'erreur de prévision réelle qui existait à une ou plusieurs occasions, avant le temps t, lorsque la configuration (S) avait une combinaison de paramètres d'état-de-système/contexte et de variables d'action se trouvant dans la même région (R) de l'espace d'état-de-système/contexte/action;
un moyen destiné à générer un signal indiquant les avantages de sélectionner une action candidate donnée en vue de la réalisation de celle-ci, ledit signal dépendant de l'amélioration attendue sur les performances de l'unité de prévision (12) évaluée par l'unité d'évaluation (14) pour ladite action candidate donnée ; et
un sélecteur d'action (4) conçu pour sélectionner l'une des actions candidates, en vue de la réalisation de celle-ci, sur la base desdits signaux indiquant les avantages de sélectionner une action candidate donnée et conçu en outre pour émettre en sortie un signal indiquant l'action candidate sélectionnée.

2. Système de sélection d'action automatisée selon la revendication 1, et comprenant un sélecteur d'action (4) conçu pour sélectionner une action en vue de la réalisation de celle-ci afin d'affecter la configuration (S), le sélecteur d'action (4) est conçu pour avoir une probabilité de p, où $0<p<1$, de sélectionner l'action de l'ensemble d'actions candidates que l'unité d'évaluation (14) évalue comme étant attendue afin de produire la meilleure amélioration en matière de performances de l'unité de prévision (12) ; dans lequel le système de sélection d'action (1) est conçu pour émettre en sortie des données définissant l'action sélectionnée par le sélecteur d'action (4).

3. Système de sélection d'action automatisée selon la revendication 1 ou 2, dans lequel le définisseur de région (2) est conçu pour diviser une région existante (R) en deux, ou plus, nouvelles régions lorsqu'un premier critère 9C1) est rencontré.

4. Système de sélection d'action automatisée selon la revendication 3, et comprenant un moyen de comptage destiné à compter le nombre d'occasions lors desquelles une première et une seconde condition sont toutes les deux rencontrées, la première condition étant qu'une action est prise, affectant la configuration (S), et la seconde condition étant que la combinaison des valeurs de variable d'action définissant ladite action et des valeurs de paramètre d'état-de-système/contexte au moment où ladite action est prise tombe dans la région (R) ;
dans lequel le premier critère (C1) est rencontré lorsque le moyen de comptage a compté jusqu'à un nombre prédéterminé (NS).

5. Système de sélection d'action automatisée selon l'une quelconque des revendications 1 à 4 et comprenant un moyen de méta-prévision (13) conçu pour évaluer l'erreur attendue sur les valeurs dudit ensemble de variables (VAR) prévues par le moyen de prévision (12), une fonction de prévision appliquée par le moyen de méta-prévision (13) dépendant de la région de l'espace d'état-de-système/contexte/action contenant la combinaison de ladite action candidate donnée avec les paramètres d'état-de-système/contexte au temps t ; dans lequel le moyen d'évaluation (14) est conçu pour évaluer l'amélioration attendue en matière de performances du moyen de prévision (12) si une action candidate donnée est réalisée en évaluant la diminution d'erreur de prévision qui est attendue en relation avec la réalisation de ladite action candidate donnée.

6. Système de sélection d'action automatisée selon l'une quelconque des revendications 1 à 5, et comprenant un moyen destiné à recevoir un retour concernant les valeurs réelles des variables prévues (VAR) résultantes de la réalisation d'une action donnée ;
le moyen de prévision (12) répond aux données de retour pour adapter la fonction de prévision qui est appliquée pour les actions candidates de la même région de l'espace d'état-de-système/contexte/action que la région contenant la combinaison de l'action qui a produit les données de retour et les valeurs de tout paramètre d'état-de-système/contexte au moment où ladite action donnée a été réalisée.

IPR2025-00944
Tesla EX1002 Page 990

7. Système de sélection d'action automatisée selon la revendication 6, et comprenant un moyen de mémoire destiné à stocker des données d'exemples de formation ; pour une action réalisée donnée, lesdits données d'exemples de formation comprenant des valeurs de variable d'action, toute valeur de paramètre d'état-de-système/contexte applicable au moment où ladite action donnée a été réalisée, et les données de retour définissant les valeurs réelles de l'ensemble de variables prévues (VAR).

8. Système de sélection d'action automatisée selon la revendication 7, dans lequel le définisseur de région (2) est conçu pour diviser une région (R) de l'espace d'état-de-système/contexte/action en deux, ou plus, nouvelles régions, lorsqu'un premier critère (C1) est rencontré, en définissant les limites des nouvelles régions de sorte qu'il y ait un équilibre entre les nombres d'exemples de formation dans chaque nouvelle région, et que la variance des exemples de formation soit minimisée dans l'espace d'état-de-système/contexte/action, ou dans un espace défini par l'ensemble de variables prévues, ou dans un espace multidimensionnel combinant l'espace d'état-de-système/contexte/action et ledit espace défini par l'ensemble de variables prévues.

9. Système de sélection d'action automatisée selon l'une quelconque des revendications 2 à 8, dans lequel le sélecteur d'action (4) est conçu pour sélectionner une action aléatoire ou l'action de l'ensemble d'actions candidates que l'unité d'évaluation (14) évalue comme étant attendue pour produire la meilleure amélioration en matières de performances de l'unité de prévision (12).

10. Système de sélection d'action automatisée selon l'une quelconque des revendications 1 à 9, dans lequel l'unité de prévision (12) est conçue pour effectuer des prévisions en appliquant des algorithmes du plus proche voisin.

11. Système de formation de machine de prévision automatisée comprenant un système de sélection d'action automatisée selon l'une des revendications 1 à 10, dans lequel la combinaison de l'identificateur de région (2) et de l'unité de prévision (12) du système de sélection d'action automatisée constitue une machine de prévision pouvant être formée ; le système de formation de machine de prévision automatisée comprenant un moyen destiné à recevoir un retour concernant les valeurs réelles des variables prévues (VAR) résultant de la réalisation d'une action donnée ; et dans lequel le moyen de prévision (12) répond aux données de retour pour adapter la fonction de prévision qui est appliquée pour les actions candidates de la même région de l'espace d'état-de-système/contexte/action que la région contenant la combinaison de l'action qui a produit les données de retour et tout paramètre d'état-de-système/contexte applicable au moment où ladite action donnée a été réalisée.

12. Système de formation de machine de prévision automatisée selon la revendication 11, et comprenant :

    un moyen de contrôle destiné à contrôler l'évolution des fonctions de prévision respectives appliquées par l'unité de prévision (12) pour les différentes régions de l'espace d'état-de-système/contexte/action; et
    un configurateur de mode de fonctionnement (16) conçu pour faire changer le fonctionnement du système de formation de machine de prévision (11) en le faisant passer d'un mode formation à un mode prévision lorsque le moyen de contrôle détermine que le taux de changement des fonctions de prévision est tombé sous un niveau seuil.

13. Système de sélection d'action automatisée selon l'une quelconque des revendications 1 à 10, conçu pour commander les actions réalisées par un robot ou autre dispositif ou système d'auto-développement, lesdites actions affectant la configuration (S).

14. Système de sélection d'action automatisée (21) selon la revendication 13, et comprenant un moyen destiné à envoyer le signal indiquant les avantages de sélectionner une action candidate donnée en vue de la réalisation de celle-ci vers une unité de prise d'action qui sélectionne les actions à réaliser par ledit robot ou autre dispositif ou système d'auto-développement.

15. Système de sélection d'action automatisée (21) selon la revendication 13 ou 14, dans lequel le système de sélection d'action (21) est séparé du robot ou autre dispositif ou système d'auto-développement, mais en communication avec celui-ci.

16. Robot d'auto-développement, ou autre dispositif ou système d'auto-développement, comprenant un système de sélection d'action automatisée (21) selon la revendication 13 ou 14.

17. Procédé de sélection d'action automatisée conçu pour générer des signaux spécifiant les valeurs d'un ensemble

IPR2025-00944
Tesla EX1002 Page 991

d'une ou de plusieurs variables d'action définissant une action qui peut être prise pour affecter une configuration (S), le procédé de sélection d'action automatisée comprenant les étapes consistant à :

recevoir des signaux indicatifs de la valeur, à un temps t, d'un ensemble de paramètres d'état-de-système/contexte (SC(t)) décrivant l'état et/ou le contexte de la configuration (S) ;
définir un ensemble de régions dans un espace d'état-de-système/contexte/action multidimensionnel, chaque dimension de l'espace d'état-de-système/contexte/action étant définie par une variable ou paramètre différent respectif des ensembles de variables d'action et paramètres d'état-de-système/contexte;
déterminer un ensemble d'actions candidates, chaque action candidate se composant d'un ensemble possible de valeurs pour les variables d'action ;
identifier la région dans l'espace d'état-de-système/contexte/action contenant la combinaison d'une action candidate donnée avec des valeurs de tout paramètre d'état-de-système/contexte au temps t ;
prévoir la valeur d'un ensemble d'une ou de plusieurs variables prévues (VAR) à un intervalle prédéterminé après le temps t, dans lequel une fonction de prévision appliquée à l'étape de prévision dépend de la région de l'espace d'état-de-système/contexte/action contenant la combinaison de cette action candidate donnée avec tout paramètre d'état-de-système/contexte au temps t ;
calculer, pour les actions candidates sélectionnées, un indicateur respectif de l'erreur attendue lors de la prévision effectuée à l'étape de prévision pour ladite action candidate sélectionnée ;
stocker des indicateurs d'erreurs de prévision réelles réalisées à l'étape de prévision pour les actions candidates respectives sélectionnées à une ou plusieurs occasions précédentes ;
évaluer l'amélioration attendue sur les performances de l'étape de prévision si une action candidate donnée est réalisée, dans lequel une évaluation effectuée à l'étape d'évaluation dépend de la région (R) dans l'espace d'état-de-système/contexte/action contenant la combinaison de cette action candidate donnée avec les valeurs, au temps t, de toute paramètre d'état-de-système/contexte, et l'étape d'évaluation est en outre conçue pour évaluer ladite amélioration attendue en comparant un indicateur de ladite erreur de prévision attendue avec un indicateur de l'erreur de prévision réelle qui existait à une ou plusieurs occasions, avant le temps t, lorsque la configuration (S) avait une combinaison de paramètres d'état-de-système/contexte et de variables d'action se trouvant dans la même région (R) de l'espace d'état-de-système/contexte/action;
générer un signal indiquant les avantages de sélectionner une action candidate donnée en vue de la réalisation de celle-ci, ledit signal dépendant de l'amélioration attendue sur les performances de l'unité de prévision (12) évaluée par l'étape d'évaluation pour ladite action candidate donnée ;
sélectionner l'une des actions candidates, en vue de la réalisation de celle-ci, sur la base desdits signaux indiquant les avantages de sélectionner une action candidate donnée ;
émettre en sortie un signal indiquant l'action candidate sélectionnée.

18. Procédé de sélection d'action automatisée selon la revendication 17, et comprenant l'étape consistant à retourner des données indicatives des valeurs réelles des variables prévues (VAR) résultant de la réalisation d'une action donnée, et des données indicatives des valeurs applicables de tout paramètre d'état-de-système/contexte au moment où ladite action donnée a été réalisée, dans lequel l'étape de prévision répond aux données de retour pour adapter la fonction de prévision qui est appliquée pour les actions candidates de la même région de l'espace d'état-de-système/contexte/action que la région contenant la combinaison de l'action qui a produit les données de retour et tout paramètre d'état-de-système/contexte applicable au moment où ladite action donnée a été réalisée.

19. Procédé de sélection d'action automatisée selon la revendication 17 ou 18, et comprenant l'étape consistant à commander les actions réalisées par un robot ou autre dispositif ou système d'auto-développement, lesdites actions affectant la configuration (S).

20. Procédé de sélection d'action automatisée selon la revendication 19, et comprenant l'étape consistant à envoyer le signal indiquant les avantages de sélectionner une action candidate donnée en vue de la réalisation de celle-ci vers une unité de prise de décision qui sélectionne les actions à réaliser par ledit robot ou dispositif ou système d'auto-développement.

21. Robot d'auto-développement, ou autre dispositif ou système d'auto-développement, formé en ayant été amené à contribuer à la réalisation d'une série d'actions, chacune des actions de la série ayant été sélectionnée à l'aide du système de sélection d'action automatisée selon l'une quelconque des revendications 13 à 15 et/ou à l'aide du procédé de sélection d'action automatisée selon la revendication 19 ou 20, dans lequel les paramètres d'état-de-système/contexte (SC(t)) résultant de la réalisation, par ledit robot ou autre dispositif d'auto-développement, d'une action donnée de ladite série ont constitué les paramètres d'état-de-système/contexte (SC(t)) utilisés par le système

IPR2025-00944
Tesla EX1002 Page 992

ou procédé de sélection d'action automatisée lors de la sélection de l'action suivante de ladite série.

22. Machine de prévision formée en ayant été amenée à prévoir une série d'ensembles de variables prévues (VAR) à l'aide du système de formation de machine de prévision automatisée selon la revendication 11 ou 12, des données indicatives des valeurs réelles des variables prévues (VAR) résultant de la réalisation d'une action donnée, ainsi que des données indicatives des valeurs applicables de tout paramètre d'état-de-système/contexte au moment où ladite action donnée a été effectuée, ayant été retournées à la machine de prévision de sorte qu'une adaptation, répondant aux données de retour, soit constituée de la fonction de prévision appliquée pour les actions candidates de la même région de l'espace d'état-de-système/contexte/action que la région contenant la combinaison de l'action qui a produit les données de retour et tout paramètre d'état-de-système/contexte applicable au moment où ladite action donnée a été réalisée.

23. Système informatique conçu pour prévoir la valeur d'un ensemble d'une ou de plusieurs variables prévues (VAR), ledit système informatique ayant été formé en ayant été amené à prévoir une série d'ensembles de variables prévues (VAR) à l'aide du système de formation de machine de prévision automatisée selon la revendication 11 ou 12, des données indicatives des valeurs réelles des variables prévues (VAR) résultant de la réalisation d'une action donnée, ainsi que des données indicatives des valeurs applicables de tout paramètre d'état-de-système/contexte au moment où ladite action donnée a été effectuée, ayant été retournées vers le système informatique de sorte qu'une adaptation, répondant aux données de retour, soit constituée de la fonction de prévision appliquée pour les actions candidates de la même région de l'espace d'état-de-système/contexte/action que la région contenant la combinaison de l'action qui a produit les données de retour et tout paramètre d'état-de-système/contexte applicable au moment où ladite action donnée a été réalisée.

24. Système expert conçu pour prévoir la valeur d'un ensemble d'une ou de plusieurs variables prévues (VAR), ledit système expert se composant d'une machine de prévision ayant été formée en ayant été amenée à prévoir une série d'ensembles de variables prévues (VAR) à l'aide du système de formation de machine de prévision automatisée selon la revendication 11 ou 12, des données indicatives des valeurs réelles des variables prévues (VAR) résultant de la réalisation d'une action donnée, ainsi que des données indicatives des valeurs applicables de tout paramètre d'état-de-système/contexte au moment où ladite action donnée a été effectuée, ayant été retournées vers ladite machine de précision de sorte qu'une adaptation, répondant aux données de retour, soit constituée de la fonction de prévision appliquée pour les actions candidates de la même région de l'espace d'état-de-système/contexte/action que la région contenant la combinaison de l'action qui a produit les données de retour et tout paramètre d'état-de-système/contexte applicable au moment où ladite action donnée a été réalisée.

IPR2025-00944
Tesla EX1002 Page 993

FIG.1

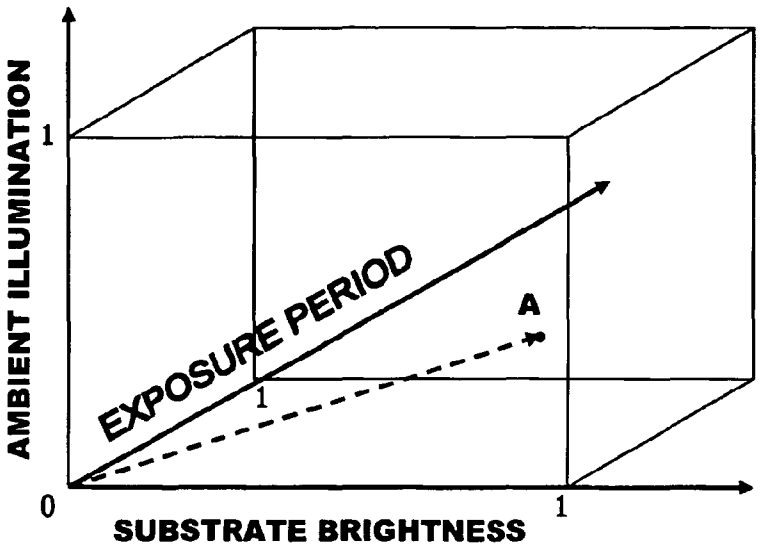
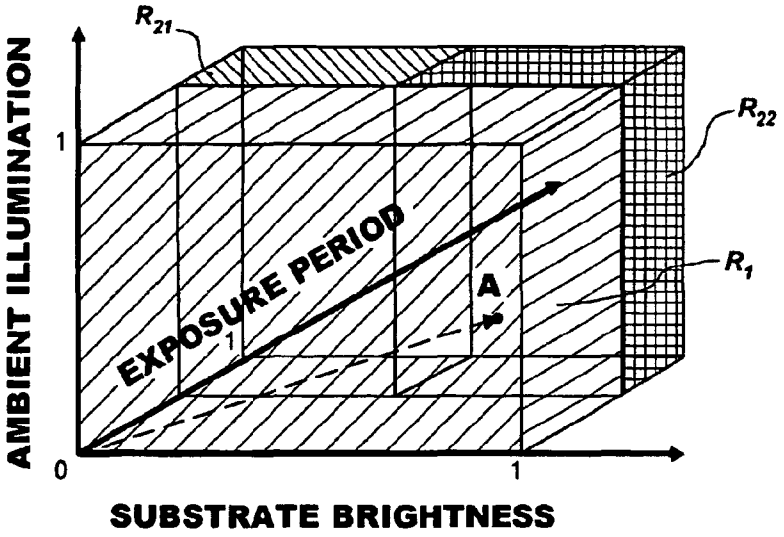


FIG.2



IPR2025-00944
Tesla EX1002 Page 994

FIG.3

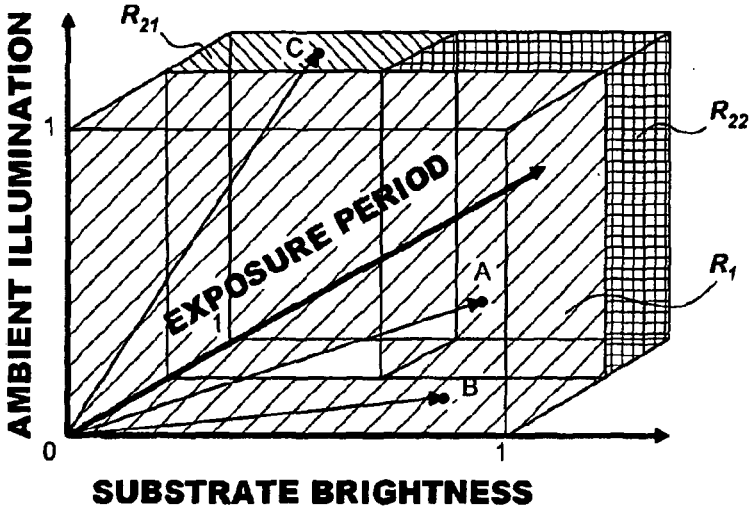
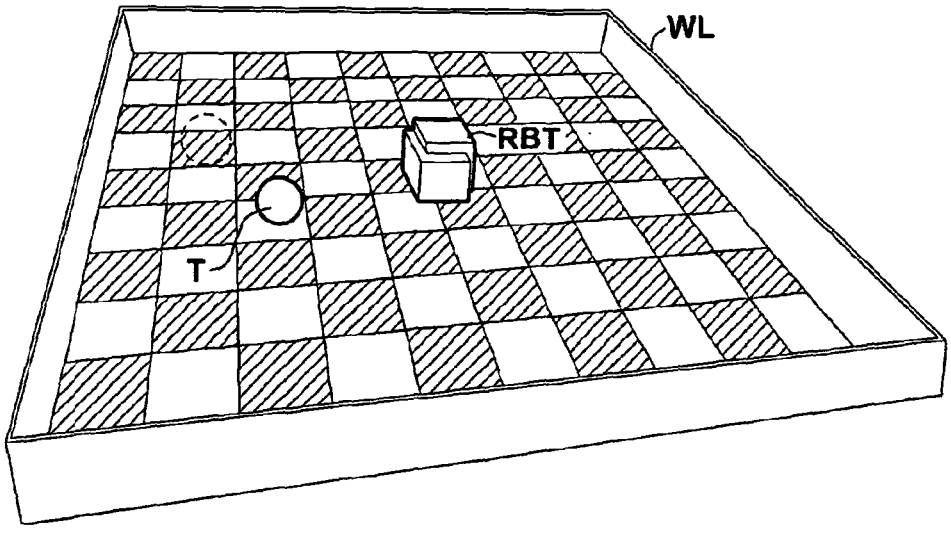
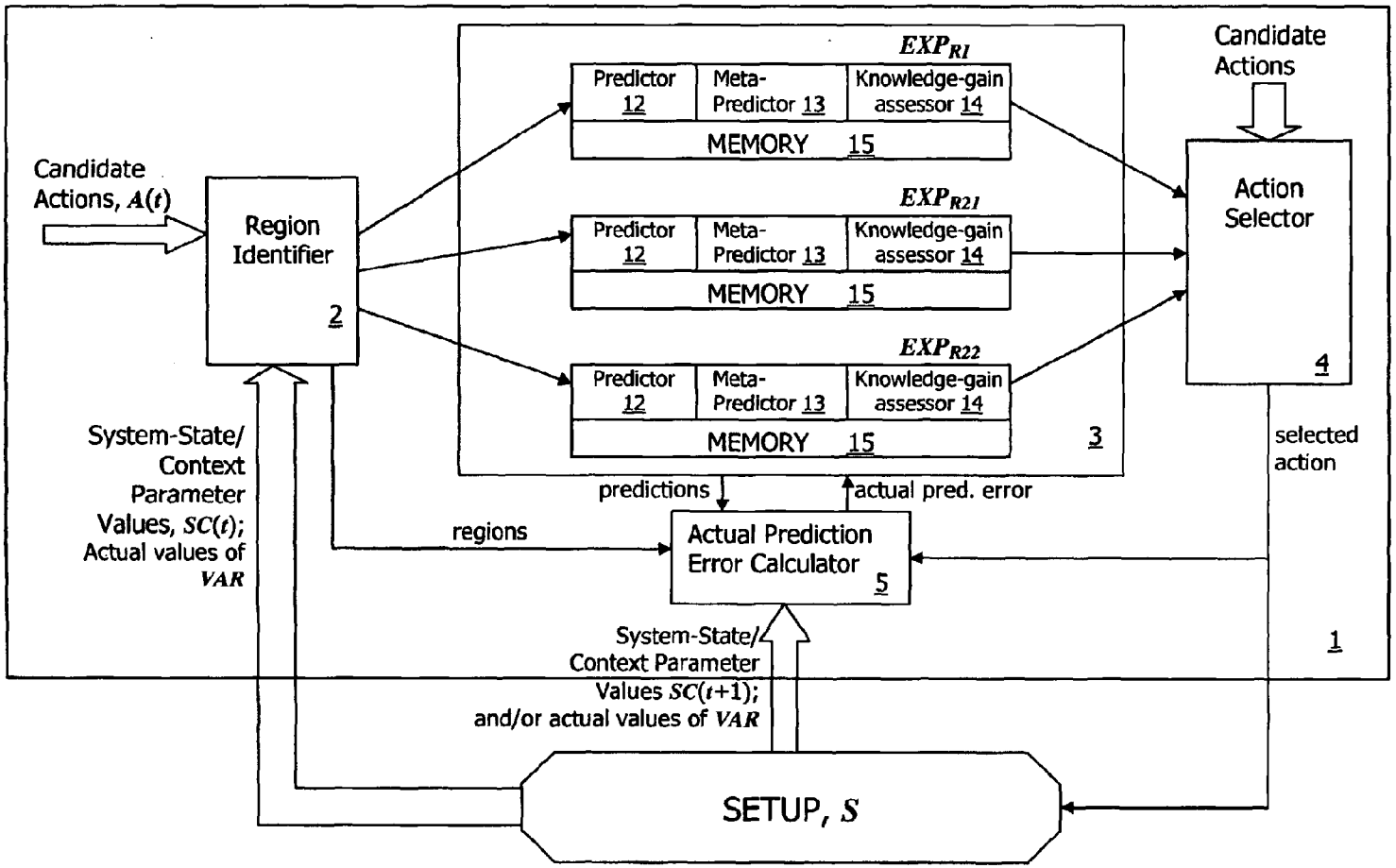


FIG.8



IPR2025-00944
Tesla EX1002 Page 995

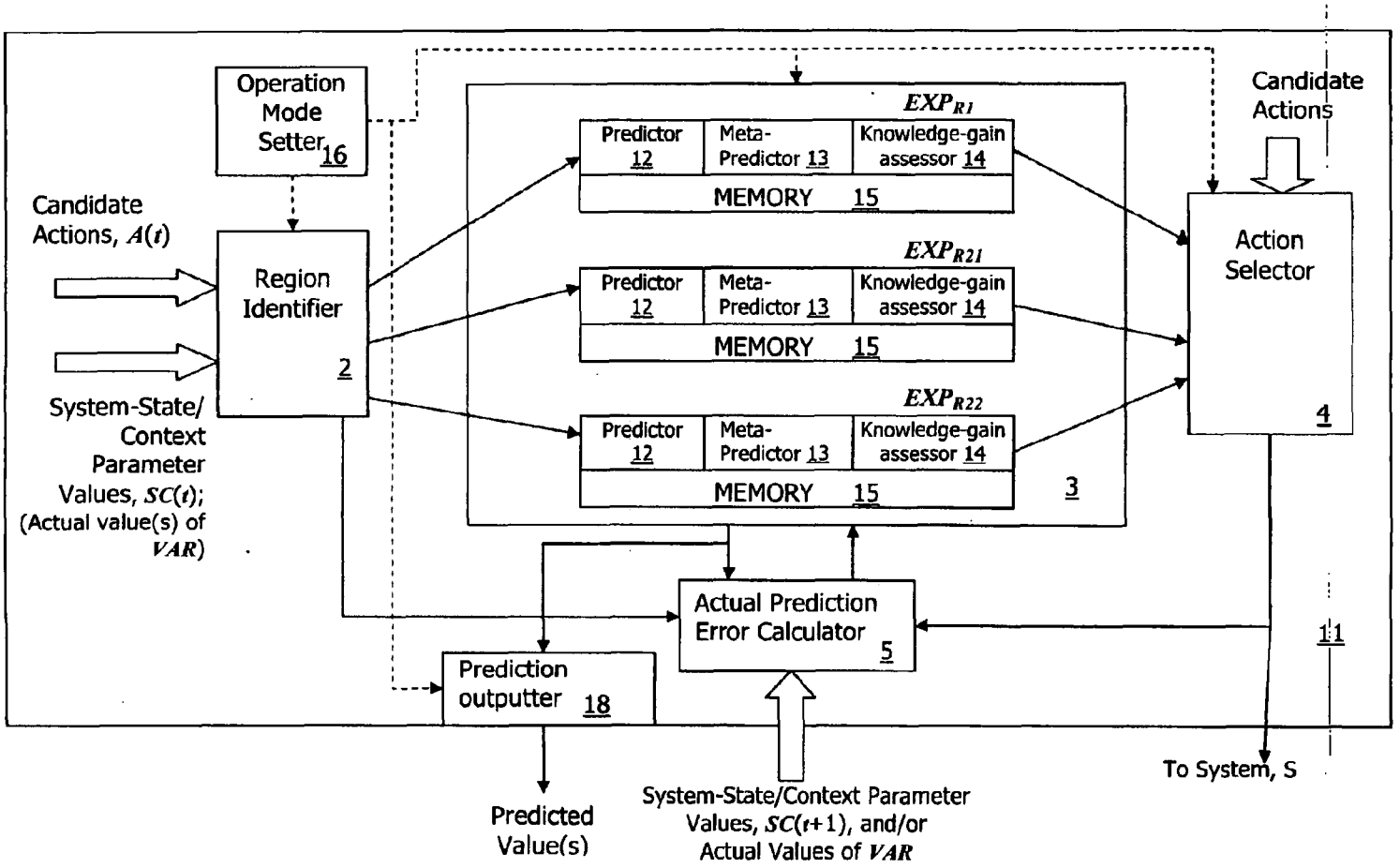
FIG.4



EP 1 622 072 B1

IPR2025-00944
Tesla EX1002 Page 996

FIG.5



EP 1 622 072 B1

IPR2025-00944
Tesla EX1002 Page 997

FIG.6

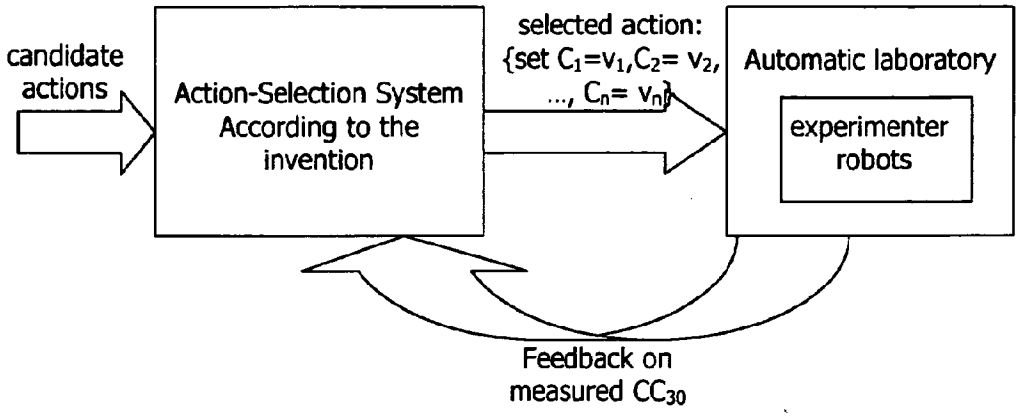
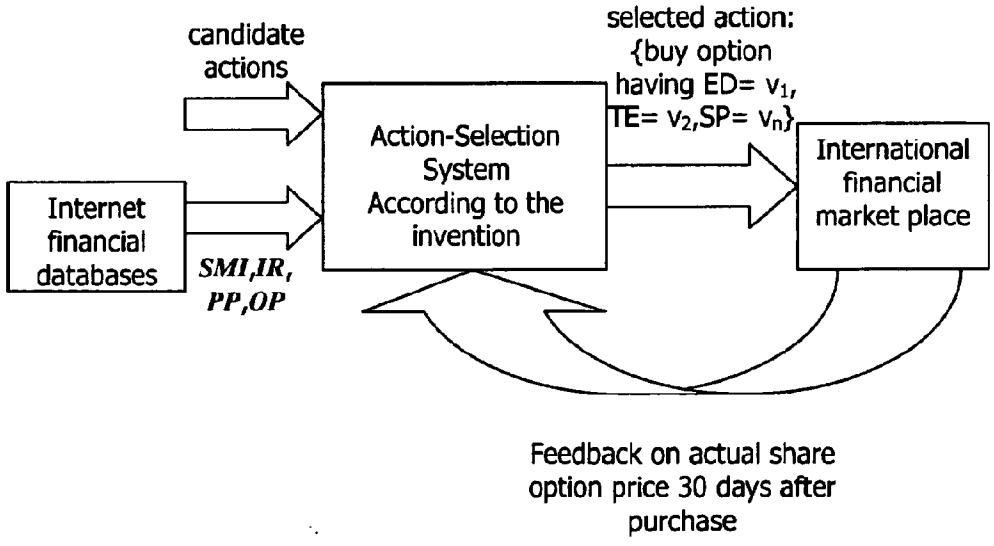
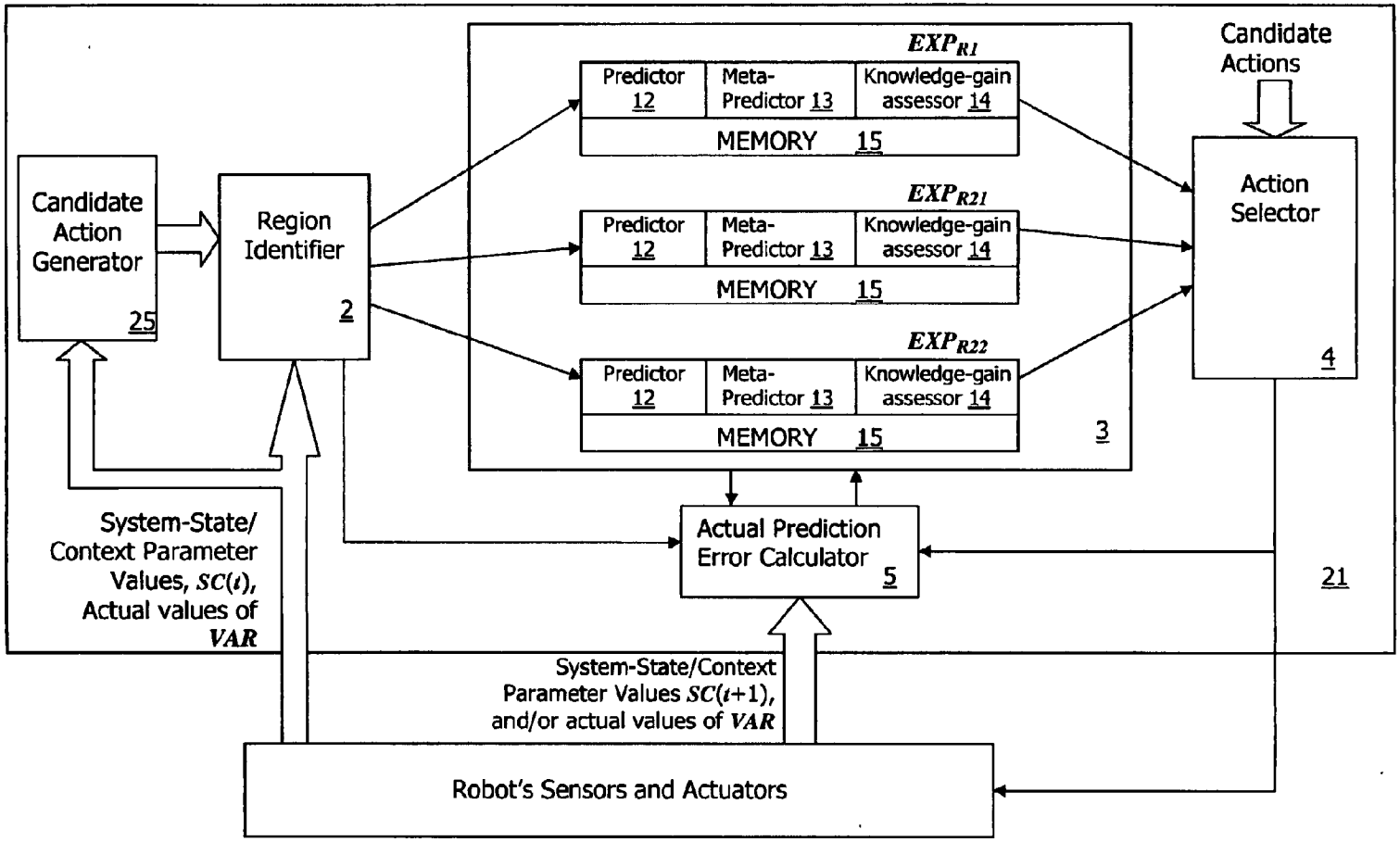


FIG.7



IPR2025-00944
Tesla EX1002 Page 998

FIG.9



EP 1 622 072 B1

IPR2025-00944
Tesla EX1002 Page 999

FIG.10

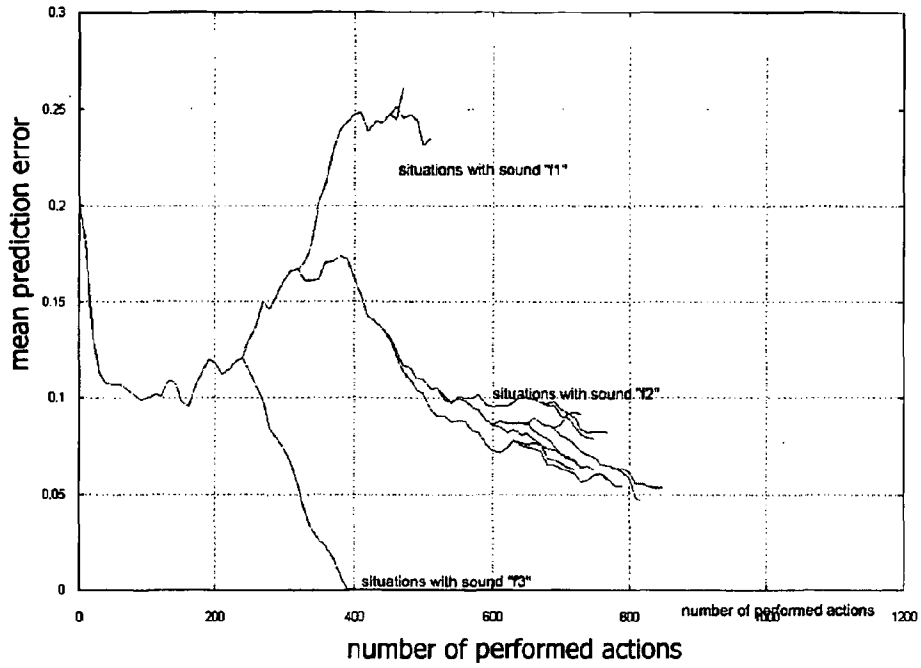
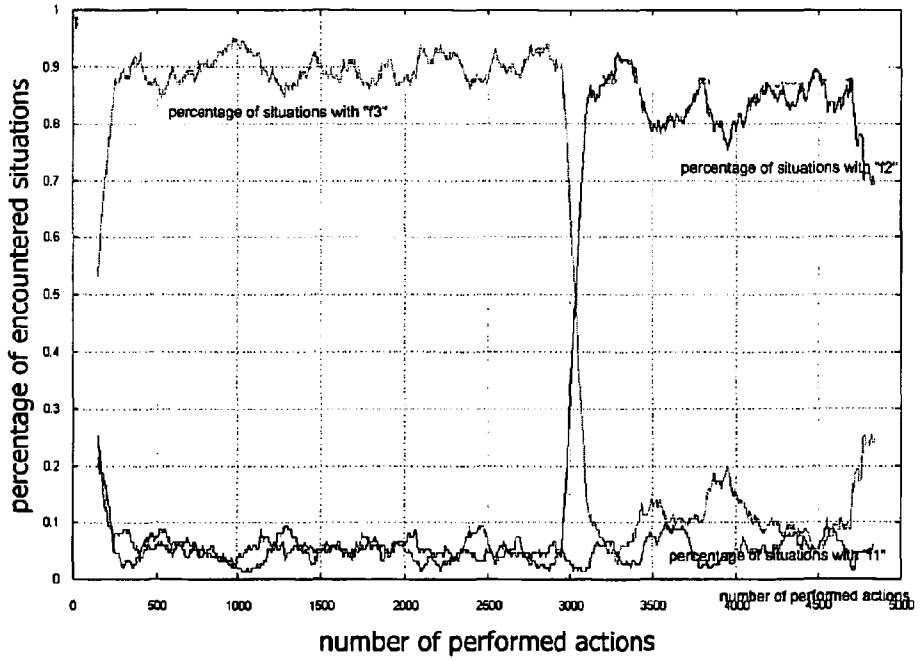
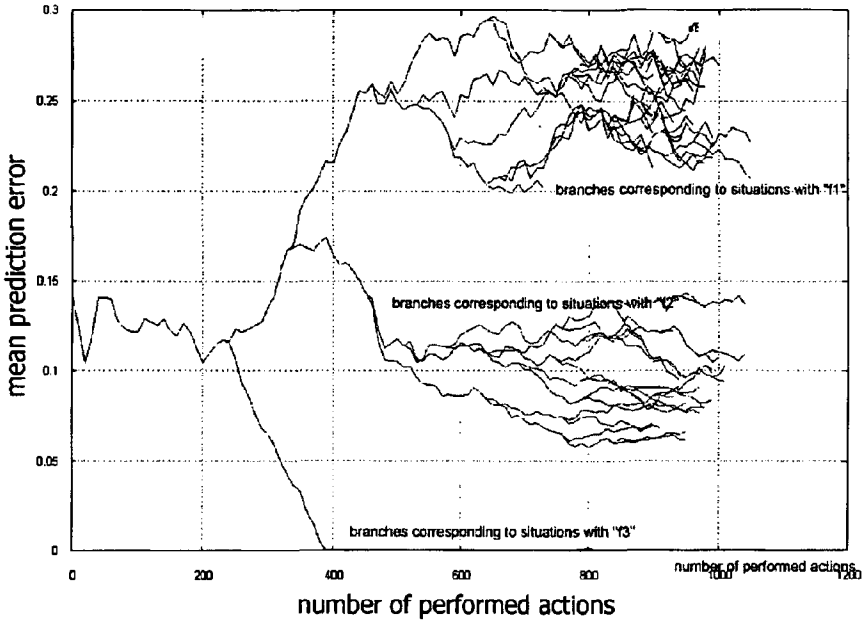


FIG.11



IPR2025-00944
Tesla EX1002 Page 1000

FIG.12



IPR2025-00944
Tesla EX1002 Page 1001

FIG.13

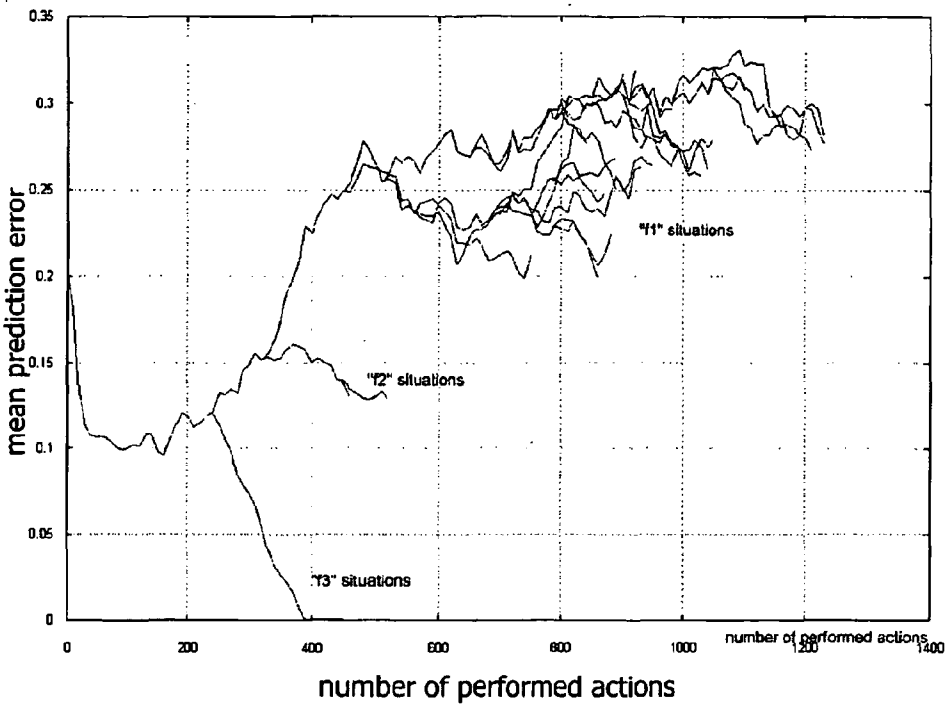
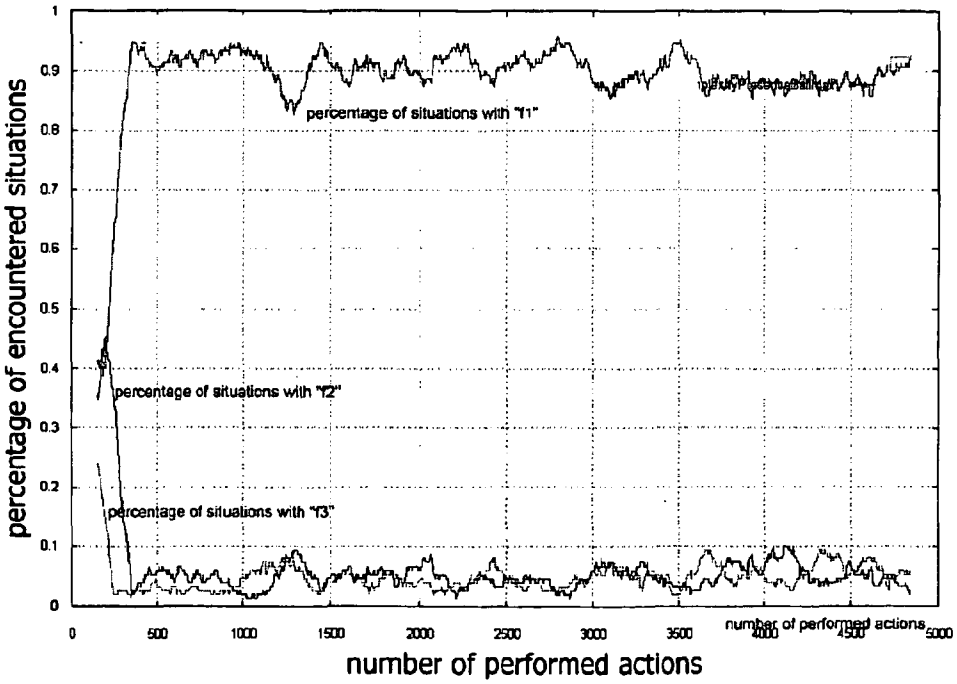


FIG.14



IPR2025-00944
Tesla EX1002 Page 1002

FIG.15

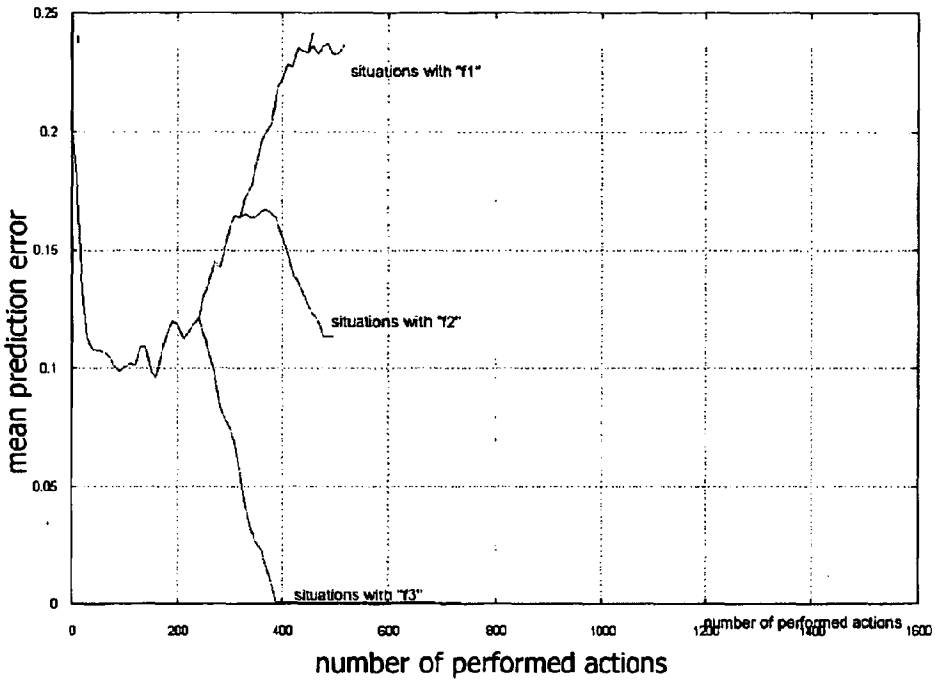
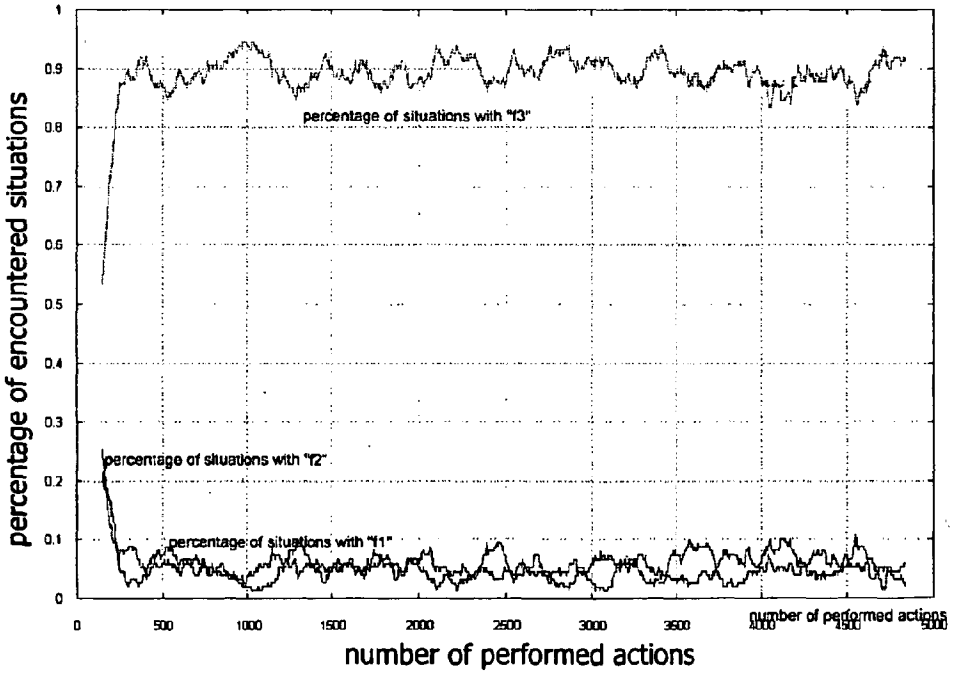
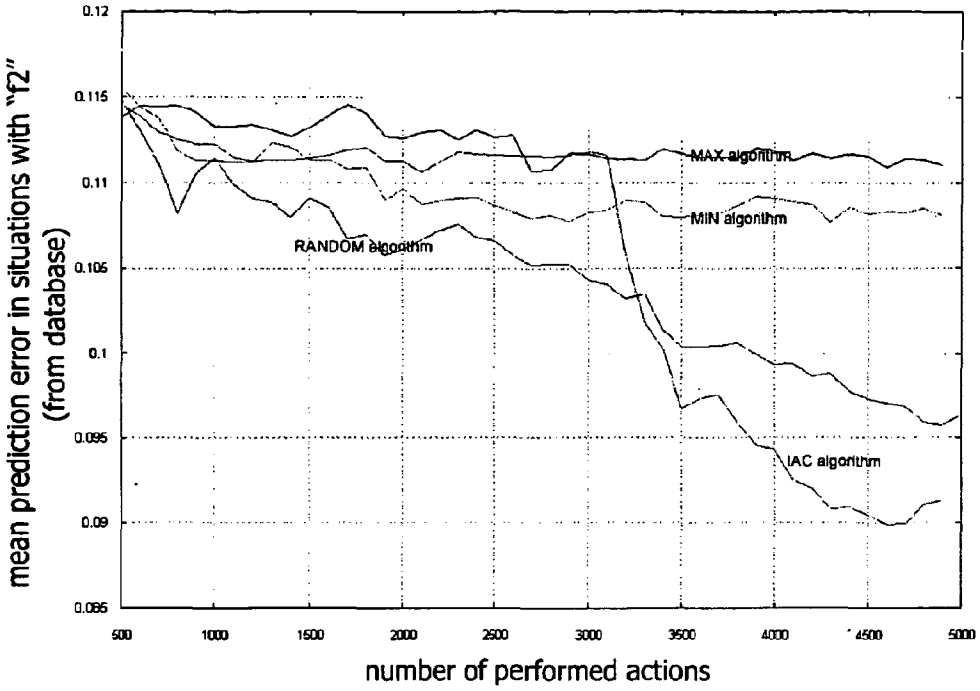


FIG.16



IPR2025-00944
Tesla EX1002 Page 1003

FIG.17



IPR2025-00944
Tesla EX1002 Page 1004

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- EP 03291359 A **[0032]**

Non-patent literature cited in the description

- **J. Denzler ; C. Brown.** Information Theoretic Sensor Data Selection for Active Object Recognition and State Estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 2001, vol. 2, 145-157 **[0022]**

IPR2025-00944
Tesla EX1002 Page 1005

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 727 089 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
29.11.2006 Bulletin 2006/48

(51) Int Cl.:
G06T 7/20 *(2006.01)* ***G05D 1/02*** *(2006.01)*

(21) Application number: **06017695.5**

(22) Date of filing: **27.11.2000**

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE TR

(30) Priority: **26.11.1999 US 167587 P**
01.09.2000 US 230166 P

(62) Document number(s) of the earlier application(s) in accordance with Art. 76 EPC:
00980706.6 / 1 257 971

(71) Applicant: **MobilEye Technologies, Ltd.**
1034 Nicosia (CY)

(72) Inventors:
- **Stein, Gideon**
  Jerusalem 96263 (IL)
- **Shashua, Amnon**
  Mevaseret Zion 90805 (IL)
- **Mano, Ofer**
  Modiin 71700 (IL)

(74) Representative: **Skone James, Robert Edmund**
Gill Jennings & Every LLP
Broadgate House
7 Eldon Street
London EC2M 7LH (GB)

Remarks:
This application was filed on 25 - 08 - 2006 as a divisional application to the application mentioned under INID code 62.

(54) System and method for estimating ego-motion of a moving vehicle using successive images recorded along the vehicle's path of motion

(57) An ego-motion determination system is disclosed for generating an estimate as to the ego-motion of a vehicle moving along a roadway. The ego-motion determination system includes an image information receiver and a processor. The image information receiver is configured to receive image information relating to a series of at least two images recorded as the vehicle moves along a roadway. The processor is configured to process the image information received by the image receiver to generate an ego-motion estimate of the vehicle, including the translation of the vehicle in the forward direction and the rotation of the vehicle around a vertical axis as between, for example, successive images.

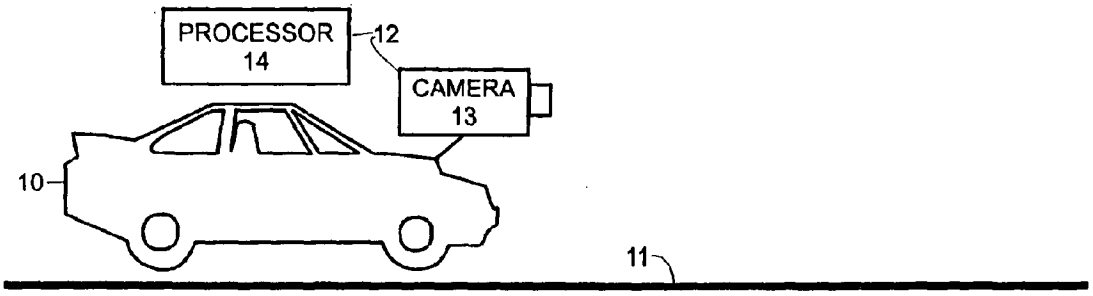


FIG. 1

EP 1 727 089 A2

IPR2025-00944
Tesla EX1002 Page 1006

Description

FIELD OF THE INVENTION

[0001] The invention relates generally to the field of systems and methods for estimating ego-motion (that is, "self-motion) of a moving vehicle, and more specifically to systems and methods that estimate ego-motion using successively-recorded images recorded along the vehicle's path ofmotion.

BACKGROUND OF THE INVENTION

[0002] Accurate estimation of the ego- ("self-") motion of a vehicle relative to a roadway is an important component in autonomous driving and computer vision-based driving assistance. Using computer vision techniques to provide assistance while driving, instead of mechanical sensors, allows for the use of the information that is recorded for use in estimating vehicle movement to also be used in detecting obstacles, identifying lanes and the like, without the need for calibration between sensors as would be necessary with mechanical sensors. This reduces cost and maintenance.

[0003] There are several problems in estimating ego-motion of a vehicle. Typically, roads have few feature points, if any. The most obvious features in a road, such as lane markings, have a generally linear structure, whereas background image structures, such as those associated with other vehicles, buildings, trees, and the like, will typically have many feature points. This will make image- or optical-flow-based estimation difficult in practice. In addition, typically images that are recorded for ego-motion estimation will contain a large amount of "outlier" information that is either not useful in estimating ego-motion, or that may result in poor estimation. For example, in estimating of ego-motion relative to a fixed structure such as a road, images of objects that are moving relative to the road, such as other moving vehicles and even moving wipers, will contribute false information for the ego-motion estimation. In addition, conditions that degrade image quality, such as raindrops and glare, will also make accurate ego-motion estimation difficult.

SUMMARY OF THE INVENTION

[0004] The invention provides a new and improved system and method for estimating ego-motion using successively-recorded images recorded along the vehicle's path of motion..

[0005] In brief summary, the invention provides an ego-motion determination system for generating an estimate as to the ego-motion of a vehicle moving along a roadway. The ego-motion determination system includes an image information receiver and a processor. The image information receiver is configured to receive image information relating to a series of at least two images recorded as the vehicle moves along a roadway. The processor is configured to process the image information received by the image receiver to generate an ego-motion estimate of the vehicle, including the translation of the vehicle in the forward direction and the rotation of the vehicle around a vertical axis as between, for example, successive images.

[0006] Several methodologies are disclosed for generating the ego-motion estimate of the vehicle.

BRIEF DESCRIPTION OF THE DRAWINGS

[0007] This invention is pointed out with particularity in the appended claims. The above and further advantages of this invention may be better understood by referring to the following description taken in conjunction with the accompanying drawings, in which:

FIG. 1 schematically depicts a vehicle moving on a roadway and including an ego-motion estimation system constructed in accordance with the invention;

FIG. 2 is a flow chart depicting operations performed by the ego-motion estimation system in determining ego-motion of the vehicle in accordance with one methodology; and

FIG. 3 is a flow chart depicting operations performed by the ego-motion estimation system in determining ego-motion of the vehicle in accordance with a second methodology.

DETAILED DESCRIPTION OF AN ILLUSTRATIVE EMBODIMENT

[0008] FIG. 1 schematically depicts a vehicle 10 moving on a roadway 11 and including an ego-motion estimation system 12 constructed in accordance with the invention. The vehicle 10 may be any kind of vehicle 10 that may move on the roadway 11, including, but not limited to automobiles, trucks, buses and the like. The ego-motion estimation system 12 includes a camera 13 and a ego-motion estimation system processor 14. The camera 13 is mounted on the vehicle 10 and is preferably pointed in a forward direction, that is, in the direction in which the vehicle would normally

IPR2025-00944
Tesla EX1002 Page 1007

move, to record successive images as the vehicle moves over the roadway. Preferably as the camera 13 records each image, it will provide the image to the ego-motion estimation system processor 14. The ego-motion estimation system processor 14, in turn, will process information that it obtains from the successive images, possibly along with other information, such as information from the vehicle's speedometer (not separately shown) to determine the ego-motion (that is, the self-motion) of the vehicle relative to the roadway 11. The ego-motion estimation system processor 14 may also be mounted in or on the vehicle 11 and may form part thereof. The ego-motion estimates generated by the ego-motion estimation system processor 14 may be used for a number of things, including, but not limited to obstacle and lane detection, autonomous driving by the vehicle, perhaps also using positioning information from, for example, the global positioning system ("GPS") and roadway mapping information from a number of sources known to those skilled in the art, and the like. Operations performed by the ego-motion estimation system processor 14 in determining ego-motion of the vehicle 10 will be described in connection with the flow charts depicted in FIGS. 2 and 3.

[0009] Before proceeding further, it would be helpful to provide background of the operations performed by the ego-motion estimation system processor 14 depicted in FIG. 1. Generally, as between two images Ψ and Ψ' the ego-motion estimation system processor 14 attempts to determine the translation $t = (t_x, t_y, t_z)^T$ and rotation $w = (w_x, w_y, w_z)^T$ (where "T" refers to the transpose operation, t_i refers to translation along the respective "X," "Y" and "Z" axes, and w_i refers to rotation around the respective axis) of the camera 13 affixed to the vehicle 10. Since the camera 13 is affixed to the vehicle 10, the translation and rotation of the camera 13 will also conform to the translation and rotation of the vehicle 10. In that case, for a point $p = (x,y)^T$ (where "x" and "y" are coordinates of a point or feature in the image) that is a projection of a point $P = (X,Y,Z)^T$ (where "X," "Y" and "Z" are coordinates of the point in three-dimensional space), the flow vector for the point, that is, the vector indicating the motion of the same point in three dimensional space from is position in the image Ψ to the image Ψ', has components (u,v)

$$u = \frac{1}{Z} S_1^T t + \left(\hat{p} \, X \, S_1 \right)^T w$$

$$v = \frac{1}{Z} S_2^T t + \left(\hat{p} \, X \, S_2 \right)^T w \qquad (1),$$

where "X" in equations (1) refers to the matrix cross-product and

$$S_1 = \begin{pmatrix} f \\ 0 \\ -x \end{pmatrix} \qquad S_2 = \begin{pmatrix} 0 \\ f \\ -y \end{pmatrix} \qquad \hat{p} = \begin{pmatrix} x/f \\ y/f \\ 1 \end{pmatrix} \qquad (2),$$

where "f" is the focal length of the camera 13, which is presumed to be known. The roadway on which the vehicle 10 is traveling is modeled as a plane. The equation for points on a plane is

$$AX + BY + BZ = 1 \qquad (3),$$

where the "X" and "Y" axes correspond to horizontal "x" and vertical "y" axes of the image plane, and the "Z" axis is perpendicular to the image plane. The camera 13 may be tilted slightly downwardly to increase the amount of the image that corresponds to the road, and reduce the amount of the image that corresponds to other features, such as other traffic, buildings, trees and the like, and in that case, the images Ψ and Ψ' will be rectified so that the images' z axis will be parallel to the plane of the roadway, as will be described below.

[0010] Dividing equation (3) by "Z" provides

IPR2025-00944
Tesla EX1002 Page 1008

$$\frac{1}{Z} = ax + by + c \qquad (4),$$

where $a=A/f$, $b=B/f$ and $c=C$. Substituting equation (4) into equations (1) results in

$$\begin{aligned} u &= (ax+by+c)S_1^T t + \left(\hat{p}\ X\ S_1\right)^T w \\ v &= (ax+by+c)S_2^T t + \left(\hat{p}\ X\ S_2\right)^T w \end{aligned} \qquad (5).$$

Expanding equations (5) results in

$$\begin{aligned} u = & -\left(ct_z + aft_x\right)x + \left(bft_x - w_z\right) + \left(fw_y + cft_x\right) + \\ & \left(\frac{w_y}{f} - at_z\right)x^2 - \left(\frac{w_x}{f} + bt_z\right)xy \end{aligned} \qquad (6)$$

$$\begin{aligned} v = & \left(w_z + aft_x\right)x + \left(-ct_z + bft_y\right)y - \left(fw_x + cft_y\right) + \\ & \left(\frac{w_y}{f} - at_z\right)xy - \left(\frac{w_x}{f} + bt_z\right)y^2 \end{aligned} \qquad (7).$$

Equations (6) and (7) are a special case (the "calibrated camera 13" case) of an eight parameter model for a camera 13 moving relative to a plane:

$$u = \alpha_1 x + \alpha_2 y + \alpha_3 + \alpha_7 x^2 + \alpha_8 xy \qquad (8)$$

$$v = \alpha_4 x + \alpha_5 y + \alpha_6 + \alpha_7 xy + \alpha_8 y^2 \qquad (9)$$

Given the flow vector (u,v), one can recover the parameters α_i, i=1,...,8, from which one can recover the motion parameters t and w.

[0011] One problem arises in connection with the methodology described above in connection with equations (1) through (9), namely, given the large number ofparameters α_i whose values need to be determined, it is difficult to devise a satisfactory method that will reject outliers. This follows from the fact that a relatively large number of optical flow vectors (u,v) will need to be used in determining the values of the parameters, which, in turn, requires a corresponding number of points in both images. In addition, it is difficult to differentiate between optical flow due to rotation around the X and Y axes and translation along the X and Z axes, respectively.

[0012] Accordingly, it is desirable to reduce the number of motion parameters to a minimum. The motion of a vehicle 10 along a roadway can be modeled as being constrained to be a translation along the Z axis, as the vehicle 10 moves

IPR2025-00944
Tesla EX1002 Page 1009

forward or in reverse, and a rotation around the X and Y axes, as the vehicle 10's path deviates from a straight-line course. In that case, equation (5) reduces to

$$u = (ax + by + c)xt_z - \frac{xy}{f}w_x + \left(f + \frac{x^2}{f}\right)w_y$$
$$v = (ax + by + c)yt_z - \left(f + \frac{y^2}{f}\right)w_x + \frac{xy}{f}w_y \quad \textbf{(10).}$$

If the images are rectified to ensure that the plane of the roadway is parallel to the camera 13's XZ plane, that is, so that the images would be as if the camera 13's optical axis is parallel to the plane of the road, then, in equations (10) a=0 and c=0, in which case

$$u = bxyt_z - \frac{xy}{f}w_x + \left(f + \frac{x^2}{f}\right)w_y$$
$$v = -by^2t_z - \left(f + \frac{y^2}{f}\right)w_x + \frac{xy}{f}w_y \quad \textbf{(11).}$$

In order to rectify the images, the camera 13 will need to be calibrated. A methodology for calibrating the camera 13 and rectifying the images will be described below.

[0013] In equations (11) there are three motion parameters, t_z (translation along the Z axis), w_x (rotation around the X axis) and w_y (rotation around the Y axis) to be determined from the flow vectors (u,v) associated with points in at least some portions of the images Ψ and Ψ'. Finding corresponding points in the images Ψ and Ψ', that is, points that are projections of the same point in three-dimensional space in the respective images is based on a "photometric constraint"

$$I(x,y,t) - I(x + u\delta t, y + v\delta t, t + \delta t) = 0 \quad \textbf{(12),}$$

which essentially states that the irradiance, or luminance, of the point $p = (x,y)^T$ in the image Ψ and and the point $p = (x + u\delta t, y + v\delta t)$ in the image Ψ', which are projections ofthe same point in three-dimensional space into the respective images, are the same. In practice, equation (12) will not hold exactly because of noise. If, for every point, the noise is modeled as zero mean Gaussian noise, equation 12 reduces to

$$P\left(I(x,y,t) - I(x + u\delta t, y + v\delta t, t + \delta t)\right) = N(\sigma^2, 0) \quad \textbf{(13)}$$

and a maximum likelihood can be sought.

[0014] Equation (13) can be computationally intensive, and, instead of using that equation, the motion parameters t_Z, w_X and w_Y can be determined directly from the images by combining the geometric constraints embodied in equation (11) with the photometric constraints embodied in equation (12). In that operation, given two consecutive images Ψ and Ψ', the goal is to determine the probability

IPR2025-00944
Tesla EX1002 Page 1010

$$P\left(\hat{m}|\Psi,\Psi'\right) \quad (14).$$

of that the motion of the vehicle 10 is $\hat{m} = (t_z, w_x, w_y)$ given the two images. The motion m that maximizes (14) is the estimate of the camera 13 motion between the two images Ψ and Ψ', and, thus, the estimate of vehicle 10 motion between the two images.

[0015] According to Bayes' rule,

$$P\left(\hat{m}|\Psi,\Psi'\right) = \frac{P\left(\Psi'|\Psi,\hat{m}\right)P\left(\hat{m}\right)}{P(\Psi')} \quad (15),$$

where $P(\Psi'|\Psi,\hat{m})$ is the probability that, given image Ψ, motion $\hat{m}$ will result in image Ψ', $P(\hat{m})$ is the *a priori* probability that the motion is $\hat{m}$, and $P(\Psi')$ is the *a priori* probability that the image is Ψ'. It will be assumed that $P(\hat{m})$, the probability that the motion is $\hat{m}$, is uniform in a small region $\hat{M}$ around the previous motion estimate, that is, the motion estimate generated as between the "i-1 st" and "i-th" images; it will be appreciated that, if the time period between the times at which images Ψ and Ψ' are recorded is sufficiently small, this assumption will hold. It will further be appreciated that, in equation (15), the denominator $P(\Psi')$ does not depend on the motion $\hat{m}$, and so it does not affect the search for a maximum.

[0016] The probability that, given image Ψ, motion $\hat{m}$ will result in image Ψ', $P(\Psi'|\Psi,\hat{m})$, can be determined by warping the image Ψ' according to the motion $\hat{m}$, thereby to generate a warped image $\hat{\Psi}'$ and determining the sum squared difference ("SSD")

$$S\left(\hat{m}\right) = \frac{1}{N}\sum_{x,y\in R}\left(\hat{\Psi}'(x,y) - \Psi(x,y)\right)^2 \quad (16),$$

between corresponding patches, or regions, R in the images that are believed to be a projection of the roadway in the two images. In equation 16, "N" is the number of points in the region R. It will be appreciated that, if the images Ψ and Ψ' are recorded at times "t" and "t+δt, respectively, the warped image $\hat{\Psi}'$ will represent the image that is assumed would be recorded at time "t" if the motion is $\hat{m}$. Using the SSD criteria (equation 16), $P(\Psi'|\Psi,\hat{m})$, the probability that image Ψ' would result from image Ψ and a motion $\hat{m}$ is given by the probability density function

$$P\left(\Psi'|\Psi,\hat{m}\right) = ce^{-\frac{S\left(\hat{m}\right)}{\sigma^2}} \quad (17),$$

where "c" is a normalization factor and "σ" is the variance of the noise, which is modeled as a zero mean Gaussian function. Since it is assumed that $P(\hat{m})$, the probability that the motion is $\hat{m}$, is uniform in a small region $\hat{M}$ around a previous motion estimate, the problem of finding the maximum likelihood motion $\hat{m}$ for a patch of image Ψ reduces to finding the maximum of the probability density function

$$P\left(\hat{m}|\Psi,\Psi'\right) = ce^{-\frac{S\left(\hat{m}\right)}{\sigma^2}} \quad (18),$$

for motion $\hat{m} \in \hat{M}$.

[0017] Since the motion for which the ego-motion estimation system processor 14 is to generate an estimate is the

IPR2025-00944
Tesla EX1002 Page 1011

translational and rotational motion ofthe vehicle 10 relative to the road, it is desirable for the ego-motion estimation system processor 14 to consider only regions of the images Ψ and Ψ' that comprise projections of the road, and ignore other regions of the images. However, it should be noted that the set R of regions, or patches, of the images that projections of the roadway in the two images Ψ and Ψ', is not known. To accommodate that, instead of attempting to maximize the function defined in equation (18), the image can be tessellated into a set of patches W_i, and a probability density $P(m|W_i,W_i')$ generated for each patch using equations (16) and (18) for the respective patches. The probability density over the entire pair of images Ψ and Ψ' will be

$$P\left(\hat{m}|\Psi,\Psi'\right) = c\frac{\sum_i P\left(\hat{m}|W_i,W_i'\right)\lambda_i\,\alpha_i\,\beta_i}{\sum_i \lambda_i} \qquad (19),$$

where λ_i and α_i are weighting functions whose values generally reflect the confidence that the "i-th" patch is a projection of the road. The value of the gradient strength β_i for a patch reflects the degree to which the patch the contains a texture, and thus will more likely to contain useful information for use in determining ego motion of the vehicle. The motion $\hat{m} \in \hat{M}$ for which equation (19) is the maximum will be deemed to correspond to the actual translational and rotational motion of the vehicle 10 as between the locations at which images Ψ and Ψ' were recorded.

[0018] The weighting function λ_i for the respective "i-th" patch is generated using patches W_i and W'_i from respective images Ψ and Ψ'. In determining the values for the weighting functions λ_i, for patches W_i, W'_i that are not of the road, the motion model reflected in equation (11) is not a good fit; instead, a better fit can be obtained using some other motion of the patch. In addition, for planar image artifacts moving on the roadway surface, such as moving shadows, the maximum of equation (18) will occur far away from the initial guess. Accordingly, the value of the weighting function λ_i for the "i-th" patch W_i, W_i' will correspond to the ratio between the best fit using the motion model in a local region ($\hat{M}$) near the initial guess and the best fit using any motion model over large search region "L." Accordingly, if

$$P_1 = \max\left(\exp\left(-\frac{S_i\left(\hat{m}\right)}{\sigma^2}\right)\right) \qquad (20)$$

for all $\hat{m} \in \hat{M}$ is the value for the best fit in a local search region, where $S_i()$ denotes the SSD over all points in the "i-th" patch, and

$$P_2 = \max\left(\exp\left(-\frac{S_i\left(\hat{m}\right)}{\sigma^2}\right)\right) \qquad (21)$$

for all $\hat{m} \in L$ the value for the best fit over all possible image motions, then

$$\lambda_i = \frac{P_1}{P_2} \qquad (22).$$

Generally, generating the value for P_2 (equation 21) can be computationally intensive. To avoid generating P_2 according to equation (21), the value for P_2 for each patch can be estimated by using the SSD as between a patch in the image

IPR2025-00944
Tesla EX1002 Page 1012

Ψ and the correspondingly-positioned patch in the image Ψ', as well as the SSD's as between the patch in the image Ψ and patches translated horizontally and vertically around the correspondingly-positioned patch in the image Ψ', for a selected number of points. That is, if the patch in image Ψ consists of points p(x,y) centered on p(a,b) (that is, points p(a-α, b-β) through p(a+α,b+β) (α and β being integers), with the patch being of dimensions $2\alpha+1$ by $2\beta+1$), P_2 is generated by using the SSD as between the patch of the same size in image Ψ' consisting of points p(x,y) centered on p(a,b), as well as SSD's as between the patch in image Ψ and patches of the same size in image Ψ' that are centered on points p(a-δ, b-δ) through p(a+δ, b+δ), a total of $(2\delta+1)^2$ patches in image Ψ'. Each patch in image Ψ' can be considered as one of the possible image motions. In one embodiment, δ is selected to be seven, in which case there will be two hundred and twenty five patches in Ψ' for which the SSD will be generated in generating the value for P_2.

[0019] Unlike weighting function λ_i, weighting function α_i for the respective "i-th" patch is generated using only patch W_i and W'_i from image Ψ'. Generally, it should be appreciated that, in three-dimensional space, there are three dominant directions for lines, namely, (i) vertical, for buildings, automobiles, and other objects that extend above the roadway surface, (ii) horizontal lines that are generally parallel to the direction of the roadway, and (iii) horizontal lines that are generally perpendicular to the direction of the roadway. In an image Ψ', vertical lines (that is, lines of type (i)) and horizontal lines that are generally perpendicular to the direction of the roadway (that is, lines of type (iii)) will project in to image Ψ' as vertical and horizontal lines, respectively. On the other hand, horizontal lines that are generally parallel to the direction of the roadway (that is, lines of type (ii)) will appear in image Ψ' as lines that pass through a common point, that is, a vanishing point. It will be appreciated that patches W'_i in image Ψ' that are projections of the roadway will predominately contain lines of type (ii) and (iii). On the other hand, patches W' that are projections of obstacles, such as automobiles will predominately contain lines of type (i) and (iii), while patches W' that are projections of, for example, buildings, fences, and the like, will contain lines of type (i) and (ii).

[0020] Accordingly, the value for weighting function α_i for patch W_i will reflect the degree to which it is deemed to contain projections of lines of type (ii) and (iii), and not projections of lines of types (i) and (iii) or types (i) and (ii). Generally, the directions of lines, if any, passing though a patch can be determined in relation to the gradients of the luminance at the various points in the patch W_i. Each point in the patch W'_i whose gradient (I_x,I_y) is above a selected threshold is considered to lie at or near a line, with the direction of the line being perpendicular to the direction of the gradient. Thus, for those points, the direction of the line associated therewith can be determined, as can whether the line is of type (i), (ii) or (iii). Thus, for each patch W'_i in the image Ψ', three sums S^n_i (n=i, ii or iii) are generated each corresponding to the sum of the magnitudes of the gradients of the points in the patch that are associated with lines of the respective type, the magnitude corresponding to $G=(I_x^2+I_y^2)^{0.5}$. A patch W'_i in image Ψ' is deemed to be:

(a) a projection of the roadway if the sums S^{ii}_i and S^{iii}_i are both large and significantly larger than sum S^i_i, since the sums indicate that the patch is associated with a line or lines that are horizontal and in the direction ofthe roadway and/or perpendicular thereto, but not a line or lines that are vertical;

(b) be a projection of an obstacle, that is, an object generally in the path of the vehicle, if sums S^i_i and S^{iii}_i are both large and significantly larger than sum S^{ii}_i, since the sums indicate that the patch is associated with a line or lines that are vertical and/or horizontal and perpendicular to the direction of the roadway, but not a line or lines that are horizontal and in the direction of the roadway; and

(c) a projection of an object to the side of the path of the vehicle if sums S^i_i and S^{ii}_i are both large and significantly larger than sum S^{iii}_i, since the sums indicate that the patch is associated with a line or lines that are vertical and/or horizontal and in the direction of the roadway.

The value of the weighting function α_i is assigned to the patch based on the degree to which the patch W_i is deemed to be a projection of the roadway (case (a) above).

[0021] It will be appreciated that, if, for a patch W'_i, the sum S^{ii}_i is relatively large, indicating that the patch is associated with a line that is horizontal and in the direction of the roadway, but sums S^i_i and S^{iii}_i are relatively small, indicating that the patch is not associated with a line that is vertical or horizontal and perpendicular to the direction of the roadway, it generally cannot be determined from the set of sums generated for the patch whether the patch is a projection of the roadway (case (a) above) or a projection of an object to the side of the path of the vehicle (case (c) above). However, since the patch is not associated with a line that is vertical, it will generally not be deemed to be a projection of an obstacle (case (b) above). In that case, an assessment as to whether the patch is a projection of the roadway (case (a) above) or a projection of an object to the side of the path of the vehicle (case (c) above) by referring to patches adjacent thereto.

[0022] As noted above, the value of the gradient strength β_i for a patch reflects the degree to which the patch the contains a texture, and thus will more likely to contain useful information for use in determining ego motion of the vehicle. The gradient strength β_i corresponds to

IPR2025-00944
Tesla EX1002 Page 1013

$$\beta_i = \left(\sum_{\hat{m} \in L} \exp\left(- \frac{S_i(\hat{m})}{\sigma^2} \right) \right)^{-1} \quad (23).$$

For relatively uniform patches, the value of the SSD $S_i(\hat{m})$ will be relatively low for all motions, in which case the value of β_i will be relatively low. On the other hand, for patches with texture, the value of the SSD will be relatively high for most motions, in which case the value of β_i. will be relatively high.

[0023] With this background, operations performed by the ego-motion estimation system processor 14 will be describe in connection with the flow chart depicted in FIG. 2. In connection with FIG. 2, it is assumed that the ego-motion estimation system processor 14 already has image Ψ, which it may have used in connection with determining the translational and rotational motion up to the location at which image Ψ was recorded. With reference to FIG. 2, after the ego-motion estimation system processor 14 has received image Ψ' (step 100), it will rectify the image according to information provided during the camera 13 calibration operation (described below) to provide that the optical axis is parallel to the plane defined by the roadway (step 101). In addition, the ego-motion estimation system processor 14 will generate an initial guess as to the translational and rotational motion, using the previous motion estimate and, perhaps information from other sensors if available (step 102). For example, the ego-motion estimation system processor 14 may make use of information from the vehicle 10's speedometer, as well as information as to the time period between the time at which image Ψ was recorded and the time at which image Ψ' was recorded, in generating the initial guess. Generally, it will be appreciated that the time period will be fixed, and will preferably the same for each successive pair of images Ψ and Ψ'. After the ego-motion estimation system processor 14 has generated the initial guess, it will use the initial guess to warp image Ψ' toward image Ψ, thereby to generate a warped image $\hat{\Psi}'$ (step 103).

[0024] After the ego-motion estimation system processor 14 has generated the warped image (step 103), it will divide the image Ψ and the warped image $\hat{\Psi}'$ into patches and, for each pair of corresponding patches in the two images Ψ and $\hat{\Psi}'$, generate the weighting value λ_i. In that operation, the ego-motion estimation system processor 14 will select a patch in the image Ψ (step 104) and generate values for P_2 (step 105), P_1 (equation 20) (step 106) and λ_i (equation 22) (step 107) as described above. In addition, the ego-motion estimation system processor 14 can generate the value for β_i (equation 23) and α_i (step 108). After the ego-motion estimation system processor 14 has generated performed steps 105 through 108 for the selected patch, it will determine whether all of the patches in image Ψ have been processed (step 109) and if not, return to step 104 to select another patch and perform steps 105 through 109 in connection therewith.

[0025] The ego-motion estimation system processor 14 will perform steps 104 through 109 in connection with each patch in the image Ψ. After the ego-motion estimation system processor 14 has performed steps 104 through 109 in connection with all of the patches in the image Ψ, it will sequence from step 109 to step 110 to search for the motion $\hat{m}$ that maximizes the value provided by equation (19) (step 110). That motion $\hat{m}$ will comprise values for translation t_z and rotation w_x, w_Y parameters that will constitute the estimate of the motion of the vehicle 10 as between the point in time at which image Ψ was recorded and the point in time at which image Ψ' is recorded. The ego-motion estimation system processor 14 can perform operations described above in connection with each successive pair of images Ψ and Ψ' to estimate the motion of the vehicle 10.

[0026] In performing steps 106 (to generate the values for P_1) and 110 (to determine the motion $\hat{m}$ that maximizes the value provided by equation (19)), the ego-motion estimation system processor 14 can perform a gradient descent that is limited to a selected cube-shaped region around the initial guess.

[0027] In determining the initial guess (step 102) for each new image Ψ', the ego-motion estimation system processor 14 can use the estimate of the motion generated for the previously-received image. In addition, the size of the region $\hat{M}$ can be adjusted adaptively.

[0028] A second methodology for determining the ego-motion of a vehicle 10 will be described in connection with FIG. 3. As with the methodology described above in connection with FIG. 2, it is assumed that the images have been rectified so that the image planes are perpendicular to the plane represented by the roadway on which the vehicle 10 is traveling, and their horizontal ("x") axes are parallel to the plane represented by the roadway. By way of background, in that case, the equation of a plane

$$AX + BY + CZ = 1 \quad (24)$$

(reference equation (3)) reduces to

IPR2025-00944
Tesla EX1002 Page 1014

$$BY = 1 \qquad (25),$$

[0029] 1 in which case $Y = \frac{1}{|B|}$ is the height of the optical axis of the camera 13 (or, more specifically, the Z axis) above the road. Since, for a point *p(x,y)* in an image that is a projection of a point *P(X,Y,Z)* in three-dimensional space, $y = f\frac{Y}{Z}$, equation (25) becomes

$$\frac{1}{Z} = by \qquad (26)$$

The brightness constraint is

$$uI_x + vI_y + I_t = 0 \qquad (27)$$

for each point, where, at each point (x,y) in the image, I_x and I_y are the horizontal and vertical components of the spatial gradient of the luminance and I_t is the time gradient of the luminance. In addition, the equations for the components *(u, v)* of the flow vector (reference equation (1)) can be written

$$u = \frac{1}{Z}S_1^T t + S_1^T [w]_X p$$
$$v = \frac{1}{Z}S_2^T t + S_2^T [w]_X p \qquad (28)$$

where $[w]_x$ is a skew-symmetric matrix, in which case

$$\frac{1}{Z}S^T t + S^T [w]_X p + I_t = 0 \qquad (29),$$

where

$$S = \begin{pmatrix} fI_x \\ fI_y \\ -xI_x - yI_y \end{pmatrix} \qquad (30).$$

For motion constrained to a plane, equation (29) reduces to

IPR2025-00944
Tesla EX1002 Page 1015

$$(ax + by + c) S^T t + S^T [w]_X p + I_t = 0 \quad (31).$$

Since the images are rectified, equation (26) will hold. In addition, since rotational motion is constrained so as to occur only around the vertical ("Y") axis, $w_x=w_z=0$. Accordingly, equation (31) will correspond to

$$byS^T t + S^T \begin{pmatrix} 0 & 0 & -w_Y \\ 0 & 0 & 0 \\ -w_Y & 0 & 0 \end{pmatrix} \begin{pmatrix} x/f \\ y/f \\ 1 \end{pmatrix} + I_t = 0 \quad (32).$$

Expanding equation 32, and using equation (30),

$$byS^T t - (fI_x + xI_z + yI_y) w_Y + I_t = 0 \quad (33)$$

Generally, the portion u_{wY} of the vertical component of the flow vector that arises due to rotation (w_Y) of the vehicle 10 will be larger than the portion u_{tX} that arises due to translation (t_X) along the "X" axis, and the portions will have a very different form. The portion u_{tX} that arises due to translation along the "X" axis will depend on the distance that the vehicle 10 moves between the times the images are recorded, and will be larger for points towards the bottom of the image than for points further up. In that case, expressions for u_{wY} and u_{iX} are

$$u_{wY} = \left(\frac{x^2}{f} + f \right) w_Y \approx f w_Y \quad (34)$$

$$u_{tX} = \frac{1}{Z} f t_X = byf t_X \quad (35)$$

Since, for small rotations around the vertical "Y," axis, the angle of rotation w_Y is approximately

$$w_Y \approx \frac{t_X}{t_Z} \quad (36),$$

$$\frac{u_{wY}}{u_{tX}} = \frac{f \frac{t_X}{t_Z}}{\frac{1}{Z} f t_X} = \frac{Z}{t_Z} \quad (37)$$

It will be appreciated that values for t_Y, the component of the translation t in the vertical direction, and w_X and w_Z, the X and Z components of rotation w, will be zero. Accordingly, after the ego-motion estimation system processor 14 receives

IPR2025-00944
Tesla EX1002 Page 1016

a new image Ψ', it will determine the values for t, and t_3, the components of the translation t in the forward (along the Z axis) and side (along the X axis) directions, and w_Y, the component of rotation around the vertical (Y) axis. In that operation, the ego-motion estimation system processor 14 will generate an initial estimate as to the motion (step 150, FIG. 3) and use that initial estimate to generate a warped image $\hat{\Psi}'$ (step 151). The ego-motion estimation system processor 14 can use information from a number of sources in connection with generating the initial estimate (step 150), including information from, for example, the vehicle 10's speedometer. Thereafter, the ego-motion estimation system processor 14 divides the image Ψ and $\hat{\Psi}'$ into patches (step 152) and determines which patches are likely to comprise images of the roadway (step 153). In determining which patches are likely to comprise images of the roadway (reference step 153), the ego-motion estimation system processor 14 can generate an SSD (equation 16) for corresponding patches in the images Ψ and $\hat{\Psi}'$ and determine that patches in the warped image $\hat{\Psi}'$ that comprise images of the roadway will be those patches with a relatively high SSD value. The ego-motion estimation system processor 14 then uses the patches identified in step 153 to minimize a cost function of the form

$$\min_{t} \sum_{x,y \in R} |At - b|^2 + \lambda_1 \sum_i \left| W_1 \left(\frac{dt_i}{dT} \right) \right|^p + \lambda_2 \left| W_2 \left(t - t_0 \right) \right|^p \qquad (38),$$

where W_1 and W_2 are weighting matrices that essentially describe confidence in the priors or smoothness values (step 154). Equation (38) can be formalized in the form of a Kalman filter, and the value of "p" can be selected to be one or two depending on whether the L_1 or L_2 norm is to be used.

[0030] As noted above, for methodologies described above in connection with both FIGS. 2 and 3, the ego-motion estimation system processor 14 will initially rectify the images as received from the camera 13. In the above description, the images I and I' are images as rectified by the ego-motion estimation system processor 14. To rectify the images as received from the camera 13, the camera 13 will need to be calibrated during a calibration operation prior to use in connection with recording images for use in estimating vehicle 10 motion as described above. Before describing operations to be performed during calibration, it would be helpful to consider the effects of incorrect calibration. If, for example, a vehicle is driving along a straight road, and if the optical axis of the camera is aligned with the direction of motion, the flow field of successively recorded images will be an expansion field with the focus of expansion located at the center of the respective images, that is, at the origin (x,y)=(0,0) of the image's rectilinear coordinate system. On the other hand, if the camera is mounted on the vehicle with a small rotation around the vertical ("Y") axis in three-dimensional space, then the focus of expansion will be displaced along the image's horizontal ("x") axis. In that case, the motion model defined by equation (11) will not account for the flow field, but will be well approximated by a forward translation and a rotational velocity w_y around the vertical ("Y") axis.

[0031] Accordingly, errors in the orientation of the camera around the vertical ("Y") axis in three dimensional space will create a bias in the rotation estimate, in which case a curved path would be estimated instead of a straight path. Similarly, errors in the camera's orientation around the horizontal ("X") axis in three dimensional space will cause a bias in the pitch estimate. Based on these observations, a calibration operation can be performed by having the camera record a sequence of images while the vehicle is being driven down a straight roadway. The vehicle's ego-motion is estimated as described above in connection with FIGS. 2 or 3, and calibration parameters are estimated that would cause the ego-motion to integrate into a straight path.

[0032] The invention provides a number of advantages. In particular, the invention provides an arrangement for determining ego-motion of a vehicle 10 on a roadway from a series of images recorded by a camera 13 mounted on the vehicle 10, at least a portion of the images comprising projections of the roadway, and without requiring mechanical sensors which are normally not provided with a vehicle 10 and that would, if provided, increase the cost and maintenance expenses thereof.

[0033] It will be appreciated that a system in accordance with the invention can be constructed in whole or in part from special purpose hardware or a general purpose computer system, or any combination thereof, any portion of which may be controlled by a suitable program. Any program may in whole or in part comprise part of or be stored on the system in a conventional manner, or it may in whole or in part be provided in to the system over a network or other mechanism for transferring information in a conventional manner. In addition, it will be appreciated that the system may be operated and/or otherwise controlled by means of information provided by an operator using operator input elements (not shown) which may be connected directly to the system or which may transfer the information to the system over a network or other mechanism for transferring information in a conventional manner.

[0034] The foregoing description has been limited to a specific embodiment of this invention. It will be apparent, however, that various variations and modifications may be made to the invention, with the attainment of some or all of the advantages of the invention. It is the object of the appended claims to cover these and such other variations and

IPR2025-00944
Tesla EX1002 Page 1017

modifications as come within the true spirit and scope of the invention.

Claims

1. A method of controlling a vehicle moving along a roadway, the method comprising:

   acquiring first and second camera images of the roadway at different times;
   providing a model that models the ego-motion of the vehicle between the times that the first and second images are acquired that is a function of only three motion parameters;
   determining values for the three motion parameters that maximize a probability function, which indicates for a given set of values for the motion parameters how probable the first image warps into the second image; and
   controlling the vehicle responsive to the determined variables.

2. A method according to claim 1 wherein controlling the vehicle comprises using an autonomous driving controller that controls the vehicle responsive to the variables.

3. A method according to claim 1 wherein controlling the vehicle comprises controlling the vehicle responsive to a computer vision-based driver assistance that generates driver assistance cues responsive to the determined variables.

4. A method according to any of claims 1-3 wherein the motion parameters comprise a translational velocity parallel to the plane of the roadway and collinear with a direction of motion of the vehicle on a straight roadway, an angular velocity perpendicular to the translational velocity and parallel to the plane of the roadway and an angular velocity perpendicular to the translational velocity and the roadway plane.

5. A method according to any of claims 1-3 wherein the motion parameters comprise a translational velocity parallel to the plane of the roadway and collinear with a direction of motion of the vehicle, on a straight roadway, a second translational velocity parallel to the plane of the roadway and perpendicular to the first translational velocity and an angular velocity perpendicular to the plane of the roadway.

6. A method according to any of the preceding claims wherein the camera has an optic axis and comprising rectifying the images so that the rectified images are acquired as it the optic axis of the camera is parallel to the translational velocity collinear the direction of motion of the vehicle on a straight roadway.

7. A method according to any of the preceding claims wherein the probability function is responsive to a plurality of probability functions, each of which indicates for a given set of values for the motion parameters, how probable a different region of the first image warps into a corresponding region in the second image; and

8. A method according to claim 7 and comprising tessellating the images to determine the different corresponding regions.

9. A method according to claim 7 or claim 8 and comprising weighting the probability functions of the different corresponding regions.

10. A method according to claim 9 wherein weighting comprises weighting responsive to features in the regions.

11. A method according to claim 9 or claim 10 wherein weighting comprises weighting responsive to texture of the regions.

12. A method according to any of claims 9-11 wherein weighting comprises weighting the probability functions for corresponding regions in the first and second images responsive to a measure as to how well the region in the second image warps to its corresponding region in the first image for an ego-motion dependent on only three motion parameters.

13. Apparatus for controlling a vehicle moving along a roadway comprising:

   a camera (13) that acquires images of the roadway at different times;
   a processor (14) that receives images acquired by the camera and:

IPR2025-00944
Tesla EX1002 Page 1018

models the ego-motion of the vehicle between the times that the images are acquired as a function of only three motion parameters;
determines values for the three motion parameters that maximize a probability function, which indicates for a given set of values for the motion parameters how probable the first image warps into the second image; and
controls the vehicle responsive to the determined variables and/or generates driver assistance cues to assist the vehicles driver.

IPR2025-00944
Tesla EX1002 Page 1019

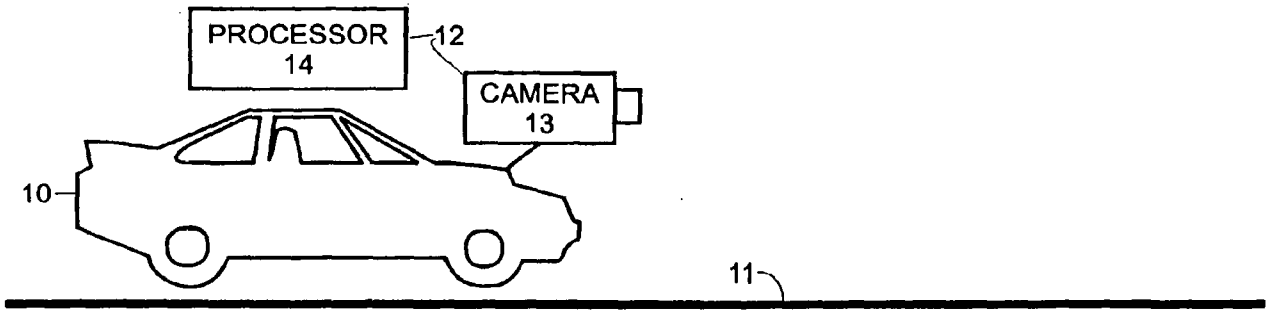
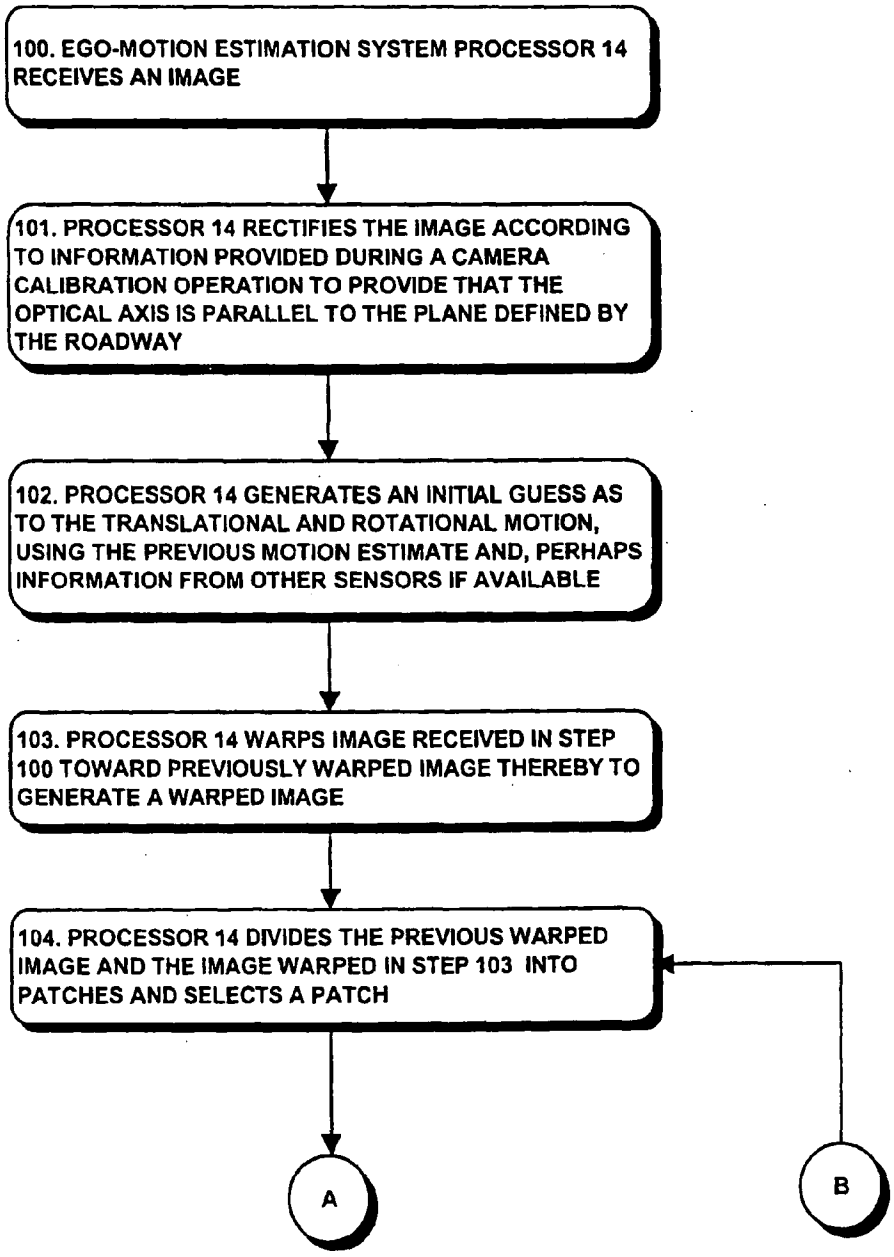


FIG. 1

IPR2025-00944
Tesla EX1002 Page 1020

FIG. 2



IPR2025-00944
Tesla EX1002 Page 1021

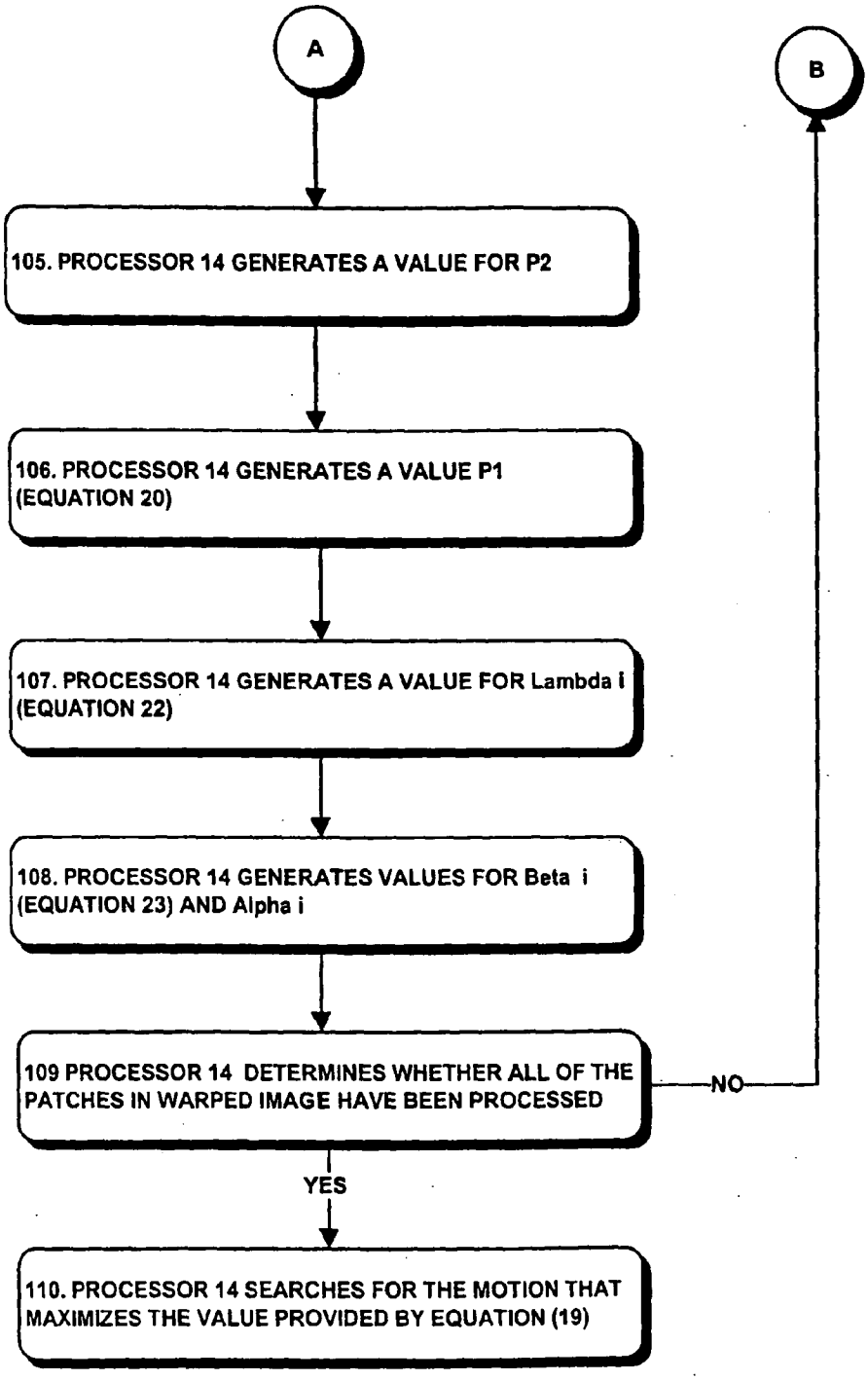
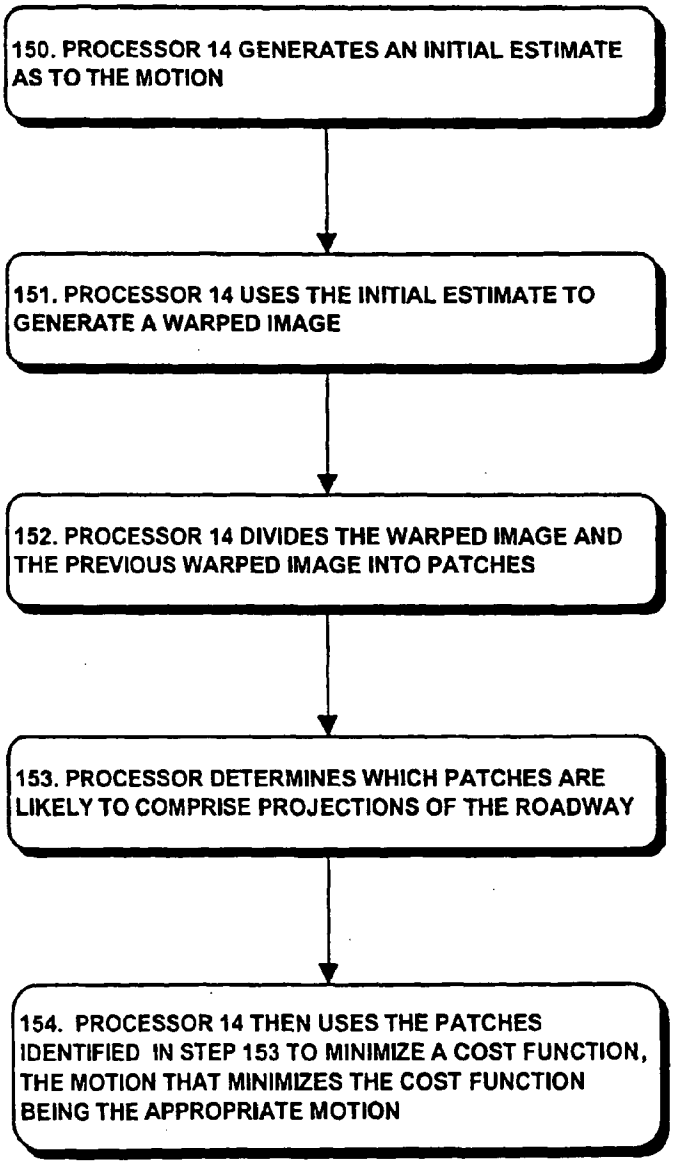


FIG. 2A

IPR2025-00944
Tesla EX1002 Page 1022

FIG. 3



IPR2025-00944
Tesla EX1002 Page 1023

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 727 089 A3

(12) EUROPEAN PATENT APPLICATION

(88) Date of publication A3:
19.09.2007 Bulletin 2007/38

(43) Date of publication A2:
29.11.2006 Bulletin 2006/48

(21) Application number: **06017695.5**

(22) Date of filing: **27.11.2000**

(51) Int Cl.:
G06T 7/20 *(2006.01)* ***G05D 1/02*** *(2006.01)*

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE TR

(30) Priority: **26.11.1999 US 167587 P**
01.09.2000 US 230166 P

(62) Document number(s) of the earlier application(s) in accordance with Art. 76 EPC:
00980706.6 / 1 257 971

(71) Applicant: **MobilEye Technologies, Ltd.**
1034 Nicosia (CY)

(72) Inventors:
- **Stein, Gideon**
  Jerusalem 96263 (IL)
- **Shashua, Amnon**
  Mevaseret Zion 90805 (IL)
- **Mano, Ofer**
  Modiin 71700 (IL)

(74) Representative: **Skone James, Robert Edmund**
Gill Jennings & Every LLP
Broadgate House
7 Eldon Street
London EC2M 7LH (GB)

(54) **System and method for estimating ego-motion of a moving vehicle using successive images recorded along the vehicle's path of motion**

(57) An ego-motion determination system is disclosed for generating an estimate as to the ego-motion of a vehicle moving along a roadway. The ego-motion determination system includes an image information receiver and a processor. The image information receiver is configured to receive image information relating to a series of at least two images recorded as the vehicle moves along a roadway. The processor is configured to process the image information received by the image receiver to generate an ego-motion estimate of the vehicle, including the translation of the vehicle in the forward direction and the rotation of the vehicle around a vertical axis as between, for example, successive images.

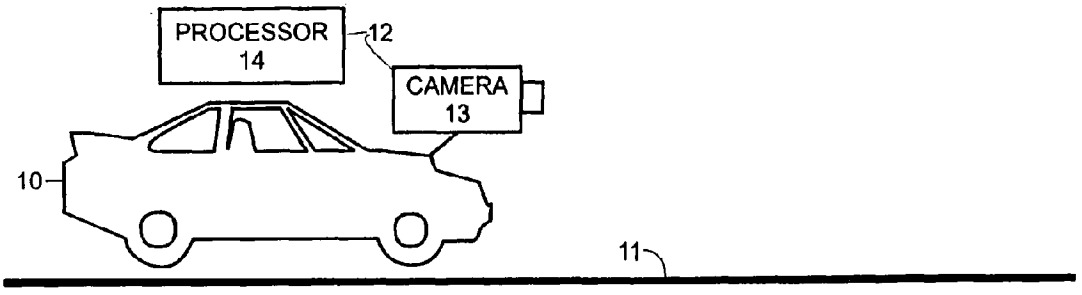


FIG. 1

EP 1 727 089 A3

IPR2025-00944
Tesla EX1002 Page 1024

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 06 01 7695

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
A	US 5 259 040 A (HANNA ET AL) 2 November 1993 (1993-11-02) * abstract; figures 1,2 * * column 4, line 33 - column 6, line 10 *	1-13	INV. G06T7/20 G05D1/02
A	EP 0 390 051 A (HONEYWELL INC) 3 October 1990 (1990-10-03) * abstract; figure 1 * * page 9, line 7 - page 11, line 8 *	1-13	
A	US 5 473 364 A (BURT PETER J [US]) 5 December 1995 (1995-12-05) * column 3, line 61 - column 5, line 19; figures 2,3 *	1-13	

TECHNICAL FIELDS SEARCHED (IPC)
G08G H04N G06T B60K G05D G01S

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
Munich	14 August 2007	Schinnerl, Arnold

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

IPR2025-00944
Tesla EX1002 Page 1025

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 06 01 7695

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

14-08-2007

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5259040 A	02-11-1993	DE 69231812 D1	07-06-2001
		DE 69231812 T2	15-11-2001
		EP 0606385 A1	20-07-1994
		JP 6511339 T	15-12-1994
		JP 3253962 B2	04-02-2002
		WO 9307585 A1	15-04-1993
EP 0390051 A	03-10-1990	JP 3048977 A	01-03-1991
		US 4969036 A	06-11-1990
US 5473364 A	05-12-1995	JP 10501386 T	03-02-1998
		WO 9534171 A1	14-12-1995

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

IPR2025-00944
Tesla EX1002 Page 1026

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 2 605 185 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
19.06.2013 Bulletin 2013/25

(51) Int Cl.:
G06K 9/00 (2006.01)

(21) Application number: **12275201.7**

(22) Date of filing: **12.12.2012**

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR
Designated Extension States:
BA ME

(30) Priority: **12.12.2011 US 201161569306 P**
19.09.2012 US 201261702755 P

(71) Applicant: **Mobileye Technologies Limited**
24854 Nicosia (CY)

(72) Inventor: **Stein, Gideon P**
96263 Jerusalem (IL)

(74) Representative: **Jehan, Robert**
Williams Powell
Staple Court
11 Staple Inn Buildings
London, WC1V 7QH (GB)

(54) Detection of obstacles at night by analysis of shadows

(57) A driver assistance systems mountable in a host vehicle while the host vehicle is moving forward with headlights on for detection of obstacles based on shadows. The driver assistance system includes a camera (12) operatively connectible to a processor (30). A first image frame and a second image frame are captured of a road. A first dark image patch and a second dark image patch include intensity values less than a threshold value. The first and the second dark image patches are tracked from the first image frame to the second image frame as corresponding images of the same portion of the road. Respective thicknesses in vertical image coordinates are measured of the first and second dark image patches responsive to the tracking.

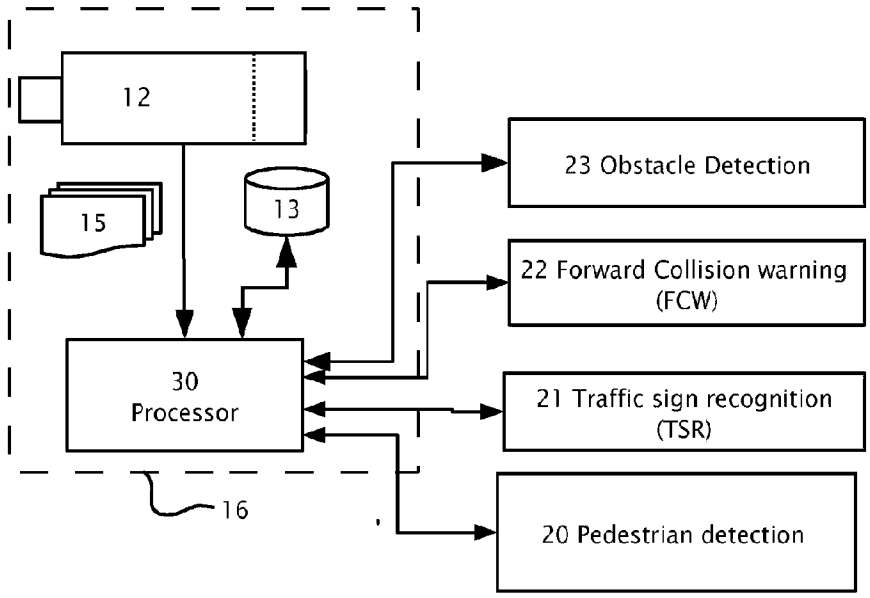


Fig. 2

EP 2 605 185 A2

IPR2025-00944
Tesla EX1002 Page 1027

Description

BACKGROUND

1. Technical Field

[0001] The present invention relates to the field of driver assistance systems (DAS) and in particular to detection of obstacles in a road using a camera.

2. Description of Related Art

[0002] Larson et al. [1] perform detection of negative obstacles using LIDAR and other three dimensional data. Belutta et al. [2] detect depth discontinuities of negative obstacles from stereo range data. Rabkin et al. [3] use stereo range data on thermal cameras because the shape of the terrain also affects the thermal properties. Muraka et al. [4] combine stereo and motion information to detect discontinuities in daytime. They assume a piecewise planar model and show results from indoor and outdoor environment. The range of detection of drop-offs is limited to a few meters since texture is required for stereo and motion correspondence. Witus et al. [5] analyze the shadows produced by light sources mounted on the host vehicle using a stereo camera. Detection range is limited by camera focal length and stereo baseline. Furthermore, depth information on the shadow edge is only available when the light sources are displaced along the stereo baseline. In contrast, a typical stereo camera for driver assistance applications has the two cameras mounted side by side and the headlights are below the cameras.

References

[0003]

1. Larson et al., "Lidar based off-road negative obstacle detection and analysis" In 14th International IEEE Conference on Intelligent Transportation Systems (ITSC), Oct. 2011

2. Belutta et al., "Terrain Perception for Demo III", In Proceedings of the IEEE Intelligent Vehicles Symposium (IV) 2000, October 2000

3. Rabkin et al. "Nighttime negative obstacle detection for off-road autonomous navigation In Unmanned Systems", Technology IX. Edited by Gerhart, Grant R.; Gage, Douglas W.; Shoemaker, Charles M.. Proceedings of the SPIE, Volume 6561, pp. 656103 (2007)

4. Muraka "Detecting obstacles and drop-offs using stereo and motion cues for safe local motion", In IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2008 Sept. 2008

5. Witus et al. Preliminary investigation into the use of stereo illumination to enhance mobile robot terrain perception In Proc. SPIE Vol. 4364, p. 290-301, Unmanned Ground Vehicle Technology III, Grant R. Gerhart; Chuck M. Shoemaker; Eds.]

BRIEF SUMMARY

[0004] The present invention seeks to provide an improved driver assistance system and method.

[0005] According to an aspect of the present invention, there is provided a computerized method performable at night by a driver assistance system mountable in a host vehicle as specified in claim 1.

[0006] According to another aspect of the present invention, there is provided a driver assistance system as specified in claim 11.

[0007] Various driver assistance systems and corresponding methods are provided for herein performable at night for detection of obstacles based on shadows. The driver assistance systems are mountable in a host vehicle while the host vehicle is moving forward with headlights on. The driver assistance systems include a camera operatively connectible to a processor. A first image frame and a second image frame are captured of a road in the field of view of the camera. The first and the second image frames are processed to locate a first dark image patch of the first image frame and a second dark image patch of the second image frame. The first image frame and the second image frame may be filtered with a threshold value of gray scale to produce the first dark image patch of the first image frame and the second dark image patch of the second image frame. The first dark image patch and the second dark image patch include intensity

IPR2025-00944
Tesla EX1002 Page 1028

values less than the threshold value. In another embodiment, the processing to find the first and second dark image patches may include searching for a transition from light to dark (e.g. high gray scale to low gray scale values) and another transition from dark to light. The first and the second dark image patches are tracked from the first image frame to the second image frame as corresponding images of the same portion of the road. Respective thicknesses in vertical image coordinates are measured of the first and second dark image patches responsive to the tracking performed from the first image frame to the second image frame.

[0008] Prior to the tracking, connectivity of picture elements may be performed within the first and second dark image patches. It may be determined that a change of the thicknesses between the first and second dark image patches is consistent with a shadow cast by an obstacle in the road from illumination from the headlights and an obstacle in the road is detected. A driver of the host vehicle may be audibly warned responsive to the detection of the obstacle.

[0009] The change in the thicknesses between the first and second image frames may be analyzed to distinguish between the detected obstacle being a positive or a negative obstacle and/ or a shape of the detected obstacle may be computed.

[0010] It may be determined that a change of the thicknesses between the first and second image frames is not consistent with a shadow cast by an obstacle from illumination from the headlights. The thicknesses may be measured in picture elements according to a total number of picture elements with intensity values less than the threshold divided by a width in picture element columns of the .first and second dark image patches.

[0011] The foregoing and/or other aspects will become apparent from the following detailed description when considered in conjunction with the accompanying drawing figures.

BRIEF DESCRIPTION OF THE DRAWINGS

[0012] Embodiments of the present invention are described below, by way of example only, with reference to the accompanying drawings, wherein:

Figures 1 and 2 illustrate a system including a camera or image sensor mounted in a host vehicle, according to an embodiment of the present invention.

Figures 3a and 3b show a simplified line drawings of a night time driving scenario, according to an embodiment of the present invention.

Figures 4a and 4b show image frames of a speed bump where a host vehicle is at different distances from the speed bump to illustrate an embodiment of the present invention.

Figures 4c and 4d show a portion of image frames as shown respectively in Figures 4a and 4b in greater detail to illustrate an embodiment of the present invention.

Figure 5a shows an example of a shadow produced , according to an embodiment of the present invention.

Figures 5b, 5c and 5d show further details of the drop in road surface as shown in Figure 5a, at different distances to the drop.

Figure 6a shows a change in asphalt color with no drop, according to an embodiment of the present invention.

Figure 6b shows a gray-scale profile for a simple step in brightness with no dark shadow patch for the change in asphalt color shown in Figure 6a.

Figures 6c and 6e show a dark stripe from a change in asphalt paving of the road shown in successive image frames, according to an embodiment of the present invention.

Figures 6d and 6f show a simple step edge of brightness in greater detail of respective

Figures 6c and 6e, according to an embodiment of the present invention.

Figure 7 which shows a graph of simulation results, according to an embodiment of the present invention.

Figures 8a and 8b shows a simplified line drawing of a night time driving scenario and an area of greater detail of the night time driving scenario respectively, according to an embodiment of the present invention.

Figure 9 shows a graph of a simulation for a drop in road surface of 0.1m over a length of 0.8m and 1.6m, according to an embodiment of the present invention.

Figures 10a-10e show a sequence of five images where a host vehicle approaches a speed bump and a shadow, according to an embodiment of the present invention.

Figures 10f-10j show greater detail of the speed bump and the shadow for respective Figures 10a-10e.

Figure 11 shows a flow chart of a method, according to an embodiment of the present invention.

Figure 12 shows a graph of the thickness of shadow at the far side of a speed bump as a function of distance to the speed bump, according to an embodiment of the present invention.

Figures 13a and 13b show two examples of soft shoulders at the road edge, according to an embodiment of the present invention.

IPR2025-00944
Tesla EX1002 Page 1029

DETAILED DESCRIPTION

[0013] Reference will now be made in detail to embodiments of the present invention, examples of which are illustrated in the accompanying drawings, wherein like reference numerals refer to the like elements throughout. The features are described below to explain the teachings herein by referring to the Figures.

[0014] Before explaining features taught herein in detail, it is to be understood that the invention is not limited in its application to the details of design and the arrangement of the components set forth in the following description or illustrated in the drawings. The invention is capable of other features or of being practiced or carried out in various ways. Also, it is to be understood that the phraseology and terminology employed herein is for the purpose of description and should not be regarded as limiting.

[0015] By way of introduction, embodiments of the present invention are directed to driver assistance systems using forward facing cameras which are becoming mainstream. These forward facing cameras may be rigidly mounted inside the driver cabin near the rear-view mirror. Cameras mounted inside the driver cabin near the rear-view mirror puts the cameras at a fixed position, significantly above the vehicle headlights. With the fixed position, the shadows produced by the headlights, when driving on dark roads, are visible in the camera image. The following detailed description describes how to use these shadows to enhance obstacle detection. In particular the description describes how the shadows can be used for detecting speed-bumps and negative obstacles at night from a moving vehicle. The information from the enhanced obstacle detection can be used to warn the driver or to prepare the vehicle for the anticipated change in road surface. The detailed description also gives a detailed analysis of the particular image motion characteristics of the shadow when the host vehicle is moving forward. The information from the detailed analysis allows the discrimination between shadows and other dark patches on the road and then, through careful analysis of the shadow over time, both the depth and shape of the obstacle can be determined. The detection range from shadow analysis may be significantly larger than what is possible from stereo or structure from motion (SFM) in dark night conditions.

[0016] As host vehicle moves forward, the image of a dark patch of a shadow, produced by the host vehicle headlights, behaves very differently to other dark patches on the road such as those due to tar patches, differences in asphalt color or shadows due to other light sources.

[0017] It will be shown below that the image of a mark on a road surface will increase as a function of the inverse of the distance (Z) squared, where Z is the distance from the camera to the mark on the road surface:

$$\frac{1}{Z^2} = Z^{-2}$$

[0018] While the size of a shadow produced by a sharp negative edge will increase only as a function of the inverse of the distance (Z):

$$\frac{1}{Z} = Z^{-1}$$

[0019] The shadow of a round edge, such as the far side of a speed bump, will increase even more slowly. By tracking the dark patch over time it is possible to differentiate between shadows and road marks.

[0020] The difference in the observed behavior can be explained as follows. A mark on the road retains its size (in world coordinates) as a host vehicle approaches. With a sharp negative obstacle, the visible part of the shadow extends from the top of the negative obstacle step to the far side of the shadow. When the host vehicle moves towards the obstacle, the lighting angle changes and the far edge of the shadow on the road moves and the extent of the shadow on the road is decreased significantly. If the edge producing the shadow is rounded the closer edge of the shadow is formed by the tangent point of the edge. As the host vehicle moves forward the tangent point moves further away and drops in height. These two factors of the edge and the host vehicle moving forward, reduce the size of the shadow on the road even further.

[0021] Referring now to the drawings, reference is now made to Figures 1 and 2 which illustrate a system **16** including a camera or image sensor **12** mounted in a host vehicle **18**, according to an aspect of the present invention. Image sensor **12**, imaging a field of view in the forward direction provides image frames **15** in real time and image frames **15** are captured by an image processor **30**. Processor **30** may be used to process image frames **15** simultaneously and/or in parallel to serve a number of advanced driver assistance systems/applications. The advanced driver assistance systems (ADAS) may be implemented using specific hardware circuitry with on board software and/or software control

IPR2025-00944
Tesla EX1002 Page 1030

algorithms in memory **302.** Image sensor **12** may be monochrome or black-white, i.e. without color separation or image sensor **12** may be color sensitive. By way of example in Figure 2, image frames **15** are used to serve pedestrian detection **20,** traffic sign recognition (TSR) **21** forward collision warning (FCW) **22** and obstacle detection **23** using shadows, according to embodiments of the present invention. Processor **30** may be used to process image frames **15** to detect and recognize an image or portions of the image in the forward field of view of camera **12.**

[0022] In some cases, image frames **15** are partitioned between different driver assistance applications and in other cases the image frames **15** may be shared between the different driver assistance applications.

[0023] Although embodiments of the present invention are presented in the context of driver assistance applications, embodiments of the present invention may be equally applicable in other real time signal processing applications and/or digital processing applications, such as communications, machine vision, audio and/or speech processing as examples.

[0024] Reference is now made to Figure 3a which shows a simplified line drawing **30a** of a night time driving scenario, according to a feature of the present invention. Camera **12** is mounted in close proximity to the host vehicle **18** windshield and host vehicle **18** headlights illuminate the scene ahead as host vehicle **18** travels along road surface **36.** Host vehicle **18** headlights produce a shadow **32** on the road when there is a obstacle. In the example shown in Figure 3a the obstacle is low object **34** on road **36.** Low object **34** may be for example a speed bump, raised manhole cover but the discussion that follows is the same for a negative obstacle such as a drop or pothole. Since the camera **12** is mounted higher than the headlights, shadow **32** can be seen by camera **12** in host vehicle **18.** Negative obstacles and speed bumps produce dark patches in image frames **15.**

[0025] In Figure 3a are a number of distances, where distance is designated as Z in the equations which are shown later on the detailed descriptions that follow. Let the coordinate system reference point (0,0) be the point on the road **36** directly below camera **12,** indicated by dotted line **31.** H_C is the height of camera **12,** H_L the height of the headlight of host vehicle **18** and H_T the height of obstacle **34.** The headlights of vehicle **18** are Z_L in front of camera **12,** that is the distance between dotted lines **31** and **33.** Z_T is the distance from camera **12** to the edge of obstacle **34,** that is the distance between dotted lines **31** and **35.** Z_{S1} and Z_{S2} are the close and far points on shadow **32** which are visible to camera **12,** the distances between dotted lines **31** and **37** and dotted lines **31** and **39** respectively.

[0026] Reference is now made to Figure 3b which shows a simplified line drawing **30b** of a night time driving scenario, according to an embodiment of the present invention. Two host vehicles **18** are shown with respective cameras **12.** From Figure 3b it can be seen that when host vehicle **18** moves forward towards obstacle **34,** the extent of shadow **32** becomes smaller.

[0027] When host vehicle **18** moves towards the obstacle **34,** the lighting angle changes and the far edge of shadow **32** on road **36** moves and the extent of shadow **32** on road **36** is decreased significantly. If the edge producing shadow **32** is rounded, the closer edge of shadow **32** is formed by the tangent point of the edge. As host vehicle **18** moves forward the tangent point moves further away and drops in height. These two factors of the edge and the host vehicle **18** moving forward, reduce the size of shadow **32** on road **36** even further.

[0028] Reference is now made to Figures 4a and 4b which show image frames of a speed bump where host vehicle **18** is at a distance away from the speed bump of 40.9m and 27.0m respectively, according to an embodiment of the present invention. Figures 4a and 4b show white reflectors **40** and a rock **42.** The shadow beyond the speed bump is only slightly wider in Figure 4b than the shadow in Figure 4a where host vehicle **18** is farther from the speed bump. Also, the lateral distance difference between reflectors **40** is much more pronounced when compared to the change in shadow width beyond the speed bump for the two image frames. Therefore, the shadow on the far side of the speed bump increases and/ or decreases in size much more slowly than the image size of the speed bump itself.

[0029] Reference is now made to Figures 4c and 4d which show a portion of image frames shown respectively in Figures 4a and 4b in greater detail, according to an embodiment of the present invention. Plotted on each of Figures 4a and 4b are respective gray scale value curves **44a** and **44b.** The further to right curves **44a/ 44b** go laterally, the greater is the gray scale value. Conversely, the further to left curves **44a/ 44b** go laterally, the lesser is the gray scale value. In Figure 4c where host vehicle **18** is farther from the speed bump, peaks in gray scale value occur by virtue of the contributions from reflection of rock **42** and the illuminated side of the speed bump along with reflection from reflectors **40.** Furthermore, it can be observed that the shadow cast by the speed bump is only slightly narrower when the host vehicle **18** is farther from the speed bump.

[0030] Changes in road **36** surface material, ice patches, oil marks and tar seams can also produce dark patches on the image of the road **36** surface. However, the size of such a patch in the image **15,** as the host vehicle **18** approaches, behaves very differently.

[0031] Reference is now made to Figure 5a shows an example of a shadow produced when there is a drop **52** in the road surface where there is a junction between older and newer asphalt paving. Reference is also made to Figures 5b, 5c and 5d shows further details of drop **52,** when the distance to drop **52** is at 22meters (m), 17.7m and 11.5m respectively. In Figures 5b-5d it can be seen that the size of the shadow of drop **52,** where newer paving makes way to old paving and when the road is illuminated by the headlights of host vehicle **18.** The shadow of drop **52** hardly changes as host vehicle **18** distance (Z) to drop **52** changes from 22m to 11.5m. Figures 5b, 5c and 5d have respective gray scale value

IPR2025-00944
Tesla EX1002 Page 1031

curves **54b**, **54c** and **54d** for a central column of pixels. The lowest gray scale value is shown by the left lateral peak in respective curves **54b**, **54c** and **54d** in the center of the shadow of drop **52.**

[0032] Reference is now made to Figure 6a which shows a change in asphalt color with no drop , according to an embodiment of the present invention. The corresponding gray-scale profile for Figure 6a is shown in Figure 6b, which shows a simple step in brightness with no dark shadow patch.

[0033] Reference is now made to Figures 6c and 6e which show a dark stripe **62** which is a change in asphalt paving of the road shown in successive image frames **15,** according to an embodiment of the present invention. The change in asphalt paving of the road is due to a double change in asphalt by virtue of a ditch which was filled in with new asphalt. In each image frame **15** their is no drop in height of the road surface and no shadow is produced. The gray-scale profiles shown in Figures 6d and 6f show a simple step edge of brightness in greater detail of respective Figures 6c and 6e, according to an embodiment of the present invention. Figures 6c and 6e show a dark stripe **62** on the road at 14.3m and 11.2m. The width of dark stripe **62** as shown in Figures 6c and 6e increase in size as a function of the inverse of distance squared (Z^{-2}) indicating that dark stripe **62** is a mark on the road and not a shadow from a negative obstacle (obstacle **34** for example).

[0034] The techniques described below, use the change in the actual shadow **32** itself due to the host vehicle **18** motion (and thus the motion of the light sources). The actual shadow is then analyzed by the camera **12** that is fixed relative to the light source and moves with it.

THE MATHEMATICAL MODEL

[0035] Referring back to Figure 3a, reference point (0,0) is the point on the road **32** directly below camera **12.** H_C is the height of the camera, H_L the height of the headlight and H_T the height of obstacle **34.** The headlights are Z_L in front of the camera. Z_T is the distance from camera **12** to the edge of obstacle **34** (dotted line **35)** and Z_{S1} and Z_{S2} are the close and far points of shadow **32** visible to camera **12.**

[0036] By similarity of triangles we get:

$$Z_{S1} = \frac{Z_T H_C}{H_C - H_T} \quad (1)$$

$$Z_{S2} = \frac{(Z_T - Z_L)H_L}{H_L - H_T} + Z_L \quad (2)$$

$$= \frac{Z_T H_L - Z_L H_T}{H_L - H_T} \quad (3)$$

[0037] Let y1 and y2 be the lower and upper edges of a shadow patch in the image:

$$y_1 = \frac{f H_C}{Z_{S1}} \quad (4)$$

$$y_2 = \frac{f H_C}{Z_{S2}} \quad (5)$$

[0038] The height or thickness of the patch in the image is then given by y1 -y2:

IPR2025-00944
Tesla EX1002 Page 1032

$$y_1 - y_2 = \frac{fH_C}{Z_{S1}} - \frac{fH_C}{Z_{S2}} \quad (6)$$

$$= \frac{fH_C(H_C - H_T)}{Z_T H_C} - \frac{fH_C(H_L - H_T)}{Z_T H_L - Z_L H_T} \quad (7)$$

[0039] For $Z_L << Z_T$ we can approximate:

$$y_1 - y_2 \approx \frac{f(H_T H_C - H_T H_L)Z_T}{Z_T^2 H_C} \quad (8)$$

$$\approx \frac{fH_T(H_C - H_L)}{Z_T H_C} \quad (9)$$

[0040] The image thickness of the shadow patch changes as a function of Z_T^{-1} and given an estimate of Z_T one can also estimate the (negative) height of obstacle **34.**

[0041] Consider now a marking on road **36** which at time to stretches from a distance Z_{S1} (to) to a distance Z_{S2} (to). Let:

$$\alpha = Z_{S2}(t_0) - Z_{S1}(t_0) \quad (10)$$

$$Z_{S3} = Z_{S1} + \alpha \quad (11)$$

[0042] The height or thickness of the marking in the image is given by:

$$y_3 - y_1 = \frac{fH_C}{Z_{S1} + \alpha} - \frac{fH_C}{Z_{S1}} \quad (12)$$

$$= \frac{fH_C\alpha}{(Z_{S1} + \alpha)Z_{S1}} \quad (13)$$

$$\approx \frac{fH_C\alpha}{Z_{S1}^2} \quad (14)$$

NUMERICAL EXAMPLE

[0043] Reference is now made to Figure 7 which shows a graph **70** for a simulation, according to an embodiment of the present invention. Graph **70** plots two curves, one for a patch or road mark (dotted line) and one for a shadow strip (solid line). The thickness of the patch or shadow strip is in pixels versus range to obstacle in meters (m). The simulation

IPR2025-00944
Tesla EX1002 Page 1033

has camera **12** mounted in a Renault Megane II where H_C = 1.25m, H_L = 0.73m and obstacle of height H_T = 0.1m. A shadow will appear just over 2 pixels thick at 50m. At 25m the shadow will double in size to about 4 pixels. A road mark on road at 25m will quadruple in size to 8 pixels.

[0044] Note that at a certain close distance (≈ 5m) the tangent line of the headlight and of the camera **12** coincide. After passing 5m the size of the shadow patch decreases rapidly.

[0045] Referring back to Figures 3a and 3b, Figures 3a and 3b show a step edge drop. The step edge drop is typical for a pot hole, road edge and steep speed bump. With the step edge drop, the tangent point which causes shadow **32** and the tangent point of the viewing angle from camera **12** can be approximated as the same point and which is fixed in space as host vehicle **18** moves. For extended smooth speed bumps this approximation breaks down, however.

[0046] Reference is now made to Figures 8a and 8b which shows a simplified line drawing **80** of a night time driving scenario and an area **83** of greater detail of the night time driving scenario respectively, according to an embodiment of the present invention. In Figures 8a and 8b, the tangent viewing ray **85** touches the speed bump **34** further away than the tangent illumination ray **87.** As host vehicle **18** moves forward, the tangent point **62** of the viewing ray **65** moves further away along the speed bump or obstacle **34.** This means that the near edge of shadow **32,** in image space, will move slower than a typical road point. The exact motion depends on the shape of speed bump **34** and can be used to measure the curvature of obstacle **34.**

[0047] It can also be seen that as host vehicle **18** moves forward, tangent point **84** of illumination ray **87** moves further away along obstacle **34** and the height H_T decreases. The result is that the far edge of shadow **32** on road **36** will move towards the host vehicle **18** even faster than predicted by a sharp edge. Shadow **32** on road **36** moving towards the host vehicle **18** even faster than predicted by the sharp edge effect will accelerate when host vehicle **18** gets close to the speed bump or obstacle **34** and at some distance shadow **32** will disappear completely.

[0048] From experimental results shown later, at far distances (above 30m), shadow **32** behaves more like a step edge and disappears at about 7.0m from camera **12.**

[0049] The analysis of rounded edges is complex. The tangent point **84** for the headlights does not correspond exactly to the tangent point **82** of camera **12** line of sight. Furthermore, as the host vehicle **18** moves closer to obstacle **34,** tangent points **82** and **84** will move further away and closer to road **36.** In order to analyze a rounded edge obstacle **34** (e.g. speed bump), the edge was approximated as a piecewise linear curve of 10 segments dropping from H_T = 0.1m down to zero over a given length:

$$Z_T(i) = Z_T(0) + i * \delta Z \qquad (15)$$

$$H_T(i) = H_T * (1 - 0.1 * i) \qquad (16)$$

where i=1...10.

[0050] Reference is now made to Figure 9 which shows a graph **90** for a simulation for a drop of 0.1 meter in road surface over a length of 0.8 meter and 1.6 meter, according to an example. Three curves are shown for thickness (pixels) of shadow versus range to speed bump or obstacle **34** in meters (m). Solid line **92** is for a sharp edge, dotted line **94** is for a round edge for the drop of 0.1m over the length of 0.8m and dotted line **96** is for another round edge for the drop of 0.1m over the length of 1.6m.

[0051] Host vehicle **18** in the simulation, approaches the speed bump or obstacle **34** from 50 meter down to 5 meter in steps of 1 meter. At each distance it was determined (for each of the two cases: 0.8m and 1.6m) which point along the piecewise linear curve was the occluding point for the headlight and produced the far edge and which was the occluding point for camera **12** and produced the near edge. As can be expected, the shorter the drop, the more it behaves like a step edge.

EXPERIMENTS AND RESULTS

[0052] Figures 10a-10e show a sequence of five image frames **15** where host vehicle **18** approaches speed bump **1002** and shadow **1006,** according to an embodiment of the present invention. Figures 10f-10j show greater detail of speed bump **1002** and shadow **1006** for respective Figures 10a-10e. In Figures 10a-10e the distance (Z) to speed bump **1002** and shadow **1006** are 66m, 42.4m, 26.4m, 13.6m and 7.9m respectively. Shadow **1006** at the far side of speed bump **1002** was first detected at over 40m and then tracked as it grew closer to host vehicle **18.** Figures 10f-10j have respective gray scale value curves **1004f-1004j** for a central column of pixels. The lowest gray scale value for the central column is shown by the lateral left peak in respective curves **1004f-1004j** in the center of shadow **1006.**

IPR2025-00944
Tesla EX1002 Page 1034

[0053] Reference is now made to Figure 11 which shows a flow chart of a method **23,** according to an embodiment of the present invention.

[0054] In step **1103,** a first image frame **15** is captured, followed by the capture of a second image frame **15** from camera **12** processor **30** of system **16.**

[0055] Optionally, image frames **15** may be warped to compensate for a change in vehicle orientation, i.e. roll, pitch and yaw, which may be determined from image frames **15** or from external sensor(s) connected to driver assistance system **16.**

[0056] The first and the second image frames **15** are then filtered (step **1105)** according to a threshold, e.g. a previously determined threshold, to produce a first dark image patch **1107** of first image frame **15** and a second dark image patch **1111** of the second image frame **15** respectively. First image patch **1107** and optionally second image patch **1111** are analyzed for connectivity of picture elements.

[0057] One or more connected components of first image patch **1107** is tracked (step **1112)** from the first image frame **15** to second dark image patch **1111** of second image frame **15.**

[0058] Tracking step **1113** may be performed by using the expected image motion for road points given host vehicle **18** speed and nearest neighbor.

[0059] The thicknesses of first image patch **1107** and second dark image patch **1111** are measured (step **1113)** in both the first image frame **15** and the second image frame **15.** By way of example, an image strip of twenty pixel columns wide, from the bottom of an image frame **15** and up to the image row corresponding to 80 meters may be analyzed for connected components. The patch or shadow thickness may be computed (step **1113)** by counting the number of pixels in the connected component divided by the strip width in columns (i.e. 20 pixels).

[0060] In decision block **1115,** the thicknesses of candidate shadows **1109** tracked across successive image frames may be accumulated. From the change of thicknesses tracked across successive image frames for first dark image patch **1107** and second dark image patch **1111** it is possible to analyze (step **1117)** if the change of thicknesses is consistent with a shadow cast by an obstacle from illumination from the headlights in order to for instance audibly warn a driver of host vehicle **18** of an obstacle **34** in road **36,** or to apply brakes without driver intervention. First image patch **1107** and second dark image patch **1111** may be shadows of an obstacle from the headlights of vehicle **18** or alternatively a darker road texture unrelated to an obstacle in the road.

[0061] Reference is now made to Figure 12 which shows a graph **1200** of the thickness of shadow **1006** at the far side of speed bump **1002** as a function of distance to speed bump **1002,** according to feature of the present invention. Dots show the results measured in the sequence of frames captured. The curves **1202** (sharp edge) and **1204** (round edge) shown the simulation of a bump 0.1m high for a sharp drop and a smooth drop respectively. The results show that in practice the shadows behave according to the model. Shadow 1006 behaves approximately like Z^{-1} and not Z^{-2}. Shadow **1006** appears to behave like a rounded drop however, further experiments may be required using ground truth from accurately measured road structures.

[0062] Reference is now made to Figures 13a and 13b which show two examples of soft shoulders at the road edge, according to an example of the present invention. The drop between the asphalt and the gravel shoulder, results in a narrow shadow line **1300a** and **1300b** respectively that are darker than both the asphalt or the gravel. Often in country roads the asphalt paved on top of gravel and there is a step transition (drop) from the asphalt to the gravel. The step transition is called a soft shoulder which is a negative obstacle and will also generate a shadow like shadow lines **1300a** and **1300b.** Soft shoulders can thus be detected by looking for a dark strip a few pixels wide at the edge between the dark asphalt and the lighter gravel. The shadow strip will be even darker than the asphalt. The width of shadow lines **1300a** and **1300b** indicate the shoulder drop. These dark shadow lines **1300a** and **1300b** have a unique feature; unlike road markings such as lane marks that get narrower in the image as they approach the horizon, the shadows like **1300a** and **1300b** on road edges stay the same width. On a straight road, the inner edge of the shadow would be a line passing through the vanishing point of the road while the outer edge is a line that passes a few pixels to the side of this vanishing point.

[0063] The term "obstacle" as used herein in reference to a road refers to a "positive obstacle" and a "negative obstacle". The term "positive obstacle" as used herein is an obstacle extends vertically upward in real space above the road surface such as a speed bump or a vertical bump caused by new asphalt paving in the direction of motion of the vehicle. The term "negative obstacle" as used herein extends vertically downward in real space below the road surface such as a hole or a drop in asphalt paving in the direction of motion of the vehicle.

[0064] The term "shape" of an obstacle, as used herein refers to the vertical contour in real space along the direction of motion of the vehicle

[0065] The indefinite articles "a", "an" is used herein, such as "a shadow", an "obstacle" has the meaning of "one or more" that is "one or more shadows" or "one or more obstacles". Although selected embodiments of the present invention have been shown and described, it is to be understood the present invention is not limited to the described features. Instead, it is to be appreciated that changes may be made to these features without departing from the principles and spirit of the invention, the scope of which is defined by the claims and the equivalents thereof.

IPR2025-00944
Tesla EX1002 Page 1035

[0066] The features of the various embodiments described herein may be combined with one another.

[0067] The disclosures in United States patent application numbers US-61/569,306 and US-61/702,755, from which this application claims priority, and in the abstract accompanying this application are incorporated herein by reference.

Claims

1. A computerized method performable at night by a driver assistance system mountable in a host vehicle while the host vehicle is moving forward with headlights on, wherein the driver assistance system includes a camera operatively connectible to a processor, the method including the steps of:

   capturing a first image frame and a second image frame of a road in the field of view of the camera;
   processing said first image frame and said second image frame to locate a first dark image patch of said first image frame and a second dark image patch of said second image frame;
   tracking said first and said second dark image patches from said first image frame to said second image frame as corresponding images of the same portion of the road; and
   measuring respective thicknesses in vertical image coordinates of said first and second dark image patches responsive to said tracking from said first image frame to said second image frame.

2. The method of claim 1, wherein said processing is performed by filtering said first image frame and said second image frame with a threshold value of gray scale to produce said first dark image patch of said first image frame and said second dark image patch of said second image frame, wherein said first dark image patch and said second dark image patch include intensity values less than said threshold value;

3. The method of claim 1 or 2, including the step of:

   prior to said tracking, verifying connectivity of picture elements within said first dark image patch.

4. The method of claim 1, 2 or 3, including the step of:

   determining that a change of said thicknesses between said first and second dark image patches is consistent with a shadow cast by an obstacle in the road from illumination from the headlights, thereby detecting said obstacle in the road.

5. The method of claim 4, including the step of:

   audibly warning a driver of the host vehicle responsive to said detecting said obstacle.

6. The method of claim 4 or 5, including the step of:

   analyzing change in said thicknesses between said first and second image frames, thereby distinguishing between the detected obstacle being a positive or a negative obstacle.

7. The method of claim 4, 5 or 6, including the step of:

   analyzing change in said thicknesses between said first and second image frames, thereby computing a shape of the detected obstacle.

8. The method of any preceding claim, including the step of:

   determining that a change of said thicknesses between said first and second image frames is not consistent with a shadow cast by an obstacle from illumination from the headlights.

9. The method of any preceding claim, including wherein said thicknesses are measured in picture elements according to a total number of picture elements with intensity values less than said threshold divided by a width in picture element columns of said .first and second dark image patches.

10. A driver assistance system including a camera and a processor configured to perform the computerized method,

IPR2025-00944
Tesla EX1002 Page 1036

according to any preceding claim.

11. A driver assistance system mountable in a host vehicle, the driver assistance system includes a camera operatively connectible to a processor, wherein the driver assistance system while the host vehicle is moving is operable to:

capture a first image frame and a second image frame of a road in the field of view of the camera;
process said first image frame and said second image frame to locate a first dark image patch of said first image frame and a second dark image patch of said second image frame;
track said first and second dark image patches from said first image frame to said second image portion of said second image frame as corresponding images of the same portion of the road; and
measure respective thicknesses in vertical image coordinates of said first and second dark image patches responsive to said tracking from said first image frame to said second image frame.

12. The driver assistance system of claim 11, wherein the system is operable to determine that a change of said thicknesses between said first and second dark image patches is consistent with a shadow cast by an obstacle in the road from illumination from the headlights, thereby detecting said obstacle in the road.

13. The driver assistance system of claim 12, wherein the system is operable to analyze change in said thicknesses between said first and second image frames to distinguish between the detected obstacle being a positive or a negative obstacle.

14. The driver assistance system of claim 12 or 13, wherein the system is operable to analyze change in said thicknesses between said first and second image frames to compute a shape of the detected obstacle.

15. The driver assistance system of any one of claims 11 to 14, wherein the system is operable to determine that a change of said thicknesses between said first and second image frames is not consistent with a shadow cast by an obstacle from illumination from the headlights.

IPR2025-00944
Tesla EX1002 Page 1037

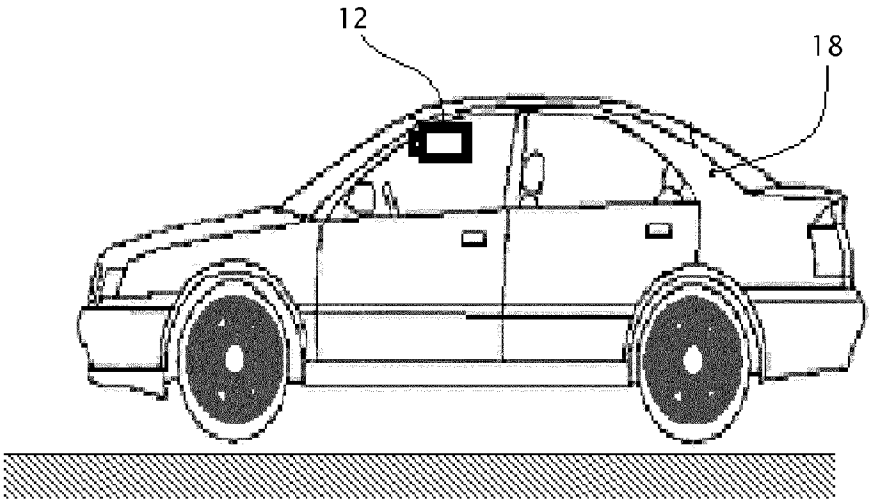


Fig. 1

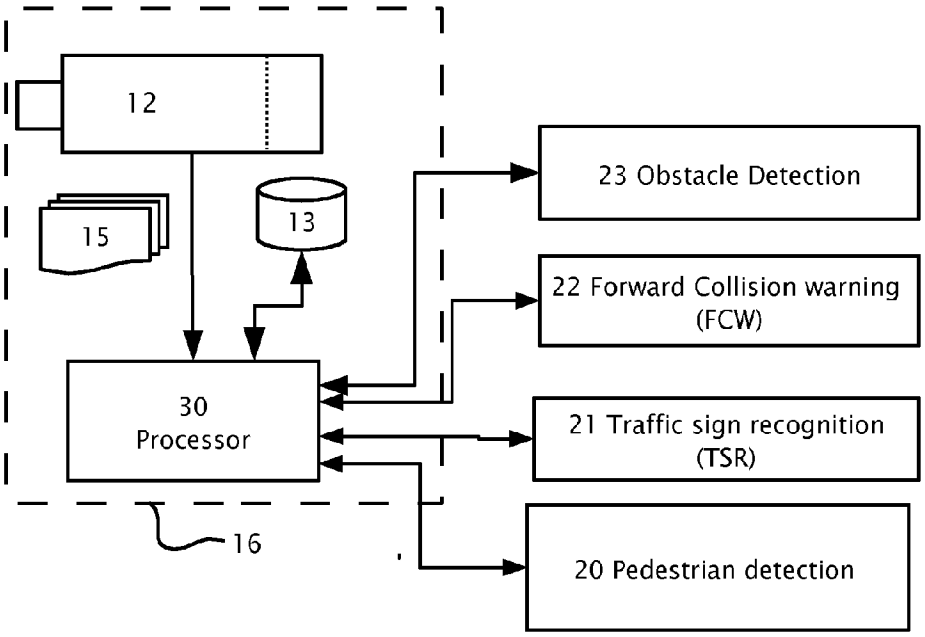


Fig. 2

IPR2025-00944
Tesla EX1002 Page 1038

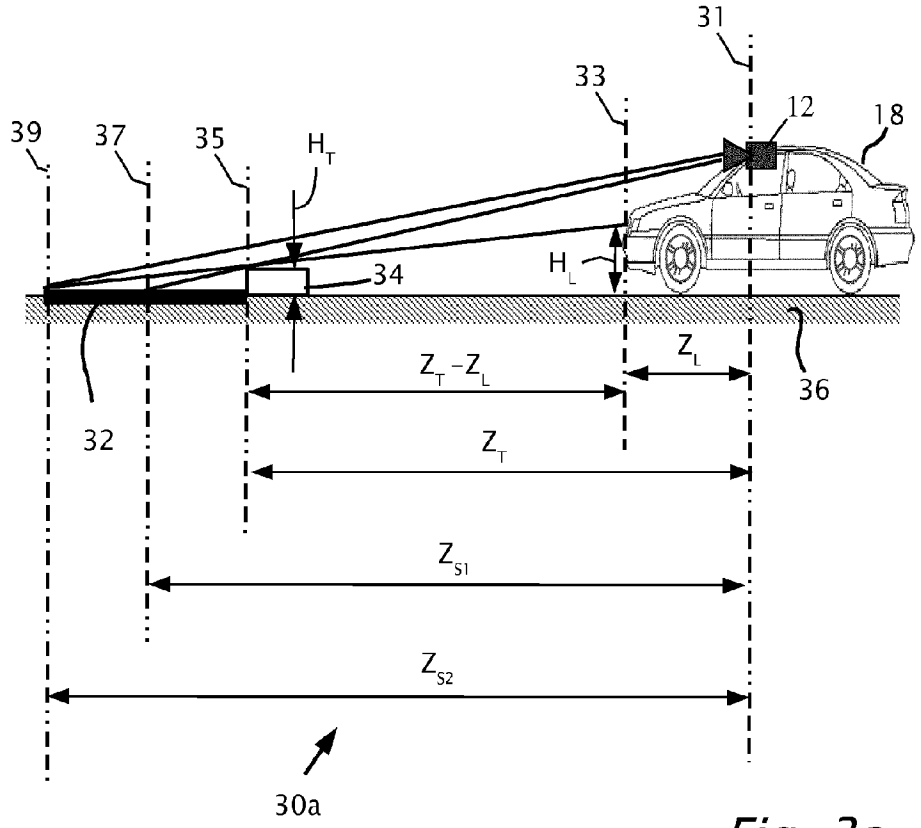


Fig. 3a

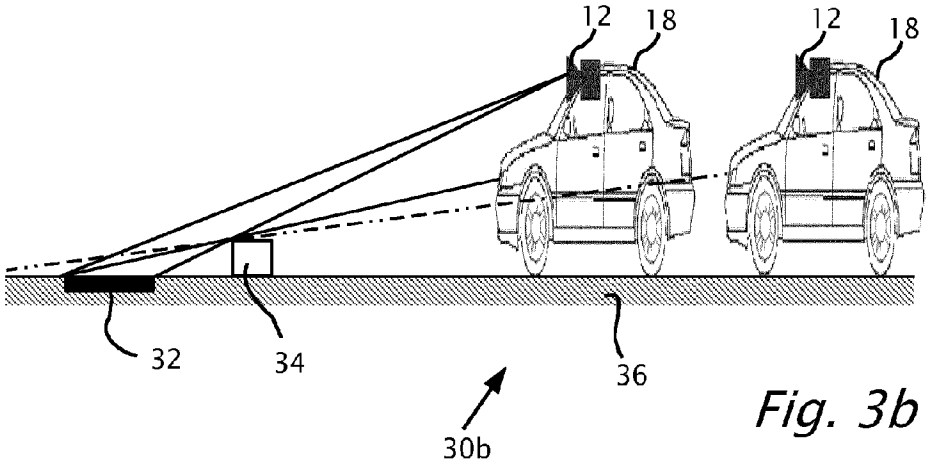


Fig. 3b

IPR2025-00944
Tesla EX1002 Page 1039

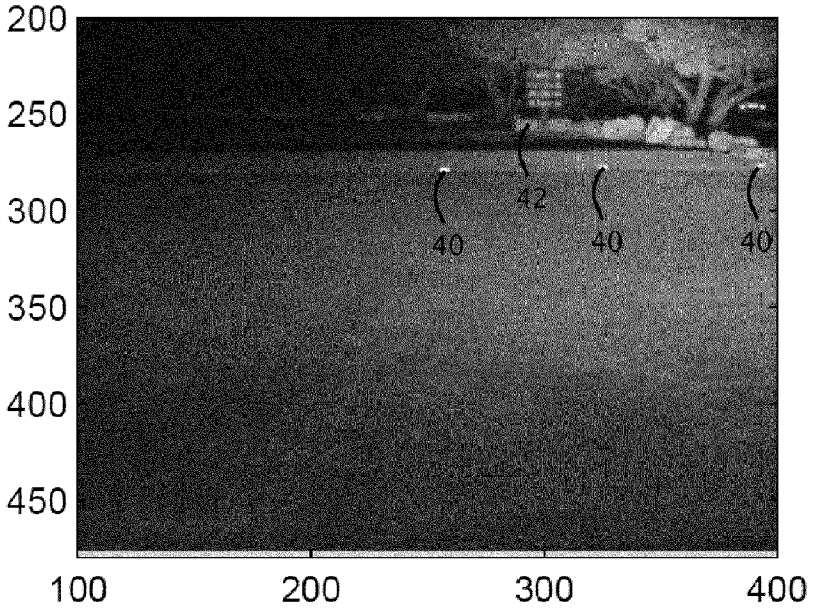


Fig. 4a

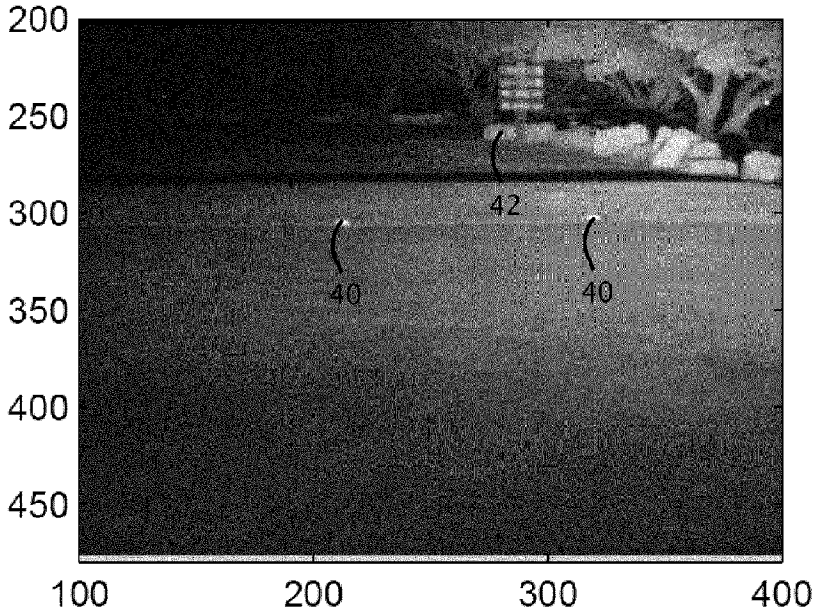


Fig. 4b

IPR2025-00944
Tesla EX1002 Page 1040

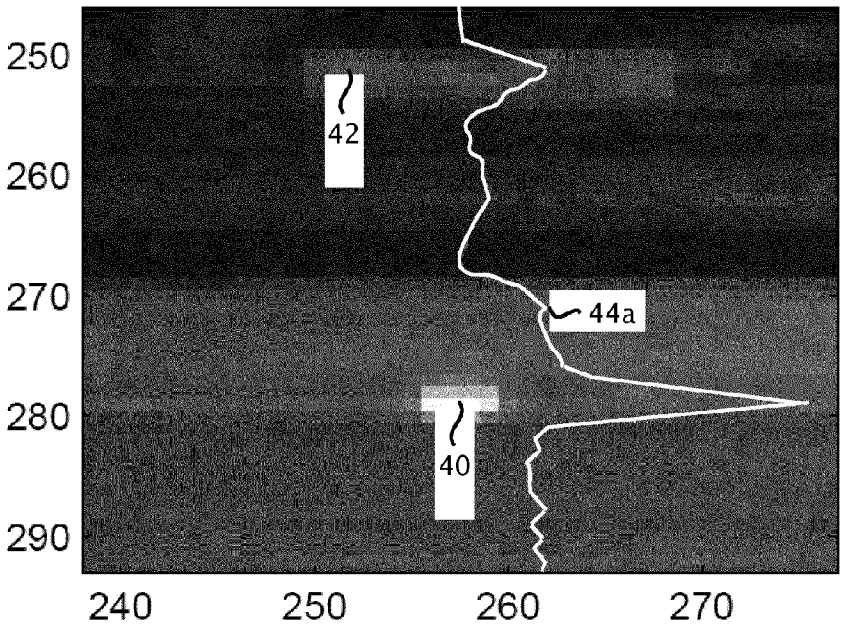


Fig. 4c

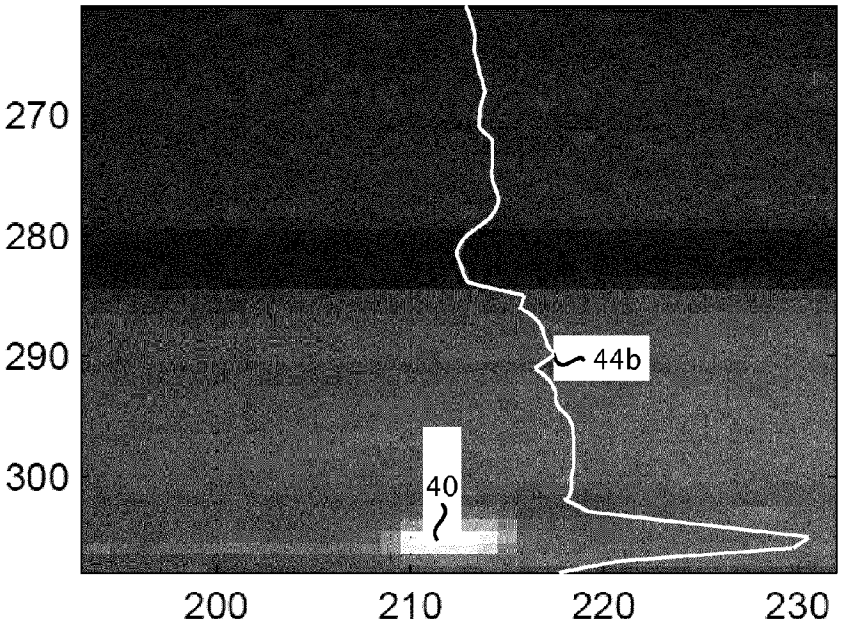


Fig. 4d

IPR2025-00944
Tesla EX1002 Page 1041

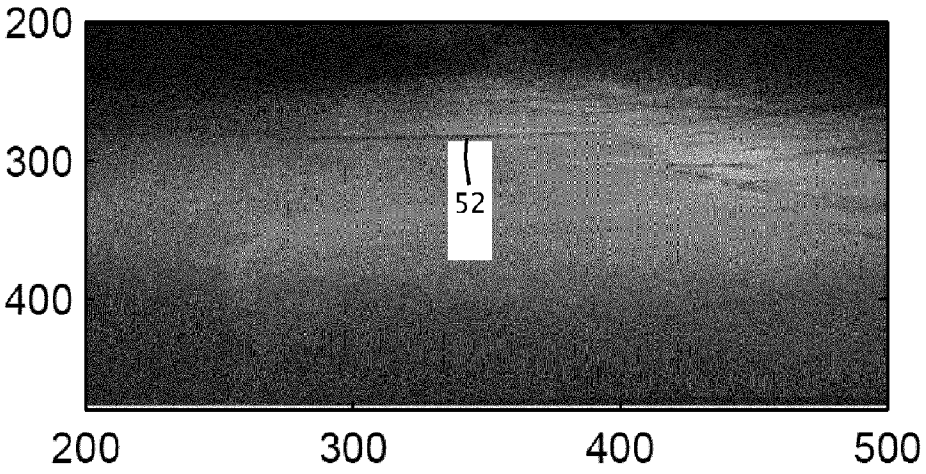


Fig. 5a

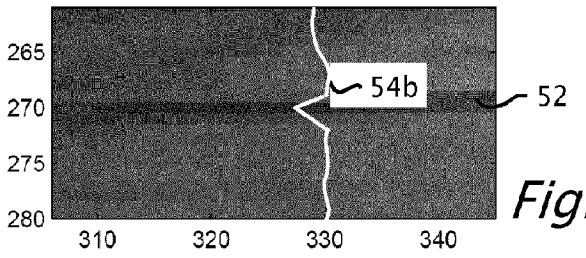


Fig. 5b

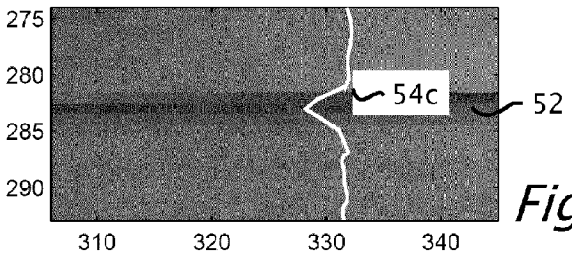


Fig. 5c

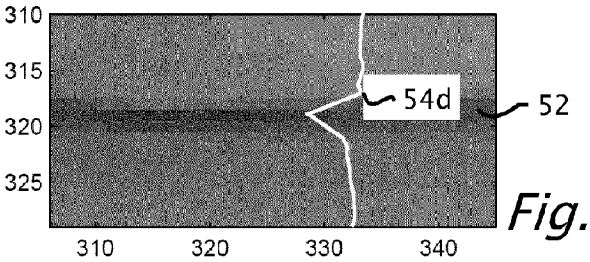


Fig. 5d

IPR2025-00944
Tesla EX1002 Page 1042

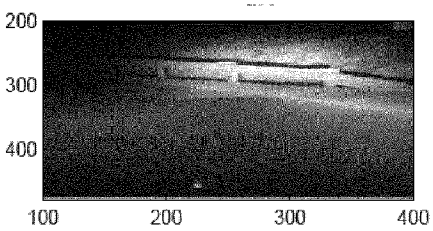


Fig. 6a

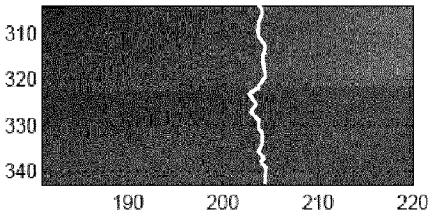


Fig. 6b

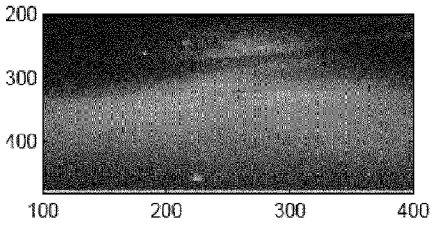


Fig. 6c

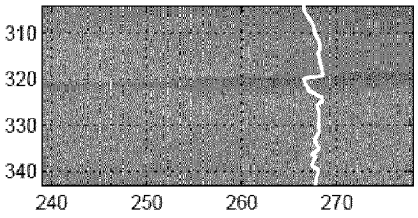


Fig. 6d

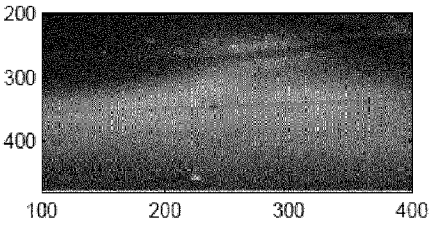


Fig. 6e

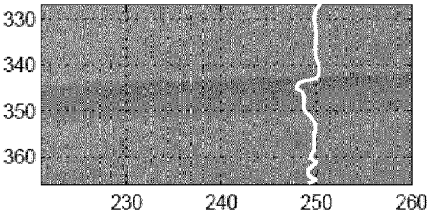


Fig. 6f

IPR2025-00944
Tesla EX1002 Page 1043

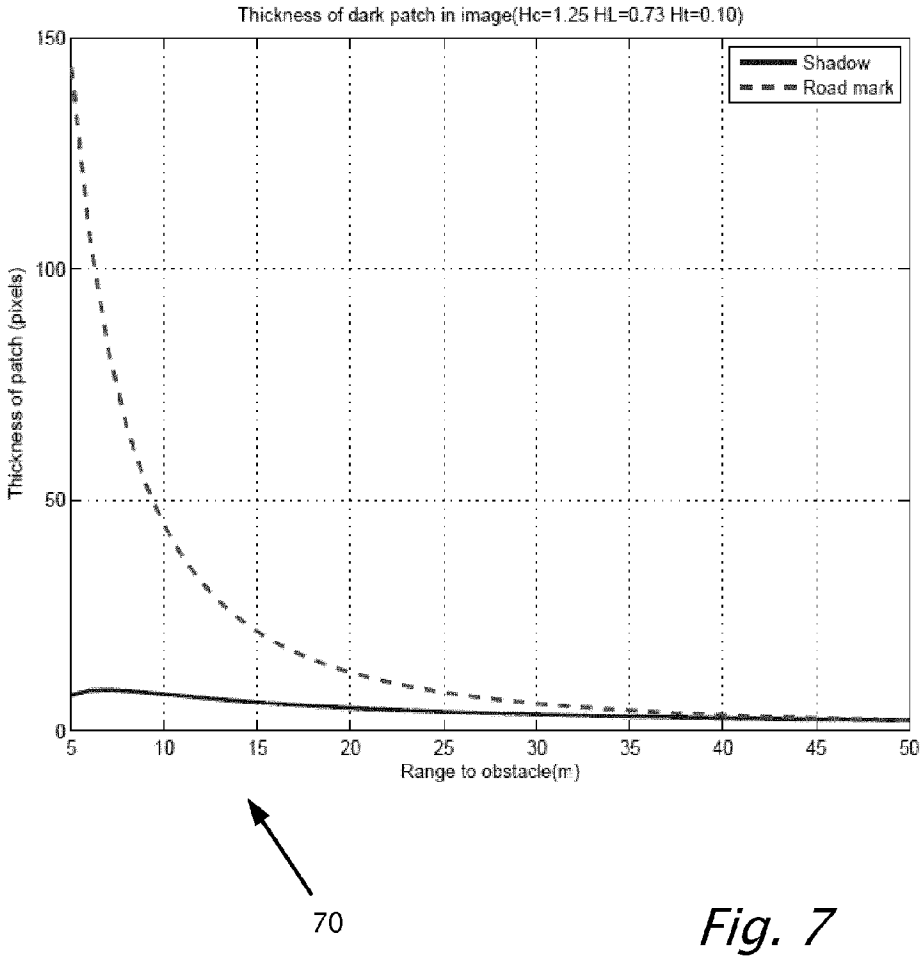


Fig. 7

IPR2025-00944
Tesla EX1002 Page 1044

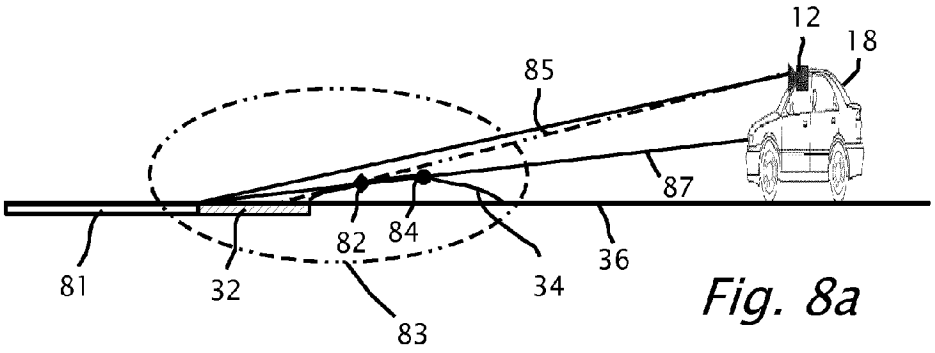


Fig. 8a

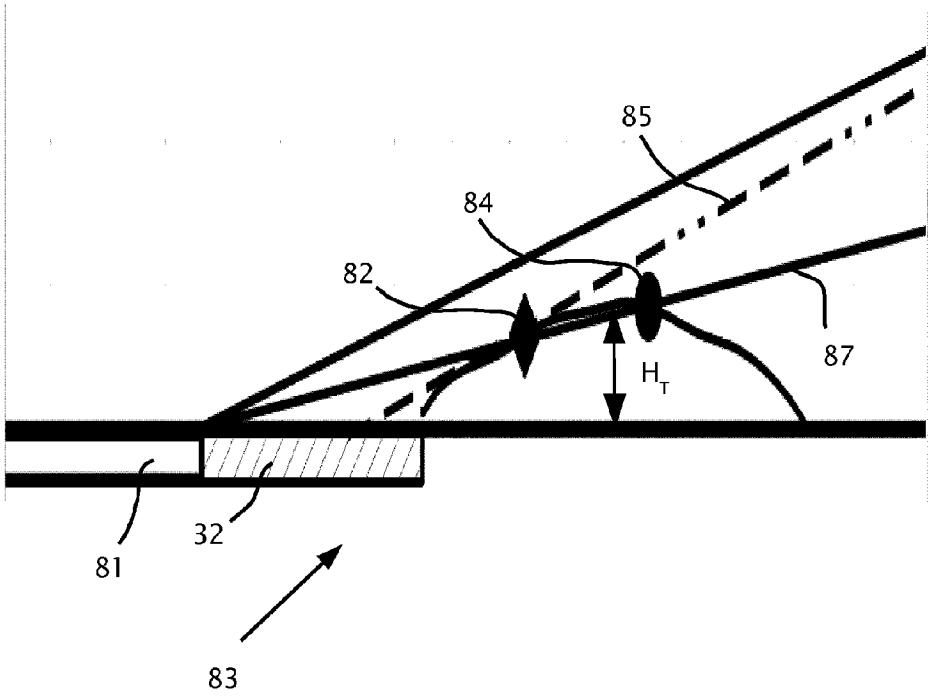


Fig. 8b

IPR2025-00944
Tesla EX1002 Page 1045

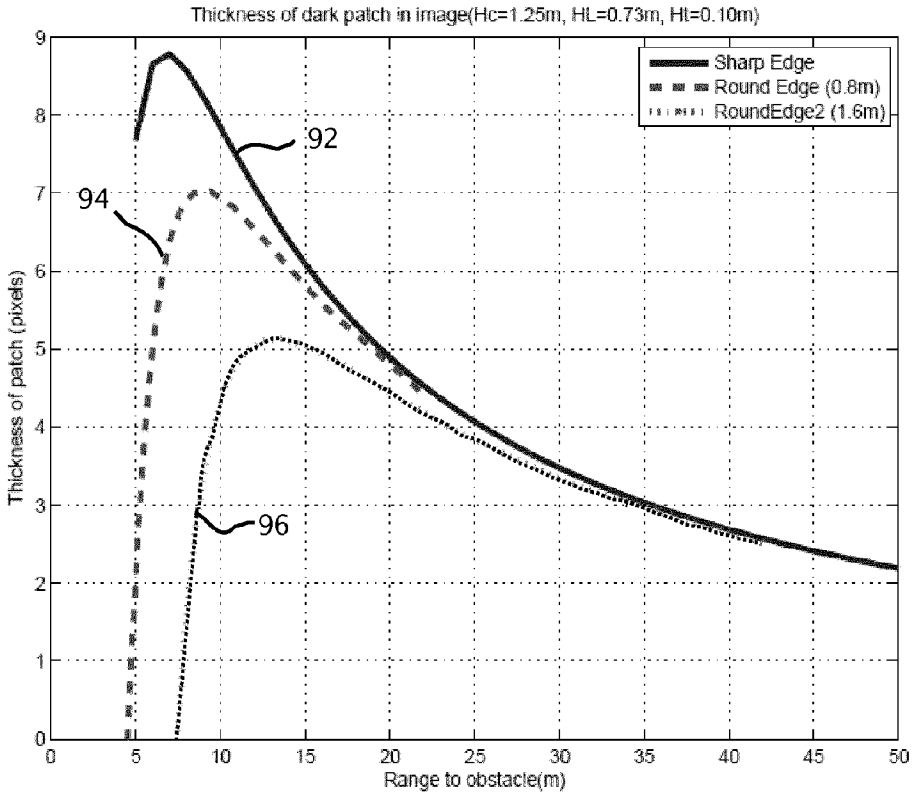


Fig. 9

IPR2025-00944
Tesla EX1002 Page 1046

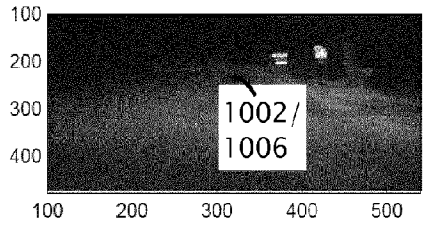


Fig. 10a

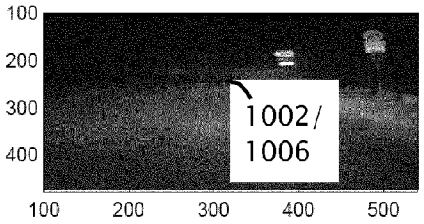


Fig. 10b

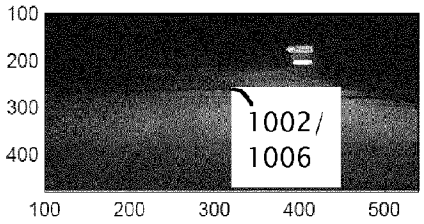


Fig. 10c

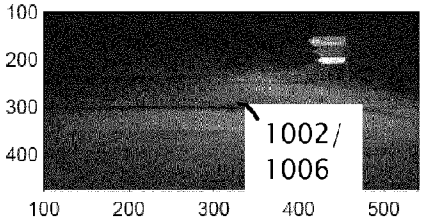


Fig. 10d

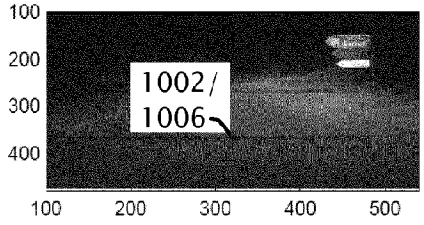


Fig. 10e

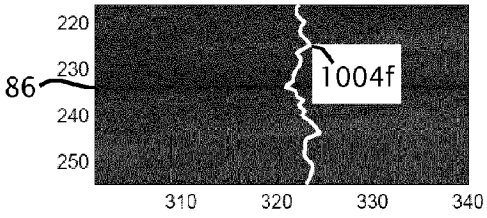


Fig. 10f

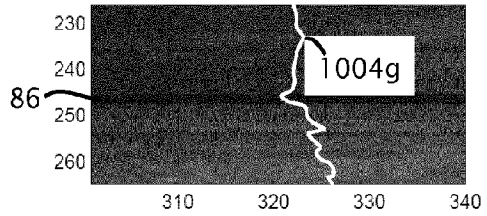


Fig. 10g

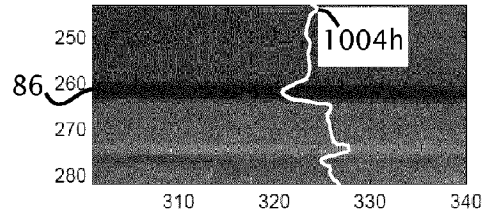


Fig. 10h

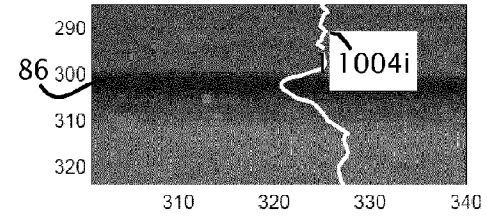


Fig. 10i

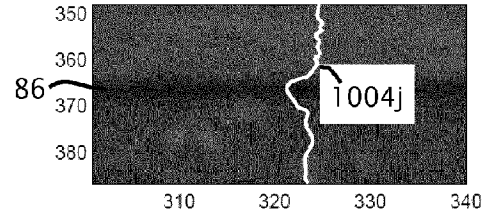


Fig. 10j

IPR2025-00944
Tesla EX1002 Page 1047

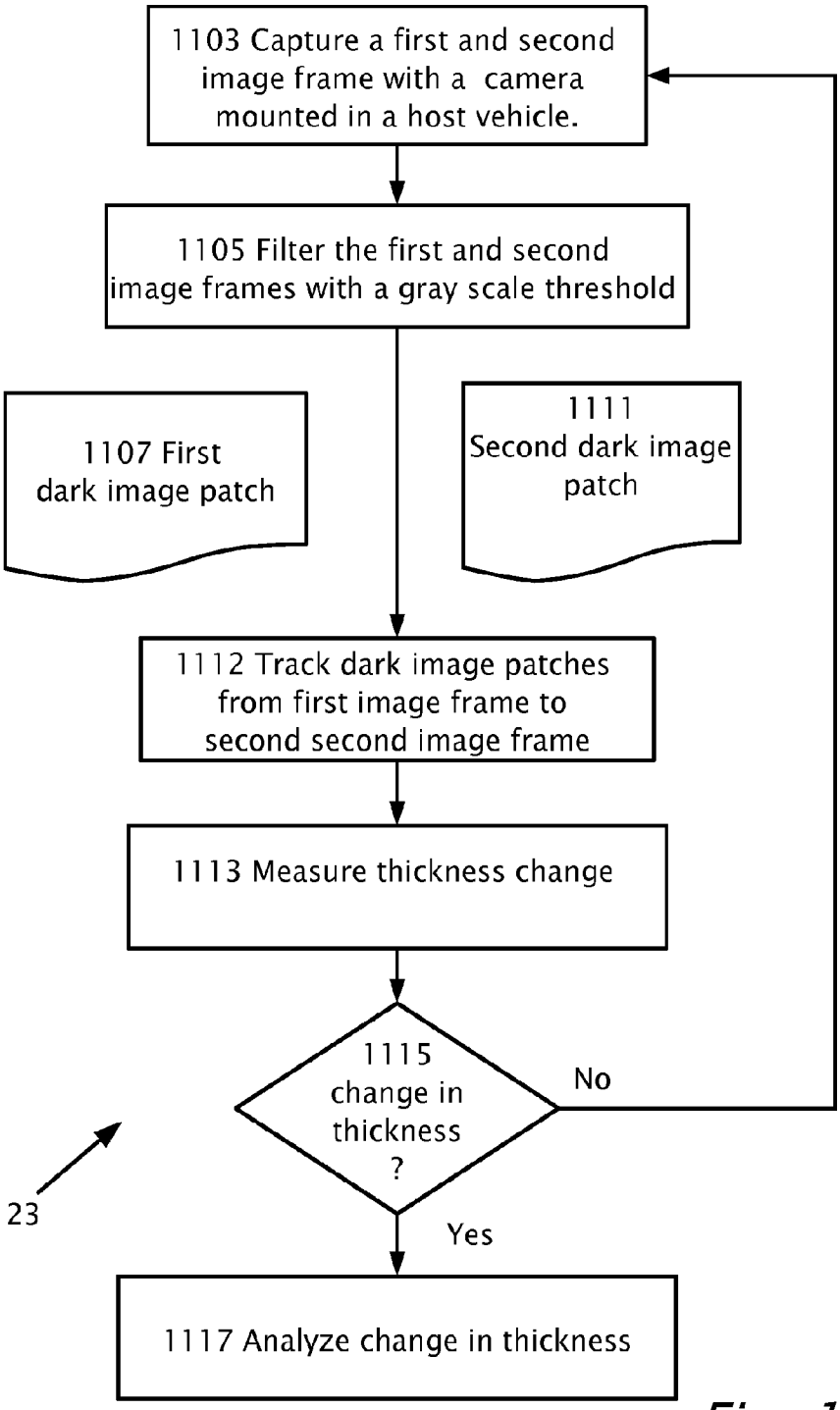


Fig. 11

IPR2025-00944
Tesla EX1002 Page 1048

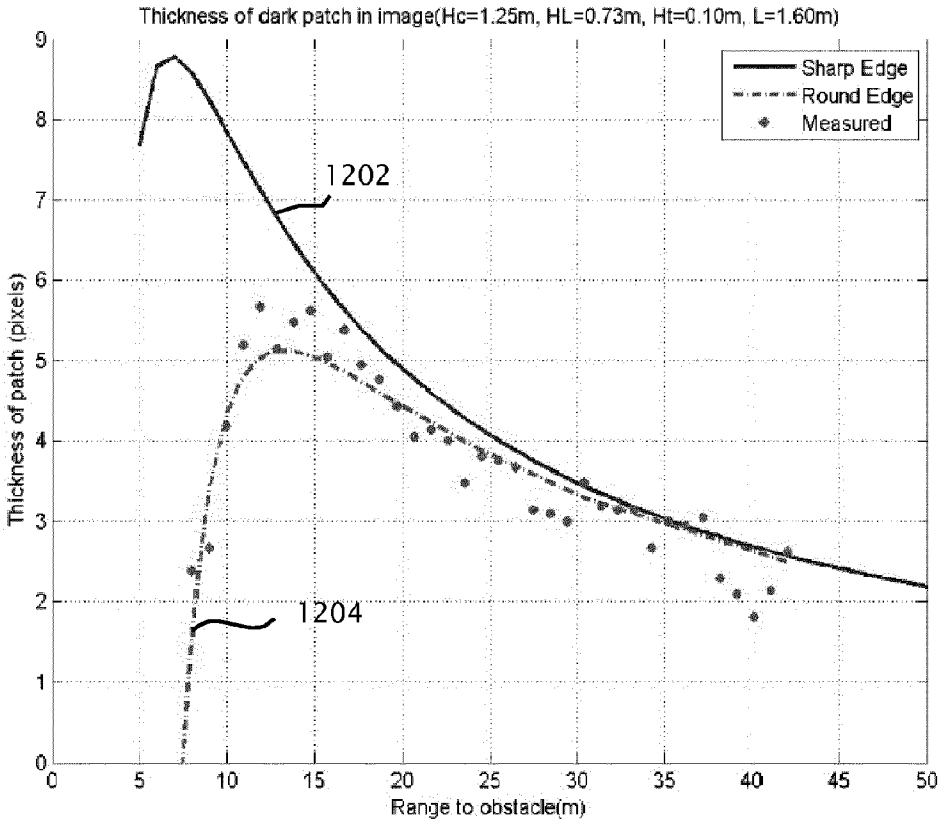


Fig. 12

IPR2025-00944
Tesla EX1002 Page 1049

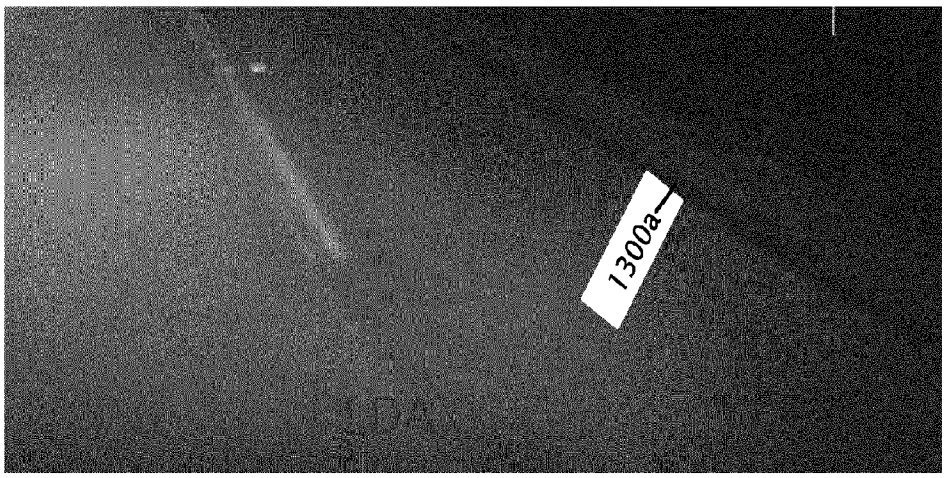


Fig. 13a

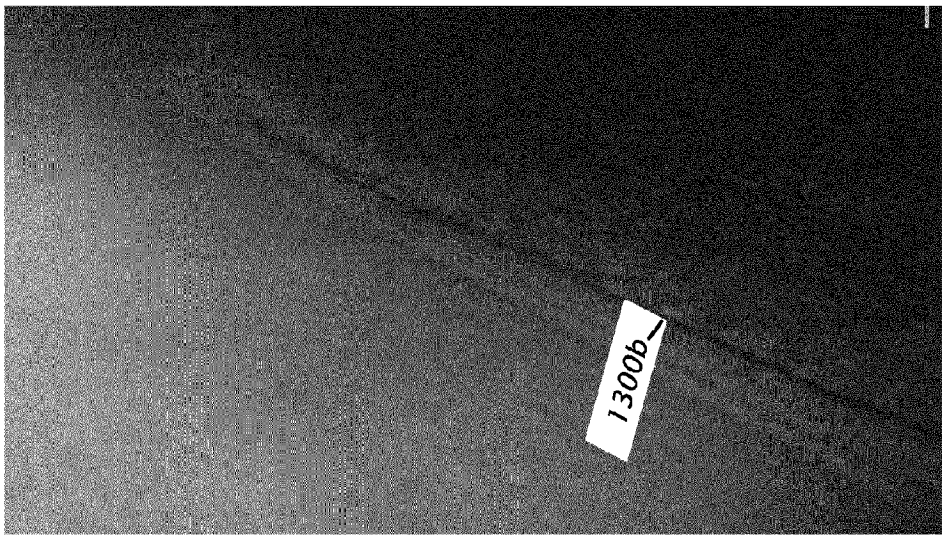


Fig. 13b

IPR2025-00944
Tesla EX1002 Page 1050

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- US 61569306 B **[0067]**
- US 61702755 B **[0067]**

Non-patent literature cited in the description

- **LARSON et al.** Lidar based off-road negative obstacle detection and analysis. *14th International IEEE Conference on Intelligent Transportation Systems,* October 2011 **[0003]**
- **BELUTTA et al.** Terrain Perception for Demo III. *Proceedings of the IEEE Intelligent Vehicles Symposium,* October 2000 **[0003]**
- Nighttime negative obstacle detection for off-road autonomous navigation In Unmanned Systems. **RABKIN et al.** Proceedings of the SPIE. 2007, vol. 6561, 656103 **[0003]**
- **MURAKA.** Detecting obstacles and drop-offs using stereo and motion cues for safe local motion. *IEEE/RSJ International Conference on Intelligent Robots and Systems,* September 2008 **[0003]**
- Preliminary investigation into the use of stereo illumination to enhance mobile robot terrain perception. **WITUS et al.** Proc. SPIE Vol. 4364, p. 290-301, Unmanned Ground Vehicle Technology III. vol. 4364, 290-301 **[0003]**

IPR2025-00944
Tesla EX1002 Page 1051

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 2 605 185 A3

(12) EUROPEAN PATENT APPLICATION

(88) Date of publication A3:
09.07.2014 Bulletin 2014/28

(43) Date of publication A2:
19.06.2013 Bulletin 2013/25

(21) Application number: **12275201.7**

(22) Date of filing: **12.12.2012**

(51) Int Cl.:
G06K 9/00 *(2006.01)*

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR
Designated Extension States:
BA ME

(30) Priority: **12.12.2011 US 201161569306 P**
19.09.2012 US 201261702755 P

(71) Applicant: **Mobileye Technologies Limited**
24854 Nicosia (CY)

(72) Inventor: **Stein, Gideon P**
96263 Jerusalem (IL)

(74) Representative: **Jehan, Robert**
Williams Powell
Staple Court
11 Staple Inn Buildings
London, WC1V 7QH (GB)

(54) **Detection of obstacles at night by analysis of shadows**

(57) A driver assistance systems mountable in a host vehicle while the host vehicle is moving forward with headlights on for detection of obstacles based on shadows. The driver assistance system includes a camera (12) operatively connectible to a processor (30). A first image frame and a second image frame are captured of a road. A first dark image patch and a second dark image patch include intensity values less than a threshold value. The first and the second dark image patches are tracked from the first image frame to the second image frame as corresponding images of the same portion of the road. Respective thicknesses in vertical image coordinates are measured of the first and second dark image patches responsive to the tracking.

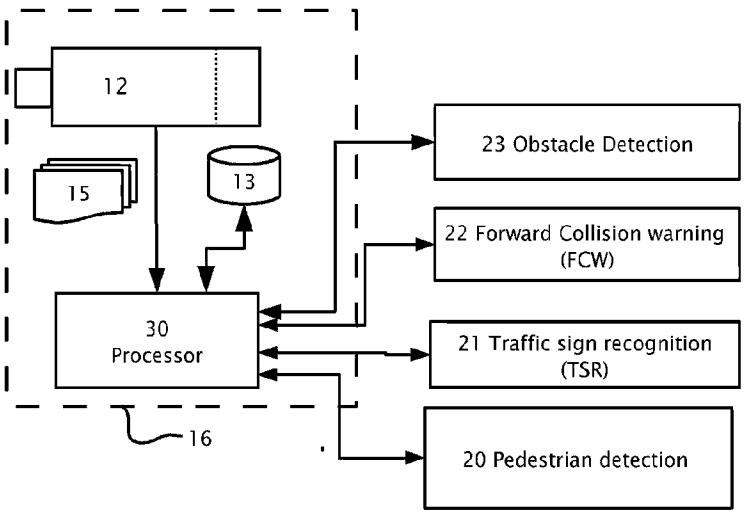


Fig. 2

EP 2 605 185 A3

Europäisches Patentamt
European Patent Office
Office européen des brevets

EUROPEAN SEARCH REPORT

Application Number
EP 12 27 5201

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
X	EP 1 708 125 A1 (VALEO VISION [FR]) 4 October 2006 (2006-10-04)	1-3,9-11	INV. G06K9/00
A	* paragraphs [0008], [0016] * * paragraphs [0018], [0021] - paragraph [0022] * * column 0024 * * paragraph [0026] - paragraph [0027] * * claim 18 *	4-8, 12-15	
A	HU TINGBO ET AL: "Negative obstacle detection from image sequences", THIRD INTERNATIONAL CONFERENCE ON DIGITAL IMAGE PROCESSING (ICDIP 2011), SPIE, 1000 20TH ST. BELLINGHAM WA 98225-6705 USA, vol. 8009, no. 1, 15 April 2011 (2011-04-15), pages 1-7, XP060015478, DOI: 10.1117/12.896288 [retrieved on 1901-01-01] * sections 2-3 *	1-15	

TECHNICAL FIELDS SEARCHED (IPC)
G06K

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
The Hague	30 May 2014	Kollreider, Klaus

CATEGORY OF CITED DOCUMENTS
X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons
& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

IPR2025-00944
Tesla EX1002 Page 1053

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 12 27 5201

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

30-05-2014

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 1708125 A1	04-10-2006	AT 425507 T	15-03-2009
		EP 1708125 A1	04-10-2006
		FR 2883818 A1	06-10-2006
		JP 2006298362 A	02-11-2006
		US 2006222208 A1	05-10-2006

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

IPR2025-00944
Tesla EX1002 Page 1054

Espacenet

Bibliographic data: EP3092599 (A1) — 2016-11-16

SYSTEMS AND METHODS FOR MIMICKING A LEADING VEHICLE

Inventor(s): CLARKE ANNA [IL]; BAGON EYAL [IL]; NIR GALI [IL] ± (CLARKE, ANNA, ; BAGON, EYAL, ; NIR, GALI, ; CLARKE, Anna, ; BAGON, Eyal, ; NIR, Gali)

Applicant(s): MOBILEYE VISION TECHNOLOGIES LTD [IL] ± (MOBILEYE VISION TECHNOLOGIES LTD, ; Mobileye Vision Technologies Ltd)

Classification: **- international:** ***B60K31/00; G06K9/00***

- cooperative: **B60K31/0066 (US); B60R1/00 (US); B60T13/662 (EP); B60T7/12 (US); B60T7/22 (EP, US); B60T8/32 (US); B60T8/321 (EP); B60W10/18 (US); B60W10/20 (US); B60W30/00 (US); B60W30/08 (US); B60W30/09 (US); B60W30/0953 (US); B60W30/12 (US); B60W30/14 (US); B60W30/143 (US); B60W30/165 (EP, US); B60W30/18154 (EP, US); B60W30/18163 (EP, US); B60W40/06 (US); B60W40/072 (US); B60W40/076 (US); B60W50/14 (US); B62D15/025 (EP); B62D15/0255 (EP); B62D15/0265 (EP, US); G01C11/04 (US); G01C21/10 (US); G01C21/26 (US); G01C21/30 (EP, US); G05D1/0088 (US); G05D1/0246 (US); G06K9/00798 (EP, US); G06K9/00805 (EP); G06K9/00825 (EP); G06T7/285 (US); G06T7/292 (US); G06T7/70 (US); G06T7/97 (US); G08G1/09623 (EP, US); G08G1/166 (EP, US); G08G1/167 (EP, US); H04N5/247 (US);** B60K2031/0025 (US); B60T2201/022 (EP, US); B60T2201/03 (EP); B60T2201/08 (EP, US); B60T2210/34 (EP, US); B60W2050/0075 (EP, US); B60W2420/403 (US); B60W2420/42 (EP, US); B60W2510/0604 (US); B60W2510/18 (US); B60W2510/20 (US); B60W2520/14 (US); B60W2552/30 (EP); B60W2554/4041 (EP, US); B60W2554/80 (US); B60W2556/10 (EP, US); B60W2556/50 (EP, US); B60W2710/18 (US); B60W2710/20 (US); B62D15/0255 (US); B62D6/00 (US); G01S5/0027 (US); G05D1/0214 (US); G05D1/0253 (US); G05D1/0278 (US); G06T2207/10028 (US); G06T2207/30256 (US); G06T2207/30261 (US); G08G1/09 (US); G08G1/22 (EP, US); H04N7/18 (US)

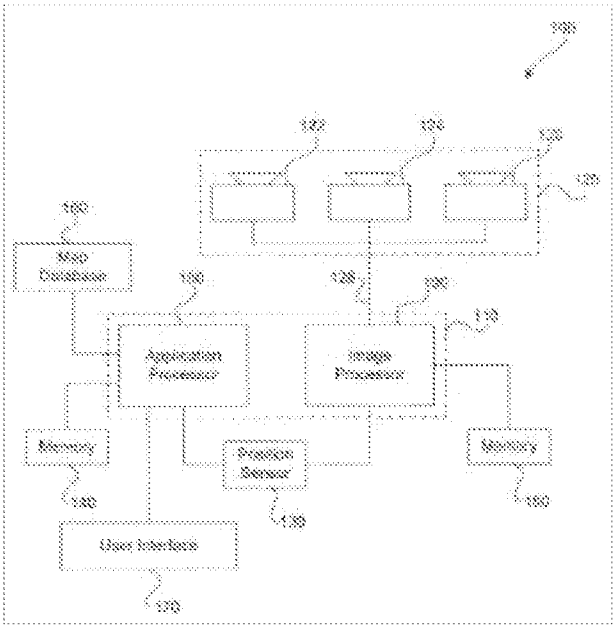
IPR2025-00944
Tesla EX1002 Page 1055

Application number: EP20140843184 20141204 Global Dossier

Priority number(s): US201361911490P 20131204 ; US201461993084P 20140514 ; US201461993111P 20140514 ; US201462015524P 20140623 ; US201462022221P 20140709 ; US201462040224P 20140821 ; US201462040269P 20140821 ; WO2014IB03151 20141204

Also published as: EP3092599 (B1) CN106030609 (A) CN106030609 (B) CN110920609 (A) CN111027420 (A) more

Abstract not available for EP3092599 (A1)
Abstract of corresponding document: US2015151725 (A1)



Systems and methods use cameras to provide autonomous navigation features. In one implementation, a driver assist navigation system may include at least one image capture device configured to acquire a plurality of images of an area in a vicinity of the vehicle; a data interface; and at least one processing device. The at least one processing device may be configured to: receive the plurality of images via the data interface; identify, based on analysis of the plurality of images, a trigger for stopping the vehicle; and based on the identified trigger, cause the vehicle to stop according to a braking profile including a first segment associated with a first deceleration rate, a second segment which includes a second deceleration rate less than the first deceleration rate, and a third segment in which a level of braking is decreased as a target stopping location is approached, as determined based on the analysis of the plurality of images.

Espacenet

Description: EP3092599 (A1) — 2016-11-16

SYSTEMS AND METHODS FOR MIMICKING A LEADING VEHICLE

Description of EP3092599 (B1)

A high quality text as facsimile in your desired language may be available amongst the following family members:

CN106030609 (A) US2015151725 (A1) WO2015083009 (A1) CN110920609 (A) CN111027420 (A) US2015151742 (A1) US2015151751 (A1) US2015151753 (A1) US2015153735 (A1) US2015367850 (A1) US2016052514 (A1) US2016167650 (A1) US2016229410 (A1) US2016347322 (A1) US2019248372 (A1) US2020094839 (A1) US2020108834 (A1) US2020331477 (A1) more

The EPO does not accept any responsibility for the accuracy of data and information originating from other authorities than the EPO; in particular, the EPO does not guarantee that they are complete, up-to-date or fit for specific purposes.

BACKGROUND

I. Technical Field

[0001] The present disclosure relates generally to autonomous vehicle navigation and, more specifically, to systems and methods that use cameras to provide autonomous vehicle navigation features.

II. Background Information

[0002] As technology continues to advance, the goal of a fully autonomous vehicle that is capable of navigating on roadways is on the horizon. Primarily, an autonomous vehicle may be able to identify its environment and navigate without input from a human operator. Autonomous vehicles may also take into account a variety of factors and make appropriate decisions based on those factors to safely and accurately reach an intended destination. For example, various objects-such as other vehicles and pedestrians-are encountered when a vehicle typically travels a roadway. Autonomous driving systems may recognize these objects in a vehicle's environment and take appropriate and timely action to avoid collisions. Additionally, autonomous driving systems may identify other indicators-such as traffic signals, traffic signs, and lane markings-that regulate vehicle movement (e.g., when the vehicle must stop and may go, a speed at which the vehicle must not exceed, where the vehicle must be positioned on the roadway, etc.). Autonomous driving systems may need to determine

when a vehicle should change lanes, turn at intersections, change roadways, etc. As is evident from these examples, many factors may need to be addressed in order to provide an autonomous vehicle that is capable of navigating safely and accurately.

[0003] " Vision-based Long-distance Lane Perception and Front Vehicle Location for Full Autonomous Vehicles on Highway Roads" by Xin Liu , et al is directed to a vision-based long-distance lane perception and front location method for decision making by full autonomous vehicles on highway roads.

[0004] " A Transportable Neural-Network Approach to Autonomous Vehicle Following" by Nasser Kehtarnavaz , et al is directed to utilizing neural networks to control the speed and steering of a vehicle for the purpose of performing autonomous vehicle following, while following a leading vehicle.

[0005] US 2011/0190972 A1 is directed to systems for detecting the presence of stationary and non-stationary objects in the vicinity of a travelling vehicle, and controlling vehicle operational parameters in response to presence of such objects.

[0006] US 8,504,233 B1 is directed to a method for controlling a vehicle in an autonomous mode, which includes obtaining lane information that provides an estimated location of a lane on a road on which the vehicle is travelling and, if the lane information has become unavailable or unreliable, using sensor(s) to monitor neighboring vehicle(s) and controlling the vehicle to maintain at least a predetermined minimum distance between the vehicle and the neighboring vehicle(s).

[0007] JP2012-168592 is directed to a method for environment recognition of a vehicle. It comprises recognizing an arrow signal of a traffic light which is located far from the vehicle. It can also comprise performing a preceding vehicle detection and means to detect lanes detection or the road surface.

SUMMARY

[0008] Embodiments consistent with the present disclosure provide systems and methods for autonomous vehicle navigation. The disclosed embodiments may use cameras to provide autonomous vehicle navigation features. For example, consistent with the disclosed embodiments, the disclosed systems may include one, two, or more cameras that monitor the environment of a vehicle and cause a navigational response based on an analysis of images captured by one or more of the cameras.

[0009] In accordance with aspects of the invention there are provided a driver assist navigation system, a vehicle, and a method for navigating a vehicle as set out in claims 1, 10, and 11 respectively.

[0010] Consistent with a disclosed embodiment, a driver assist navigation system is provided for a vehicle. The system may include at least one image capture device configured to acquire a plurality of images of an area in a vicinity of the vehicle; a data interface; and at least one processing device. The at least one processing device may be configured to: receive the plurality of images via the data interface; determine from the plurality of images a first lane constraint on a first side of the vehicle; determine from the plurality of images a second lane constraint on a second side of the vehicle opposite to the first side of the vehicle, wherein the first and second lane constraints define a lane within which the vehicle travels and wherein a first distance corresponds

Tesla EX1002 Page 1058 IPR2025-00944

to a distance between the first side of the vehicle and the first lane constraint and a second distance corresponds to a distance between the second side of the vehicle and the second lane constraint; determine, based on the plurality of images, whether a lane offset condition exists on the first side of the vehicle; if a lane offset condition exists on the first side of the vehicle, cause the vehicle to travel within the first and second lane constraints such that the first distance is greater than the second distance; determine, based on the plurality of images, whether a lane offset condition exists on the second side of the vehicle; and if a lane offset condition exists on the second side of the vehicle, cause the vehicle to travel within the first and second lane constraints such that the first distance is less than the second distance.

[0011] Consistent with the disclosed embodiment, the at least one processing device may be further configured to cause the first distance to be substantially the same as the second distance if, based on the plurality of images, a lane offset condition is determined to exist on both the first and second sides of the vehicle. The at least one processing device may be configured to determine, based on the plurality of images, whether a lane offset condition is substantially non-existent, and if the lane offset condition is substantially non-existent, cause the vehicle to travel within the first and second lane constraints such that the first distance is substantially equal to the second distance. If the driver assist navigation system determines that a lane offset condition exists, the at least one processing device may be configured to cause the first distance to be at least 1.25 times greater than the second distance. Alternatively, the at least one processing device may be further configured to cause the first distance to be at least 2 times greater than the second distance. If the driver assist navigation system determines that the lane offset condition includes a curved road, the at least one processing device may be configured to cause the vehicle to travel within the first and second lane constraints such that the first distance is toward an outside of the curve relative to the vehicle, and the second distance is toward an inside of the curve relative to the vehicle. If the driver assist navigation system determines that the lane offset condition includes a presence of one or more objects on one side of the lane, the at least one processing device may be configured to cause the vehicle to travel within the first and second lane constraints such that the first distance, which is greater than the second distance, is provided toward the one or more objects. The one or more objects may include parked cars, a wall, and/or pedestrians, and may have a height of greater than 10 cm from a road surface. At least one of the first and second lane constraints may include a line marked on a road surface, an edge of a road surface, or a determined midpoint of a road surface width. Determination of whether a lane offset condition exists on the first side of the vehicle may be performed based on a position of a target object on the first side of the vehicle. The causing of the vehicle to travel within the first and second lane constraints such that the first distance is greater than the second distance may be overridden based on input from a driver of the vehicle.

[0012] Consistent with another disclosed embodiment, a vehicle may include a first vehicle side; a second vehicle side opposite the first vehicle side; at least one image capture device configured to acquire a plurality of images of an area in a vicinity of the vehicle; a data interface; and at least one processing device. The at least one processing device may be configured to: receive the plurality of images via the data interface; determine from the plurality of images a first lane constraint on the first vehicle side; determine from the plurality of images a second lane constraint on the second vehicle side, wherein the first and second lane constraints define a lane within which the vehicle travels and wherein a first distance corresponds to a distance between the first vehicle side and the first lane constraint and a second distance

corresponds to a distance between the second vehicle side and the second lane constraint; determine, based on the plurality of images, whether a lane offset condition exists on the first vehicle side; if a lane offset condition exists on the first vehicle side, cause the vehicle to travel within the first and second lane constraints such that the first distance is greater than the second distance; determine, based on the plurality of images, whether a lane offset condition exists on the second vehicle side; and if a lane offset condition exists on the second vehicle side, cause the vehicle to travel within the first and second lane constraints such that the first distance is less than the second distance.

[0013] Consistent with the disclosed embodiment, the at least one processing device may be further configured to cause the first distance to be substantially the same as the second distance if, based on the plurality of images, a lane offset condition is determined to exist on both the first and second vehicle sides. The at least one processing device may be configured to determine, based on the plurality of images, whether a lane offset condition is substantially non-existent, and if the lane offset condition is substantially non-existent, cause the vehicle to travel within the first and second lane constraints such that the first distance is substantially equal to the second distance. The lane offset condition may include a curved road, and the at least one processing device may be configured to cause the vehicle to travel within the first and second lane constraints such that the first distance is toward an outside of the curve relative to the vehicle, and the second distance is toward an inside of the curve relative to the vehicle. The lane offset condition may include a presence of one or more objects on one side of the lane, and the at least one processing device may be configured to cause the vehicle to travel within the first and second lane constraints such that the first distance, which is greater than the second distance, is provided toward the one or more objects. At least one of the first and second lane constraints may include an edge of a road surface and/or a determined midpoint of a road surface width. Determination of whether a lane offset condition exists on the first side of the vehicle may be performed based on a position of a target object on the first side of the vehicle. The causing of the vehicle to travel within the first and second lane constraints such that the first distance is greater than the second distance may be overridden based on input from a driver of the vehicle.

[0014] Consistent with another disclosed embodiment, a method is provided for navigating a vehicle. The method may include acquiring, using at least one image capture device, a plurality of images of an area in the vicinity of the vehicle; determining from the plurality of images a first lane constraint on a first side of the vehicle; determining from the plurality of images a second lane constraint on a second side of the vehicle opposite to the first side of the vehicle, wherein the first and second lane constraints define a lane within which the vehicle travels and wherein a first distance corresponds to a distance between the first side of the vehicle and the first lane constraint and a second distance corresponds to a distance between the second side of the vehicle and the second lane constraint; determining, based on the plurality of images, whether a lane offset condition exists on the first side of the vehicle; if a lane offset condition exists on the first side of the vehicle, causing the vehicle to travel within the first and second lane constraints such that the first distance is greater than the second distance; determining, based on the plurality of images, whether a lane offset condition exists on the second side of the vehicle; and if a lane offset condition exists on the second side of the vehicle, causing the vehicle to travel within the first and second lane constraints such that the first distance is less than the second distance.

IPR2025-00944
Tesla EX1002 Page 1060

[0015] Consistent with a disclosed embodiment, a driver assist navigation system is provided for a vehicle. The system may include at least one image capture device configured to acquire a plurality of images of an area in a vicinity of the vehicle; a data interface; and at least one processing device. The at least processing device may be configured to: receive the plurality of images via the data interface; determine from the plurality of images a current lane of travel from among a plurality of available travel lanes; and cause the vehicle to change lanes if the current lane of travel is not the same as a predetermined default travel lane.

[0016] Consistent with the disclosed embodiment, the predetermined default travel lane may be a right-most lane among the plurality of available travel lanes. The at least one processing device may be configured to activate a turn signal prior to causing the vehicle to change lanes. The at least one processing device may be configured to cause an audible announcement prior to causing the vehicle to change lanes. The at least one processing device may be configured to cause an audible indicator prior to causing the vehicle to change lanes. The predetermined default travel lane may be determined based on the plurality of images of the area in the vicinity of the vehicle, based on data received from a global positioning system, and/or based on an input received from a user.

[0017] Consistent with another disclosed embodiment, a vehicle may include a body; at least one image capture device configured to acquire a plurality of images of an area in a vicinity of the vehicle; a data interface; and at least one processing device. The at least one processing device may be configured to: receive the plurality of images via the data interface; determine from the plurality of images a current lane of travel from among a plurality of available travel lanes; and cause the vehicle to change lanes if the current lane of travel is not the same as a predetermined default travel lane.

[0018] Consistent with the disclosed embodiment, the predetermined default travel lane may be a right-most lane among the plurality of available travel lanes. The at least one processing device may be configured to activate a turn signal prior to causing the vehicle to change lanes. The at least one processing device may be configured to cause an audible announcement prior to causing the vehicle to change lanes. The at least one processing device may be configured to cause an audible indicator prior to causing the vehicle to change lanes. The predetermined default travel lane may be determined based on the plurality of images of the area in the vicinity of the vehicle, based on data received from a global positioning system, and/or based on an input received from a user.

[0019] Consistent with another disclosed embodiment, a method is provided for navigating a vehicle. The method may include acquiring, using at least one image capture device, a plurality of images of an area in a vicinity of the vehicle; determining from the plurality of images a current lane of travel from among a plurality of available travel lanes; and causing the vehicle to change lanes if the current lane of travel is not the same as a predetermined default travel lane.

[0020] Consistent with the disclosed embodiment, the predetermined default travel lane may be a right-most lane among the plurality of available travel lanes. A turn signal may be activated prior to causing the vehicle to change lanes. The predetermined default travel lane may be determined based on the plurality of images of the area in the vicinity of the vehicle.

Tesla EX1002 Page 1061 IPR2025-00944

IPR2025-00944
Tesla EX1002 Page 1062

[0021] Consistent with a disclosed embodiment, a driver assist navigation system is provided for a vehicle. The system may include at least one image capture device configured to acquire a plurality of images of an area in a vicinity of the vehicle; a data interface; and at least one processing device. The at least one processing device may be configured to: receive the plurality of images via the data interface; recognize a curve to be navigated based on map data and vehicle position information; determine an initial target velocity for the vehicle based on at least one characteristic of the curve as reflected in the map data; adjust a velocity of the vehicle to the initial target velocity; determine, based on the plurality of images, one or more observed characteristics of the curve; determine an updated target velocity based on the one or more observed characteristics of the curve; and adjust the velocity of the vehicle to the updated target velocity.

[0022] Consistent with the disclosed embodiment, the at least one characteristic of the curve may include an expected radius of curvature for the curve as determined from the map data. The initial target velocity may be determined based on the expected radius of curvature for the curve and a lateral acceleration limit. The lateral acceleration limit may be between about 0.2 m/sec<2>and 3.0 m/sec<2>. Alternatively, the lateral acceleration limit may be between about 1.5 m/sec<2>and 3.0 m/sec<2>. The lateral acceleration limit may be adjusted based on a road width determined based on the plurality of images. The one or more observed characteristics of the curve may include an actual radius of curvature for the curve as determined through analysis of the plurality of images. The updated target velocity may be determined based on the actual radius of curvature for the curve and a lateral acceleration limit. The one or more observed characteristics of the curve may include a rate of change in an actual radius of curvature for the curve as determined through analysis of the plurality of images. The one or more observed characteristics of the curve may include a degree of banking associated with the curve as determined through analysis of the plurality of images. The one or more observed characteristics of the curve may include information, determined through analysis of the plurality of images, relating to the curve conveyed by one or more road signs. The initial target velocity may also be based on a focal range of the at least one image capture device. The initial target velocity may also be based on whether a traffic light is determined to be located along the curve. The initial target velocity may also be based on whether a traffic light is determined to be located within a predetermined distance after the curve.

[0023] Consistent with another disclosed embodiment, a vehicle may include a body; at least one image capture device configured to acquire a plurality of images of an area in a vicinity of the vehicle; a data interface; and at least one processing device. The at least one processing device may be configured to: receive the plurality of images via the data interface; recognize a curve to be navigated based on map data and vehicle position information; determine an initial target velocity for the vehicle based on at least one characteristic of the curve as reflected in the map data; adjust a velocity of the vehicle to the initial target velocity; determine, based on the plurality of images, one or more observed characteristics of the curve; determine an updated target velocity based on the one or more observed characteristics of the curve; and adjust the velocity of the vehicle to the updated target velocity.

[0024] Consistent with the disclosed embodiment, the at least one characteristic of the curve may include an expected radius of curvature for the curve as determined from the map data. The initial target velocity may be determined based on the expected radius of curvature for the curve and a lateral acceleration limit. The lateral acceleration

limit may be adjusted based on a road width determined based on the plurality of images. The one or more observed characteristics of the curve may include an actual radius of curvature for the curve as determined through analysis of the plurality of images. The updated target velocity may be determined based on the actual radius of curvature for the curve and a lateral acceleration limit. The one or more observed characteristics of the curve may include a rate of change in an actual radius of curvature for the curve as determined through analysis of the plurality of images. The one or more observed characteristics of the curve may include a degree of banking associated with the curve as determined through analysis of the plurality of images. The one or more observed characteristics of the curve may include information, determined through analysis of the plurality of images, relating to the curve conveyed by one or more road signs. The initial target velocity may also be based on a focal range of the at least one image capture device. The initial target velocity may also be based on whether a traffic light is determined to be located along the curve. The initial target velocity may also be based on whether a traffic light is determined to be located within a predetermined distance after the curve.

[0025] Consistent with another disclosed embodiment, a method is provided for navigating a vehicle. The method may include acquiring, using at least one image capture device, a plurality of images of an area in a vicinity of the vehicle; recognizing a curve to be navigated based on map data and vehicle position information; determining an initial target velocity for the vehicle based on at least one characteristic of the curve as reflected in the map data; adjusting a velocity of the vehicle to the initial target velocity; determining, based on the plurality of images, one or more observed characteristics of the curve; determining an updated target velocity based on the one or more observed characteristics of the curve; and adjusting the velocity of the vehicle to the updated target velocity.

[0026] Consistent with a disclosed embodiment, a driver assist navigation system is provided for a primary vehicle. The system may include at least one image capture device configured to acquire a plurality of images of an area in a vicinity of the primary vehicle; a data interface; and at least one processing device. The at least one processing device may be configured to: receive the plurality of images via the data interface; determine, from at least some of the plurality of images, a first lane constraint on a first side of the primary vehicle; determine, from at least some of the plurality of images, a second lane constraint on a second side of the primary vehicle opposite to the first side of the primary vehicle, wherein the first and second lane constraints define a lane within which the primary vehicle travels; cause the primary vehicle to travel within the first and second lane constraints; locate in the plurality of images a leading vehicle; determine, based on the plurality of images, at least one action taken by the leading vehicle; and cause the primary vehicle to mimic the at least one action of the leading vehicle.

[0027] Consistent with the disclosed embodiment, at least one of the first and second lane constraints may include a line marked on a road surface. At least one of the first and second lane constraints may include an edge of a road surface. At least one of the first and second lane constraints may include a determined midpoint of a road surface width. The at least one action taken by the leading vehicle may include a turn, and the at least one processing device may cause the primary vehicle to execute a similar turn beginning at a location at which the leading vehicle initiated its turn. The at least one action taken by the leading vehicle may include an acceleration. The at least one action taken by the leading vehicle may include braking. The at least one action taken by the

IPR2025-00944
Tesla EX1002 Page 1063

leading vehicle may include providing a first offset distance on a side of the leading vehicle adjacent to the first lane constraint that is different from a second offset distance on a side of the leading vehicle adjacent to the second lane constraint. The at least one processing device may be further configured to determine, based on the plurality of images or based on map data and position information, whether the leading vehicle is turning at an intersection, and if the leading vehicle is determined to be turning at an intersection, decline to mimic the turn of the leading vehicle. The at least one processing device may be configured to determine, based on the plurality of images or based on map data and position information, whether the leading vehicle is turning at an intersection, determine whether the leading vehicle is changing lanes within the intersection, mimic the turn of the leading vehicle if the leading vehicle is determined to be turning at an intersection and not changing lanes within the intersection, and decline to mimic the turn of the leading vehicle if the leading vehicle is determined to be turning at an intersection, but changing lanes within the intersection. The at least one processing device may be configured to acquire a first leading vehicle based on the plurality of images, acquire a second leading vehicle based on the plurality of images, and cause the primary vehicle to mimic the first leading vehicle if, based on the plurality of images, the first leading vehicle is determined to be continuing on a path having less of a turning radius than a path followed by the second leading vehicle. The at least one processing device may be configured to determine position information for the leading vehicle and compare the position information to predetermined map data, create a navigation history for the leading vehicle by tracking movements of the leading vehicle relative to the predetermined map data, and determine whether to cause the primary vehicle to mimic the at least one action of the leading vehicle based on the navigation history of the leading vehicle. The at least one processing device may be configured to cause the primary vehicle to mimic a turn of the leading vehicle, even where no turn is detected based on the plurality of images, when the navigation history indicates that the leading vehicle has successfully navigated at least one prior turn.

[0028] Consistent with another disclosed embodiment, a primary vehicle may include a body; at least one image capture device configured to acquire a plurality of images of an area in a vicinity of the primary vehicle; a data interface; and at least one processing device. The at least one processing device may be configured to: receive the plurality of images via the data interface; determine, from at least some of the plurality of images, a first lane constraint on a first side of the primary vehicle; determine, from at least some of the plurality of images, a second lane constraint on a second side of the primary vehicle opposite to the first side of the primary vehicle, wherein the first and second lane constraints define a lane within which the primary vehicle travels; cause the primary vehicle to travel within the first and second lane constraints; locate in the plurality of images a leading vehicle; determine, based on the plurality of images, at least one action taken by the leading vehicle; and cause the primary vehicle to mimic the at least one action of the leading vehicle.

[0029] Consistent with the disclosed embodiment, at least one of the first and second lane constraints may include a line marked on a road surface. At least one of the first and second lane constraints may include an edge of a road surface. At least one of the first and second lane constraints may include a determined midpoint of a road surface width. The at least one action taken by the leading vehicle may include a turn, and the at least one processing device may cause the primary vehicle to execute a similar turn beginning at a location at which the leading vehicle initiated its turn. The at least one action taken by the leading vehicle may include an acceleration. The at least one action

Tesla EX1002 Page 1064 IPR2025-00944

taken by the leading vehicle may include braking. The at least one action taken by the leading vehicle may include providing a first offset distance on a side of the leading vehicle adjacent to the first lane constraint that is different from a second offset distance on a side of the leading vehicle adjacent to the second lane constraint. The at least one processing device may be further configured to determine, based on the plurality of images or based on map data and position information, whether the leading vehicle is turning at an intersection, and if the leading vehicle is determined to be turning at an intersection, decline to mimic the turn of the leading vehicle. The at least one processing device may be configured to determine, based on the plurality of images or based on map data and position information, whether the leading vehicle is turning at an intersection, determine whether the leading vehicle is changing lanes within the intersection, mimic the turn of the leading vehicle if the leading vehicle is determined to be turning at an intersection and not changing lanes within the intersection, and decline to mimic the turn of the leading vehicle if the leading vehicle is determined to be turning at an intersection, but changing lanes within the intersection. The at least one processing device may be configured to acquire a first leading vehicle based on the plurality of images, acquire a second leading vehicle based on the plurality of images, and cause the primary vehicle to mimic the first leading vehicle if, based on the plurality of images, the first leading vehicle is determined to be continuing on a path having less of a turning radius than a path followed by the second leading vehicle. The at least one processing device may be configured to determine position information for the leading vehicle and compare the position information to predetermined map data, create a navigation history for the leading vehicle by tracking movements of the leading vehicle relative to the predetermined map data, and determine whether to cause the primary vehicle to mimic the at least one action of the leading vehicle based on the navigation history of the leading vehicle. The at least one processing device may be configured to cause the primary vehicle to mimic a turn of the leading vehicle, even where no turn is detected based on the plurality of images, when the navigation history indicates that the leading vehicle has successfully navigated at least one prior turn.

[0030] Consistent with another disclosed embodiment, a method is provided for navigating a primary vehicle. The method ma include acquiring, using at least one image capture device, a plurality of images of an area in a vicinity of the primary vehicle; determining, from at least some of the plurality of images, a first lane constraint on a first side of the primary vehicle; determining, from at least some of the plurality of images, a second lane constraint on a second side of the primary vehicle opposite to the first side of the primary vehicle, wherein the first and second lane constraints define a lane within which the primary vehicle travels; causing the primary vehicle to travel within the first and second lane constraints; locating in the plurality of images a leading vehicle; determining, based on the plurality of images, at least one action taken by the leading vehicle; and causing the primary vehicle to mimic the at least one action of the leading vehicle.

[0031] Consistent with a disclosed embodiment, a driver assist navigation system is provided for a user vehicle. They system may include at least one image capture device configured to acquire a plurality of images of an area in a vicinity of the user vehicle; a data interface; and at least one processing device. The at least one processing device may be configured to: receive the plurality of images via the data interface; determine from the plurality of images a first lane constraint on a first side of the user vehicle; determine from the plurality of images a second lane constraint on a second side of the user vehicle opposite to the first side of the user vehicle, wherein the

Tesla EX1002 Page 1065 IPR2025-00944

first and second lane constraints define a lane within which the user vehicle travels; acquire, based on the plurality of images, a target vehicle; enable the user vehicle to pass the target vehicle if the target vehicle is determined to be in a lane different from the lane in which the user vehicle is traveling; monitor a position of the target vehicle based on the plurality of images; and cause the user vehicle to abort the pass before completion of the pass, if the target vehicle is determined to be entering the lane in which the user vehicle is traveling.

[0032] Consistent with the disclosed embodiment, causing the user vehicle to abort the pass may include causing an application of brakes in the user vehicle. Causing the user vehicle to abort the pass may include causing an audible announcement. Causing the user vehicle to abort the pass may include causing the user vehicle to stop accelerating. Monitoring the position of the target vehicle may include estimating a speed of the target vehicle. At least one of the first and second lane constraints may include a line marked on a road surface, an edge of a road surface, and/or a determined midpoint of a road surface width.

[0033] Consistent with another disclosed embodiment, a user vehicle may include at least one image capture device configured to acquire a plurality of images of an area in a vicinity of the user vehicle; a data interface; and at least one processing device. The at least one processing device may be configured to: receive the plurality of images via the data interface; determine from the plurality of images a first lane constraint on a first side of the user vehicle; determine from the plurality of images a second lane constraint on a second side of the user vehicle opposite to the first side of the user vehicle, wherein the first and second lane constraints define a lane within which the user vehicle travels; acquire, based on the plurality of images, a target vehicle; enable the user vehicle to pass the target vehicle if the target vehicle is determined to be in a lane different from the lane in which the user vehicle is traveling; monitor a position of the target vehicle based on the plurality of images; and cause the user vehicle to abort the pass before completion of the pass, if the target vehicle is determined to be entering the lane in which the user vehicle is traveling.

[0034] Consistent with the disclosed embodiment, causing the user vehicle to abort the pass may include causing an application of brakes in the user vehicle. Causing the user vehicle to abort the pass may include causing an audible announcement. Causing the user vehicle to abort the pass may include causing the user vehicle to stop accelerating. Monitoring the position of the target vehicle may include estimating a speed of the target vehicle. At least one of the first and second lane constraints may include a line marked on a road surface, an edge of a road surface, and/or a determined midpoint of a road surface width.

[0035] Consistent with another disclosed embodiment, a method is provided for navigating a user vehicle. The method may include acquiring, using at least one image capture device, a plurality of images of an area in a vicinity of the user vehicle; determining from the plurality of images a first lane constraint on a first side of the user vehicle; determining from the plurality of images a second lane constraint on a second side of the user vehicle opposite to the first side of the user vehicle, wherein the first and second lane constraints define a lane within which the user vehicle travels; acquiring, based on the plurality of images, a target vehicle; enabling the user vehicle to pass the target vehicle if the target vehicle is determined to be in a lane different from the lane in which the user vehicle is traveling; monitoring a position of the target vehicle based on the plurality of images; and causing the user vehicle to abort the pass

before completion of the pass, if the target vehicle is determined to be entering the lane in which the user vehicle is traveling.

[0036] Consistent with the disclosed embodiment, causing the user vehicle to abort the pass may include causing an application of brakes in the user vehicle. Causing the user vehicle to abort the pass may include causing the user vehicle to stop accelerating. At least one of the first and second lane constraints may include a line marked on a road surface, an edge of a road surface, and/or a determined midpoint of a road surface width.

[0037] Consistent with a disclosed embodiment, a driver assist navigation system is provided for a user vehicle. The system may include at least one image capture device configured to acquire a plurality of images of an area in a vicinity of the user vehicle; a data interface; and at least one processing device. The at least one processing device may be configured to: receive the plurality of images via the data interface; determine from the plurality of images a first lane constraint on a first side of the user vehicle; determine from the plurality of images a second lane constraint on a second side of the user vehicle opposite to the first side of the user vehicle, wherein the first and second lane constraints define a lane within which the user vehicle travels; determine, based on the plurality of images, whether an encroaching vehicle is approaching from a side of the user vehicle; and cause the user vehicle to maintain a current velocity and to travel within the first and second lane constraints such that a first distance, on the side of the user vehicle that the encroaching vehicle is approaching from, is greater than a second distance.

[0038] Consistent with the disclosed embodiment, the at least one processing device may be further configured to cause the user vehicle to take evasive action if analysis of the plurality of images indicates that the encroaching vehicle has crossed at least one collision threshold. The evasive action may include changing lanes, altering course, and/or braking. The at least one collision threshold may include a minimum predetermined distance to maintain between the user vehicle and the encroaching vehicle. The at least one collision threshold may include a time to collision. The at least one processing device may be configured to cause the user vehicle to return to a center of the lane, within which the user vehicle originally traveled, after completion of the evasive action. The at least one image capture device may have a field of view of at least 140 degrees. The at least one processing device may be configured to cause the user vehicle to resume a preset speed after braking in response to an encroaching vehicle. The at least one processing device may be configured to cause the user vehicle to return to a lane location within the first and second lane constraints such that the first distance between the user vehicle and the first lane constraint is substantially equal to the second distance between the user vehicle and the second lane constraint after altering course in response to an encroaching vehicle.

[0039] Consistent with another disclosed embodiment, a user vehicle may include at least one image capture device configured to acquire a plurality of images of an area in a vicinity of the user vehicle; a data interface; and at least one processing device. The at least one processing device may be configured to: receive the plurality of images via the data interface; determine from the plurality of images a first lane constraint on a first side of the user vehicle; determine from the plurality of images a second lane constraint on a second side of the user vehicle opposite to the first side of the user vehicle, wherein the first and second lane constraints define a lane within which the user vehicle travels; determine, based on the plurality of images, whether an encroaching

IPR2025-00944
Tesla EX1002 Page 1067

vehicle is approaching from a side of the user vehicle; and cause the user vehicle to maintain a current velocity and to travel within the first and second lane constraints such that a first distance, on the side of the user vehicle that the encroaching vehicle is approaching from, is greater than a second distance.

[0040] Consistent with the disclosed embodiment, the at least one processing device may be further configured to cause the user vehicle to take evasive action if analysis of the plurality of images indicates that the encroaching vehicle has crossed at least one collision threshold. The evasive action may include changing lanes, altering course, and/or braking. The at least one collision threshold may include a minimum predetermined distance to maintain between the user vehicle and the encroaching vehicle. The at least one collision threshold may include a time to collision. The at least one processing device may be configured to cause the user vehicle to return to a center of the lane, within which the user vehicle originally traveled, after completion of the evasive action. The at least one image capture device may have a field of view of at least 140 degrees. The at least one processing device may be configured to cause the user vehicle to resume a preset speed after braking in response to an encroaching vehicle. The at least one processing device may be configured to cause the user vehicle to return to a lane location within the first and second lane constraints such that the first distance between the user vehicle and the first lane constraint is substantially equal to the second distance between the user vehicle and the second lane constraint after altering course in response to an encroaching vehicle.

[0041] Consistent with another disclosed embodiment, a method for navigating a user vehicle is provided. The method may include acquiring, using at least one image capture device, a plurality of images of an area in a vicinity of the user vehicle; determining from the plurality of images a first lane constraint on a first side of the user vehicle; determining from the plurality of images a second lane constraint on a second side of the user vehicle opposite to the first side of the user vehicle, wherein the first and second lane constraints define a lane within which the user vehicle travels; determining, based on the plurality of images, whether an encroaching vehicle is approaching from a side of the user vehicle; and causing the user vehicle to maintain a current velocity and to travel within the first and second lane constraints such that a first distance, on the side of the user vehicle that the encroaching vehicle is approaching from, is greater than a second distance.

[0042] Consistent with a disclosed embodiment, a driver assist navigation system is provided for a primary vehicle. They system may include at least one image capture device configured to acquire a plurality of images of an area in a vicinity of the primary vehicle; a data interface; and at least one processing device. The at least processing device may be configured to: receive the plurality of images via the data interface; identify a target object within the plurality of images; monitor, via the plurality of images, a motion of the target object and a distance between the primary vehicle and the target object; determine an indicator of an intercept time between the primary vehicle and the target object based on the monitored motion and the distance between the primary vehicle and the target object; and cause a response in the primary vehicle based on a comparison of the intercept time to a plurality of predetermined intercept thresholds.

[0043] Consistent with the disclosed em" odiment, the target object may be a target vehicle traveling in a lane different from the primary vehicle. The intercept time may be based on an indicator of speed of the tracked object, as determined based on the plurality of images. The intercept time may be based on an indicator of an acceleration

IPR2025-00944
Tesla EX1002 Page 1068

of the tracked object, as determined based on the plurality of images. The at least one processing device may be further configured to determine whether the target object is traveling in the same lane as the primary vehicle, and forego the response in the primary vehicle if the target object is determined to be traveling in a lane different from the primary vehicle. Determination of whether the target object is traveling in the same lane as the primary vehicle may be based on the plurality of images. Determination of whether the target object is traveling in the same lane as the primary vehicle may be based on at least one of predetermined map data or yaw rate of the primary vehicle. The response may be based on an averaged motion profile associated with the target object if the indicator of intercept time indicates that the target object is outside a first intercept threshold relative to the vehicle. The response may be based on actual motion of the target object if the indicator of intercept time indicates that the target object is between the first intercept threshold and a second intercept threshold relative to the vehicle. The response may include an emergency avoidance action if the indicator of intercept time indicates that the target object is within the second intercept threshold. The emergency avoidance action may include changing lanes and/or emergency braking.

[0044] Consistent with another disclosed embodiment, a primary vehicle may include a body; at least one image capture device configured to acquire a plurality of images of an area in a vicinity of the primary vehicle; a data interface; and at least one processing device. The at least one processing device may be configured to: receive the plurality of images via the data interface; identify a target object within the plurality of images; monitor, via the plurality of images, a motion of the target object and a distance between the primary vehicle and the target object; determine an indicator of an intercept time between the primary vehicle and the target object based on the monitored motion and the distance between the primary vehicle and the target object; and cause a response in the primary vehicle based on a comparison of the intercept time to a plurality of predetermined intercept thresholds.

[0045] Consistent with the disclosed embodiment, the target object may be a target vehicle traveling in a lane different from the primary vehicle. The intercept time may be based on an indicator of speed of the tracked object, as determined based on the plurality of images. The intercept time may be based on an indicator of an acceleration of the tracked object, as determined based on the plurality of images. The at least one processing device may be further configured to determine whether the target object is traveling in the same lane as the primary vehicle, and forego the response in the primary vehicle if the target object is determined to be traveling in a lane different from the primary vehicle. Determination of whether the target object is traveling in the same lane as the primary vehicle may be based on the plurality of images. Determination of whether the target object is traveling in the same lane as the primary vehicle may be based on at least one of predetermined map data or yaw rate of the primary vehicle. The response may be based on an averaged motion profile associated with the target object if the indicator of intercept time indicates that the target object is outside a first intercept threshold relative to the vehicle. The response may be based on actual motion of the target object if the indicator of intercept time indicates that the target object is between the first intercept threshold and a second intercept threshold relative to the vehicle. The response may include an emergency avoidance action if the indicator of intercept time indicates that the target object is within the second intercept threshold. The emergency avoidance action may include changing lanes and/or emergency braking.

IPR2025-00944
Tesla EX1002 Page 1069

[0046] Consistent with another disclosed embodiment, a method for navigating a primary vehicle may include acquiring, via at least one image capture device, a plurality of images of an area in a vicinity of the primary vehicle; identifying a target object within the plurality of images; monitoring, based on the plurality of images, a motion of the target object and a distance between the primary vehicle and the target object; determining an indicator of an intercept time between the primary vehicle and the target object based on the monitored motion and the distance between the primary vehicle and the target object; and causing a response in the primary vehicle based on a comparison of the intercept time to a plurality of predetermined intercept thresholds.

[0047] Consistent with a disclosed embodiment, a driver assist navigation system is provided for a vehicle. The system may include at least one image capture device configured to acquire a plurality of images of an area in a vicinity of the vehicle; a data interface; and at least one processing device. The at least one processing device may be configured to: receive the plurality of images via the data interface; identify, based on analysis of the plurality of images, a trigger for stopping the vehicle; and based on the identified trigger, cause the vehicle to stop according to a braking profile including a first segment associated with a first deceleration rate, a second segment which includes a second deceleration rate less than the first deceleration rate, and a third segment in which a level of braking is decreased as a target stopping location is approached, as determined based on the analysis of the plurality of images.

[0048] Consistent with the disclosed embodiment, reaching of the target stopping location may be accompanied by a release in braking pressure. The trigger may include an object. The object may include another vehicle in front of the vehicle. The trigger may include a traffic light signal, a stop sign, and/or a speed limit sign. The at least one processing device may be further configured to provide an audible announcement identifying the trigger. The at least one processing device may be configured to cause an audible announcement indicating that the vehicle is braking.

[0049] Consistent with another disclosed embodiment, a vehicle may include a body; at least one image capture device configured to acquire a plurality of images of an area in a vicinity of the vehicle; a data interface; and at least one processing device. The at least one processing device may be configured to: receive the plurality of images via the data interface; identify, based on analysis of the plurality of images, a trigger for stopping the vehicle; and based on the identified trigger, cause the vehicle to stop according to a braking profile including a first segment associated with a first deceleration rate, a second segment which includes a second deceleration rate less than the first deceleration rate, and a third segment in which a level of braking is decreased as a target stopping location is approached, as determined based on the analysis of the plurality of images.

[0050] Consistent with the disclosed embodiment, reaching of the target stopping location may be accompanied by a release in braking pressure. The trigger may include an object. The object may include another vehicle in front of the vehicle. The trigger may include a traffic light signal, a stop sign, and/or a speed limit sign. The at least one processing device may be further configured to provide an audible announcement identifying the trigger. The at least one processing device may be configured to cause an audible announcement indicating that the vehicle is braking.

[0051] Consistent with another disclosed embodiment, a method is provided for navigating a vehicle. The method may include acquiring, via at least one image capture

IPR2025-00944
Tesla EX1002 Page 1070

plurality of images of an area in a vicinity of the vehicle; identifying, based on analysis of the plurality of images, a trigger for stopping the vehicle; and based on the identified trigger, causing the vehicle to stop according to a braking profile including a first segment associated with a first deceleration rate, a second segment which includes a second deceleration rate less than the first deceleration rate, and a third segment in which a level of braking is decreased as a target stopping location is approached, as determined based on the analysis of the plurality of images.

[0052] Consistent with other disclosed embodiments, non-transitory computer-readable storage media may store program instructions, which are executed by at least one processing device and perform any of the methods described herein.

[0053] The foregoing general description and the following detailed description are exemplary and explanatory.

BRIEF DESCRIPTION OF THE DRAWINGS

[0054] The accompanying drawings, which are incorporated in and constitute a part of this disclosure, illustrate various disclosed embodiments. In the drawings:
FIG. 1 is a diagrammatic representation of an exemplary system consistent with the disclosed embodiments.
FIG. 2A is a diagrammatic side view representation of an exemplary vehicle including a system consistent with the disclosed embodiments.
FIG. 2B is a diagrammatic top view representation of the vehicle and system shown in FIG. 2A consistent with the disclosed embodiments.
FIG. 2C is a diagrammatic top view representation of another embodiment of a vehicle including a system consistent with the disclosed embodiments.
FIG. 2D is a diagrammatic top view representation of yet another embodiment of a vehicle including a system consistent with the disclosed embodiments.
FIG. 2E is a diagrammatic top view representation of yet another embodiment of a vehicle including a system consistent with the disclosed embodiments.
FIG. 2F is a diagrammatic representation of exemplary vehicle control systems consistent with the disclosed embodiments.
FIG. 3A is a diagrammatic representation of an interior of a vehicle including a rearview mirror and a user interface for a vehicle imaging system consistent with the disclosed embodiments.
FIG. 3B is an illustration of an example of a camera mount that is configured to be positioned behind a rearview mirror and against a vehicle windshield consistent with the disclosed embodiments.
FIG. 3C is an illustration of the camera mount shown in FIG. 3B from a different perspective consistent with the disclosed embodiments.
FIG. 3D is an illustration of an example of a camera mount that is configured to be positioned behind a rearview mirror and against a vehicle windshield consistent with the disclosed embodiments.
FIG. 4 is an exemplary block diagram of a memory configured to store instructions for performing one or more operations consistent with the disclosed embodiments.
FIG. 5A is a flowchart showing an exemplary process for causing one or more navigational responses based on monocular image analysis consistent with disclosed embodiments.
FIG. 5B is a flowchart showing an exemplary process for detecting one or more vehicles and/or pedestrians in a set of images consistent with the disclosed embodiments.

IPR2025-00944
Tesla EX1002 Page 1071

IPR2025-00944
Tesla EX1002 Page 1072

FIG. 5C is a flowchart showing an exemplary process for detecting road marks and/or lane geometry information in a set of images consistent with the disclosed embodiments.
FIG. 5D is a flowchart showing an exemplary process for detecting traffic lights in a set of images consistent with the disclosed embodiments.
FIG. 5E is a flowchart showing an exemplary process for causing one or more navigational responses based on a vehicle path consistent with the disclosed embodiments.
FIG. 5F is a flowchart showing an exemplary process for determining whether a leading vehicle is changing lanes consistent with the disclosed embodiments.
FIG. 6 is a flowchart showing an exemplary process for causing one or more navigational responses based on stereo image analysis consistent with the disclosed embodiments.
FIG. 7 is a flowchart showing an exemplary process for causing one or more navigational responses based on an analysis of three sets of images consistent with the disclosed embodiments.
FIG. 8A is a diagrammatic representation of an exemplary vehicle traveling within a first and a second lane constraint in which a lane offset condition exists on the first vehicle side consistent with the disclosed embodiments.
FIG. 8B is a diagrammatic representation of an exemplary vehicle traveling within a first and a second lane constraint in which a lane offset condition exists on the second vehicle side consistent with the disclosed embodiments.
FIG. 8C is a diagrammatic representation of an exemplary vehicle traveling within a first and a second lane constraint on a curved road consistent with the disclosed embodiments.
FIG. 8D is a diagrammatic representation of an exemplary vehicle traveling within a first and a second lane constraint in which a lane offset condition exists on both the first vehicle side and on the second vehicle side consistent with the disclosed embodiments.
FIG. 9 is a diagrammatic representation of the memory of an exemplary navigation system consistent with the disclosed embodiments.
FIG. 10 is a flowchart of an exemplary method for navigating a vehicle consistent with the disclosed embodiments.
FIG 11 is a diagrammatic representation of an exemplary vehicle including a navigation system traveling within a default lane consistent with the disclosed embodiments.
FIG 12A is a diagrammatic representation of the memory of an exemplary navigation system consistent with the disclosed embodiments.
FIG 12B is a flowchart of an exemplary process for navigating a vehicle consistent with the disclosed embodiments.
FIGS. 13A and 13B are diagrammatic representations of an exemplary vehicle approaching, and navigating, a curve with one or more characteristics consistent with the disclosed embodiments.
FIG. 14 is an exemplary block diagram of a memory configured to store instructions for performing one or more operations consistent with the disclosed embodiments.
FIG. 15 is a flow chart showing an exemplary process for controlling the velocity of a vehicle based on a detected curve and observed characteristics of the curve consistent with disclosed embodiments.
FIGS. 16A-16D are diagrammatic representations of a primary vehicle mimicking one or more actions of a leading vehicle consistent with the disclosed embodiments.
FIG. 17 is an exemplary block diagram of a memory configured to store instructions for performing one or more operations consistent with the disclosed embodiments.
FIG. 18 is a flow chart showing an exemplary process for causing a primary vehicle to

mimic one or more actions of a leading vehicle consistent with disclosed embodiments.
FIG. 19 is a flow chart showing an exemplary process for causing a primary vehicle to decline to mimic a turn of a leading vehicle consistent with disclosed embodiments.
FIG. 20 is a flow chart showing an exemplary process for causing a primary vehicle to mimic or decline a turn of a leading vehicle consistent with disclosed embodiments.
FIG. 21 is a flow chart showing another exemplary process for causing a primary vehicle to mimic one or more actions of a leading vehicle consistent with disclosed embodiments.
FIG. 22 is a flow chart showing an exemplary process for causing a primary vehicle to mimic one or more actions of a leading vehicle based on a navigation history consistent with disclosed embodiments.
FIG. 23 is a diagrammatic representation of an exemplary vehicle traveling within a first and a second lane constraint on a roadway along on which another vehicle is traveling consistent with the disclosed embodiments.
FIG. 24A is a diagrammatic representation of a memory storing instructions for navigating a vehicle among encroaching vehicles consistent with the disclosed embodiments.
FIG. 24B is a flowchart of an exemplary process for navigating a vehicle among encroaching vehicles consistent with the disclosed embodiments.
FIG. 25 is a diagrammatic representation of an exemplary vehicle traveling within a first and a second lane constraint on a roadway along on which another vehicle is traveling consistent with the disclosed embodiments.
FIG. 26 is a diagrammatic representation of a memory storing instructions for detecting and responding to traffic laterally encroaching on a vehicle consistent with the disclosed embodiments.
FIG. 27 is a flowchart of an exemplary process for detecting and responding to traffic laterally encroaching on a vehicle consistent with the disclosed embodiments.
FIGS. 28A and 28B are diagrammatic representations of causing a response in a primary vehicle consistent with the disclosed embodiments.
FIG. 29 is an exemplary block diagram of a memory configured to store instructions for performing one or more operations consistent with the disclosed embodiments.
FIG. 30 is a flow chart showing an exemplary process for causing a response in a primary vehicle consistent with disclosed embodiments.
FIG. 31 is a flow chart showing an exemplary process for declining to cause a response in a primary vehicle consistent with disclosed embodiments.
FIG. 32A is a diagrammatic representation of an exemplary vehicle in which an object is present in the vicinity of vehicle consistent with the disclosed embodiments.
FIG. 32B illustrates an exemplary braking profile, consistent with the disclosed embodiments.
FIG. 33 is a diagrammatic representation of the memory of an exemplary navigation system consistent with the disclosed embodiments.
FIG. 34 is a flowchart of an exemplary method for implementing a braking profile while navigating a vehicle consistent with the disclosed embodiments.

DETAILED DESCRIPTION

[0055] The following detailed description refers to the accompanying drawings. Wherever possible, the same reference numbers are used in the drawings and the following description to refer to the same or similar parts. While several illustrative embodiments are described herein, modifications, adaptations and other implementations are possible. For example, substitutions, additions or modifications may be made to the components illustrated in the drawings, and the illustrative

IPR2025-00944
Tesla EX1002 Page 1073

methods described herein may be modified by substituting, reordering, removing, or adding steps to the disclosed methods. Accordingly, the following detailed description is not limited to the disclosed embodiments and examples. Instead, the proper scope is defined by the appended claims.

[0056] Disclosed embodiments provide systems and methods that use cameras to provide autonomous navigation features. In various embodiments, the system may include one, two or more cameras that monitor the environment of a vehicle. In one embodiment, the system may monitor and adjust the free space between a vehicle and the boundaries of the lane within which the vehicle is traveling. In another embodiment, the system may select a particular lane as a default lane for the vehicle to use while traveling. In another embodiment, the system may control the speed of the vehicle in different scenarios, such as while making a turn. In yet another embodiment, the system may mimic the actions of a leading vehicle. In yet another embodiment, the system may monitor a target vehicle and enable the vehicle to pass the target vehicle under certain conditions (e.g., if the target vehicle is traveling in a lane different from the lane within which the vehicle is traveling. In still yet another embodiment, the system may produce a natural feeling response to a laterally encroaching vehicle, such as a vehicle attempting to move into the lane within which the vehicle is traveling.

[0057] FIG. 1 is a block diagram representation of a system 100 consistent with the exemplary disclosed embodiments. System 100 may include various components depending on the requirements of a particular implementation. In some embodiments, system 100 may include a processing unit 110, an image acquisition unit 120, a position sensor 130, one or more memory units 140, 150, a map database 160, and a user interface 170. Processing unit 110 may include one or more processing devices. In some embodiments, processing unit 110 may include an applications processor 180, an image processor 190, or any other suitable processing device. Similarly, image acquisition unit 120 may include any number of image acquisition devices and components depending on the requirements of a particular application. In some embodiments, image acquisition unit 120 may include one or more image capture devices (e.g., cameras), such as image capture device 122, image capture device 124, and image capture device 126. System 100 may also include a data interface 128 communicatively connecting processing device 110 to image acquisition device 120. For example, data interface 128 may include any wired and/or wireless link or links for transmitting image data acquired by image accusation device 120 to processing unit 110.

[0058] Both applications processor 180 and image processor 190 may include various types of processing devices. For example, either or both of applications processor 180 and image processor 190 may include a microprocessor, preprocessors (such as an image preprocessor), graphics processors, a central processing unit (CPU), support circuits, digital signal processors, integrated circuits, memory, or any other types of devices suitable for running applications and for image processing and analysis. In some embodiments, applications processor 180 and/or image processor 190 may include any type of single or multi-core processor, mobile device microcontroller, central processing unit, etc. Various processing devices may be used, including, for example, processors available from manufacturers such as Intel®, AMD®, etc. and may include various architectures (e.g., x86 processor, ARM®, etc.).

[0059] In some embodiments, applications processor 180 and/or image processor 190 may include any of the EyeQ series of processor chips available from Mobileye®.

IPR2025-00944
Tesla EX1002 Page 1074

These processor designs each include multiple processing units with local memory and instruction sets. Such processors may include video inputs for receiving image data from multiple image sensors and may also include video out capabilities. In one example, the EyeQ2® uses 90nm-micron technology operating at 332Mhz. The EyeQ2® architecture consists of two floating point, hyper-thread 32-bit RISC CPUs (MIPS32® 34K® cores), five Vision Computing Engines (VCE), three Vector Microcode Processors (VMP®), Denali 64-bit Mobile DDR Controller, 128-bit internal Sonics Interconnect, dual 16-bit Video input and 18-bit Video output controllers, 16 channels DMA and several peripherals. The MIPS34K CPU manages the five VCEs, three VMP™ and the DMA, the second MIPS34K CPU and the multi-channel DMA as well as the other peripherals. The five VCEs, three VMP® and the MIPS34K CPU can perform intensive vision computations required by multi-function bundle applications. In another example, the EyeQ3®, which is a third generation processor and is six times more powerful that the EyeQ2®, may be used in the disclosed embodiments.

[0060] While FIG. 1 depicts two separate processing devices included in processing unit 110, more or fewer processing devices may be used. For example, in some embodiments, a single processing device may be used to accomplish the tasks of applications processor 180 and image processor 190. In other embodiments, these tasks may be performed by more than two processing devices.

[0061] Processing unit 110 may comprise various types of devices. For example, processing unit 110 may include various devices, such as a controller, an image preprocessor, a central processing unit (CPU), support circuits, digital signal processors, integrated circuits, memory, or any other types of devices for image processing and analysis. The image preprocessor may include a video processor for capturing, digitizing and processing the imagery from the image sensors. The CPU may comprise any number of microcontrollers or microprocessors. The support circuits may be any number of circuits generally well known in the art, including cache, power supply, clock and input-output circuits. The memory may store software that, when executed by the processor, controls the operation of the system. The memory may include databases and image processing software. The memory may comprise any number of random access memories, read only memories, flash memories, disk drives, optical storage, tape storage, removable storage and other types of storage. In one instance, the memory may be separate from the processing unit 110. In another instance, the memory may be integrated into the processing unit 110.

[0062] Each memory 140, 150 may include software instructions that when executed by a processor (e.g., applications processor 180 and/or image processor 190), may control operation of various aspects of system 100. These memory units may include various databases and image processing software. The memory units may include random access memory, read only memory, flash memory, disk drives, optical storage, tape storage, removable storage and/or any other types of storage. In some embodiments, memory units 140, 150 may be separate from the applications processor 180 and/or image processor 190. In other embodiments, these memory units may be integrated into applications processor 180 and/or image processor 190.

[0063] Position sensor 130 may include any type of device suitable for determining a location associated with at least one component of system 100. In some embodiments, position sensor 130 may include a GPS receiver. Such receivers can determine a user position and velocity by processing signals broadcasted by global positioning system satellites. Position information from position sensor 130 may be made available to

IPR2025-00944
Tesla EX1002 Page 1075

applications processor 180 and/or image processor 190.

[0064] User interface 170 may include any device suitable for providing information to or for receiving inputs from one or more users of system 100. In some embodiments, user interface 170 may include user input devices, including, for example, a touchscreen, microphone, keyboard, pointer devices, track wheels, cameras, knobs, buttons, etc. With such input devices, a user may be able to provide information inputs or commands to system 100 by typing instructions or information, providing voice commands, selecting menu options on a screen using buttons, pointers, or eye-tracking capabilities, or through any other suitable techniques for communicating information to system 100.

[0065] User interface 170 may be equipped with one or more processing devices configured to provide and receive information to or from a user and process that information for use by, for example, applications processor 180. In some embodiments, such processing devices may execute instructions for recognizing and tracking eye movements, receiving and interpreting voice commands, recognizing and interpreting touches and/or gestures made on a touchscreen, responding to keyboard entries or menu selections, etc. In some embodiments, user interface 170 may include a display, speaker, tactile device, and/or any other devices for providing output information to a user.

[0066] Map database 160 may include any type of database for storing map data useful to system 100. In some embodiments, map database 160 may include data relating to the position, in a reference coordinate system, of various items, including roads, water features, geographic features, businesses, points of interest, restaurants, gas stations, etc. Map database 160 may store not only the locations of such items, but also descriptors relating to those items, including, for example, names associated with any of the stored features. In some embodiments, map database 160 may be physically located with other components of system 100. Alternatively or additionally, map database 160 or a portion thereof may be located remotely with respect to other components of system 100 (e.g., processing unit 110). In such embodiments, information from map database 160 may be downloaded over a wired or wireless data connection to a network (e.g., over a cellular network and/or the Internet, etc.).

[0067] Image capture devices 122, 124, and 126 may each include any type of device suitable for capturing at least one image from an environment. Moreover, any number of image capture devices may be used to acquire images for input to the image processor. Some embodiments may include only a single image capture device, while other embodiments may include two, three, or even four or more image capture devices. Image capture devices 122,124, and 126 will be further described with reference to FIGS. 2B-2E , below.

[0068] System 100, or various components thereof, may be incorporated into various different platforms. In some embodiments, system 100 may be included on a vehicle 200, as shown in FIG. 2A . For example, vehicle 200 may be equipped with a processing unit 110 and any of the other components of system 100, as described above relative to FIG. 1 . While in some embodiments vehicle 200 may be equipped with only a single image capture device (e.g., camera), in other embodiments, such as those discussed in connection with FIGS. 2B-2E , multiple image capture devices may be used. For example, either of image capture devices 122 and 124 of vehicle 200, as shown in FIG. 2A , may be part of an ADAS (Advanced Driver Assistance Systems)

imaging set.

[0069] The image capture devices included on vehicle 200 as part of the image acquisition unit 120 may be positioned at any suitable location. In some embodiments, as shown in FIGS. 2A-2E and 3A-3C , image capture device 122 may be located in the vicinity of the rearview mirror. This position may provide a line of sight similar to that of the driver of vehicle 200, which may aid in determining what is and is not visible to the driver. Image capture device 122 may be positioned at any location near the rearview mirror, but placing image capture device 122 on the driver side of the mirror may further aid in obtaining images representative of the driver's field of view and/or line of sight.

[0070] Other locations for the image capture devices of image acquisition unit 120 may also be used. For example, image capture device 124 may be located on or in a bumper of vehicle 200. Such a location may be especially suitable for image capture devices having a wide field of view. The line of sight of bumper-located image capture devices can be different from that of the driver and, therefore, the bumper image capture device and driver may not always see the same objects. The image capture devices (e.g., image capture devices 122, 124, and 126) may also be located in other locations. For example, the image capture devices may be located on or in one or both of the side mirrors of vehicle 200, on the roof of vehicle 200, on the hood of vehicle 200, on the trunk of vehicle 200, on the sides of vehicle 200, mounted on, positioned behind, or positioned in front of any of the windows of vehicle 200, and mounted in or near light figures on the front and/or back of vehicle 200, etc.

[0071] In addition to image capture devices, vehicle 200 may include various other components of system 100. For example, processing unit 110 may be included on vehicle 200 either integrated with or separate from an engine control unit (ECU) of the vehicle. Vehicle 200 may also be equipped with a position sensor 130, such as a GPS receiver and may also include a map database 160 and memory units 140 and 150.

[0072] FIG. 2A is a diagrammatic side view representation of an exemplary vehicle imaging system consistent with the disclosed embodiments. FIG. 2B is a diagrammatic top view illustration of the embodiment shown in FIG. 2A . As illustrated in FIG. 2B , the disclosed embodiments may include a vehicle 200 including in its body a system 100 with a first image capture device 122 positioned in the vicinity of the rearview mirror and/or near the driver of vehicle 200, a second image capture device 124 positioned on or in a bumper region (e.g., one of bumper regions 210) of vehicle 200, and a processing unit 110.

[0073] As illustrated in FIG. 2C , image capture devices 122 and 124 may both be positioned in the vicinity of the rearview mirror and/or near the driver of vehicle 200. Additionally, while two image capture devices 122 and 124 are shown in FIGS. 2B and 2C , it should be understood that other embodiments may include more than two image capture devices. For example, in the embodiments shown in FIGS. 2D and 2E , first, second, and third image capture devices 122, 124, and 126, are included in the system 100 of vehicle 200.

[0074] As illustrated in FIG. 2D , image capture device 122 may be positioned in the vicinity of the rearview mirror and/or near the driver of vehicle 200, and image capture devices 124 and 126 may be positioned on or in a bumper region (e.g., one of bumper regions 210) of vehicle 200. And as shown in FIG. 2E , image capture devices 122, 124, and 126 may be positioned in the vicinity of the rearview mirror and/or near the

IPR2025-00944
Tesla EX1002 Page 1077

driver seat of vehicle 200. The disclosed embodiments are not limited to any particular number and configuration of the image capture devices, and the image capture devices may be positioned in any appropriate location within and/or on vehicle 200.

[0075] It is to be understood that the disclosed embodiments are not limited to vehicles and could be applied in other contexts. It is also to be understood that disclosed embodiments are not limited to a particular type of vehicle 200 and may be applicable to all types of vehicles including automobiles, trucks, trailers, and other types of vehicles.

[0076] The first image capture device 122 may include any suitable type of image capture device. Image capture device 122 may include an optical axis. In one instance, the image capture device 122 may include an Aptina M9V024 WVGA sensor with a global shutter. In other embodiments, image capture device 122 may provide a resolution of 1280x960 pixels and may include a rolling shutter. Image capture device 122 may include various optical elements. In some embodiments one or more lenses may be included, for example, to provide a desired focal length and field of view for the image capture device. In some embodiments, image capture device 122 may be associated with a 6mm lens or a 12mm lens. In some embodiments, image capture device 122 may be configured to capture images having a desired field-of-view (FOV) 202, as illustrated in FIG. 2D . For example, image capture device 122 may be configured to have a regular FOV, such as within a range of 40 degrees to 56 degrees, including a 46 degree FOV, 50 degree FOV, 52 degree FOV, or greater. Alternatively, image capture device 122 may be configured to have a narrow FOV in the range of 23 to 40 degrees, such as a 28 degree FOV or 36 degree FOV. In addition, image capture device 122 may be configured to have a wide FOV in the range of 100 to 180 degrees. In some embodiments, image capture device 122 may include a wide angle bumper camera or one with up to a 180 degree FOV.

[0077] The first image capture device 122 may acquire a plurality of first images relative to a scene associated with the vehicle 200. Each of the plurality of first images may be acquired as a series of image scan lines, which may be captured using a rolling shutter. Each scan line may include a plurality of pixels.

[0078] The first image capture device 122 may have a scan rate associated with acquisition of each of the first series of image scan lines. The scan rate may refer to a rate at which an image sensor can acquire image data associated with each pixel included in a particular scan line.

[0079] Image capture devices 122, 124, and 126 may contain any suitable type and number of image sensors, including CCD sensors or CMOS sensors, for example. In one embodiment, a CMOS image sensor may be employed along with a rolling shutter, such that each pixel in a row is read one at a time, and scanning of the rows proceeds on a row-by-row basis until an entire image frame has been captured. In some embodiments, the rows may be captured sequentially from top to bottom relative to the frame.

[0080] The use of a rolling shutter may result in pixels in different rows being exposed and captured at different times, which may cause skew and other image artifacts in the captured image frame. On the other hand, when the image capture device 122 is configured to operate with a global or synchronous shutter, all of the pixels may be exposed for the same amount of time and during a common exposure period. As a

Tesla EX1002 Page 1078 IPR2025-00944

result, the image data in a frame collected from a system employing a global shutter represents a snapshot of the entire FOV (such as FOV 202) at a particular time. In contrast, in a rolling shutter application, each row in a frame is exposed and data is capture at different times. Thus, moving objects may appear distorted in an image capture device having a rolling shutter. This phenomenon will be described in greater detail below.

[0081] The second image capture device 124 and the third image capturing device 126 may be any type of image capture device. Like the first image capture device 122, each of image capture devices 124 and 126 may include an optical axis. In one embodiment, each of image capture devices 124 and 126 may include an Aptina M9V024 WVGA sensor with a global shutter. Alternatively, each of image capture devices 124 and 126 may include a rolling shutter. Like image capture device 122, image capture devices 124 and 126 may be configured to include various lenses and optical elements. In some embodiments, lenses associated with image capture devices 124 and 126 may provide FOVs (such as FOVs 204 and 206) that are the same as, or narrower than, a FOV (such as FOV 202) associated with image capture device 122. For example, image capture devices 124 and 126 may have FOVs of 40 degrees, 30 degrees, 26 degrees, 23 degrees, 20 degrees, or less.

[0082] Image capture devices 124 and 126 may acquire a plurality of second and third images relative to a scene associated with the vehicle 200. Each of the plurality of second and third images may be acquired as a second and third series of image scan lines, which may be captured using a rolling shutter. Each scan line or row may have a plurality of pixels. Image capture devices 124 and 126 may have second and third scan rates associated with acquisition of each of image scan lines included in the second and third series.

[0083] Each image capture device 122, 124, and 126 may be positioned at any suitable position and orientation relative to vehicle 200. The relative positioning of the image capture devices 122, 124, and 126 may be selected to aid in fusing together the information acquired from the image capture devices. For example, in some embodiments, a FOV (such as FOV 204) associated with image capture device 124 may overlap partially or fully with a FOV (such as FOV 202) associated with image capture device 122 and a FOV (such as FOV 206) associated with image capture device 126.

[0084] Image capture devices 122, 124, and 126 may be located on vehicle 200 at any suitable relative heights. In one instance, there may be a height difference between the image capture devices 122, 124, and 126, which may provide sufficient parallax information to enable stereo analysis. For example, as shown in FIG. 2A , the two image capture devices 122 and 124 are at different heights. There may also be a lateral displacement difference between image capture devices 122, 124, and 126, giving additional parallax information for stereo analysis by processing unit 110, for example. The difference in the lateral displacement may be denoted by dx, as shown in FIGS. 2C and 2D . In some embodiments, fore or aft displacement (e.g., range displacement) may exist between image capture devices 122, 124, and 126. For example, image capture device 122 may be located 0.5 to 2 meters or more behind image capture device 124 and/or image capture device 126. This type of displacement may enable one of the image capture devices to cover potential blind spots of the other image capture device(s).

Tesla EX1002 Page 1079 IPR2025-00944

[0085] Image capture devices 122 may have any suitable resolution capability (e.g., number of pixels associated with the image sensor), and the resolution of the image sensor(s) associated with the image capture device 122 may be higher, lower, or the same as the resolution of the image sensor(s) associated with image capture devices 124 and 126. In some embodiments, the image sensor(s) associated with image capture device 122 and/or image capture devices 124 and 126 may have a resolution of 640 x 480, 1024 x 768, 1280 x 960, or any other suitable resolution.

[0086] The frame rate (e.g., the rate at which an image capture device acquires a set of pixel data of one image frame before moving on to capture pixel data associated with the next image frame) may be controllable. The frame rate associated with image capture device 122 may be higher, lower, or the same as the frame rate associated with image capture devices 124 and 126. The frame rate associated with image capture devices 122, 124, and 126 may depend on a variety of factors that may affect the timing of the frame rate. For example, one or more of image capture devices 122, 124, and 126 may include a selectable pixel delay period imposed before or after acquisition of image data associated with one or more pixels of an image sensor in image capture device 122, 124, and/or 126. Generally, image data corresponding to each pixel may be acquired according to a clock rate for the device (e.g., one pixel per clock cycle). Additionally, in embodiments including a rolling shutter, one or more of image capture devices 122, 124, and 126 may include a selectable horizontal blanking period imposed before or after acquisition of image data associated with a row of pixels of an image sensor in image capture device 122, 124, and/or 126. Further, one or more of image capture devices 122, 124, and/or 126 may include a selectable vertical blanking period imposed before or after acquisition of image data associated with an image frame of image capture device 122, 124, and 126.

[0087] These timing controls may enable synchronization of frame rates associated with image capture devices 122,124, and 126, even where the line scan rates of each are different. Additionally, as will be discussed in greater detail below, these selectable timing controls, among other factors (e.g., image sensor resolution, maximum line scan rates, etc.) may enable synchronization of image capture from an area where the FOV of image capture device 122 overlaps with one or more FOVs of image capture devices 124 and 126, even where the field of view of image capture device 122 is different from the FOVs of image capture devices 124 and 126.

[0088] Frame rate timing in image capture device 122,124, and 126 may depend on the resolution of the associated image sensors. For example, assuming similar line scan rates for both devices, if one device includes an image sensor having a resolution of 640 x 480 and another device includes an image sensor with a resolution of 1280 x 960, then more time will be required to acquire a frame of image data from the sensor having the higher resolution.

[0089] Another factor that may affect the timing of image data acquisition in image capture devices 122, 124, and 126 is the maximum line scan rate. For example, acquisition of a row of image data from an image sensor included in image capture device 122, 124, and 126 will require some minimum amount of time. Assuming no pixel delay periods are added, this minimum amount of time for acquisition of a row of image data will be related to the maximum line scan rate for a particular device. Devices that offer higher maximum line scan rates have the potential to provide higher frame rates than devices with lower maximum line scan rates. In some embodiments, one or more of image capture devices 124 and 126 may have a maximum line scan

IPR2025-00944
Tesla EX1002 Page 1080

rate that is higher than a maximum line scan rate associated with image capture device 122. In some embodiments, the maximum line scan rate of image capture device 124 and/or 126 may be 1.25, 1.5, 1.75, or 2 times or more than a maximum line scan rate of image capture device 122.

[0090] In another embodiment, image capture devices 122, 124, and 126 may have the same maximum line scan rate, but image capture device 122 may be operated at a scan rate less than or equal to its maximum scan rate. The system may be configured such that one or more of image capture devices 124 and 126 operate at a line scan rate that is equal to the line scan rate of image capture device 122. In other instances, the system may be configured such that the line scan rate of image capture device 124 and/or image capture device 126 may be 1.25, 1.5, 1.75, or 2 times or more than the line scan rate of image capture device 122.

[0091] In some embodiments, image capture devices 122, 124, and 126 may be asymmetric. That is, they may include cameras having different fields of view (FOV) and focal lengths. The fields of view of image capture devices 122, 124, and 126 may include any desired area relative to an environment of vehicle 200, for example. In some embodiments, one or more of image capture devices 122, 124, and 126 may be configured to acquire image data from an environment in front of vehicle 200, behind vehicle 200, to the sides of vehicle 200, or combinations thereof.

[0092] Further, the focal length associated with each image capture device 122,124, and/or 126 may be selectable (e.g., by inclusion of appropriate lenses etc.) such that each device acquires images of objects at a desired distance range relative to vehicle 200. For example, in some embodiments image capture devices 122, 124, and 126 may acquire images of close-up objects within a few meters from the vehicle. Image capture devices 122, 124, and 126 may also be configured to acquire images of objects at ranges more distant from the vehicle (e.g., 25 m, 50 m, 100 m, 150 m, or more). Further, the focal lengths of image capture devices 122, 124, and 126 may be selected such that one image capture device (e.g., image capture device 122) can acquire images of objects relatively close to the vehicle (e.g., within 10 m or within 20 m) while the other image capture devices (e.g., image capture devices 124 and 126) can acquire images of more distant objects (e.g., greater than 20 m, 50 m, 100 m, 150 m, etc.) from vehicle 200.

[0093] According to some embodiments, the FOV of one or more image capture devices 122, 124, and 126 may have a wide angle. For example, it may be advantageous to have a FOV of 140 degrees, especially for image capture devices 122, 124, and 126 that may be used to capture images of the area in the vicinity of vehicle 200. For example, image capture device 122 may be used to capture images of the area to the right or left of vehicle 200 and, in such embodiments, it may be desirable for image capture device 122 to have a wide FOV (e.g., at least 140 degrees).

[0094] The field of view associated with each of image capture devices 122, 124, and 126 may depend on the respective focal lengths. For example, as the focal length increases, the corresponding field of view decreases.

[0095] Image capture devices 122, 124, and 126 may be configured to have any suitable fields of view. In one particular example, image capture device 122 may have a horizontal FOV of 46 degrees, image capture device 124 may have a horizontal FOV of

Tesla EX1002 Page 1081 IPR2025-00944

23 degrees, and image capture device 126 may have a horizontal FOV in between 23 and 46 degrees. In another instance, image capture device 122 may have a horizontal FOV of 52 degrees, image capture device 124 may have a horizontal FOV of 26 degrees, and image capture device 126 may have a horizontal FOV in between 26 and 52 degrees. In some embodiments, a ratio of the FOV of image capture device 122 to the FOVs of image capture device 124 and/or image capture device 126 may vary from 1.5 to 2.0. In other embodiments, this ratio may vary between 1.25 and 2.25.

[0096] System 100 may be configured so that a field of view of image capture device 122 overlaps, at least partially or fully, with a field of view of image capture device 124 and/or image capture device 126. In some embodiments, system 100 may be configured such that the fields of view of image capture devices 124 and 126, for example, fall within (e.g., are narrower than) and share a common center with the field of view of image capture device 122. In other embodiments, the image capture devices 122, 124, and 126 may capture adjacent FOVs or may have partial overlap in their FOVs. In some embodiments, the fields of view of image capture devices 122, 124, and 126 may be aligned such that a center of the narrower FOV image capture devices 124 and/or 126 may be located in a lower half of the field of view of the wider FOV device 122.

[0097] FIG. 2F is a diagrammatic representation of exemplary vehicle control systems, consistent with the disclosed embodiments. As indicated in FIG. 2F , vehicle 200 may include throttling system 220, braking system 230, and steering system 240. System 100 may provide inputs (e.g., control signals) to one or more of throttling system 220, braking system 230, and steering system 240 over one or more data links (e.g., any wired and/or wireless link or links for transmitting data). For example, based on analysis of images acquired by image capture devices 122, 124, and/or 126, system 100 may provide control signals to one or more of throttling system 220, braking system 230, and steering system 240 to navigate vehicle 200 (e.g., by causing an acceleration, a turn, a lane shift, etc.). Further, system 100 may receive inputs from one or more of throttling system 220, braking system 230, and steering system 24 indicating operating conditions of vehicle 200 (e.g., speed, whether vehicle 200 is braking and/or turning, etc.). Further details are provided in connection with FIGS. 4-7 , below.

[0098] As shown in FIG. 3A , vehicle 200 may also include a user interface 170 for interacting with a driver or a passenger of vehicle 200. For example, user interface 170 in a vehicle application may include a touch screen 320, knobs 330, buttons 340, and a microphone 350. A driver or passenger of vehicle 200 may also use handles (e.g., located on or near the steering column of vehicle 200 including, for example, turn signal handles), buttons (e.g., located on the steering wheel of vehicle 200), and the like, to interact with system 100. In some embodiments, microphone 350 may be positioned adjacent to a rearview mirror 310. Similarly, in some embodiments, image capture device 122 may be located near rearview mirror 310. In some embodiments, user interface 170 may also include one or more speakers 360 (e.g., speakers of a vehicle audio system). For example, system 100 may provide various notifications (e.g., alerts) via speakers 360.

[0099] FIGS. 3B-3D are illustrations of an exemplary camera mount 370 configured to be positioned behind a rearview mirror (e.g., rearview mirror 310) and against a vehicle windshield, consistent with disclosed embodiments. As shown in FIG. 3B , camera mount 370 may include image capture devices 122, 124, and 126. Image capture devices 124 and 126 may be positioned behind a glare shield 380, which may be flush

Tesla EX1002 Page 1082 IPR2025-00944

against the vehicle windshield and include a composition of film and/or anti-reflective materials. For example, glare shield 380 may be positioned such that it aligns against a vehicle windshield having a matching slope. In some embodiments, each of image capture devices 122, 124, and 126 may be positioned behind glare shield 380, as depicted, for example, in FIG. 3D . The disclosed embodiments are not limited to any particular configuration of image capture devices 122, 124, and 126, camera mount 370, and glare shield 380. FIG. 3C is an illustration of camera mount 370 shown in FIG. 3B from a front perspective.

[0100] As will be appreciated by a person skilled in the art having the benefit of this disclosure, numerous variations and/or modifications may be made to the foregoing disclosed embodiments. For example, not all components are essential for the operation of system 100. Further, any component may be located in any appropriate part of system 100 and the components may be rearranged into a variety of configurations while providing the functionality of the disclosed embodiments. Therefore, the foregoing configurations are examples and, regardless of the configurations discussed above, system 100 can provide a wide range of functionality to analyze the surroundings of vehicle 200 and navigate vehicle 200 in response to the analysis.

[0101] As discussed below in further detail and consistent with various disclosed embodiments, system 100 may provide a variety of features related to autonomous driving and/or driver assist technology. For example, system 100 may analyze image data, position data (e.g., GPS location information), map data, speed data, and/or data from sensors included in vehicle 200. System 100 may collect the data for analysis from, for example, image acquisition unit 120, position sensor 130, and other sensors. Further, system 100 may analyze the collected data to determine whether or not vehicle 200 should take a certain action, and then automatically take the determined action without human intervention. For example, when vehicle 200 navigates without human intervention, system 100 may automatically control the braking, acceleration, and/or steering of vehicle 200 (e.g., by sending control signals to one or more of throttling system 220, braking system 230, and steering system 240). Further, system 100 may analyze the collected data and issue warnings and/or alerts to vehicle occupants based on the analysis of the collected data.

[0102] Further, consistent with disclosed embodiments, the functionality provided by system 100 may cause vehicle 200 to take different actions to navigate vehicle 200 within a lane and/or relative to other vehicles and/or objects. For example, system 100 may adjust the positioning of vehicle 200 relative to a lane within which vehicle 200 is traveling and/or relative to objects positioned near vehicle 200, select a particular lane for vehicle 200 to use while traveling, and take action in response to an encroaching vehicle, such as a vehicle attempting to move into the lane within which vehicle 200 is traveling. Additionally, system 100 may control the speed of vehicle 200 in different scenarios, such as when vehicle 200 is making a turn. System 100 may further cause vehicle 200 to mimic the actions of a leading vehicle or monitor a target vehicle and navigate vehicle 200 so that it passes the target vehicle. Additional details regarding the various embodiments that are provided by system 100 are provided below.

Forward-Facing Multi-Imaging System

[0103] As discussed above, system 100 may provide drive assist functionality that uses a multi-camera system. The multi-camera system may use one or more cameras facing

in the forward direction of a vehicle. In other embodiments, the multi-camera system may include one or more cameras facing to the side of a vehicle or to the rear of the vehicle. In one embodiment, for example, system 100 may use a two-camera imaging system, where a first camera and a second camera (e.g., image capture devices 122 and 124) may be positioned at the front and/or the sides of a vehicle (e.g., vehicle 200). The first camera may have a field of view that is greater than, less than, or partially overlapping with, the field of view of the second camera. In addition, the first camera may be connected to a first image processor to perform monocular image analysis of images provided by the first camera, and the second camera may be connected to a second image processor to perform monocular image analysis of images provided by the second camera. The outputs (e.g., processed information) of the first and second image processors may be combined. In some embodiments, the second image processor may receive images from both the first camera and second camera to perform stereo analysis. In another embodiment, system 100 may use a three-camera imaging system where each of the cameras has a different field of view. Such a system may, therefore, make decisions based on information derived from objects located at varying distances both forward and to the sides of the vehicle. References to monocular image analysis may refer to instances where image analysis is performed based on images captured from a single point of view (e.g., from a single camera). Stereo image analysis may refer to instances where image analysis is performed based on two or more images captured with one or more variations of an image capture parameter. For example, captured images suitable for performing stereo image analysis may include images captured: from two or more different positions, from different fields of view, using different focal lengths, along with parallax information, etc.

[0104] For example, in one embodiment, system 100 may implement a three camera configuration using image capture devices 122-126. In such a configuration, image capture device 122 may provide a narrow field of view (e.g., 34 degrees, or other values selected from a range of about 20 to 45 degrees, etc.), image capture device 124 may provide a wide field of view (e.g., 150 degrees or other values selected from a range of about 100 to about 180 degrees), and image capture device 126 may provide an intermediate field of view (e.g., 46 degrees or other values selected from a range of about 35 to about 60 degrees). In some embodiments, image capture device 126 may act as a main or primary camera. Image capture devices 122-126 may be positioned behind rearview mirror 310 and positioned substantially side-by-side (e.g., 6 cm apart). Further, in some embodiments, as discussed above, one or more of image capture devices 122-126 may be mounted behind glare shield 380 that is flush with the windshield of vehicle 200. Such shielding may act to minimize the impact of any reflections from inside the car on image capture devices 122-126.

[0105] In another embodiment, as discussed above in connection with FIGS. 3B and 3C , the wide field of view camera (e.g., image capture device 124 in the above example) may be mounted lower than the narrow and main field of view cameras (e.g., image devices 122 and 126 in the above example). This configuration may provide a free line of sight from the wide field of view camera. To reduce reflections, the cameras may be mounted close to the windshield of vehicle 200, and may include polarizers on the cameras to damp reflected light.

[0106] A three camera system may provide certain performance characteristics. For example, some embodiments may include an ability to validate the detection of objects by one camera based on detection results from another camera. In the three camera configuration discussed above, processing unit 110 may include, for example, three

processing devices (e.g., three EyeQ series of processor chips, as discussed above), with each processing device dedicated to processing images captured by one or more of image capture devices 122-126.

[0107] In a three camera system, a first processing device may receive images from both the main camera and the narrow field of view camera, and perform vision processing of the narrow FOV camera to, for example, detect other vehicles, pedestrians, lane marks, traffic signs, traffic lights, and other road objects. Further, the first processing device may calculate a disparity of pixels between the images from the main camera and the narrow camera and create a 3D reconstruction of the environment of vehicle 200. The first processing device may then combine the 3D reconstruction with 3D map data or with 3D information calculated based on information from another camera.

[0108] The second processing device may receive images from main camera and perform vision processing to detect other vehicles, pedestrians, lane marks, traffic signs, traffic lights, and other road objects. Additionally, the second processing device may calculate a camera displacement and, based on the displacement, calculate a disparity of pixels between successive images and create a 3D reconstruction of the scene (e.g., a structure from motion). The second processing device may send the structure from motion based 3D reconstruction to the first processing device to be combined with the stereo 3D images.

[0109] The third processing device may receive images from the wide FOV camera and process the images to detect vehicles, pedestrians, lane marks, traffic signs, traffic lights, and other road objects. The third processing device may further execute additional processing instructions to analyze images to identify objects moving in the image, such as vehicles changing lanes, pedestrians, etc.

[0110] In some embodiments, having streams of image-based information captured and processed independently may provide an opportunity for providing redundancy in the system. Such redundancy may include, for example, using a first image capture device and the images processed from that device to validate and/or supplement information obtained by capturing and processing image information from at least a second image capture device.

[0111] In some embodiments, system 100 may use two image capture devices (e.g., image capture devices 122 and 124) in providing navigation assistance for vehicle 200 and use a third image capture device (e.g., image capture device 126) to provide redundancy and validate the analysis of data received from the other two image capture devices. For example, in such a configuration, image capture devices 122 and 124 may provide images for stereo analysis by system 100 for navigating vehicle 200, while image capture device 126 may provide images for monocular analysis by system 100 to provide redundancy and validation of information obtained based on images captured from image capture device 122 and/or image capture device 124. That is, image capture device 126 (and a corresponding processing device) may be considered to provide a redundant sub-system for providing a check on the analysis derived from image capture devices 122 and 124 (e.g., to provide an automatic emergency braking (AEB) system).

[0112] One of skill in the art will recognize that the above camera configurations, camera placements, number of cameras, camera locations, etc., are examples only.

Tesla EX1002 Page 1085 IPR2025-00944

These components and others described relative to the overall system may be assembled and used in a variety of different configurations without departing from the scope of the disclosed embodiments. Further details regarding usage of a multi-camera system to provide driver assist and/or autonomous vehicle functionality follow below.

[0113] FIG. 4 is an exemplary functional block diagram of memory 140 and/or 150, which may be stored/programmed with instructions for performing one or more operations consistent with the disclosed embodiments. Although the following refers to memory 140, one of skill in the art will recognize that instructions may be stored in memory 140 and/or 150.

[0114] As shown in FIG. 4 , memory 140 may store a monocular image analysis module 402, a stereo image analysis module 404, a velocity and acceleration module 406, and a navigational response module 408. The disclosed embodiments are not limited to any particular configuration of memory 140. Further, application processor 180 and/or image processor 190 may execute the instructions stored in any of modules 402-408 included in memory 140. One of skill in the art will understand that references in the following discussions to processing unit 110 may refer to application processor 180 and image processor 190 individually or collectively. Accordingly, steps of any of the following processes may be performed by one or more processing devices.

[0115] In one embodiment, monocular image analysis module 402 may store instructions (such as computer vision software) which, when executed by processing unit 110, performs monocular image analysis of a set of images acquired by one of image capture devices 122, 124, and 126. In some embodiments, processing unit 110 may combine information from a set of images with additional sensory information (e.g., information from radar) to perform the monocular image analysis. As described in connection with FIGS. 5A-5D below, monocular image analysis module 402 may include instructions for detecting a set of features within the set of images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, hazardous objects, and any other feature associated with an environment of a vehicle. Based on the analysis, system 100 (e.g., via processing unit 110) may cause one or more navigational responses in vehicle 200, such as a turn, a lane shift, a change in acceleration, and the like, as discussed below in connection with navigational response module 408.

[0116] In one embodiment, stereo image analysis module 404 may store instructions (such as computer vision software) which, when executed by processing unit 110, performs stereo image analysis of first and second sets of images acquired by a combination of image capture devices selected from any of image capture devices 122, 124, and 126. In some embodiments, processing unit 110 may combine information from the first and second sets of images with additional sensory information (e.g., information from radar) to perform the stereo image analysis. For example, stereo image analysis module 404 may include instructions for performing stereo image analysis based on a first set of images acquired by image capture device 124 and a second set of images acquired by image capture device 126. As described in connection with FIG. 6 below, stereo image analysis module 404 may include instructions for detecting a set of features within the first and second sets of images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, hazardous objects, and the like. Based on the analysis, processing unit 110 may cause one or more navigational responses in vehicle 200, such as a turn, a lane shift, a change in acceleration, and the like, as discussed below in connection with

Tesla EX1002 Page 1086 IPR2025-00944

navigational response module 408.

[0117] In one embodiment, velocity and acceleration module 406 may store software configured to analyze data received from one or more computing and electromechanical devices in vehicle 200 that are configured to cause a change in velocity and/or acceleration of vehicle 200. For example, processing unit 110 may execute instructions associated with velocity and acceleration module 406 to calculate a target speed for vehicle 200 based on data derived from execution of monocular image analysis module 402 and/or stereo image analysis module 404. Such data may include, for example, a target position, velocity, and/or acceleration, the position and/or speed of vehicle 200 relative to a nearby vehicle, pedestrian, or road object, position information for vehicle 200 relative to lane markings of the road, and the like. In addition, processing unit 110 may calculate a target speed for vehicle 200 based on sensory input (e.g., information from radar) and input from other systems of vehicle 200, such as throttling system 220, braking system 230, and/or steering system 240 of vehicle 200. Based on the calculated target speed, processing unit 110 may transmit electronic signals to throttling system 220, braking system 230, and/or steering system 240 of vehicle 200 to trigger a change in velocity and/or acceleration by, for example, physically depressing the brake or easing up off the accelerator of vehicle 200.

[0118] In one embodiment, navigational response module 408 may store software executable by processing unit 110 to determine a desired navigational response based on data derived from execution of monocular image analysis module 402 and/or stereo image analysis module 404. Such data may include position and speed information associated with nearby vehicles, pedestrians, and road objects, target position information for vehicle 200, and the like. Additionally, in some embodiments, the navigational response may be based (partially or fully) on map data, a predetermined position of vehicle 200, and/or a relative velocity or a relative acceleration between vehicle 200 and one or more objects detected from execution of monocular image analysis module 402 and/or stereo image analysis module 404. Navigational response module 408 may also determine a desired navigational response based on sensory input (e.g., information from radar) and inputs from other systems of vehicle 200, such as throttling system 220, braking system 230, and steering system 240 of vehicle 200. Based on the desired navigational response, processing unit 110 may transmit electronic signals to throttling system 220, braking system 230, and steering system 240 of vehicle 200 to trigger a desired navigational response by, for example, turning the steering wheel of vehicle 200 to achieve a rotation of a predetermined angle. In some embodiments, processing unit 110 may use the output of navigational response module 408 (e.g., the desired navigational response) as an input to execution of velocity and acceleration module 406 for calculating a change in speed of vehicle 200.

[0119] FIG. 5A is a flowchart showing an exemplary process 500A for causing one or more navigational responses based on monocular image analysis, consistent with disclosed embodiments. At step 510, processing unit 110 may receive a plurality of images via data interface 128 between processing unit 110 and image acquisition unit 120. For instance, a camera included in image acquisition unit 120 (such as image capture device 122 having field of view 202) may capture a plurality of images of an area forward of vehicle 200 (or to the sides or rear of a vehicle, for example) and transmit them over a data connection (e.g., digital, wired, USB, wireless, Bluetooth, etc.) to processing unit 110. Processing unit 110 may execute monocular image analysis module 402 to analyze the plurality of images at step 520, as described in further detail in connection with FIGS. 5B-5D below. By performing the analysis,

Tesla EX1002 Page 1087 IPR2025-00944

processing unit 110 may detect a set of features within the set of images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, and the like.

[0120] Processing unit 110 may also execute monocular image analysis module 402 to detect various road hazards at step 520, such as, for example, parts of a truck tire, fallen road signs, loose cargo, small animals, and the like. Road hazards may vary in structure, shape, size, and color, which may make detection of such hazards more challenging. In some embodiments, processing unit 110 may execute monocular image analysis module 402 to perform multi-frame analysis on the plurality of images to detect road hazards. For example, processing unit 110 may estimate camera motion between consecutive image frames and calculate the disparities in pixels between the frames to construct a 3D-map of the road. Processing unit 110 may then use the 3D-map to detect the road surface, as well as hazards existing above the road surface.

[0121] At step 530, processing unit 110 may execute navigational response module 408 to cause one or more navigational responses in vehicle 200 based on the analysis performed at step 520 and the techniques as described above in connection with FIG. 4 . Navigational responses may include, for example, a turn, a lane shift, a change in acceleration, and the like. In some embodiments, processing unit 110 may use data derived from execution of velocity and acceleration module 406 to cause the one or more navigational responses. Additionally, multiple navigational responses may occur simultaneously, in sequence, or any combination thereof. For instance, processing unit 110 may cause vehicle 200 to shift one lane over and then accelerate by, for example, sequentially transmitting control signals to steering system 240 and throttling system 220 of vehicle 200. Alternatively, processing unit 110 may cause vehicle 200 to brake while at the same time shifting lanes by, for example, simultaneously transmitting control signals to braking system 230 and steering system 240 of vehicle 200.

[0122] FIG. 5B is a flowchart showing an exemplary process 500B for detecting one or more vehicles and/or pedestrians in a set of images, consistent with disclosed embodiments. Processing unit 110 may execute monocular image analysis module 402 to implement process 500B. At step 540, processing unit 110 may determine a set of candidate objects representing possible vehicles and/or pedestrians. For example, processing unit 110 may scan one or more images, compare the images to one or more predetermined patterns, and identify within each image possible locations that may contain objects of interest (e.g., vehicles, pedestrians, or portions thereof). The predetermined patterns may be designed in such a way to achieve a high rate of "false hits" and a low rate of "misses." For example, processing unit 110 may use a low threshold of similarity to predetermined patterns for identifying candidate objects as possible vehicles or pedestrians. Doing so may allow processing unit 110 to reduce the probability of missing (e.g., not identifying) a candidate object representing a vehicle or pedestrian.

[0123] At step 542, processing unit 110 may filter the set of candidate objects to exclude certain candidates (e.g., irrelevant or less relevant objects) based on classification criteria. Such criteria may be derived from various properties associated with object types stored in a database (e.g., a database stored in memory 140). Properties may include object shape, dimensions, texture, position (e.g., relative to vehicle 200), and the like. Thus, processing unit 110 may use one or more sets of criteria to reject false candidates from the set of candidate objects.

Tesla EX1002 Page 1088 IPR2025-00944

IPR2025-00944
Tesla EX1002 Page 1089

[0124] At step 544, processing unit 110 may analyze multiple frames of images to determine whether objects in the set of candidate objects represent vehicles and/or pedestrians. For example, processing unit 110 may track a detected candidate object across consecutive frames and accumulate frame-by-frame data associated with the detected object (e.g., size, position relative to vehicle 200, etc.). Additionally, processing unit 110 may estimate parameters for the detected object and compare the object's frame-by-frame position data to a predicted position.

[0125] At step 546, processing unit 110 may construct a set of measurements for the detected objects. Such measurements may include, for example, position, velocity, and acceleration values (relative to vehicle 200) associated with the detected objects. In some embodiments, processing unit 110 may construct the measurements based on estimation techniques using a series of time-based observations such as Kalman filters or linear quadratic estimation (LQE), and/or based on available modeling data for different object types (e.g., cars, trucks, pedestrians, bicycles, road signs, etc.). The Kalman filters may be based on a measurement of an object's scale, where the scale measurement is proportional to a time to collision (e.g., the amount of time for vehicle 200 to reach the object). Thus, by performing steps 540-546, processing unit 110 may identify vehicles and pedestrians appearing within the set of captured images and derive information (e.g., position, speed, size) associated with the vehicles and pedestrians. Based on the identification and the derived information, processing unit 110 may cause one or more navigational responses in vehicle 200, as described in connection with FIG. 5A , above.

[0126] At step 548, processing unit 110 may perform an optical flow analysis of one or more images to reduce the probabilities of detecting a "false hit" and missing a candidate object that represents a vehicle or pedestrian. The optical flow analysis may refer to, for example, analyzing motion patterns relative to vehicle 200 in the one or more images associated with other vehicles and pedestrians, and that are distinct from road surface motion. Processing unit 110 may calculate the motion of candidate objects by observing the different positions of the objects across multiple image frames, which are captured at different times. Processing unit 110 may use the position and time values as inputs into mathematical models for calculating the motion of the candidate objects. Thus, optical flow analysis may provide another method of detecting vehicles and pedestrians that are nearby vehicle 200. Processing unit 110 may perform optical flow analysis in combination with steps 540-546 to provide redundancy for detecting vehicles and pedestrians and increase the reliability of system 100.

[0127] FIG. 5C is a flowchart showing an exemplary process 500C for detecting road marks and/or lane geometry information in a set of images, consistent with disclosed embodiments. Processing unit 110 may execute monocular image analysis module 402 to implement process 500C. At step 550, processing unit 110 may detect a set of objects by scanning one or more images. To detect segments of lane markings, lane geometry information, and other pertinent road marks, processing unit 110 may filter the set of objects to exclude those determined to be irrelevant (e.g., minor potholes, small rocks, etc.). At step 552, processing unit 110 may group together the segments detected in step 550 belonging to the same road mark or lane mark. Based on the grouping, processing unit 110 may develop a model to represent the detected segments, such as a mathematical model.

[0128] At step 554, processing unit 110 may construct a set of measurements associated with the detected segments. In some embodiments, processing unit 110

may create a projection of the detected segments from the image plane onto the real-world plane. The projection may be characterized using a 3rd-degree polynomial having coefficients corresponding to physical properties such as the position, slope, curvature, and curvature derivative of the detected road. In generating the projection, processing unit 110 may take into account changes in the road surface, as well as pitch and roll rates associated with vehicle 200. In addition, processing unit 110 may model the road elevation by analyzing position and motion cues present on the road surface. Further, processing unit 110 may estimate the pitch and roll rates associated with vehicle 200 by tracking a set of feature points in the one or more images.

[0129] At step 556, processing unit 110 may perform multi-frame analysis by, for example, tracking the detected segments across consecutive image frames and accumulating frame-by-frame data associated with detected segments. As processing unit 110 performs multi-frame analysis, the set of measurements constructed at step 554 may become more reliable and associated with an increasingly higher confidence level. Thus, by performing steps 550-556, processing unit 110 may identify road marks appearing within the set of captured images and derive lane geometry information. Based on the identification and the derived information, processing unit 110 may cause one or more navigational responses in vehicle 200, as described in connection with FIG. 5A , above.

[0130] At step 558, processing unit 110 may consider additional sources of information to further develop a safety model for vehicle 200 in the context of its surroundings. Processing unit 110 may use the safety model to define a context in which system 100 may execute autonomous control of vehicle 200 in a safe manner. To develop the safety model, in some embodiments, processing unit 110 may consider the position and motion of other vehicles, the detected road edges and barriers, and/or general road shape descriptions extracted from map data (such as data from map database 160). By considering additional sources of information, processing unit 110 may provide redundancy for detecting road marks and lane geometry and increase the reliability of system 100.

[0131] FIG. 5D is a flowchart showing an exemplary process 500D for detecting traffic lights in a set of images, consistent with disclosed embodiments. Processing unit 110 may execute monocular image analysis module 402 to implement process 500D. At step 560, processing unit 110 may scan the set of images and identify objects appearing at locations in the images likely to contain traffic lights. For example, processing unit 110 may filter the identified objects to construct a set of candidate objects, excluding those objects unlikely to correspond to traffic lights. The filtering may be done based on various properties associated with traffic lights, such as shape, dimensions, texture, position (e.g., relative to vehicle 200), and the like. Such properties may be based on multiple examples of traffic lights and traffic control signals and stored in a database. In some embodiments, processing unit 110 may perform multi-frame analysis on the set of candidate objects reflecting possible traffic lights. For example, processing unit 110 may track the candidate objects across consecutive image frames, estimate the real-world position of the candidate objects, and filter out those objects that are moving (which are unlikely to be traffic lights). In some embodiments, processing unit 110 may perform color analysis on the candidate objects and identify the relative position of the detected colors appearing inside possible traffic lights.

[0132] At step 562, processing unit 110 may analyze the geometry of a junction. The

IPR2025-00944
Tesla EX1002 Page 1090

IPR2025-00944
Tesla EX1002 Page 1091

analysis may be based on any combination of: (i) the number of lanes detected on either side of vehicle 200, (ii) markings (such as arrow marks) detected on the road, and (iii) descriptions of the junction extracted from map data (such as data from map database 160). Processing unit 110 may conduct the analysis using information derived from execution of monocular analysis module 402. In addition, Processing unit 110 may determine a correspondence between the traffic lights detected at step 560 and the lanes appearing near vehicle 200.

[0133] As vehicle 200 approaches the junction, at step 564, processing unit 110 may update the confidence level associated with the analyzed junction geometry and the detected traffic lights. For instance, the number of traffic lights estimated to appear at the junction as compared with the number actually appearing at the junction may impact the confidence level. Thus, based on the confidence level, processing unit 110 may delegate control to the driver of vehicle 200 in order to improve safety conditions. By performing steps 560-564, processing unit 110 may identify traffic lights appearing within the set of captured images and analyze junction geometry information. Based on the identification and the analysis, processing unit 110 may cause one or more navigational responses in vehicle 200, as described in connection with FIG. 5A , above.

[0134] FIG. 5E is a flowchart showing an exemplary process 500E for causing one or more navigational responses in vehicle 200 based on a vehicle path, consistent with the disclosed embodiments. At step 570, processing unit 110 may construct an initial vehicle path associated with vehicle 200. The vehicle path may be represented using a set of points expressed in coordinates (x, z),and the distancedibetween two points in the set of points may fall in the range of 1 to 5 meters. In one embodiment, processing unit 110 may construct the initial vehicle path using two polynomials, such as left and right road polynomials. Processing unit 110 may calculate the geometric midpoint between the two polynomials and offset each point included in the resultant vehicle path by a predetermined offset (e.g., a smart lane offset), if any (an offset of zero may correspond to travel in the middle of a lane). The offset may be in a direction perpendicular to a segment between any two points in the vehicle path. In another embodiment, processing unit 110 may use one polynomial and an estimated lane width to offset each point of the vehicle path by half the estimated lane width plus a predetermined offset (e.g., a smart lane offset).

[0135] At step 572, processing unit 110 may update the vehicle path constructed at step 570. Processing unit 110 may reconstruct the vehicle path constructed at step 570 using a higher resolution, such that the distancedkbetween two points in the set of points representing the vehicle path is less than the distancedidescribed above. For example, the distancedkmay fall in the range of 0.1 to 0.3 meters. Processing unit 110 may reconstruct the vehicle path using a parabolic spline algorithm, which may yield a cumulative distance vectorScorresponding to the total length of the vehicle path (i.e., based on the set of points representing the vehicle path).

[0136] At step 574, processing unit 110 may determine a look-ahead point (expressed in coordinates as (xi , zi)) based on the updated vehicle path constructed at step 572. Processing unit 110 may extract the look-ahead point from the cumulative distance vectorS, and the look-ahead point may be associated with a look-ahead distance and look-ahead time. The look-ahead distance, which may have a lower bound ranging from 10 to 20 meters, may be calculated as the product of the speed of vehicle 200 and the look-ahead time. For example, as the speed of vehicle 200 decreases, the look-ahead distance may also decrease (e.g., until it reaches the lower bound). The look-

ahead time, which may range from 0.5 to 1.5 seconds, may be inversely proportional to the gain of one or more control loops associated with causing a navigational response in vehicle 200, such as the heading error tracking control loop. For example, the gain of the heading error tracking control loop may depend on the bandwidth of a yaw rate loop, a steering actuator loop, car lateral dynamics, and the like. Thus, the higher the gain of the heading error tracking control loop, the lower the look-ahead time.

[0137] At step 576, processing unit 110 may determine a heading error and yaw rate command based on the look-ahead point determined at step 574. Processing unit 110 may determine the heading error by calculating the arctangent of the look-ahead point, e.g., arctan (xi/zi). Processing unit 110 may determine the yaw rate command as the product of the heading error and a high-level control gain. The high-level control gain may be equal to: (2 / look-ahead time), if the look-ahead distance is not at the lower bound. Otherwise, the high-level control gain may be equal to: (2 * speed of vehicle 200 / look-ahead distance).

[0138] FIG. 5F is a flowchart showing an exemplary process 500F for determining whether a leading vehicle is changing lanes, consistent with the disclosed embodiments. At step 580, processing unit 110 may determine navigation information associated with a leading vehicle (e.g., a vehicle traveling ahead of vehicle 200). For example, processing unit 110 may determine the position, velocity (e.g., direction and speed), and/or acceleration of the leading vehicle, using the techniques described in connection with FIGS. 5A and 5B , above. Processing unit 110 may also determine one or more road polynomials, a look-ahead point (associated with vehicle 200), and/or a snail trail (e.g., a set of points describing a path taken by the leading vehicle), using the techniques described in connection with FIG. 5E , above.

[0139] At step 582, processing unit 110 may analyze the navigation information determined at step 580. In one embodiment, processing unit 110 may calculate the distance between a snail trail and a road polynomial (e.g., along the trail). If the variance of this distance along the trail exceeds a predetermined threshold (for example, 0.1 to 0.2 meters on a straight road, 0.3 to 0.4 meters on a moderately curvy road, and 0.5 to 0.6 meters on a road with sharp curves), processing unit 110 may determine that the leading vehicle is likely changing lanes. In the case where multiple vehicles are detected traveling ahead of vehicle 200, processing unit 110 may compare the snail trails associated with each vehicle. Based on the comparison, processing unit 110 may determine that a vehicle whose snail trail does not match with the snail trails of the other vehicles is likely changing lanes. Processing unit 110 may additionally compare the curvature of the snail trail (associated with the leading vehicle) with the expected curvature of the road segment in which the leading vehicle is traveling. The expected curvature may be extracted from map data (e.g., data from map database 160), from road polynomials, from other vehicles' snail trails, from prior knowledge about the road, and the like. If the difference in curvature of the snail trail and the expected curvature of the road segment exceeds a predetermined threshold, processing unit 110 may determine that the leading vehicle is likely changing lanes.

[0140] In another embodiment, processing unit 110 may compare the leading vehicle's instantaneous position with the look-ahead point (associated with vehicle 200) over a specific period of time (e.g., 0.5 to 1.5 seconds). If the distance between the leading vehicle's instantaneous position and the look-ahead point varies during the specific period of time, and the cumulative sum of variation exceeds a predetermined threshold (for example, 0.3 to 0.4 meters on a straight road, 0.7 to 0.8 meters on a moderately

IPR2025-00944
Tesla EX1002 Page 1092

curvy road, and 1.3 to 1.7 meters on a road with sharp curves), processing unit 110 may determine that the leading vehicle is likely changing lanes. In another embodiment, processing unit 110 may analyze the geometry of the snail trail by comparing the lateral distance traveled along the trail with the expected curvature of the snail trail. The expected radius of curvature may be determined according to the calculation: (δx<2>+ δx<2>) / 2 / (δx), where δxrepresents the lateral distance traveled and δzrepresents the longitudinal distance traveled. If the difference between the lateral distance traveled and the expected curvature exceeds a predetermined threshold (e.g., 500 to 700 meters), processing unit 110 may determine that the leading vehicle is likely changing lanes. In another embodiment, processing unit 110 may analyze the position of the leading vehicle. If the position of the leading vehicle obscures a road polynomial (e.g., the leading vehicle is overlaid on top of the road polynomial), then processing unit 110 may determine that the leading vehicle is likely changing lanes. In the case where the position of the leading vehicle is such that, another vehicle is detected ahead of the leading vehicle and the snail trails of the two vehicles are not parallel, processing unit 110 may determine that the (closer) leading vehicle is likely changing lanes.

[0141] At step 584, processing unit 110 may determine whether or not leading vehicle 200 is changing lanes based on the analysis performed at step 582. For example, processing unit 110 may make the determination based on a weighted average of the individual analyses performed at step 582. Under such a scheme, for example, a decision by processing unit 110 that the leading vehicle is likely changing lanes based on a particular type of analysis may be assigned a value of "1" (and "0" to represent a determination that the leading vehicle is not likely changing lanes). Different analyses performed at step 582 may be assigned different weights, and the disclosed embodiments are not limited to any particular combination of analyses and weights.

[0142] FIG. 6 is a flowchart showing an exemplary process 600 for causing one or more navigational responses based on stereo image analysis, consistent with disclosed embodiments. At step 610, processing unit 110 may receive a first and second plurality of images via data interface 128. For example, cameras included in image acquisition unit 120 (such as image capture devices 122 and 124 having fields of view 202 and 204) may capture a first and second plurality of images of an area forward of vehicle 200 and transmit them over a digital connection (e.g., USB, wireless, Bluetooth, etc.) to processing unit 110. In some embodiments, processing unit 110 may receive the first and second plurality of images via two or more data interfaces. The disclosed embodiments are not limited to any particular data interface configurations or protocols.

[0143] At step 620, processing unit 110 may execute stereo image analysis module 404 to perform stereo image analysis of the first and second plurality of images to create a 3D map of the road in front of the vehicle and detect features within the images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, road hazards, and the like. Stereo image analysis may be performed in a manner similar to the steps described in connection with FIGS. 5A-5D , above. For example, processing unit 110 may execute stereo image analysis module 404 to detect candidate objects (e.g., vehicles, pedestrians, road marks, traffic lights, road hazards, etc.) within the first and second plurality of images, filter out a subset of the candidate objects based on various criteria, and perform multi-frame analysis, construct measurements, and determine a confidence level for the remaining candidate objects. In performing the steps above, processing unit 110 may consider information from both the first and second plurality of images, rather than information from one set of images

Tesla EX1002 Page 1093 IPR2025-00944

alone. For example, processing unit 110 may analyze the differences in pixel-level data (or other data subsets from among the two streams of captured images) for a candidate object appearing in both the first and second plurality of images. As another example, processing unit 110 may estimate a position and/or velocity of a candidate object (e.g., relative to vehicle 200) by observing that the object appears in one of the plurality of images but not the other or relative to other differences that may exist relative to objects appearing if the two image streams. For example, position, velocity, and/or acceleration relative to vehicle 200 may be determined based on trajectories, positions, movement characteristics, etc. of features associated with an object appearing in one or both of the image streams.

[0144] At step 630, processing unit 110 may execute navigational response module 408 to cause one or more navigational responses in vehicle 200 based on the analysis performed at step 620 and the techniques as described above in connection with FIG. 4 . Navigational responses may include, for example, a turn, a lane shift, a change in acceleration, a change in velocity, braking, and the like. In some embodiments, processing unit 110 may use data derived from execution of velocity and acceleration module 406 to cause the one or more navigational responses. Additionally, multiple navigational responses may occur simultaneously, in sequence, or any combination thereof.

[0145] FIG. 7 is a flowchart showing an exemplary process 700 for causing one or more navigational responses based on an analysis of three sets of images, consistent with disclosed embodiments. At step 710, processing unit 110 may receive a first, second, and third plurality of images via data interface 128. For instance, cameras included in image acquisition unit 120 (such as image capture devices 122, 124, and 126 having fields of view 202, 204, and 206) may capture a first, second, and third plurality of images of an area forward and/or to the side of vehicle 200 and transmit them over a digital connection (e.g., USB, wireless, Bluetooth, etc.) to processing unit 110. In some embodiments, processing unit 110 may receive the first, second, and third plurality of images via three or more data interfaces. For example, each of image capture devices 122, 124, 126 may have an associated data interface for communicating data to processing unit 110. The disclosed embodiments are not limited to any particular data interface configurations or protocols.

[0146] At step 720, processing unit 110 may analyze the first, second, and third plurality of images to detect features within the images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, road hazards, and the like. The analysis may be performed in a manner similar to the steps described in connection with FIGS. 5A-5D and 6 , above. For instance, processing unit 110 may perform monocular image analysis (e.g., via execution of monocular image analysis module 402 and based on the steps described in connection with FIGS. 5A-5D , above) on each of the first, second, and third plurality of images. Alternatively, processing unit 110 may perform stereo image analysis (e.g., via execution of stereo image analysis module 404 and based on the steps described in connection with FIG. 6 , above) on the first and second plurality of images, the second and third plurality of images, and/or the first and third plurality of images. The processed information corresponding to the analysis of the first, second, and/or third plurality of images may be combined. In some embodiments, processing unit 110 may perform a combination of monocular and stereo image analyses. For example, processing unit 110 may perform monocular image analysis (e.g., via execution of monocular image analysis module 402) on the first plurality of images and stereo image analysis (e.g., via execution of stereo image

Tesla EX1002 Page 1094 IPR2025-00944

analysis module 404) on the second and third plurality of images. The configuration of image capture devices 122, 124, and 126-including their respective locations and fields of view 202, 204, and 206-may influence the types of analyses conducted on the first, second, and third plurality of images. The disclosed embodiments are not limited to a particular configuration of image capture devices 122, 124, and 126, or the types of analyses conducted on the first, second, and third plurality of images.

[0147] In some embodiments, processing unit 110 may perform testing on system 100 based on the images acquired and analyzed at steps 710 and 720. Such testing may provide an indicator of the overall performance of system 100 for certain configurations of image capture devices 122, 124, and 126. For example, processing unit 110 may determine the proportion of "false hits" (e.g., cases where system 100 incorrectly determined the presence of a vehicle or pedestrian) and "misses."

[0148] At step 730, processing unit 110 may cause one or more navigational responses in vehicle 200 based on information derived from two of the first, second, and third plurality of images. Selection of two of the first, second, and third plurality of images may depend on various factors, such as, for example, the number, types, and sizes of objects detected in each of the plurality of images. Processing unit 110 may also make the selection based on image quality and resolution, the effective field of view reflected in the images, the number of captured frames, the extent to which one or more objects of interest actually appear in the frames (e.g., the percentage of frames in which an object appears, the proportion of the object that appears in each such frame, etc.), and the like.

[0149] In some embodiments, processing unit 110 may select information derived from two of the first, second, and third plurality of images by determining the extent to which information derived from one image source is consistent with information derived from other image sources. For example, processing unit 110 may combine the processed information derived from each of image capture devices 122, 124, and 126 (whether by monocular analysis, stereo analysis, or any combination of the two) and determine visual indicators (e.g., lane markings, a detected vehicle and its location and/or path, a detected traffic light, etc.) that are consistent across the images captured from each of image capture devices 122, 124, and 126. Processing unit 110 may also exclude information that is inconsistent across the captured images (e.g., a vehicle changing lanes, a lane model indicating a vehicle that is too close to vehicle 200, etc.). Thus, processing unit 110 may select information derived from two of the first, second, and third plurality of images based on the determinations of consistent and inconsistent information.

[0150] Navigational responses may include, for example, a turn, a lane shift, a change in acceleration, and the like. Processing unit 110 may cause the one or more navigational responses based on the analysis performed at step 720 and the techniques as described above in connection with FIG. 4 . Processing unit 110 may also use data derived from execution of velocity and acceleration module 406 to cause the one or more navigational responses. In some embodiments, processing unit 110 may cause the one or more navigational responses based on a relative position, relative velocity, and/or relative acceleration between vehicle 200 and an object detected within any of the first, second, and third plurality of images. Multiple navigational responses may occur simultaneously, in sequence, or any combination thereof.

IPR2025-00944
Tesla EX1002 Page 1095

Smart Lane Offset

[0151] System 100 may provide driver assist functionality that monitors and adjusts the free space between vehicle 200 and boundaries (e.g., lane boundaries) within which vehicle 200 is traveling. A lane may refer to a designated or intended travel path of a vehicle and may have marked (e.g., lines on a road) or unmarked boundaries (e.g., an edge of a road, a road barrier, guard rail, parked vehicles, etc.). For instance, as a default, system 100 may maximize or increase the free space between vehicle 200 and the current lane boundaries. Further, under certain conditions, system 100 may offset vehicle 200 toward one side of the current lane of travel without leaving the lane. For example, where vehicle 200 is traveling around a curve, system 100 may decrease the offset towards the inside of the curve (e.g., adjust the position of vehicle 200 so that it is closer to the inside of the curve). As another example, where obstacles-such as parked cars, pedestrians, or cyclists-are present on one side of the lane, system 100 may increase the offset on the side of vehicle 200 where the obstacles are located (e.g., adjust the position of vehicle 200 so that it is further from the obstacles).

[0152] FIG. 8A illustrates vehicle 200 traveling on a roadway 800 in which the disclosed systems and methods for identifying lane constraints and operating vehicle 200 within the lane constraints may be used. As shown, vehicle 200 may have a first vehicle side 802, which may be a first distance 805 from a first lane constraint 810. Similarly, vehicle 200 may have a second vehicle side 812 opposite from first vehicle side 802, and second vehicle side 812 may be a second distance 815 from second lane constraint 820. In this manner, first lane constraint 810 and second lane constraint 820 may define a lane within which vehicle 200 may travel.

[0153] Processing unit 110 may be configured to determine first lane constraint 810 and second lane constraint 820 based on a plurality of images acquired by image capture device 122-126 that processing unit 110 may receive via data interface 128. According to some embodiments, first lane constraint 810 and/or second lane constraint 820 may be identified by visible lane boundaries, such as dashed or solid lines marked on a road surface. Additionally or alternatively, first lane constraint 810 and/or second lane constraint 820 may include an edge of a road surface or a barrier. Additionally or alternatively, first lane constraint 810 and/or second lane constrain 820 may include markers (e.g., Botts' dots). According to some embodiments, processing unit 110 may determine first lane constraint 810 and/or second lane constraint 820 by identifying a midpoint of a road surface width. According to some embodiments, if processing unit 110 identifies only first lane constraint 810, processing unit 110 may estimate second lane constraint 820, such as based on an estimated lane width or road width. Processing unit 110 may identify lane constraints in this manner when, for example, lines designating road lanes are not painted or otherwise labeled.

[0154] Detection of first lane constraint 810 and/or second lane constraint 820 may include processing unit 110 determining their 3D models in camera coordinate system. For example, the 3D models of first lane constraint 810 and/or second lane constraint 820 may be described by a third-degree polynomial. In addition to 3D modeling of first lane constraint 810 and/or second lane constraint 820, processing unit 110 may perform multi-frame estimation of host motion parameters, such as the speed, yaw and pitch rates, and acceleration of vehicle 200. Optionally, processing unit 110 may detect static and moving vehicles and their position, heading, speed, and acceleration, all relative to vehicle 200. Processing unit 110 may further determine a road elevation model to transform all of the information acquired from the plurality of images into 3D

IPR2025-00944
Tesla EX1002 Page 1096

space.

[0155] Generally, as a default condition, vehicle 200 may travel relatively centered within first and second lane constraints 810 and/or 820. However, in some circumstances, environmental factors may make this undesirable or unfeasible, such as when objects or structures are present on one side of the road. Thus, there may be circumstances in which it may be desirable or advantageous for vehicle 200 to be closer to one lane constraint or the other. Processing unit 110 may also be configured to determine whether such circumstances, called lane offset conditions, exist based on the plurality of images.

[0156] To determine whether lane offset conditions exist, processing unit 110 may be configured to determine the presence of one or more objects 825 in the vicinity of vehicle 200. For example object 825 may comprise another vehicle, such as a car, truck or motorcycle traveling on roadway 800. Processing unit 110 may determine, for each object 825, an offset profile. The offset profile may include a determination of whether in its current and predicted position object 825 will be within a predefined range of vehicle 200. If processing unit 110 determines that object 825 is or will be within a predefined range of vehicle 200, processing unit 110 may determine whether there is enough space within first lane constraint 810 and second lane constraint 820, for vehicle 200 to bypass object 825. If there is not enough space, processing unit 110 may execute a lane change. If there is enough space, processing unit 110 may determine whether there is enough time to bypass object 825 before the distance between vehicle 200 and object 825 closes. If there is enough time, processing unit 110 may initiate and or activate the offset maneuver to bypass object 825. The offset maneuver may be executed by processing unit 110 so that the movement of vehicle 200 is smooth and comfortable for the driver. For example, the offset slew rate may be approximately 0.15 to 0.75 m/sec. The offset maneuver may be executed so that the maximum amplitude of the offset maneuver will occur when the gap between vehicle 200 and object 825 closes. If there is not enough time, processing unit 110 may initiate braking (e.g., by transmitting electronic signals to braking system 230) and then initiate offset maneuver.

[0157] After each offset profile is set for each object 825, processing unit may combine offset profiles such that the maximum lateral distance is kept from each object 825 that vehicle 200 is bypassing. That is, offset maneuvers may be modified and/or executed such that the lateral distance between vehicle 200 and each object 825 that vehicle 200 is bypassing is maximized.

[0158] Additionally or alternatively, object 825 may be a parked or stationary vehicle, a wall, or a person, such as pedestrian or cycling traffic. Processing unit 110 may determine a lane offset condition exists based on certain characteristics of object 825. For example, processing unit 110 may determine object 825 constitutes a lane offset condition if the height of object 825 exceeds a predetermined threshold, such as 10 cm from a road surface. In this manner, processing unit 110 may be configured to filter out objects that would not present a lane offset condition, such as small pieces of debris.

[0159] Additionally or alternatively, processing unit 110 may be configured to determine whether a lane offset condition exists on first vehicle side 802 based on a position of a target object 830 on vehicle first side 802. Target object 830 may include one or more objects that constitute a lane offset condition. Additionally or alternatively, target object 830 may not constitute a lane offset condition. For example, target object 830 may be a

Tesla EX1002 Page 1097 IPR2025-00944

street sign or other object a sufficient distance from the travel lane or occurring with sufficient infrequency to not constitute a lane offset condition. Processing unit 110 may be configured to conclude a lane offset condition does not exist if such a target object 830 is detected.

[0160] Processing unit 110 may be configured to cause vehicle 200 to travel such that first distance 805 and second distance 810 are substantially different to avoid or stay farther away from object 825. For example, if a lane offset condition exists on first vehicle side 802, processing unit 110 may be configured to cause vehicle 200 to travel within first lane constraint 810 and second lane constraint 820 such that first distance 805 is greater than second distance 815. This is illustrated, for example, in FIG. 8A . Alternatively, if a lane offset condition exists on second vehicle side 812, processing unit 110 may create or execute an offset profile that is configured to cause vehicle 200 to travel within first lane constraint 810 and second lane constraint 820 such that first distance 805 is less than second distance 815. This is illustrated, for example, in FIG 8B .

[0161] As another example, a parked vehicle may be located on one side of vehicle 200 and a pedestrian may be located on the other side of vehicle 200. In such a situation, system 100 may determine that is desirable to reduce the distance between vehicle 200 and the parked vehicle, in order to provide more space between vehicle 200 and the pedestrian. Accordingly, processing unit 110 may cause vehicle 200 to travel such that the distance between the parked vehicle is reduced.

[0162] The identity of first vehicle side 802 and second vehicle side 812 need not be static, such that at different points in time, processing unit 110 may identify first vehicle side 802 as the driver's side or as the passenger's side of vehicle 200. For example, the identity of first vehicle side 802 may change depending on the presence of one or more objects 825 on that side of vehicle 200, such that if one lane offset condition exists based on object 825, it exists on first vehicle side 802. In such circumstances, processing unit 110 may create or execute an offset profile that is configured to cause vehicle 200 to travel within first lane constraint 810 and second lane constraint 820 such that first distance 805, which is greater than second distance 815, is provided toward object 825.

[0163] According to some embodiments, when a lane offset condition exists on first vehicle side 802, processing unit 110 may create an offset profile configured to cause vehicle 200 to travel within first lane constraint 810 and second lane constraint 820 such that first distance 805 is at least 1.25 times greater than second distance 815. Additionally or alternatively, when a lane offset condition exists on first vehicle side 802, processing unit 110 may create an offset profile that is configured to cause vehicle 200 to travel within first lane constraint 810 and second lane constraint 820 such that first distance 805 is at least 2 times greater than second distance 815.

[0164] In some embodiments, as shown in FIG. 8C , processing unit 110 may determine a lane offset condition exists on both first vehicle side 802 and on second vehicle side 812. Processing unit 110 may then create an offset profile for this offset condition that is configured to cause vehicle 200 to travel within first lane constraint 810 and second lane constraint 820 such that first distance 805 is substantially the same as second distance 815.

[0165] Processing unit 110 may be configured so that the offset profiles for moving

Tesla EX1002 Page 1098 IPR2025-00944

objects 825 as well as stationary objects 825 are prioritized. For example, processing unit 110 may prioritize offset profiles to keep a predefined distance between vehicle 200 and moving objects 825 and/or static objects 825. Processing unit 200 may prioritize these offset conditions over those offset conditions based on certain lane characteristics, such as a curve. The last priority lane offset condition may include mimicking a lead vehicle.

[0166] Processing unit 110 may factor offset profiles or offset conditions into a desired look-ahead point, which may be located at the middle of the lane defined by first lane constraint 810 and second lane constraint 820. Processing unit 110 may compute and implement a yaw rate command based on the look-ahead point and the offset profiles. The yaw rate command may be used to compute a torque command based on the yaw rate command and the yaw rate of vehicle 200 measured by using a gyroscope. Processing unit 110 may transmit the torque command to a steering system (e.g., steering system 240) of vehicle 200.

[0167] According to some embodiments, input from a driver of vehicle 200, such as via user interface 170, may override processing unit 110. For example, a user input may override the control of processing unit 110 to cause vehicle 200 to travel within first lane constraint 810 and second lane constraint 820 such that first distance 805 is greater than second distance 815.

[0168] Additionally or alternatively, processing unit 110 may be configured to determine, based on the plurality of images captured by image capture device 122, whether a lane offset condition is substantially non-existent. This may include circumstances in which one or more objects 825 below a certain size threshold are detected, one or more objects 825 detected are outside of a predetermined distance from the lane constraint, and/or when no objects 825 are detected. Under such circumstances, processing unit 110 may be configured to cause vehicle 200 to travel within first lane constraint 810 and second lane constraint 820 such that first distance 805 is equal to or substantially equal to second distance 815.

[0169] A lane offset condition may include a curved road, as illustrated in FIG. 8D . Processing unit 110 may be configured to cause vehicle 200 to travel within first lane constraint 810 and second lane constraint 820 such that first distance 805 is toward the outside of the curved road relative to vehicle 200 and second distance 815 is toward an inside of the curved road relative to vehicle 200. For example, first distance 805 may be less than second distance 815.

[0170] FIG. 9 is an exemplary block diagram of memory 140 and/or 150, which may stare instructions for performing one or more operations consistent with disclosed embodiments. As illustrated in FIG. 9 , memory 140 may store one or more modules for performing the lane offset condition detection and responses described herein. For example, memory 140 may store a lane constraint module 910 and a lane offset module 920. Application processor 180 and/or image processor 190 may execute the instructions stored in any of modules 910 and 920 included in memory 140. One of skill in the art will understand that references in the following discussions to processing unit 110 may refer to application processor 180 and image processor 190 individually or collectively. Accordingly, steps of any of the following processes may be performed by one or more processing devices.

[0171] Lane constraint module 910 may store instructions which, when executed by

Tesla EX1002 Page 1099 IPR2025-00944

processing unit 110, may detect and define first lane constraint 810 and second lane constraint 820. For example, lane offset module 910 may process the plurality of images received from at least one image capture device 122-124 to detect first lane constraint 810 and second lane constraint 820. As discussed above, this may include identifying painted lane lines and/or measuring a midpoint of a road surface.

[0172] Lane offset module 920 may store instructions which, when executed by processing unit 110, may detect the presence of a lane offset condition and/or identify whether lane offset condition is present on first vehicle side 802 and/or second vehicle side 812. For example, lane offset module 920 may process the plurality of images to detect the presence of object(s) 825 on first vehicle side 802 and/or second vehicle side 812. Lane offset module 920 may process the plurality of images to detect a curve in the road on which vehicle 200 is traveling. Additionally or alternatively, lane offset module 920 may receive information from another module or other system indicative of the presence of object(s) 825 and/or a curved lane. Lane offset module 920 may execute control to change the position of vehicle 200 to first lane constraint 810 and/or second lane constraint 820. For example, if lane offset module 920 determines a lane offset condition only on first vehicle side 802, lane offset module 920 may move vehicle 200 closer to second lane constraint 820. On the other hand, if lane offset module 920 determines a lane offset condition only on first vehicle side 812, lane offset module 920 may move vehicle 200 closer to first lane constraint 810.

[0173] Fig. 10 illustrates a process 1000 for navigating vehicle 200, consistent with disclosed embodiments. Process 1000 may identify lane constraints that define a lane of travel for vehicle travel. Process 1000 include causing vehicle 200 to travel closer to one or the other lane constraint, such as first lane constraint 810 or second lane constraint 820, in response to detecting a lane offset condition. In this manner, process 1000 may be used to drive vehicle 200 within lane constraints as well as account for environmental factors that may make it more desirable to shift vehicle 200 within first lane constraint 810 and second lane constraint 820.

[0174] At step 1010, process 1000 may include acquiring, using at least one image capture device 122, 124, and/or 126, a plurality of images of an area in the vicinity of vehicle 200. For example, processing unit 110 may receive the plurality of images may through data interface 128. For example, processing unit 110 may be configured to determine the presence of one or more objects 825 on the side of the road. For example, object 825 may be a parked or stationary vehicle, a wall, or a person, such as pedestrian or cycling traffic. Processing unit 110 may determine if a lane offset condition exists based on certain characteristics of object 825. For example, processing unit 110 may determine that object 825 constitutes a lane offset condition if the height of object 825 exceeds a predetermined threshold, such as 10 cm from a road surface. In this manner, processing unit 110 may be configured to filter out objects that would not present a lane offset condition, such as small pieces of debris.

[0175] Additionally or alternatively, processing unit 110 may be configured to determine whether a lane offset condition exists on first vehicle side 802 based on a position of a target object 830 on vehicle first side 802. Target object 830 may be one or more objects that constitute a lane offset condition. Additionally or alternatively, target object 830 may not constitute a lane offset condition. For example, target object 830 may be a road barrier or a street sign. Processing unit 110 may be configured to conclude a lane offset condition does not exist if such a target object 830 is detected.

IPR2025-00944
Tesla EX1002 Page 1100

[0176] At step 1020, process 1000 may determine from the plurality of images a first lane constraint on first vehicle side 802 and a second lane constraint on second vehicle side 812. For example, processing unit 110 may be configured to determine first lane constraint 810 and second lane constraint 820 based on the plurality of images received via data interface 128.

[0177] At step 1030, process 1000 may determine whether a lane offset condition exists on first vehicle side 802. If a lane offset condition does exist on first vehicle side 802, at step 1040, process 1000 may include causing vehicle 200 to travel within first lane constraint 810 and second lane constraint 820 such that second distance 815 is less than first distance 805. Processing unit 110 may be configured to cause vehicle 200 to travel such that first distance 805 and second distance 810 are substantially different to avoid or stay farther away from object 825. For example, if a lane offset condition exists on first vehicle side 802, processing unit 110 may be configured to cause vehicle 200 to travel within first lane constraint 810 and second lane constraint 820 such that first distance 805 is greater than second distance 815. This is illustrated, for example, in FIG. 8A . Alternatively, if a lane offset condition exists on second vehicle side 812, processing unit 110 may be configured to cause vehicle 200 to travel within first lane constraint 810 and second lane constraint 820 such that first distance 805 is less than second distance 815. This is illustrated, for example, in FIG 8B . For example, to navigate vehicle 200 according to process 1000, processing unit 110 may transmit electronic signals to one or more of throttling system 220, braking system 230, and/or steering system 240 of vehicle 200.

[0178] If a lane offset condition does not exist on first vehicle side 802, in step 1050, method may include determining whether lane offset condition exists on second vehicle side 812. If a lane offset condition exists on second vehicle side 812, the method may include causing vehicle 200 to travel between first lane constraint 810 and second lane constraint 820 such that first distance 805 is less than second distance 815.

Navigating a Vehicle to a Default Lane

[0179] System 100 may provide driver assist functionality that monitors the location of vehicle 200 in a current lane of travel and moves vehicle 200 into a predetermined default travel lane if the current lane of travel is not the predetermined default travel lane. The predetermined default travel lane may be defined by a user input and/or as a result of processing images received from at least one image capture device 122-126.

[0180] System 100 may also provide notification of an impending lane change of vehicle 200 to the predetermined default travel lane to the driver and passengers of vehicle 200 and/or other drivers. For example, system 100 may activate a turn signal positioned at or near one of bumper regions 210. Additionally or alternatively, system 100 may sound an audible notification, such as through speakers 360. Vehicle 200 may use system 100 to select a particular lane as predetermined default travel lane when vehicle is traveling. When the current lane within which vehicle 200 is traveling-e.g., the current lane of travel-differs from the predetermined default travel lane, system 100 may cause vehicle 200 to make a lane change, e.g., into the predetermined default travel lane.

[0181] FIG. 11 illustrates vehicle 200 traveling in a current lane of travel 1110 of roadway 800. According to some embodiments, vehicle 200 may have a predetermined default travel lane 1120. For example, predetermined default travel lane 1120 may be

IPR2025-00944
Tesla EX1002 Page 1101

the right-most lane of roadway 800. Navigation system 100 may be configured to determine whether vehicle 200 is traveling in predetermined default travel lane 1120 and, if not, cause vehicle 200 to change to predetermined default travel lane 1120. For example, processing unit 110 may be configured to compare current lane of travel 1110 with predetermined default travel lane 1120 and cause vehicle 200 to return to predetermined default travel lane 1120.

[0182] According to some embodiments, processing unit 110 may detect an object 1130 in predetermined default travel lane 1120 based on a plurality of images acquired by one or more of image capture devices 122-126. Processing unit 110 may detect the position and speed of vehicle 200 relative to object 1130 by analyzing the images using techniques described in connection with FIGS. 5A-5D , above. According to some embodiments, processing unit 110 may determine whether vehicle 200 should pass object 1130. This may be desirable if, for example, the speed and/or acceleration of vehicle 200 relative to object 1130 exceeds a certain threshold.

[0183] FIG. 12A is an exemplary block diagram of memory 140, which may store instructions for performing one or more operations consistent with disclosed embodiments. As illustrated in FIG. 12A , memory 140 may store one or more modules for performing the default travel lane and current travel lane detection and responses described herein. For example, memory 140 may store a current lane detection module 1210, a default lane detection module 1220, and an action module 1230.

[0184] Current lane detection module 1210 may store instructions which, when executed by processing unit 110, may detect and define current lane of travel 1110. For example, processing unit 110 may execute current lane detection module 1212 to process the plurality of images received from at least one image capture device 122-124 and detect current lane of travel 1110. This may include detecting the relative position of vehicle 200 with respect to the edge of roadway 800, a midpoint of roadway 800, and/or lane markers, such as painted lane lines. Processing unit 110 may detect the position of vehicle 200 relative to roadway 800 by analyzing the images using techniques described in connection with FIGS. 5A-5D , above.

[0185] Default lane detection module 1220 may store instructions which, when executed by processing unit 110, may identify predetermined default travel lane 1120 and compare default lane detection module 1220 to current lane of travel 1110 identified by current lane detection module 1210. Default lane detection module 1220 may identify predetermined default travel lane 1120 based upon a user input and/or the plurality of images received from at least one image capture device 122-124. For example, a user may select a default lane through voice commands received by speakers 360 or by making a selection on a menu displayed on touch screen 320.

[0186] In some embodiments, predetermined default travel lane 1120 may be the right-most travel lane, or it may be the lane that has the more desirable features. For example, default lane detection module 1220 may include instructions to select the lane with less traffic, the lane in which vehicles are traveling closest to speed of vehicle 200, and/or a lane closest to an approaching exit or street to which vehicle 200 will turn. In some embodiments, predetermined default travel lane 1120 may be determined based on data received from position sensor 130, such as a global positioning system.

[0187] Action module 1230 may store instructions which, when executed by processing unit 110, may cause a response to the identification of current lane of travel 1110 and

IPR2025-00944
Tesla EX1002 Page 1102

predetermined default travel lane 1120. For example, processing unit 110 may execute action module 1230 to cause vehicle 200 to change lanes if current lane of travel 1110 is not the same as predetermined default travel lane 1120. Additionally or alternatively, action module 1230 may initiate a notification of the lane change, such as activating turn signal 210 and or causing an audible announcement.

[0188] Action module 1230 may include instructions that cause processing unit 110 to determine whether vehicle 200 should move out of predetermined default travel lane 1120. For example, action module 1230 may determine that object 1130 in predetermined default travel lane 1120 should be bypassed by vehicle 200.

[0189] Action module 1230 may further include instructions that cause processing unit 110 to determine whether it is safe to make a lane change. For example, prior to executing a lane change, processing unit 110 may evaluate determine one or more characteristics of roadway 800 based on, for example, analysis of the plurality of images conducted using techniques described in connection with FIGS. 5A-5D , above. If the analysis indicates conditions are not safe to change lanes, processing unit 110 may determine that vehicle 200 should stay in the current lane.

[0190] Action module 1230 may include additional instructions for manipulating vehicle 200 to maneuver out of and return to predetermined default travel lane 1120, such as for bypassing object 1130. When executing these instructions, processing unit 110 may estimate a location of predetermined default travel lane 1120 using estimated motion of vehicle 200. This may be advantageous in circumstances in which predetermined default travel lane 1120 does not remain detectable by image capture devices 122-124.

[0191] Processing unit 110 may be configured to execute an offset profile that offsets vehicle 200 from predetermined default travel lane 1120 by a full lane width, which may be estimated based on the distance between detected lane markers. The offset maneuver defined by offset profile may be executed by processing unit 110 so that the movement of vehicle 200 is smooth and comfortable for the driver. For example, the offset slew rate may be approximately 0.15 to 0.75 m/sec. During the offset maneuver, processing unit 110 may execute instructions from action module 1230 that switch to following the center of current lane of travel 1110. These instructions may include gradually transitioning from following the predicted lane with offset to following current lane of travel 1110. A gradual transition may be accomplished by defining desired look-ahead point so that it is not allowed to change in time more than the recommended slew rate, which may be around 0.25 m/sec.

[0192] The instructions for transitioning vehicle 200 from predetermined default travel lane 1120 may be used to transition vehicle 200 back into predetermined default travel lane. For example, processing unit 110 may be configured to execute an offset profile that offsets vehicle 200 from current lane of travel 1110 by a full lane width, which may be estimated based on the distance between detected lane markers. The offset maneuver defined by offset profile may be executed by processing unit 110 so that the movement of vehicle 200 is smooth and comfortable for the driver. For example, the offset slew rate may be approximately 0.15 to 0.75 m/sec. During the offset maneuver, processing unit 110 may execute instructions from action module 1230 that switch to following the center of predetermined default travel lane 1120. These instructions may include gradually transitioning from following the predicted lane with offset to following predetermined default travel lane 1120. A gradual transition may be accomplished by defining desired look-ahead point so that it is not allowed to change in time more than

IPR2025-00944
Tesla EX1002 Page 1103

the recommended slew rate, which may be around 0.25 m/sec.

[0193] At one or more of the steps for transitioning vehicle 200 between two lanes, processing unit 110 may compute and implement a yaw rate command based on the look-ahead point and the offset profiles. The yaw rate command may be used to compute a torque command based on the yaw rate command and the yaw rate of vehicle 200 measured using a gyroscope. Processing unit 110 may transmit a torque command to a steering system (e.g., steering system 240) of vehicle 200.

[0194] FIG. 12B illustrates a process 1250 for navigating vehicle 200 to a default lane consistent with disclosed embodiments. According to some embodiments, process 1250 may be implemented by one or more components of navigation system 100, such as at least one processing unit 110. Process 1250 may identify current lane of travel 1110 of vehicle 200 and compare it with predetermined default travel lane 1120 to determine whether vehicle 200 is traveling within predetermined default travel lane 1120. If vehicle 200 is not traveling in predetermined default travel lane 1120, processing unit 110 may cause vehicle 200 to change lanes so that it navigates to default travel lane 1120.

[0195] At step 1260, one or more of image capture devices 122-124 may acquire a plurality of images of an area in a vicinity of vehicle 200. Processing unit 110 may receive the plurality of images via data interface 128.

[0196] At step 1270, processing unit 110 may execute current lane detection module 1210 detect and define current lane of travel 1110. For example, processing unit 110 may process the plurality of images to, for example, detect the relative position of vehicle 200 with respect to the edge of roadway 100, a midpoint of roadway 100, and/or painted lane markers. Processing unit 110 may detect the position of vehicle 200 relative to roadway 800 by analyzing the images using techniques described in connection with FIGS. 5A-5D , above.

[0197] At step 1280, processing unit 110 may execute default lane detection module 1220 to identify predetermined default travel lane 1120 and compare default lane detection module 1220 to current lane of travel 1110 identified by current lane detection module 1210. Default lane detection module 1220 may identify predetermined default travel lane 1120 based upon a user input and/or the plurality of images received from at least one image capture device 122-124.

[0198] For example, as discussed above, processing unit 110 may determine predetermined default travel lane 1120 based on any number of factors. Predetermined default travel lane 1120 may be the right-most lane among the plurality of travel lanes 1110 and 1120. Additionally or alternatively, predetermined default travel lane 1120 may be determined based on an input received from a user via user interface 170, by processing the plurality of images, and/or based on data received from position sensor 130, such as a global positioning system. Predetermined default travel lane 1120 may be determined once, once per vehicle trip, or on regular intervals. For example, predetermined default travel lane 1120 may be determined dynamically in response to conditions at a given location.

[0199] If current lane 1110 is not the same as predetermined default travel lane 1120, processing unit 110 may execute action module 1230 to cause vehicle 100 to navigate to default travel lane 1120. For example, processing unit 110 may transmit electronic

IPR2025-00944
Tesla EX1002 Page 1104

signals to one or more of throttling system 220, braking system 230, and/or steering system 240 of vehicle 200 to trigger the desired response. For example, processing unit 110 may cause steering system 240 to turn the steering wheel of vehicle 200 to achieve a rotation of a predetermined angle.

[0200] Further, prior to executing a lane change, processing unit 110 may determine whether it is safe (e.g., there are no vehicles or objects in the way) to change lanes. As discussed above, processing unit 110 may evaluate one or more characteristics of roadway 800 based on, for example, analysis of the plurality of images conducted using techniques described in connection with FIGS. 5A-5D , above. If conditions are not safe to change lanes, processing unit 110 may determine that vehicle 200 should stay in the current lane.

[0201] In some embodiments, execution of action module 1230 may initiate a notification of the lane change. For example, processing unit 110 may cause a signal to activate a turn signal and/or cause an audible announcement via speakers 360. Accordingly, such notification may occur prior to vehicle 200 changing lanes. In other embodiments, execution of action module 1230 may initiate a notification to a driver of vehicle 200 that vehicle 200 is not traveling in a default lane without causing the lane change to occur. Accordingly, in such an embodiment, the driver of vehicle 200 may determine whether to manually take action and steer vehicle 200 to change lanes.

[0202] In some embodiments, system 100 may notify the user of vehicle 200 and/or other drivers in the vicinity of vehicle 200 that vehicle 200 will be changing lanes if vehicle 200 will be moved to predetermined default travel lane 1120. For example, processing unit 110 may activate a turn signal of vehicle 200 prior to causing vehicle 200 to change lanes. Additionally or alternatively, an audible announcement to notify other drivers in the vicinity may be caused by processing unit 110. The audible announcement may be transmitted via, for example, Bluetooth, to receivers included in other nearby vehicles. Additionally or alternatively, at least one processing unit 110 may cause an audible indicator to notify the driver of vehicle 200 prior to causing vehicle 200 to switch lanes. In this manner, driver of vehicle 200 can anticipate the control system 100 will exert on vehicle 200.

Controlling Velocity of a Turning Vehicle

[0203] System 100 may provide driver assist functionality that controls the velocity (i.e., speed and/or direction) of vehicle 200 in different scenarios, such as while making a turn (e.g., through a curve). For example, system 100 may be configured to use a combination of map, position, and/or visual data to control the velocity of vehicle 200 when turning. In particular, system 100 may control the velocity of vehicle 200 as it approaches a curve, while navigating a curve, and/or as it exits a curve. System 100 may consider different factors to control the velocity depending on the position of vehicle 200 relative to the curve at a given time. For example, system 100 may initially cause vehicle 200 to slow down in response to learning that a curve is located a certain distance ahead of vehicle 200. While vehicle 200 turns through the curve, however, system 100 may cause vehicle 200 to accelerate and/or charge direction based on an analysis of the characteristics of the curve.

[0204] FIGS. 13A and 13B are diagrammatic representations of a vehicle (e.g., vehicle 200) approaching, and navigating, a curve with one or more characteristics on road 1300, consistent with the disclosed embodiments. As illustrated in FIG. 13A , visual

Tesla EX1002 Page 1105 IPR2025-00944

indicators, such as road sign 1302 located near road 1300, may inform vehicle 200 that a curve lies ahead. In response to the warning provided by road sign 1302, a driver of vehicle 200 and/or a system 100 providing driver assist functionality of vehicle 200 may adjust the velocity of vehicle 200 (e.g., by slowing down and/or steering) to safely navigate the curve.

[0205] As vehicle 200 approaches the curve, system 100 may detect the lane width 1306 and the curve radius 1308 by using image capture devices 122 and/or 124 (which may be cameras) of vehicle 200. Based on the curve radius 1308, the lane width 1306, and other characteristics associated with the curve detected using image capture devices 122 and/or 124, system 100 may adjust the velocity of vehicle 200. Specifically, system 100 may adjust the velocity as vehicle 200 approaches the curve and/or while vehicle 200 navigates the curve. System 100 may also adjust the velocity as vehicle 200 exits the curve, in response to, for example, traffic light 1304 appearing after the curve. Further detail regarding adjustments to the velocity of vehicle 200 and detection of the lane width 1306, the radius of curvature or curve radius 1308, and characteristics associated with the curve are described in connection with FIGS. 14 and 15 , below.

[0206] In some embodiments, system 100 may recognize a curve to be navigated based on map data and/or vehicle position information (e.g., GPS data). System 100 may determine an initial target velocity for the vehicle based on one or more characteristics of the curve as reflected in the map data. System 100 may adjust a velocity of the vehicle to the initial target velocity and determine, based on one or more images acquired by one or more of image capture devices 122-126, one or more observed characteristics of the curve. System 100 may determine an updated target velocity based on the one or more observed characteristics of the curve and adjust the velocity of the vehicle to the updated target velocity.

[0207] FIG. 14 is an exemplary block diagram of a memory configured to store instructions for performing one or more operations consistent with the disclosed embodiments. As indicated in FIG. 14 , memory 140 may store a curve recognition module 1402, a characteristic observation module 1404, and a velocity module 1406. The disclosed embodiments are not limited to any particular configuration of memory 140. Further, processing unit 110 may execute the instructions stored in any of modules 1402-1406 included in memory 140.

[0208] In one embodiment, curve recognition module 1402 may store software instructions (such as computer vision software) which, when executed by processing unit 110, detects a curve by, for example, analyzing one or more sets of images and/or using map and/or position data (e.g., data stored in map database 160). The image sets may be acquired by one or more of image capture devices 122, 124, and 126. Based on the analysis, system 100 (e.g., via processing unit 110) may cause a change in velocity of vehicle 200. For example, processing unit 110 may cause vehicle 200 to reduce its speed by a predetermined amount and/or adjust its direction by a particular angle if processing unit 110 detects the presence of a curve located within a minimum threshold distance ahead of vehicle 200.

[0209] In one embodiment, characteristic observation module 1404 may store software instructions (such as computer vision software) which, when executed by processing unit 110, observes characteristics associated with a curve detected by system 100 (e.g., via execution of curve recognition module 1402). Processing unit 110 may

Tesla EX1002 Page 1106 IPR2025-00944

observe characteristics associated with a detected curve by analyzing one or more sets of images acquired by one or more of image capture devices 122, 124, and 126 and/or using map and/or position data (e.g., data stored in map database 160). The analysis may yield information associated with the curve, such as curve radius 1308, lane width 1306, a degree of curvature, a rate of change in curvature, a degree of banking, a length or arc length of the curve, and the like. These characteristics may include estimated values (e.g., curve characteristics calculated using preexisting data and/or mathematical models based on such data) and actual values (e.g., curve characteristics calculated by analyzing captured images of the curve). Based on the analysis, processing unit 110 may cause a change in velocity of vehicle 200. For example, processing unit 110 may cause vehicle 200 to reduce its speed and/or adjust its direction by a particular angle in view of a high rate of change in curvature associated with a curve located 25 meters ahead of vehicle 200.

[0210] In one embodiment, velocity module 1406 may store software instructions configured to analyze data from one or more computing and electromechanical devices configured to determine a target velocity and cause a change in velocity of vehicle 200 to the target velocity. For example, processing unit 110 may execute velocity module 1406 to calculate a target velocity for vehicle 200 based on data derived from execution of curve recognition module 1402 and characteristic observation module 1404. Such data may include, for example, an initial target position and initial velocity, an updated target position and updated velocity, a final target position and final velocity, and the like. In addition, processing unit 110 may calculate a velocity for vehicle 200 based on input from other systems of vehicle 200, such as a throttling system 220, braking system 230, and/or steering system 240 of vehicle 200. Based on the calculated target velocity, processing unit 110 may transmit electronic signals (e.g., via a controller area network bus ("CAN bus")) to throttling system 220, braking system 230, and/or steering system 240 of vehicle 200 to trigger a change in the speed and/or direction of vehicle 200 by, for example, physically depressing the brake, easing up off the accelerator of vehicle 200, or steering vehicle 200 in a particular direction (e.g., at a particular angle). An increase in speed may be associated with a corresponding acceleration (e.g., between 0.2 m/sec<2>and 3.0 m/sec<2>). Conversely, a decrease in speed may be associated with a corresponding deceleration. In addition, the acceleration or deceleration may be based on the road type (e.g., highway, city street, country road, etc.), the presence of any speed constraints nearby (e.g., sharp curves, traffic lights), as well as a road width or lane width 1306. For purposes of this disclosure, deceleration may refer to an acceleration having a negative value.

[0211] FIG. 15 is a flow chart showing an exemplary process 1500 for controlling the velocity of a vehicle based on a detected curve and observed characteristics of the curve and/or map data regarding the curve consistent with disclosed embodiments. At step 1510, processing unit 110 may receive a plurality of images via data interface 128 between processing unit 110 and image acquisition unit 120. For instance, a camera included in image acquisition unit 120 (such as image capture device 122) may capture a plurality of images and transmit them over a digital connection (e.g., USB, wireless, Bluetooth, etc.) to processing unit 110. In some embodiments, processing unit 110 may receive more than one plurality of images via a plurality of data interfaces. For example, processing unit 110 may receive a plurality of images from each of image capture devices 122, 124, 126, each of which may have an associated data interface for communicating data to processing unit 110. The disclosed embodiments are not limited to any particular data interface configurations or protocols.

Tesla EX1002 Page 1107 IPR2025-00944

[0212] At step 1520, processing unit 110 may execute curve recognition module 1502 to detect a curve located ahead of vehicle 200 based on, for example, map and/or position data (e.g., data stored in map database 160). Such data may indicate the location of a curve that is present on a given road. The data may indicate the absolute position of a particular curve (e.g., via GPS coordinates) or the relative position of the curve (e.g., by describing the curve relative to vehicle 200, road sign 1302, traffic light 1304, and/or other descriptive landmarks). In some embodiments, processing unit 110 may execute curve recognition module 1502 to detect the curve by analyzing the plurality of images. The analysis may be conducted using techniques described in connection with FIGS. 5A-5D and 6 , above.

[0213] At step 1530, processing unit 110 may execute velocity module 1406 to determine an initial target velocity for vehicle 200 based on, for example, data derived from execution of curve recognition module 1402. The initial target velocity may reflect a speed and a direction for vehicle 200 to safely enter an upcoming curve and may be determined using techniques described in connection with FIG. 14 , above. For example, velocity module 1406 may determine the initial target velocity based on known or estimated characteristics associated with the upcoming curve, such as the curve radius 1308 and the degree of banking. Velocity module 1406 may calculate the initial target velocity in accordance with mathematical models and/or formulae and based on various constraints. In some embodiments, for example, velocity module 1406 may determine the initial target velocity based on the presence of traffic lights (such as traffic light 1304) before, during, or after a curve, weather conditions (e.g., light rain, heavy rain, windy, clear, etc.), road conditions (e.g., paved road, unpaved road, potholes, etc.), posted speed limits, the presence of vehicles or pedestrians nearby, etc. The initial target velocity may also be based on a lateral acceleration limit associated with vehicle 200. In some embodiments, the lateral acceleration limit may be in the range of 0.2 m/sec<2>and 3.0 m/sec<2>and may be adjusted depending on the curve characteristics, the road width, or lane width 1306. Additionally, the initial target velocity may be based on a focal range associated with one or more of cameras, such as image capture devices 122 and 124.

[0214] For example, velocity module 1406 may determine the initial target velocity based on the radius of curvature of the road ahead of vehicle 200. Velocity module 1406 may determine the radius of curvature at various distances from vehicle 200 and store the information in a vectorR(1:n). In one embodiment, velocity module 1406 may determine the radius of curvature at: (1) the current position of vehicle 200, (2) a distanceDendfrom vehicle 200, whereDendrepresents the distance ahead of vehicle 200 at which knowledge of the road characteristics allows system 100 to control navigational responses in vehicle 200 such that the driver may feel comfortable, and (3) at incremental distances from vehicle 200, such as every 1 meter, etc. The value of distanceDendmay depend on the current speed of vehicle 200,Vego .For example, the greater the value ofVego(e.g., the faster vehicle 200 is traveling), the greater the value ofDend(e.g., the longer the distance may be taken into account to provide system 100 with sufficient time to provide a navigational response, such as comfortably reducing the speed of vehicle 200 based on an approaching sharp turn or curve). In one embodiment,Dendmay be calculated according to the following equation:Dend= (Vego<2>/2)/(0.7/40*Vego+0.5). Thus, the vectorD(1:n) = [1:1:n], wheren= floor(Dend).

[0215] In some embodiments, velocity module 1406 may determine the initial target velocity based on a lateral acceleration constraint,amax, e.g., the maximum allowed lateral acceleration,amaxmay fall in the range of 1.5 to 3.5 meters / sec<2>. Based

IPR2025-00944
Tesla EX1002 Page 1108

onamaxand the radius of curvature vectorR(discussed above), velocity module 1406 may limit the speed of vehicle 200 (e.g., at different points or distances from the current position of vehicle 200) to a speedvlim, calculated as:vlim(1:n)=sqrt(amax*R(1:n)). In addition, a deceleration constraint that follows from the determination ofvlimmay be calculated as:declim= min{((vlim(1:n)<2>-Vego<2>)/2), max{dmin,D(1:n)-2*Vego}}, wheredminis equal to 40 meters. The value ofdminmay represent a point at which deceleration feels comfortable for the driver of vehicle 200.

[0216] In some embodiments, velocity module 1406 may determine the initial target velocity based on a curvature slew rate constraint,slewmax, which may be based on the curvature rate of change.slewmaxmay fall in the range of 0.5 to 1.5 meters / sec<3>. Velocity module 1406 may determinevlimaccording the following equation:vlim (1:n) = (slewmax/abs{diff{1/R(1:n)}})<1/3>. Further, velocity module 1406 may determinedeclimaccording to the equation:declim= min{((vlim(1:n)<2>-Vego<2>)/2), max {dmin , D(1:n) - 2 *Vego}}, wheredminis equal to 40 meters.

[0217] In some embodiments, velocity module 1406 may determine a comfort constraint such that the driver of vehicle 200 is comfortable while vehicle 200 is undergoing various navigational responses, such as making a turn or navigating a curve. Velocity module 1406 may determinevlimaccording to the following equation:vlim (1:n) = 4/500*R(1:n) + 22.5. Further, velocity module 1406 may determinedeclimaccording to the equation:declim= min{((vlim(1:n)<2>-Vego<2>)/2), max {dmin , D(1:n) - 2*Vego}}, wheredminis equal to 40 meters.

[0218] Velocity module 1406 may merge the three constraints determined above-the lateral acceleration constraint, the curvature slew rate constraint, and comfort constraint-by using the smallest value ofdeclimcalculated in association with each constraint. After merging the constraints, velocity module 1406 may arrive at the initial target velocity by determining how much to adjust the speed of vehicle 200, denotedvcom, according to the following equation:vcom=vcom(prev) +declim*Δt,wherevcom(prev)represents the prior speed of vehicle 200 (e.g., a previously determined speed, the speed of vehicle 200 prior to any adjustment by velocity module 1406, etc.), and Δtrepresents the time between the determination ofvcomandvcom(prev).

[0219] At step 1540, processing unit 110 may execute velocity module 1406 to adjust a velocity of vehicle 200 to the initial target velocity determined at step 1530. Based on the current velocity of vehicle 200 and the initial target velocity determined at step 1530, processing unit 110 may transmit electronic signals to throttling system 220, braking system 230, and/or steering system 240 of vehicle 200 to change the velocity of vehicle 200 to the initial target velocity. Further, one or more actuators may control throttling system 220, braking system 230, and/or steering system 240. For example, processing unit 110 may transmit electronic signals that cause system 100 to physically depress the brake by a predetermined amount or ease partially off the accelerator of vehicle 200. Further, processing unit 110 may transmit electronic signals that cause system 100 to steer vehicle 200 in a particular direction. In some embodiments, the magnitude of acceleration (or de-acceleration) may be based on the difference between the current velocity of vehicle 200 and the initial target velocity. Where the difference is large (e.g., the current velocity is 10 km/hr greater than the initial target velocity), for example, system 100 may change the current velocity by braking in a manner that maximizes deceleration and results in vehicle 200 achieving the initial target velocity in the shortest amount of time safely possible. Alternatively, system 100

IPR2025-00944
Tesla EX1002 Page 1109

may apply the brakes in a manner that minimizes deceleration and results in vehicle 200 achieving the initial target velocity gradually. More generally, system 100 may change the current velocity to the initial target velocity according to any particular braking profile (e.g., a braking profile calling for a high level of braking for the first two seconds, and a low level of braking for the subsequent three seconds). The disclosed embodiments are not limited to any particular braking profile or manner of braking.

[0220] At step 1550, processing unit 110 may execute characteristic observation module 1404 to determine one or more characteristics of a curve located ahead of vehicle 200 based on, for example, analysis of the plurality of images conducted using techniques described in connection with FIGS. 5A-5D , above. The characteristics may include, for example, estimated and/or actual values associated with curve radius 1308, a road width or lane width 1306, a degree of curvature, a rate of change in curvature, a degree of banking, and the like. The characteristics may also include, for example, information conveyed by one or more road signs (such as road sign 1302). Furthermore, the characteristics may include the length or arc length of the detected curve, which may be a function of the speed of vehicle 200 as it approaches the curve. Processing unit 110 may execute characteristic observation module 1404 to calculate values associated with these characteristics (e.g., curve radius 1308, the degree of curvature, the rate of change in curvature, the arc length of the curve, etc.) based on mathematical models and/or formulae. In some embodiments, processing unit 110 may execute characteristic observation module 1404 to construct a mathematical model expressing the curvature of the curve as a function of distance (e.g., distance traveled). Such a model may be subject to constraints such as a maximum lateral acceleration constraint (e.g., in the range of 1.5 to 3.0 m/sec<2>) and/or a maximum lateral acceleration derivative constraint (e.g., in the range of 0.8 to 1.2 m/sec<2>). These constraints may yield a maximum speed for vehicle 200 as a function of distance traveled.

[0221] At step 1560, processing unit 110 may execute velocity module 1406 to calculate an updated target velocity based on the characteristics determined at step 1550. The updated target velocity may reflect an updated speed and/or direction for vehicle 200 to safely perform any combination of: (i) entering an upcoming curve, (ii) navigating through a curve, or (iii) exiting a curve. The updated target velocity may be determined using techniques described in connection with FIG. 14 and step 1530, above. For example, velocity module 1406 may determine the updated target velocity based on characteristics associated with a curve and mathematical models and/or formulae.

[0222] At step 1570, processing unit 110 may execute velocity module 1406 to adjust the velocity of vehicle 200 to the updated target velocity determined at step 1560. This adjustment may be accomplished using techniques described in connection with step 1540, above. For example, based on the current velocity of vehicle 200 and the updated target velocity determined at step 1560, processing unit 110 may transmit electronic signals (e.g., via a CAN bus) to throttling system 220, braking system 230, and/or steering system 240 of vehicle 200 to change the velocity of vehicle 200 to the updated target velocity. At this point, the current velocity of vehicle 200 is not necessarily the same as the initial target velocity determined at step 1530. For example, processing unit 110 may have determined the updated target velocity at step 1560 while in the process of adjusting the velocity of vehicle 200 to the initial target velocity. This may be true in the case where the difference in the velocity of vehicle 200 and the initial target velocity is large (e.g., greater than 10 km/hr) and the associated

Tesla EX1002 Page 1110 IPR2025-00944

acceleration is small.

[0223] In some embodiments, system 100 may make regular updates to the target velocity based on continued observations of characteristics of a curve (e.g., based on image data depicting portions of the curve that vehicle 200 is approaching). Additionally, after system 100 has navigated vehicle 200 through a curve, system 100 may cause vehicle 200 to accelerate to a new target velocity suitable for traveling a segment of the road that does not have curves, as determined by analyzing map data, positional data, and/or image data acquired by system 100.

Mimicking A Leading Vehicle

[0224] System 100 may provide driver assist functionality that causes vehicle 200 to mimic (or decline to mimic) a leading vehicle in different scenarios, such as when the leading vehicle switches lanes, accelerates, or makes a turn. For example, system 100 may be configured to detect the leading vehicle by analyzing a plurality of images and determine one or more actions taken by the leading vehicle. In some scenarios, such as when the leading vehicle is turning at an intersection, system 100 may be configured to cause vehicle 200 to decline to mimic the turn. In other scenarios, such as when the leading vehicle is turning at an intersection without changing lanes within the intersection, system 100 may be configured to cause vehicle 200 to mimic the turn. System 100 may also be configured to cause vehicle 200 to mimic one or more actions of the leading vehicle based on a navigation history of the leading vehicle.

[0225] FIGS. 16A and 16B are diagrammatic representations of a primary vehicle 200a mimicking one or more actions of a leading vehicle 200b on road 1600a, consistent with the disclosed embodiments. As illustrated in FIGS. 16A and 16B , primary vehicle 200a may be trailing leading vehicle 200b while traveling in the same lane of road 1600a. The lane may have a left edge 1620, a right edge 1630, and a midpoint 1610, with w1and w2indicating the left and right halves of the lane, respectively. Primary vehicle 200a and leading vehicle 200b may be positioned in the lane such that neither vehicle is centered on the midpoint 1610. As shown in FIG. 16A , for example, primary vehicle 200a may be centered to the right of midpoint 1610 such that the distance c1between primary vehicle 200a and left edge 1620 is greater than the distance c2between primary vehicle 200a and right edge 1630. In contrast, leading vehicle 200b may be centered to the left of midpoint 1610 such that the distance c3between leading vehicle 200b and left edge 1620 is less than the distance c4between leading vehicle 200b and right edge 1630. Different configurations of the positions of primary vehicle 200a and leading vehicle 200b are possible. For example, primary vehicle 200a and leading vehicle 200b may be positioned anywhere on road 1600a. Further, although FIG. 16A depicts leading vehicle 200b as having a camera system, the disclosed embodiments are not limited to a configuration in which leading vehicle 200b includes a camera system.

[0226] System 100 of primary vehicle 200a may cause primary vehicle 200a to mimic one or more actions of leading vehicle 200b. As shown in FIG. 16B , for example, primary vehicle 200a may mimic a lane shift performed by leading vehicle 200b. After mimicking the lane shift, distance c1may be equal to distance c3, and distance c2may be equal to distance c4. In some embodiments, primary vehicle 200a may mimic a lane shift performed by leading vehicle 200b in a way such that distances c1and c3, and c2and c4, are unequal. For example, in the case where primary vehicle 200a mimics a leftward lane shift of primary vehicle 200b and is positioned (after the lane shift) to the

IPR2025-00944
Tesla EX1002 Page 1111

left of primary vehicle 200a, c1will be less than c3and c2will be greater than c4. Alternatively, where primary vehicle 200a mimics a leftward lane shift of primary vehicle 200b and is positioned to the right of primary vehicle 200a, c1will be greater than c3and c2will be less than c4.

[0227] FIGS. 16C and 16D are diagrammatic representations of a primary vehicle 200a mimicking one or more actions of a leading vehicle 200b on road 1600b, consistent with the disclosed embodiments. As illustrated in FIG. 16C , primary vehicle 200a may be traveling behind leading vehicle 200b, which may make a left turn at an intersection. System 100 of primary vehicle 200a may cause primary vehicle 200a to mimic the left turn of leading vehicle 200b, as shown in FIG. 16D . System 100 may cause primary vehicle 200a to mimic the left turn based on various factors, such as whether leading vehicle 200b changes lanes while turning within the intersection, the navigation history associated with leading vehicle 200b, the relative difference in speed between primary vehicle 200a and leading vehicle 200b, and the like. Additionally, primary vehicle 200a may mimic leading vehicle 200b under various conditions, such as based on whether leading vehicle 200b is traveling along the same route or the same portion of a route that primary vehicle 200a is traveling along. Further detail regarding scenarios in which system 100 may cause primary vehicle 200a to mimic one or more actions of leading vehicle 200b are described in connection with FIGS. 17-22 , below.

[0228] In some embodiments, system 100 may be configured to determine whether to mimic a particular movement of a leading vehicle. For example, system 100 may determine that certain movements of the leading vehicle do not need to be mimicked because the movements do not affect the

[0229] course of the leading vehicle. For example, system 100 may determine that small changes in movements, such as moving within a particular lane, do not need to be mimicked. In other embodiments, system 100 may implement a smoothing operation to filter out small movements of a leading vehicle. The smoothing operation may cause the primary vehicle to implement more significant and/or important movements (e.g., changing lanes) of a leading vehicle while filtering out smaller movements not affecting the overall course of the leading vehicle.

[0230] FIG. 17 is an exemplary block diagram of a memory configured to store instructions for performing one or more operations consistent with the disclosed embodiments. As shown in FIG. 17 , memory 140 may store a lane constraint module 1702, an action detection module 1704, and an action response module 1706. The disclosed embodiments are not limited to any particular configuration of memory 140. Further, processing unit 110 may execute the instructions stored in any of modules 1702-1706 included in memory 140.

[0231] In one embodiment, lane constraint module 1702 may store software instructions (such as computer vision software) which, when executed by processing unit 110, determine one or more constraints associated with a lane that primary vehicle 200a is traveling in. The lane constraints may include a midpoint 1610, a left edge 1620, and a right edge 1630 of the lane. The lane constraints may also include one or more lines or other symbols marked on the surface of road 1600a or 1600b. Processing unit 110 may execute lane constraint module 1702 to determine the constraints by, for example, analyzing one or more sets of images and/or using map and/or position data (e.g., data stored in map database 160). The image sets may be acquired by one or more of image capture devices 122, 124, and 126. Based on the analysis, system 100

Tesla EX1002 Page 1112 IPR2025-00944

(e.g., via processing unit 110) may cause primary vehicle 200a to travel within a lane defined by left edge 1620 and right edge 1630. In some embodiments, system 100 may cause primary vehicle 200a to move away from left edge 1620 or right edge 1630 if one of distance c1or c2is less than a predetermined threshold.

[0232] In one embodiment, action detection module 1704 may store software instructions (such as computer vision software) which, when executed by processing unit 110, detects one or more actions taken by leading vehicle 200b traveling in front of primary vehicle 200a. Leading vehicle 200b may be traveling in the same or different lane as primary vehicle 200a. Processing unit 110 may execute action detection module 1704 to detect one or more actions taken by leading vehicle 200b by, for example, analyzing one or more sets of images acquired by one or more of image capture devices 122, 124, and 126. Based on the analysis, processing unit 110 may determine, for example, that leading vehicle 200b has shifted lanes, made a turn, accelerated, decelerated, applied its brakes, and/or the like. As another example, leading vehicle 200b may perform a maneuver such that a first offset distance on a side of leading vehicle 200b adjacent to a first lane constraint is different from a second offset distance on a side of leading vehicle 200b adjacent to a second lane constraint. For example, processing unit 110 may perform the analysis based on the techniques described in connection with FIGS. 5A-5D , above.

[0233] In one embodiment, action response module 1706 may store software instructions (such as computer vision software) which, when executed by processing unit 110, determines one or more actions for primary vehicle 200a to take based on one or more actions taken by leading vehicle 200b detected by system 100 (e.g., via execution of action detection module 1704). For example, processing unit 110 may execute action response module 1706 to determine whether to mimic one or more actions of leading vehicle 200b. This determination may be based on the nature of leading vehicle 200b's actions (e.g., turn, lane shift, etc.), information associated with primary vehicle 200a (e.g., speed, distance c1and c2from the lane edges, etc.), road and environmental conditions (e.g., potholes, heavy rain or wind, etc.), and the like. In the case where processing unit 110 determines to mimic one or more actions of leading vehicle 200b, processing unit 110 may accomplish this by, for example, transmitting electronic signals (e.g., via a CAN bus) to throttling system 220, braking system 230, and/or steering system 240 of primary vehicle 200a to trigger a turn, lane shift, change in speed, and/or change in direction of primary vehicle 200a. Processing unit 110 may use the techniques described in connection with FIGS. 4-7 , above, to cause primary vehicle 200a to mimic the one or more actions. To mimic the one or more actions, system 100 may provide control signals to one or more of throttling system 220, braking system 230, and steering system 240 to navigate vehicle 200 (e.g., by causing an acceleration, a turn, a lane shift, etc.). Further, one or more actuators may control throttling system 220, braking system 230, and/or steering system 240.

[0234] FIG. 18 is a flow chart showing an exemplary process for causing primary vehicle 200a to mimic one or more actions of leading vehicle 200b consistent with disclosed embodiments. At step 1810, processing unit 110 may receive a plurality of images via data interface 128 between processing unit 110 and image acquisition unit 120. For instance, a camera included in image acquisition unit 120 (such as image capture device 122) may capture a plurality of images and transmit them over a digital connection (e.g., USB, wireless, Bluetooth, etc.) to processing unit 110. In some embodiments, processing unit 110 may receive more than one plurality of images via a plurality of data interfaces. For example, processing unit 110 may receive a plurality of

Tesla EX1002 Page 1113 IPR2025-00944

images from each of image capture devices 122, 124, 126, each of which may have an associated data interface for communicating data to processing unit 110. The disclosed embodiments are not limited to any particular data interface configurations or protocols.

[0235] At steps 1820 and 1830, processing unit 110 may execute lane constraint module 1702 to determine constraints for each side of primary vehicle 200a associated with a lane that primary vehicle 200a is traveling in. Processing unit 110 may determine the constraints by analyzing the images received at step 1810. Thus, at step 1820, processing unit 110 may determine a first lane constraint on a first side of primary vehicle 200a (e.g., the distance between the left side of primary vehicle 200a and left edge 1620 of the lane, c1) and, at step 1830, a second lane constraint on a second side opposite the first side (e.g., the distance between the right side of primary vehicle 200a and right edge 1620 of the lane, c2). In some embodiments, processing unit 110 may determine the constraints by using map and/or position data (e.g., data stored in map database 160) indicating the position of primary vehicle 200a relative to midpoint 1610, left edge 1620, and right edge 1630 of the lane. Processing unit 110 may determine the constraints based on both an analysis of the images received at step 1810 and map and/or position data; doing so may increase the confidence level associated with the constraints.

[0236] At step 1840, processing unit 110 may execute lane constraint module 1702 to cause primary vehicle 200a to travel within the lane constraints determined at steps 1820 and 1830. For example, the left side of primary vehicle 200a may be positioned to the left of left edge 1620, or the right side of primary vehicle 200a may be positioned to the right of right edge 1630. In such a scenario, processing unit 110 may transmit electronic signals (e.g., via a CAN bus) to throttling system 220, braking system 230, and/or steering system 240 of primary vehicle 200a to cause primary vehicle 200a to adjust its position and travel within the lane constraints.

[0237] At step 1850, processing unit 110 may execute action detection module 1704 to locate a leading vehicle 200b within the images received at step 1810 by analyzing the images. Processing unit 110 may detect the presence of vehicles in the images using the techniques described in connection with FIGS. 5A and 5B , above. For each detected vehicle, processing unit 110 may construct a multi-frame model of the detected vehicle's position, velocity (e.g., speed and direction), and acceleration relative to primary vehicle 200a. Processing unit 110 may use the techniques described in connection with FIGS. 5A and 5B , above, to construct such a model. In addition, processing unit 110 may determine a lane of travel associated with each detected vehicle, based on, for example, lane or other road markings and lane constraints determined at steps 1820 and 1830. Processing unit 110 may determine leading vehicle 200b to be the detected vehicle that is closest to, and is traveling in the same lane as, primary vehicle 200a. At step 1860, processing unit 110 may execute action detection module 1704 to determine one or more actions taken by the leading vehicle 200b detected within the images at step 1850. For example, processing unit 110 may determine that leading vehicle 200b has shifted lanes, made a turn, accelerated, decelerated, and/or the like. In addition, processing unit 110 may analyze the images to determine characteristics associated with leading vehicle 200b, such as its speed and position (e.g., relative to primary vehicle 200a). These determinations may be based on the techniques described in connection with FIGS. 5A and 5B , above. In some embodiments, processing unit 110 may analyze the images to determine characteristics associated with an action performed by leading vehicle 200b. For example, where leading vehicle 200b makes a turn, processing unit 110 may determine

position information for leading vehicle 200b right before it begins the turn, position information for leading vehicle 200b as it completes the turn, a turning radius, and a speed profile describing changes in speed (e.g., over time) for leading vehicle 200b as it makes the turn.

[0238] At step 1870, processing unit 110 may execute action response module 1706 to cause primary vehicle 200a to mimic the one or more actions of leading vehicle 200b determined at step 1850. For example, processing unit 110 may transmit electronic signals (e.g., via a CAN bus) to throttling system 220, braking system 230, and/or steering system 240 of primary vehicle 200a to trigger a turn, lane shift, change in speed, and/or change in direction of primary vehicle 200a. In addition, processing unit 110 may construct a set of path points to guide changes in position and speed for primary vehicle 200a. The path points may be based on lane marks (e.g., midpoint 1610, left edge 1620, right edge 1630, etc.) represented by a polynomial (e.g., of the third-degree) and target coordinates associated with a target vehicle, such as leading vehicle 200b. Processing unit 110 may adjust the coefficients of the polynomial so that the path points pass through the target coordinates. Processing unit 110 may also offset the path points (e.g., relative to left edge 1620 and right edge 1630) based on a predetermined offset amount. As processing unit 110 adjusts the polynomial coefficients and offsets the path points, processing unit 110 may also calculate local curvature information for each of the resulting segments formed between the path points.

[0239] In some embodiments, processing unit 110 (e.g., via execution of action response module 1706) may cause primary vehicle 200a to mimic leading vehicle 200b within certain position, speed, and/or acceleration constraints associated with the two vehicles. For example, the constraints may require a predetermined minimum distance (e.g., a safety distance) between primary vehicle 200a and leading vehicle 200b or a predetermined maximum difference in speed and/or acceleration between the vehicles. Thus, if mimicking an action performed by leading vehicle 200b would violate one or more constraints (e.g., mimicking an acceleration would result in primary vehicle 200a being too close to leading vehicle 200b), processing unit 110 may cause primary vehicle 200a to decline to mimic the one or more actions of leading vehicle 200b. In addition, even while mimicking an action performed by leading vehicle 200b (e.g., a lane shift), processing unit 110 may cause primary vehicle 200a to follow leading vehicle 200b's speed profile and maintain a predetermined distance or a safety distance from leading vehicle 200b. Processing unit 110 may accomplish this by causing primary vehicle 200a to: (i) continuously close the distance gap from leading vehicle 200b when the gap is greater than the predetermined minimum distance, (ii) continuously close the gap between the speeds of primary vehicle 200a and leading vehicle 200b, and (iii) match the acceleration of leading vehicle 200b to the extent it is consistent with predetermined constraints. To control the extent to which of these actions should be performed, processing unit 110 may associate each action with weights in real-time based on the current and predicted distance to leading vehicle 200b. Thus, for example, as the distance to leading vehicle 200b decreases more weight may be put on decelerating rather than matching the acceleration of leading vehicle 200b.

[0240] As another example, in the case where primary vehicle 200a is approaching leading vehicle 200b at a speed greater than the speed of leading vehicle 200b, processing unit 110 may cause primary vehicle 200a to gradually slow down such that, when the distance between the two vehicles equals a predetermined safety distance

IPR2025-00944
Tesla EX1002 Page 1115

such as a minimum following distance, the speed of the two vehicles is the same. In another example, where a leading vehicle 200b abruptly appears in front of primary vehicle 200a, causing the distance between the two vehicles to be less than the minimum following distance, processing unit 110 may cause primary vehicle 200a to gradually slow down to avoid approaching leading vehicle 200b any closer while at the same increasing the distance between the vehicles until it reaches the minimum following distance.

[0241] FIG. 19 is a flow chart showing an exemplary process 1900 for causing primary vehicle 200a to decline to mimic a turn of leading vehicle 200b, consistent with disclosed embodiments. At step 1910, processing unit 110 may receive a plurality of images via data interface 128 between processing unit 110 and image acquisition unit 120, as described in connection with step 1810 of FIG. 18 , above. At step 1920, processing unit 110 may execute action detection module 1704 to locate a leading vehicle 200b within the images received at step 1910 by analyzing the images, as described in connection with step 1850 of FIG. 18 , above. At step 1925, processing unit 110 may execute action detection module 1704 to determine whether leading vehicle 200b is turning at an intersection, using the techniques described in connection with step 1860 of FIG. 18 , above. If leading vehicle 200b is determined to be turning at the intersection (step 1925, yes), processing unit 110 may cause primary vehicle 200a to decline to mimic the turn of leading vehicle 200b at step 1930. Otherwise (step 1925, no), process 1900 concludes.

[0242] FIG. 20 is a flow chart showing an exemplary process 2000 for causing primary vehicle 200a to mimic or decline a turn of leading vehicle 200b, consistent with disclosed embodiments. At step 2010, processing unit 110 may receive a plurality of images via data interface 128 between processing unit 110 and image acquisition unit 120, as described in connection with step 1810 of FIG. 18 , above. At step 2020, processing unit 110 may execute action detection module 1704 to locate a leading vehicle 200b within the images received at step 2010 by analyzing the images, as described in connection with step 1850 of FIG. 18 , above. At step 2025, processing unit 110 may execute action detection module 1704 to determine whether leading vehicle 200b is turning at an intersection, using the techniques described in connection with step 1860 of FIG. 18 , above. If leading vehicle 200b is determined to be turning at the intersection (step 2025, yes), at step 2035, processing unit 110 may execute action detection module 1704 to determine whether leading vehicle 200b is changing lanes within the intersection, using the techniques described in connection with step 1860 of FIG. 18 , above. Otherwise (step 2025, no), process 2000 concludes. If leading vehicle 200b is determined to be changing lanes within the intersection (step 2035, yes), processing unit 110 may cause primary vehicle 200a to decline to mimic the turn of leading vehicle 200b at step 2050. Otherwise (step 2035, no), processing unit 110 may cause primary vehicle 200a to mimic the turn of leading vehicle 200b at step 2040.

[0243] FIG. 21 is a flow chart showing another exemplary process 2100 for causing primary vehicle 200a to mimic one or more actions of a first leading vehicle, consistent with disclosed embodiments. At step 2110, processing unit 110 may receive a plurality of images via data interface 128 between processing unit 110 and image acquisition unit 120, as described in connection with step 1810 of FIG. 18 , above. At steps 2120 and 2130, processing unit 110 may execute action detection module 1704 to locate a first and second leading vehicle within the images received at step 2110 by analyzing the images, as described in connection with step 1850 of FIG. 18 , above. At step 2135, processing unit 110 may execute action detection module 1704 to determine

IPR2025-00944
Tesla EX1002 Page 1116

whether the first leading vehicle path has a lower turning radius than the second leading vehicle path, using the techniques described in connection with step 1860 of FIG. 18 , above. If the first leading vehicle path is determined to have a lower turning radius than the second leading vehicle path (step 2135, yes), processing unit 110 may cause primary vehicle 200a to mimic the turn of the first leading vehicle at step 2140. Otherwise (step 2135, no), process 2100 concludes.

[0244] FIG. 22 is a flow chart showing an exemplary process 2200 for causing primary vehicle 200a to mimic one or more actions of leading vehicle 200b based on a navigation history, consistent with disclosed embodiments. At step 2210, processing unit 110 may receive a plurality of images via data interface 128 between processing unit 110 and image acquisition unit 120, as described in connection with step 1810 of FIG. 18 , above. At step 2220, processing unit 110 may execute action detection module 1704 to locate a leading vehicle 200b within the images received at step 2210 by analyzing the images, as described in connection with step 1850 of FIG. 18 , above.

[0245] At step 2230, processing unit 110 may determine position information for leading vehicle 200b by analyzing the images received at step 2210. Position information for leading vehicle 200b may indicate the position of leading vehicle 200b relative to primary vehicle 200a. Processing unit 110 may analyze the images to derive position information for leading vehicle 200b and one or more actions taken by leading vehicle 200b by using the techniques described in connection with FIGS. 5A-5D , above.

[0246] At step 2240, processing unit 110 may compare the position information for leading vehicle 200b determined at step 2230 to predetermined map data (e.g., data stored in map database 160). For example, by comparing the position information to predetermined map data, processing unit 110 may derive position coordinates for leading vehicle 200b. In addition, processing unit 110 may use the predetermined map data as an indicator of a confidence level associated with the position information determined at step 2230.

[0247] At step 2250, processing unit 110 may create a navigation history associated with leading vehicle 200b by, for example, tracking the actions taken by leading vehicle 200b. Processing unit 110 may associate the actions taken by leading vehicle 200b (e.g., determined from analysis of the images at step 2230) with predetermined map data from step 2240. Thus, the navigation history may indicate that leading vehicle 200b performed certain actions (e.g., a turn, a lane shift, etc.) at certain position coordinates.

[0248] At step 2260, processing unit 110 may execute action response module 1706 to determine whether to cause primary vehicle 200a to mimic one or more actions of leading vehicle 200b based on the navigation history created at step 2250 and using the techniques described in connection with step 1870 of FIG. 18 , above. For example, processing unit 110 (e.g., via action response module 1706) may cause primary vehicle 200a to mimic a turn of leading vehicle 200b, even where no turn is detected from analysis of the images (e.g., at step 2230), when the navigation history indicates that leading vehicle 200b has successfully navigated at least one prior turn. More generally, processing unit 110 may cause primary vehicle 200a to mimic an action taken by leading vehicle 200b in cases where the navigation history indicates the leading vehicle 200b had successfully performed such an action.

Navigating a Vehicle to Pass Another Vehicle

Tesla EX1002 Page 1117 IPR2025-00944

[0249] System 100 may provide driver assist functionality that monitors the vicinity around vehicle 200 and aborts lane changes in the event that the lane change is deemed to be at risk of causing a collision, such as with another vehicle (e.g., a target vehicle). For example, system 100 may enable vehicle 200 to complete a pass of a target vehicle if the target vehicle is determined to be in a different lane than vehicle 200. If, before completion of a pass, system 100 determines that the target vehicle is entering the lane in which vehicle 200 is traveling, system 100 may cause vehicle 200 to abort the pass. For example, aborting the pass may include causing vehicle 200 to brake or causing vehicle 200 to stop accelerating.

[0250] As illustrated in FIG. 23 , vehicle 200 may travel on roadway 2300. Vehicle 200 may be equipped with system 100 and may implement any of the disclosed embodiments for identifying lane constraints and operating vehicle 200 within the lane constraints. System 100 may also facilitate lane changes made by vehicle 200. For example, system 100 may determine whether circumstances allow vehicle 200 to safely change lanes and, if the circumstances change during the lane change, cause vehicle 200 to abort lane change.

[0251] As shown in FIG. 23 and as previously discussed with respect to FIG. 8 , vehicle 200 may include first vehicle side 802, which may be first distance 805 from a first lane constraint 2302. Similarly, vehicle 200 may include second vehicle side 812 opposite from first vehicle side 802, and second vehicle side 812 may be second distance 815 from a second lane constraint 2304. In this manner, first lane constraint 810 and second lane constraint 2304 may define a lane 2330 within which vehicle 200 may travel.

[0252] System 100 may determine first lane constraint 2302 and second lane constraint 2304 based on a plurality of images acquired by one or more of image capture devices 122-126. According to some embodiments, first lane constraint 2302 and/or second lane constraint 2304 may be identified by visible lane boundaries, including lines marked on a road surface. Additionally or alternatively, first lane constraint 2302 and/or second lane constraint 2304 may include an edge of a road surface. According to some embodiments, system 100 may determine first lane constraint 2302 and/or second lane constraint 2304 by identifying a midpoint of a road surface width 2320 (or any other suitable road feature that vehicle 200 may use as a navigation reference). System 100 may identify lane constraints in this manner when, for example, lines designating road lanes are not painted or otherwise labeled.

[0253] When passing a target vehicle 2310, the lane change of vehicle 200 may be complicated by a change in the trajectory of target vehicle 2330. Thus, system 100 may identify target vehicle 2310 based on an analysis of a plurality of images acquired by one or more of image capture devices 122-126. System 100 may also identify second lane 2340 in which target vehicle 2310 is traveling. Further, system 100 may determine whether second lane 2340 is different than lane 2330 in which vehicle 200 is traveling. If second lane 2340 is different than lane 2330, system 100 may be configured to enable vehicle 200 to pass target vehicle 2310.

[0254] System 100 may also monitor target vehicle 2310 based on the plurality of images. For example, system 100 may monitor a position of target vehicle 2310, such as monitoring its position relative to vehicle 200, relative to first lane constraint 2302 and/or second lane constraint 2304, an absolute position, or relative to another

Tesla EX1002 Page 1118 IPR2025-00944

reference point. Additionally or alternatively, monitoring a position of target vehicle 2310 may include estimating a speed of target vehicle 2310.

[0255] If system 100 determines that target vehicle 2310 is entering lane 2330 in which vehicle 200 is traveling, system 100 may cause vehicle 200 to abort the pass of target vehicle 2310. For example, system 100 may cause vehicle 200 to stop accelerating or brake. Additionally or alternatively, system 100 may issue an audible announcement, such as to notify driver of vehicle 200 that the lane change is being aborted.

[0256] FIG. 24A is an exemplary block diagram of memory 140, which may store instructions for performing one or more operations consistent with disclosed embodiments. As illustrated in FIG. 24A , memory 140 may store one or more modules for identification of the target vehicle and responses described herein. For example, memory 140 may store a lane constraint module 2400, a target vehicle acquisition module 2410, and an action module 2415.

[0257] In one embodiment, lane constraint module 2400 may store instructions which, when executed by processing unit 110, may detect and define first lane constraint 2302 and second lane constraint 2304. For example, processing unit 110 may execute lane offset module 2400 to process the plurality of images received from at least one image capture device 122-124 and detect first lane constraint 2302 and second lane constraint 2304. As discussed above, this may include identifying painted lane lines or markers and/or measuring a midpoint of a road surface. Processing unit 110 may perform the analysis based on the techniques described in connection with FIGS. 5A-5D , above.

[0258] In some embodiments, target acquisition module 2410 may store instructions which, when executed by processing unit 110, may detect the presence of target vehicle 2310 and monitor the position of target vehicle 2310. For example, target acquisition module 2410 may process the plurality of images to detect and monitor target vehicle 2310. Additionally or alternatively, when executing target acquisition module 2410, processing unit 110 may receive information from another module or other system indicative of the presence and/or location of target vehicle 2310. Processing unit 110 may perform the analysis based on the techniques described in connection with FIGS. 5A-5D , above.

[0259] Processing unit 110 may be configured to determine the presence of one or more objects such as target vehicle 2310 in the vicinity of vehicle 200. For example target vehicle 2310 may comprise another vehicle, such as a car, truck or motorcycle traveling near vehicle 200. Processing unit 110 may determine, for each target vehicle 2310, an offset profile. The offset profile may include a determination of whether in its current and predicted position, target vehicle 2310 will be within a predefined range of vehicle 200. If processing unit 110 determines that target vehicle 2310 is or will be within a predefined range of vehicle 200, processing unit may determine whether there is enough space within first lane constraint 810 and second lane constraint 820, for vehicle 200 to bypass target vehicle 2310. If there is not enough space, processing unit 110 may execute a lane change. If there is enough space, processing unit 110 may determine whether there is enough time to bypass target vehicle 2310 before the distance between vehicle 200 and target vehicle 2310 closes. If there is enough time, processing unit 110 may initiate and or activate the offset maneuver to bypass target vehicle 2310. The offset maneuver may be executed by processing unit 110 so that the movement of vehicle 200 is smooth and comfortable for the driver. For example, the

offset slew rate may be approximately 0.15 to 0.75 m/sec. The offset maneuver may be executed so that the maximum amplitude of the offset maneuver will occur when the gap between vehicle 200 and object 825 closes. If there is not enough time, processing unit 110 may initiate braking and then initiate an offset maneuver.

[0260] Further, processing unit 110 may execute action module 2415 to abort the plan of vehicle 200 to pass target vehicle 2310. If target vehicle 2310 is not in lane 2330 of vehicle 200, instructions included in action module 1230 may enable vehicle 200 to pass target vehicle 2330. However, if execution of target vehicle acquisition module 2410 indicates that target vehicle 2310 is changing lanes into lane 2330 of vehicle 200, processing unit 110 may execute instructions included in action module 2415 to abort the lane change of vehicle 200. For example, processing unit 110 may transmit electronic signals to throttling system 220, braking system 230, and/or steering system 240 of vehicle 200 to abort the lane change. Additionally or alternatively, processing unit may cause an audible notification to sound.

[0261] FIG. 24B illustrates a flowchart of an exemplary process 2420 for navigating a vehicle among encroaching vehicles, consistent with disclosed embodiments. Process 2420 may identify lane constraints that define a lane of travel for a vehicle (e.g., vehicle 200) to travel, identify and monitor a target vehicle (e.g., vehicle 2310), and enable the vehicle to pass the target vehicle or abort the pass if it is determined that the target vehicle is entering a lane in which the vehicle is traveling.

[0262] At step 2430, at least one of image capture device 122, 124, and/or 126 may acquire a plurality of images of an area in the vicinity of vehicle 200. For example, a camera included in image acquisition unit 120 (such as image capture device 122) may capture a plurality of images and transmit them over a digital connection (e.g., USB, wireless, Bluetooth, etc.) to processing unit 110. In some embodiments, processing unit 110 may receive more than one plurality of images via a plurality of data interfaces. For example, processing unit 110 may receive a plurality of images from each of image capture devices 122, 124, 126, each of which may have an associated data interface for communicating data to processing unit 110. The disclosed embodiments are not limited to any particular data interface configurations or protocols.

[0263] Next, at step 2440, processing unit 110 may determine a first lane constraint on first vehicle side 802 and a second lane constraint on second vehicle side 812 using the plurality of images received via data interface 128. For example, as part of implementing step 2440, processing unit 110 may execute instructions of lane constraint module 910. Further, as part of determining the first and second lane constraints, processing unit 110 may use one or more of the processes discussed above in connection with FIGS. 5A-5D .

[0264] At step 2450, processing unit 110 may acquire target vehicle 2310. To acquire target vehicle 2310, processing unit 110 may execute target vehicle acquisition module 2410. For example, processing unit 110 may execute instructions to analyze acquired images to identify second lane 2340 in which target vehicle 2310 is traveling and determine whether target vehicle 2310 is traveling in lane 2330 in which vehicle 200 is traveling. This determination may be based on the techniques described in connection with FIGS. 5A-5D , above.

[0265] At step 2455, processing unit 110 may determine whether second lane 2340 in which target vehicle 2310 is traveling is different than lane 2330 in which vehicle 200 is

IPR2025-00944
Tesla EX1002 Page 1120

traveling.

[0266] At step 2460, processing unit 110 may enable vehicle 200 to pass target vehicle 2310 if second lane 2340 is different than lane 2330 in which vehicle 200 is traveling. To enable vehicle 200 to pass target vehicle 2310, processing unit may execute action module 2415.

[0267] Next, at step 2465, processing unit 110 may execute target acquisition module 2410 to monitor target vehicle 2310. This may include estimating a velocity of target vehicle 2310 and determining the location of target vehicle 2310, such as with respect to vehicle 200.

[0268] Step 2470 may include determining whether target vehicle 2310 has entered lane 2330. If processing unit 110 determines that target vehicle 2310 has not entered lane 2330, at step 2480, processing unit 110 may allow vehicle 200 to complete the pass.

[0269] At step 2490, if processing unit 110 determines that target vehicle 2310 has entered lane 2330 (or is entering lane 2330 or is otherwise on a trajectory to bring vehicle 2310 into lane 2330), processing unit 110 may execute action module 2415 to abort the pass. This may include, for example, issuing an audible notification to driver of vehicle 200 via speakers 360. Additionally or alternatively, step 2490 may include executing instructions included in action module 2415 to cause processing unit 110 to transmit electronic signals to throttling system 220, braking system 230, and/or steering system 240 of vehicle 200 to abort lane change.

Navigating a Vehicle to Avoid Encroaching Traffic

[0270] System 100 may provide driver assist functionality that monitors the vicinity around vehicle 200 and responds to the presence of another vehicle encroaching upon vehicle 200. Monitoring the vicinity around vehicle 200 may include monitoring a collision threshold. For example a collision threshold may include a time to collision or a minimum predetermined distance between vehicle 200 and other traffic. System 100 may further cause vehicle 200 to take evasive action to avoid encroaching traffic. For example, evasive action may include braking, changing lanes, or otherwise altering the course of vehicle 200.

[0271] As illustrated in FIG. 25 , vehicle 200 may travel on roadway 2500. Vehicle 200 may include system 100 to detect and respond to encroaching traffic. For example, vehicle 200 may determine that another vehicle is entering into the lane in which vehicle 200 is traveling or otherwise determine that another vehicle is crossing a collision threshold. Vehicle 200 may react to an encroaching vehicle, such as by altering course, braking, or accelerating.

[0272] As discussed with respect to FIG. 8 and as shown in FIG. 25 , vehicle 200 may include first vehicle side 802, which may be first distance 805 from a first lane constraint 2502. Similarly, vehicle 200 may include second vehicle side 812 opposite from first vehicle side 802, and second vehicle side 812 may be second distance 815 from a second lane constraint 2504. In this manner, first lane constraint 2502 and second lane constraint 2504 may define a lane 2505 within which vehicle 200 may travel.

Tesla EX1002 Page 1121 IPR2025-00944

[0273] System 100 may determine first lane constraint 2502 and second lane constraint 2504 based on a plurality of images acquired by one or more of image capture devices 122-126. According to some embodiments, first lane constraint 2502 and/or second lane constraint 2504 may be identified by visible lane boundaries, such a line marked on a road surface. Additionally or alternatively, first lane constraint 2502 and/or second lane constraint 2504 may include an edge of a road surface. According to some embodiments, system 100 may determine first lane constraint 2502 and/or second lane constraint 2502 by identifying a midpoint 2520 of a road surface width (or any other suitable road feature vehicle 200 may use as a navigation reference). System 100 may identify lane constraints in this manner when, for example, lines designating road lanes are not painted or otherwise labeled.

[0274] Detection of first lane constraint 2502 and/or second lane constraint 2504 may include processing unit 110 determining their 3D models in a camera coordinate system. For example, the 3D models of first lane constraint 2502 and/or second lane constraint 2504 may be described by a third-degree polynomial. In addition to 3D modeling of first lane constraint 2502 and/or second lane constraint 2504, processing unit 110 may estimate motion parameters, such as the speed, yaw and pitch rates, and acceleration of vehicle 200. Optionally, processing unit may detect static and moving vehicles and their position, heading, speed, and acceleration, all relative to vehicle 200. Processing unit 110 may determine a road elevation model to transform the information acquired from the plurality of images into 3D space.

[0275] The driver of vehicle 200 cannot assume that other traffic traveling in adjacent or near-adjacent lanes will only change lanes if it is safe for that traffic to shift into lane 2505 in which vehicle 200 is driving. Thus, it may be desirable to consider and monitor traffic traveling in nearby lanes to detect lateral encroachment into lane 2505 of vehicle 200, so that vehicle 200 may react to lateral encroachment to avoid a collision. System 100 may identify an encroaching vehicle 2510 based on the plurality of images and determine that encroaching vehicle 2510 is approaching vehicle 200. System 100 may then cause an action to account for the encroaching vehicle. For example, system 100 may cause vehicle 200 to maintain a current velocity and travel within first lane constraint 2502 and second lane constraint 2504 such that a first distance 2530, which is on first vehicle side 802 on which encroaching vehicle 2510 is approaching, is greater than a second distance 2540.

[0276] Processing unit 110 may be configured to determine the presence of one or more encroaching vehicles 2510 in the vicinity of vehicle 200. Processing unit 110 may determine, for each encroaching vehicle 2510 an offset profile. The offset profile may include a determination of whether in its current and predicted position, encroaching vehicle 2510 will be within a predefined range of vehicle 200. If processing unit 110 determines that encroaching vehicle 2510 is or will be within a predefined range of vehicle 200, processing unit may determine whether there is enough space within first lane constraint 2502 and second lane constraint 2504, for vehicle 200 to bypass encroaching vehicle 2510. If there is not enough space, processing unit 110 may execute a lane change. If there is enough space, processing unit 110 may determine whether there is enough time to bypass encroaching vehicle 2510 before the distance between vehicle 200 and encroaching vehicle 2510 closes.

[0277] For example, if there is enough time, processing unit 110 may initiate and or activate the offset maneuver to bypass encroaching vehicle 2510. The offset maneuver may be executed by processing unit 110 so that the movement of vehicle 200 is

IPR2025-00944
Tesla EX1002 Page 1122

smooth and comfortable for the driver. For example, the offset slew rate may be approximately 0.15 to 0.75 m/sec. The offset maneuver may be executed so that the maximum amplitude of the offset maneuver will occur when the gap between vehicle 200 and encroaching vehicle 2510 closes. If there is not enough time, processing unit 110 may initiate braking (e.g., via transmitting electronic signals to braking system 230) and then initiate offset maneuver (e.g., by transmitting electronic signals to throttling system 220 and/or steering system 240).

[0278] System 100 may determine, based on a plurality of images acquired by one or more of image capture devices 122-126, that encroaching vehicle 2510 has crossed at least one collision threshold. For example, the at least one collision threshold may include a minimum predetermined distance 2550 between vehicle 200 and encroaching vehicle 2510. For example, encroaching vehicle 2510, as illustrated in FIG. 25 , is shown exceeding minimum predetermined distance 2550. Additionally or alternatively, the at least one collision threshold may include a time to collision of vehicle 200 and encroaching vehicle 2510. The time-to-collision threshold may be determined based on the distance between vehicle 200 and encroaching vehicle 2510, the velocity of vehicle 200 and/or encroaching vehicle 2510, the lateral velocity of vehicle 200 and/or encroaching vehicle 2510, and/or the acceleration of vehicle 200 and/or encroaching vehicle 2510. A collision threshold based on a time to collision may be set to allow enough time for system 100 to effectively conduct evasive action to avoid collision.

[0279] If system 100 determines that at least one collision threshold has been crossed, system 100 may cause vehicle 200 to take evasive action. For example, this may include causing other subsystems of vehicle 200 to operate. Thus, system 100 may operate to cause vehicle 200 to increase speed by communicating with throttling system 220, braking system 230, and/or steering system 240 to change the speed and/or direction of vehicle 200.

[0280] The evasive action may include causing vehicle 200 to change lanes. This may be a desirable response if, for example, vehicle 200 is traveling on a multilane highway and system 100 determines that changing lanes to move away from encroaching vehicle 2510 will avoid a collision. Evasive action may include vehicle 200 altering its course. Additionally or alternatively, evasive action may include altering course of vehicle 200. For example, to take evasive action, processing unit 110 may transmit electronic signals to throttling system 220, braking system 230, and/or steering system 240.

[0281] System 100 may also be configured to cause vehicle 200 to take an action after the evasive action has been completed. For example, system 100 may cause vehicle 200 to return to midpoint 2520 of lane 2505 in which vehicle 200 is traveling. Additionally or alternatively, system 100 may cause vehicle 200 to resume a preset speed after braking in response to encroaching vehicle 2510.

[0282] FIG. 26 is an exemplary block diagram of memory 140, which may store instructions for detecting and responding to traffic laterally encroaching on a vehicle consistent with disclosed embodiments. As illustrated in FIG. 26 , memory 140 may store a lane constraint module 2610, an encroaching vehicle acquisition module 2620, and an action module 2630.

[0283] In one embodiment, lane constraint module 2610 may store instructions which, when executed by processing unit 110, may detect and define first lane constraint 2502

and second lane constraint 2504. For example, processing unit 110 may execute lane offset constraint module 2610 to process the plurality of images received from at least one image capture device 122-126 and detect first lane constraint 2502 and second lane constraint 2504. As discussed above, this may include identifying painted lane lines and/or measuring a midpoint of a road surface. Processing unit 110 may perform the analysis based on the techniques described in connection with FIGS. 5A-5D , above.

[0284] In one embodiment, encroaching vehicle acquisition module 2620 may store instructions which, when executed by processing unit 110, may detect the presence of encroaching vehicle 2510 and monitor encroaching vehicle 2510. For example, encroaching vehicle acquisition module 2620 may include instructions for determining a relative location of encroaching vehicle 2510 to vehicle 200. In some embodiments, this may include monitoring the velocity of encroaching vehicle 200. According to some embodiments, encroaching vehicle acquisition module 2620 may determine whether encroaching vehicle 2510 has exceeded a collision threshold. A collision threshold may be, for example, a minimum predetermined distance between vehicle 200 and encroaching vehicle 2510, If encroaching vehicle 2510 comes closer to vehicle 200 than the minimum predetermined distance, processing unit 110 may determine that encroaching vehicle 2510 is indeed encroaching and execute instructions to avoid collision with encroaching vehicle 2510.

[0285] Additionally or alternatively, a collision threshold may include a time to collision. The time-to-collision threshold may be determined based on the distance between vehicle 200 and encroaching vehicle 2510, the velocity of vehicle 200 and/or encroaching vehicle 2510, the lateral velocity of vehicle 200 and/or encroaching vehicle 2510, and/or the acceleration of vehicle 200 and/or encroaching vehicle 2510. A collision threshold based on a time to collision may be set to allow enough time for processing unit 110 and system 100 to effectively conduct evasive action to avoid collision.

[0286] Further, processing unit 110 may execute action module 2630 to respond to the detection of encroaching vehicle 2510 and/or the crossing of a collision threshold. For example, action module 2630 may include instructions to cause vehicle 200 to maintain a current velocity and to travel within first lane constraint 2502 and second lane constraint 2504 such that first distance 2530 is greater than second distance 2540. The first distance 2530 may be on the same side of vehicle 200 as encroaching vehicle 2510, so that vehicle 200 is farther away from encroaching vehicle 2510 while still within first lane constraint 2502 and second lane constraint 2504. If processing unit 110 determines that a collision threshold has been crossed, processing unit 110 may execute action module 2630 to cause vehicle 200 to take evasive action, such as altering course, braking, and/or changing lanes. Action module 2630 may also include instructions that cause vehicle 200 to resume course after completing evasive action. For example, processing unit 110 may cause vehicle 200 to resume a preset speed after braking or return to the center of lane 2505.

[0287] FIG. 27 illustrates a flowchart of an exemplary process 2720 detecting and responding to traffic laterally encroaching on a vehicle, consistent with disclosed embodiments. Process 2720 may identify lane constraints that define a lane of travel for vehicle travel, determine whether an encroaching vehicle 2510 is approaching, and cause vehicle 200 to maintain current velocity and to travel within first lane constraint 2501 and second lane constraint 2504 such that first distance 2530, on the side of

IPR2025-00944
Tesla EX1002 Page 1124

vehicle 200 that encroaching vehicle 2510 is approaching, is greater than second distance 2540.

[0288] At step 2730, at least one image capture device 122,124, and/or 126 may acquire a plurality of images of an area in the vicinity of vehicle 200. For example, a camera included in image acquisition unit 120 (such as image capture device 122) may capture a plurality of images and transmit them over a digital connection (e.g., USB, wireless, Bluetooth, etc.) to processing unit 110. In some embodiments, processing unit 110 may receive more than one plurality of images via a plurality of data interfaces. For example, processing unit 110 may receive a plurality of images from each of image capture devices 122, 124, 126, each of which may have an associated data interface for communicating data to processing unit 110. The disclosed embodiments are not limited to any particular data interface configurations or protocols.

[0289] Next, at step 2740, processing unit 110 may determine a first lane constraint on first vehicle side 802 and a second lane constraint on second vehicle side 812 using the plurality of images received via data interface 128. For example, as part of implementing step 2740, processing unit 110 may execute instructions of lane constraint module 2610. Further, as part of determining the first and second lane constraints, processing unit 110 may use one or more of the processes discussed above in connection with FIGS. 5A-5D .

[0290] At step 2750, processing unit 110 may determine whether encroaching vehicle 2510 is approaching. To make this determination, processing unit 110 may execute encroaching vehicle acquisition module 2620. Encroaching vehicle acquisition module 2620 may include instructions to process images to detect other traffic. This may include identifying vehicles within a certain range of vehicle 200, such as those vehicles that are adjacent to vehicle 200. These adjacent vehicles may be encroaching vehicles 2510.

[0291] Once encroaching vehicle 2510 is identified, processing unit may determine more information regarding vehicle 2510. For example, at step 2555, processing unit 110 may determine whether target vehicle 2330 has exceeded a collision threshold. A collision threshold may be, for example, a minimum predetermined distance between vehicle 200 and encroaching vehicle 2510. If encroaching vehicle 2510 comes closer to vehicle 200 than the minimum predetermined distance, processing unit 110 may determine that encroaching vehicle 2510 is indeed encroaching and execute instructions to avoid collision with encroaching vehicle 2510.

[0292] Additionally or alternatively, a collision threshold may be a time to collision. The time-to-collision threshold may be determined based on the distance between vehicle 200 and encroaching vehicle 2510, the velocity of vehicle 200 and/or encroaching vehicle 2510, the lateral velocity of vehicle 200 and/or encroaching vehicle 2510, and/or the acceleration of vehicle 200 and/or encroaching vehicle 2510. A collision threshold based on a time to collision may be set to allow enough time for processing unit 110 and system 100 to effectively conduct evasive action to avoid collision.

[0293] At step 2760, processing unit 110 may cause vehicle 200 to maintain current velocity and to travel within first lane constraint 2502 and second lane constraint 2504. To cause vehicle 200 to operate in this manner, processing unit 110 may execute action module 2630. Additionally or alternatively, step 2760 may include executing instructions on action module 1230 to cause processing unit 110 to transmit electronic

Tesla EX1002 Page 1125 IPR2025-00944

signals to the accelerator 2610, brakes 2620, and/or steering system of vehicle 200 to cause vehicle 200 to maintain current velocity and to travel within first and second lane constraints 810 and 820.

[0294] If, at step 2755, processing unit 110 determined that a collision threshold has been exceeded, the processing unit 110 may take evasive action. The evasive action may include causing vehicle 200 to change lanes. This may be a desirable response if, for example, vehicle 200 is traveling on a multilane highway and system 100 determines that changing lanes to move away from encroaching vehicle 2510 will avoid a collision. Evasive action may include vehicle 200 altering its course. Additionally or alternatively, evasive action may include altering course of vehicle 200. For example, processing unit 110 may undertake evasive action by transmitting electronic signals to throttling system 220, braking system 230, and/or steering system 240 of vehicle 200. After the evasive action is completed, at step 2780, processing unit 110 may control vehicle 200 to resume its prior navigation (e.g., to resume navigation to a desired location).

Multi-Threshold Reaction Zone for Vehicle Navigation

[0295] System 100 may provide driver assist functionality that causes a response in vehicle 200, such as braking, accelerating, switching lanes, turning, and/or other navigational responses. For example, if system 100 detects the presence of another vehicle ahead of vehicle 200 that is traveling at a rate of speed slower than vehicle 200, system 100 may cause vehicle 200 to reduce its speed by braking. System 100 may apply a braking profile that accounts for the immediacy of the risk (e.g., the risk of collision) to provide the driver with a natural driving sensation.

[0296] System 100 may cause various different responses depending on the situation and a determined risk of collision with another vehicle, for example. System 100 may take no action in response to changes in speed or direction of a target vehicle ahead if the target vehicle is sufficiently far ahead and the determined risk of collision is low. Where the target vehicle may be closer or when the risk of collision rises above a predetermined level, system 100 may cause a response, such as braking, etc. When the risk of collision is even higher (e.g., where system 100 determines that a collision would be imminent without evasive action), system 100 may cause vehicle 100 to take an evasive action, such as maximum or near maximum braking, change of direction, etc.

[0297] FIGS. 28A and 28B are diagrammatic representations of a primary vehicle 200a on a road 2800, consistent with disclosed embodiments. In some instances, road 2800 may be divided into two lanes, lane 2810 and lane 2820. A target object, such as target vehicle 200b, may be traveling ahead of primary vehicle 200a in the same lane (e.g., lane 2820 as shown in FIGS. 28A and 28B) or in a different lane. As shown in the example depicted in FIG. 28A , primary vehicle 200a and target vehicle 200b may be separated by a distance d1. At a given time after the scene depicted in FIG. 28A , primary vehicle 200a and target vehicle 200b may be separated by a distance d2as shown in FIG. 28B , where d2is less than d1. For example, primary vehicle 200a may be traveling at a rate of speed greater than target vehicle 200b, causing the distance between the vehicles to decrease from d1to d2.

[0298] As another example, primary vehicle 200a and target vehicle 200b may be traveling at the same rate of speed until target vehicle 200b reduces its speed by

Tesla EX1002 Page 1126 IPR2025-00944

braking, or primary vehicle 200a increases its speed by accelerating, causing the distance between the vehicles to decrease from d1to d2. As another example, target vehicle 200b may encroach from a side of primary vehicle 200a. In some circumstances, if target vehicle 200b is ahead or a sufficient distance away from primary vehicle 200a (e.g., five meters or more, 10 meters or more, etc.) and moves toward primary vehicle 200a, system 100 may not take any action. However, if target vehicle 200b continues to move closer to primary vehicle 200a, then system 100 may cause primary vehicle 200a to brake, change speed, and/or change lanes. For example, if target vehicle 200b is within a predetermined threshold (e.g., within five meters), primary vehicle 200a may then take action. Different scenarios involving the relative speeds and positions of primary vehicle 200a and target vehicle 200b are possible, and the disclosed embodiments are not limited to any particular scenario.

[0299] In some embodiments, the multi-threshold reaction zone may provide a natural feeling driving experience to the user. For example, as discussed above, if the risk of collision is low, then system 100 may disregard movements by target vehicle 200b. As a result, primary vehicle 200a may not respond to movements or actions of target vehicle 200b when there is not a risk of a collision (e.g., primary vehicle 200a would not need to mimic every motion of target vehicle 200b when it's 100 meters ahead). However, if target vehicle 200b is close to primary vehicle 200a and/or if system 100 determines that the risk of collision is high, then a responsive maneuver may be taken such that the user feels safer by increasing space between primary vehicle 200a and target vehicle 200b. If system 100 determines that a crash may be imminent, then system 100 may take a more dramatic action without causing the user to be surprised, because the action may be necessary to avoid a collision.

[0300] Accordingly, in some embodiments, system 100 may associate the different reaction zones with different degrees of responsive actions. For example, in some embodiments, system 100 may take no action when target vehicle 200b is within a first threshold (e.g., when target vehicle 200b is far away from primary vehicle 200a), system 100 may take a medium action with target vehicle 200b is within a second threshold (e.g., when target vehicle 200b is becoming closer to primary vehicle 200a), and system 100 may take a more dramatic action when target vehicle 200 is within a third threshold (e.g., when target vehicle 200b is sufficiently close to primary vehicle 200a that there is a risk of a collision). In a first reaction zone associated with the first threshold, system 100 may take no action when the estimated distance-to-target associated with target vehicle 200b is greater than twice a predetermined safety distance. In some embodiments, the first reaction zone may apply where the estimated distance-to-target associated with target vehicle 200b falls in the range of 1.5 to 2.5 times the predetermined safety distance.

[0301] The estimated distance-to-target may be the current distance between primary vehicle 200a and target vehicle 200b plus the expected change of this distance over a prediction time (e.g., in the range of 0.1 to 1.0 seconds, depending on factors associated with system 100 such as velocity tracking control loops, actuators, and vehicle dynamics). The expected change of distance may be calculated as: (vtarget-vprimary)*tprediction+ ((atarget-aprimary)*tprediction<1/2>), wherevtargetandvprimarycorrespond to the speed of target vehicle 200b and primary vehicle 200a, respectively,atargetandaprimarycorrespond to the acceleration of target vehicle 200b and primary vehicle 200a, respectively, andtpredictioncorresponds to the prediction time. Choosing a prediction time approximating the overall reaction time of system 100 may account for the fact that desired navigational responses for primary

vehicle 200a may not be implemented immediately, and thereby may provide the driver and/or passengers of primary vehicle 200a with a smooth, natural driving experience.

[0302] The predetermined safety distance may be defined as: max {dstatic , ddynamic}.dstatic ,which may fall in the range of 2 to 10 meters, may represent a desired safety distance between primary vehicle 200a and another vehicle (such as target vehicle 200b) while the vehicles are stopping and/or moving at a low rate of speed.ddynamicmay represent a desired safety distance when the vehicles are moving at a speed greater than a low rate of speed, and may be calculated as:tsafety*min (vtarget , vprimary), wherevtargetandvprimarycorrespond to the speed of target vehicle 200b and primary vehicle 200a, respectively.tsafetymay fall in the range of 0.5 to 2.5 seconds and may be adjusted by the driver of primary vehicle 200a via user interface 170 according to the driver's preferences. Thus, the driver may be able to control the distance that system 100 maintains between primary vehicle 200a and other vehicles.

[0303] In a second reaction zone associated with the second threshold, system 100 may place target vehicle 200b in a safety zone and apply different weights to different maneuvers performed by target vehicle 200b. For example, as target vehicle 200b may approach a lower boundary of the safety zone, e.g., the estimated distance-to-target associated with target vehicle 200b may approach the predetermined safety distance (discussed above in connection with the first reaction zone), system 100 may assign more weight to deceleration maneuvers (e.g., mimicked more) and less weight to acceleration maneuvers (e.g., mimicked less). In some embodiments, the lower boundary may fall in the range of 0.5 to 1.5 times the predetermined safety distance, depending on driver preferences. Conversely, as target vehicle 200b approaches an upper boundary of the safety zone, e.g., the estimated distance-to-target associated with target vehicle 200b may approach twice the predetermined safety distance, system 100 may assign less weight to deceleration maneuvers (e.g., mimicked less) and more weight to acceleration maneuvers (e.g., mimicked more). In some embodiments, the lower boundary may fall in the range of 1.5 to 2.5 times the predetermined safety distance, depending on driver preferences.

[0304] In a third reaction zone associated with the third threshold, system 100 may place target vehicle 200b in a danger zone when the estimated distance-to-target associated with target vehicle 200b is less than the predetermined safety distance. In this scenario, system 100 may cause primary vehicle 200a to mimic (e.g., immediately) any deceleration maneuvers (while disregarding any acceleration maneuvers) performed by target vehicle 200b. In some embodiments, the first reaction zone may apply where the estimated distance-to-target associated with target vehicle 200b falls in the range of 0.5 to 1.5 times the predetermined safety distance, depending on driver preferences.

[0305] In some embodiments, system 100 may cause a response in primary vehicle 200a, such as braking. As indicated in FIG. 28B , for example, the distance between the primary vehicle 200a and target vehicle 200b may decrease from d1to d2. Depending on the distance d2and, by extension, the time before collision, system 100 may cause primary vehicle 200a to reduce its speed so as to avoid a collision with target vehicle 200b. In addition, system 100 may apply one or more braking profiles when causing primary vehicle 200a to reduce its speed, also based on, for example, the time before collision. Thus, if the time before collision (and/or the distance d2) exceeds a first predetermined threshold, system 100 may cause primary vehicle 200a to brake in a gradual manner, change lanes, and/or change speed, whereas if the time

Tesla EX1002 Page 1128 IPR2025-00944

before collision (and/or the distance d2) is below a second predetermined threshold, system 100 may cause primary vehicle 200a to brake, change lanes, and/or change speed in a more rapid manner.

[0306] FIG. 29 is an exemplary block diagram of a memory configured to store instructions for performing one or more operations, consistent with the disclosed embodiments. As shown in FIG. 29 , memory 140 may store a target monitoring module 2902, an intercept time module 2904, and an action response module 2906. The disclosed embodiments are not limited to any particular configuration of memory 140. Further, processing unit 110 may execute the instructions stored in any of modules 2902-2906 included in memory 140.

[0307] In one embodiment, target monitoring module 2902 may store software instructions (such as computer vision software) which, when executed by processing unit 110, identifies and monitors a target object (such as target vehicle 200b) appearing in a plurality of images representing an environment in the vicinity of primary vehicle 200. For purposes of this disclosure, the target object may also be referred to as target vehicle 200b, although the target object may be a pedestrian, road hazard (e.g., large debris), and the like. The plurality of images may be acquired by one or more of image capture devices 122, 124, and 126. Processing unit 110 may execute target monitoring module 2902 to identify and monitor target vehicle 200b by, for example, analyzing the plurality of images using the techniques described in connection with FIGS. 5A-5D , above. Thus, target monitoring module 2902 may track different variables, such as the motion (e.g., direction of travel and trajectory), speed, and acceleration associated with target vehicle 200b, as well as the distance between primary vehicle 200a and the target vehicle 200b. Such information may be tracked continuously or at predetermined intervals.

[0308] In one embodiment, intercept time module 2904 may store software instructions which, when executed by processing unit 110, determines an indicator of an intercept time between primary vehicle 200a and target vehicle 200b. The intercept time may refer to an amount of time before primary vehicle 200a makes contact (e.g., collides) with target vehicle 200b. Processing unit 110 may execute intercept time module 2904 to determine an indicator of the intercept time based on an analysis of images acquired by one or more of image capture devices 122, 124, and 126. For example, intercept time module 2904 may make the determination using a series of time-based observations, e.g., of the position, speed, and/or acceleration of target vehicle 200b (relative to primary vehicle 200a), such as Kalman filters or linear quadratic estimation (LQE). The Kalman filters may be based on a measurement of the scale of target vehicle 200b, where the scale measurement is proportional to the intercept time.

[0309] In one embodiment, action response module 2906 may store software instructions (such as computer vision software) which, when executed by processing unit 110, causes one or more responses in primary vehicle 200a. The responses may be based on the identification and monitoring of target vehicle 200b (e.g., performed via execution of target monitoring module 2902) and/or the indicator of an intercept time between primary vehicle 200a and target vehicle 200b (e.g., determined via execution of intercept time module 2904). Processing unit 110 may cause one or more responses in primary vehicle 200a by, for example, transmitting electronic signals (e.g., via a CAN bus) to throttling system 220, braking system 230, and/or steering system 240 of primary vehicle 200a to brake, accelerate, or trigger a turn, lane shift, and/or change in direction of primary vehicle 200a. Processing unit 110 may use extended Kalman

Tesla EX1002 Page 1129 IPR2025-00944

Filters to estimate data associated with target vehicle 200b and primary vehicle 200a, such as: speed, acceleration, turn rate, pitch angle, pitch rate, position and heading of target vehicle 200b relative to primary vehicle 200a, the size of target vehicle 200b, road geometry, and the like. For example, processing unit 110 may receive measurements from components on primary vehicle 200a, such as an odometer, an inertial measurement unit (IMU), image processor 190, and the like. Thus, processing unit 110 may calculate data associated with target vehicle 200b and/or primary vehicle 200a using state equations derived from kinematic equations of motion, rigidity constraints, and motion assumptions, where inputs to the equations may include measurements from components on primary vehicle 200a. Processing unit 110 may also use the techniques described in connection with FIGS. 4-7 , above, to cause one or more responses in primary vehicle 200a.

[0310] FIG. 30 is a flow chart showing an exemplary process 3000 for causing a response in a primary vehicle, consistent with disclosed embodiments. At step 3010, processing unit 110 may receive a plurality of images via data interface 128 between processing unit 110 and image acquisition unit 120. For instance, one or more cameras included in image acquisition unit 120 may capture a plurality of images of an area forward of primary vehicle 200a (or to the sides or rear of a vehicle, for example) and transmit them over a data connection (e.g., digital, wired, USB, wireless, Bluetooth, etc.) to processing unit 110.

[0311] At step 3020, processing unit 110 may execute target monitoring module 2902 to identify a target object (such as target vehicle 200b) within the plurality of images. For example, target monitoring module 2902 may detect the presence of vehicles in the images using the techniques described in connection with FIGS. 5A and 5B, above. Multiple vehicles (and/or objects) may be detected in the same lane or in a different lane as primary vehicle 200a, and at varying distances from primary vehicle 200a, based on image data, position data (e.g., GPS location information), map data, the yaw rate of primary vehicle 200a, lane or other road markings, speed data, and/or data from sensors included in primary vehicle 200a. In a scenario where multiple vehicles are detected, target monitoring module 2902 may determine target vehicle 200b to be the vehicle traveling in the same lane as primary vehicle 200a and/or located closest to primary vehicle 200a.

[0312] At step 3030, processing unit 110 may execute target monitoring module 2902 to monitor target vehicle 200b identified at step 3020. For example, target monitoring module 2902 may track information associated with target vehicle 200b, such as the motion (e.g., direction of travel and trajectory), speed, and acceleration, as well as the distance between primary vehicle 200a and the target vehicle 200b. Target monitoring module 2902 may track such information continuously, or at predetermined intervals, by analyzing the plurality of images using the techniques described in connection with FIGS. 5A-5D , above. At step 3040, processing unit 110 may execute intercept time module 2904 to determine an indicator of an intercept time between primary vehicle 200a and target vehicle 200b. Intercept time module 2904 may make the determination by tracking the position, speed, and/or acceleration of target vehicle 200b (relative to primary vehicle 200a) over one or more periods of time. Intercept time module 2904 may use the techniques described in connection with FIGS. 5A and 5B , above to track such information.

[0313] At step 3050, processing unit 110 may execute action response module 2906 to cause one or more responses in primary vehicle 200a. For example, processing unit

IPR2025-00944
Tesla EX1002 Page 1130

110 may transmit electronic signals (e.g., via a CAN bus) to throttling system 220, braking system 230, and/or steering system 240 of primary vehicle 200a to brake, accelerate, or trigger a turn, lane shift, and/or change in direction of primary vehicle 200a. Further, one or more actuators may control throttling system 220, braking system 230, and/or steering system 240. For example, processing unit 110 may transmit electronic signals that cause system 100 to physically depress the brake by a predetermined amount or ease partially off the accelerator of vehicle 200. Further, processing unit 110 may transmit electronic signals that cause system 100 to steer vehicle 200 in a particular direction. Such responses may be based on the monitoring of target vehicle 200b (performed at step 3030) and/or the indicator of an intercept time between primary vehicle 200a and target vehicle 200b (determined at step 3040). In some embodiments, the intercept time may be compared with one or more predetermined thresholds, e.g., a first and second intercept threshold. For example, the response may include an emergency avoidance action if the indicator of intercept time indicates that a target object is within a particular intercept threshold (e.g., within the second intercept threshold). The emergency avoidance action may include changing lanes and/or emergency braking.

[0314] The thresholds may be expressed as units of time. Thus, for example, if the indicator of intercept time indicates that target vehicle 200b is outside of the first threshold (e.g., the time before collision between primary vehicle 200a and target vehicle 200b exceeds the first threshold), then action response module 2906 may cause a response in primary vehicle 200a based on an averaged motion profile associated with target vehicle 200b. The averaged motion profile associated with target vehicle 200b may represent an averaged position and/or averaged speed of target vehicle 200b for a particular threshold. In some embodiments, if target vehicle 200b is a certain distance and/or traveling a certain speed, system 100 may elect to take no action. For example, if target vehicle 200b is ahead or a sufficient distance away from primary vehicle 200a (e.g., five meters, 10 meters, etc.), system 100 may not take any action unless target vehicle 200b becomes closer to primary vehicle 200.

[0315] If the indicator of intercept time indicates that target vehicle 200b is between the first and second thresholds (e.g., the time before collision between primary vehicle 200a and target vehicle 200b falls in between the first and second thresholds), then action response module 2906 may cause a response in primary vehicle 200a based on the actual motion associated with target vehicle 200b. The actual motion may represent a position and/or speed of target vehicle 200b (e.g., not averaged or delayed) If the indicator of intercept time indicates that target vehicle 200b is within the second threshold (e.g., the time before collision between primary vehicle 200a and target vehicle 200b is below the second threshold), then action response module 2906 may cause an emergency action in primary vehicle 200a (such as a lane change or emergency braking).

[0316] FIG. 31 is a flow chart showing an exemplary process 3100 for declining to cause a response in a primary vehicle, consistent with disclosed embodiments. At step 3110, processing unit 110 may receive a plurality of images via data interface 128 between processing unit 110 and image acquisition unit 120, as described in connection with step 3010 of FIG. 30 , above.

[0317] At step 3120, processing unit 110 may execute target monitoring module 2902 to locate a target object (e.g., target vehicle 200b) within the images received at step 3010 by analyzing the images, as described in connection with step 3020 of FIG. 30 ,

IPR2025-00944
Tesla EX1002 Page 1131

above. At step 3125, processing unit 110 may execute target monitoring module 2902 to determine whether the target object (e.g., target vehicle 200b) is traveling in a lane different from primary vehicle 200a, using the techniques described in connection with step 3030 of FIG. 30 , above. If target vehicle 200b is determined to be traveling in a lane different from primary vehicle 200a (step 3125, yes), processing unit 110 may decline to cause a response in primary vehicle 200a at step 3130. Alternatively, if system 100 determines that target vehicle 200b does not constitute an imminent risk of collision (e.g., target vehicle 200b a nonintersecting trajectory with primary vehicle 200a or a potential time to collision beyond a threshold such that the target object does not pose an imminent risk at the present time), processing unit 110 may decline to cause a response in primary vehicle 200a at step 3130. Otherwise (step 3125, no), process 3100 may conclude and no response occurs.

Stopping an Autonomously Driven Vehicle Using a Multi-Segment Braking Profile

[0318] System 100 may provide driver assist functionality that monitors and responds to environmental conditions to control the braking of vehicle 200. For instance, system 100 may detect, based on a plurality of images, an object that triggers the stopping of vehicle 200. For example, this object may include another vehicle in the vicinity of vehicle 200, a traffic light, and/or a traffic sign. System 100 may use a braking profile to slow and/or stop vehicle 200. According to some embodiments, the braking profile may include multiple segments, where deceleration is executed at different stages. For example, the braking profile may cause system 100 to stop vehicle 200 at a decreasing rate of deceleration.

[0319] By providing a braking profile that includes multiple segments, system 100 may cause vehicle 200 to operate in a manner that feels natural to a driver and/or passengers. For example, when a driver stops a vehicle at a traffic light, the brakes are typically not applied at a constant rate from the moment they are applied until the moment when the vehicle stops. Operating brakes in such a manner result in an abrupt stop at the end. Instead, the driver may apply the brakes gently at first, then with increasing pressure, and then as the vehicle reaches its target stopping location, the driver may decrease the level of braking until it reaches zero (or near zero) at the moment the car stops. Operating the vehicle in this manner may provide a natural stopping progression, i.e., the driver may cause-a progressive deceleration at first so the vehicle does not experience an abrupt initial deceleration and then at the end, so that vehicle does not come to an abrupt stop.

[0320] System 100 may provide autonomous vehicle navigation that may implement braking according to an approach that feels natural to the driver and/or passengers of vehicle 200. For example, in some embodiments, system 100 may cause one or more of image capture devices 122-126 to acquire a plurality of images of an area in a vicinity of vehicle 200. Processing device 110 may receive the receive the plurality of images via data interface 128 and identify, based on analysis of the plurality of images, a trigger (e.g., an object, a traffic light, and/or a traffic sign) for stopping vehicle 200. Based on the identified trigger, system 100 may cause vehicle 200 to stop according to a braking profile. In some embodiments, the braking profile may include a plurality of segments (e.g., including a first segment associated with a first deceleration rate, a second segment which includes a second deceleration rate less than the first deceleration rate, and a third segment in which a level of braking is decreased as a target stopping location is approached, as determined based on the analysis of the plurality of images.

IPR2025-00944
Tesla EX1002 Page 1132

[0321] FIG. 32A illustrates vehicle 200 traveling on a roadway 3210 that includes an object 3220. System 100 may identify a trigger (e.g., object 3220) for stopping vehicle 200 in respect to detection of the trigger. For example, system 100 may first slow vehicle 200 after the trigger is detected and bring vehicle 200 to a stop before reaching object 3220.

[0322] Processing unit 110 may be configured to determine a trigger for stopping vehicle 200 based on a plurality of images acquired by one or more of image capture devices 122-126. Processing unit 110 may receive the images from image capture devices 122-126 via data interface 128. According to some embodiments, the trigger may include object 3220. For example, object 3220 may be another vehicle on roadway 2810. According to some embodiments, processing unit 110 may identify a trigger if object 3220 is within a certain vicinity of vehicle 200. For example, identification of the trigger may depend on a relative distance between object 3220 and vehicle 200. Further, this identification may optionally consider the relative velocity and/or relative acceleration between object 3220 and vehicle 200.

[0323] Additionally or alternatively, the trigger may include a traffic sign, such as traffic sign 3240 shown in FIG. 32A . For example, traffic sign 3240 may be a stop sign and/or a speed limit sign. According to some embodiments, processing unit 110 may process the plurality of images to determine the content of traffic sign 3240 as part of determining whether a trigger exists. For example, processing unit 110 may determine that a trigger is present based on a speed limit specified by traffic sign 3240 if vehicle 200 is exceeding the speed limit. As another example, processing unit 110 may determine a trigger exists based on a traffic light 2830 if the yellow and/or red light is on.

[0324] Based on the plurality of images, processing unit 110 may determine a target stopping location 3250 at which vehicle 200 will stop once a braking profile is executed. Reaching target stopping location 3250 may be accompanied by a release in braking pressure of vehicle 200. Target stopping location 3250 may be based on the location of traffic light 3230 and/or traffic sign 3240 such that vehicle 200 complies with traffic light 3230 and/or traffic sign 3240. Alternatively, target stopping location 3250 may be based on object 3220 such that vehicle 200 stops before reaching object 3220. Accordingly, target stopping location 3250 may be based on the location of object 3220 to avoid a collision between vehicle 200 and object 3220. In some embodiments, target stopping location 3250 may be updated based on new information obtained from the plurality of images, such as newly identified targets, newly visible obstacles, newly visible signs, etc.

[0325] FIG. 32B illustrates an exemplary braking profile, consistent with disclosed embodiments. In the example shown in FIG. 32B , the braking profile includes four segments. However, the number of segments is exemplary and a braking profile consistent with the disclosed embodiments may include any appropriate number of segments (e.g., 2, 3, 4, 5, 6, etc., segments).

[0326] In segment 1, system 100 may prepare for a possible braking scenario. For example, in this zone, system 100 may receive information (e.g., map data and/or image data) about an upcoming intersection (e.g., an intersection with traffic lights). System 100 may gradually reduce the speed of vehicle 200 (e.g., the speed of vehicle 200 may be gradually reduced to about 70-50 km/hr, depending on the characteristics

Tesla EX1002 Page 1133 IPR2025-00944

of the intersection). System 100 may cause the speed adjustment, for example, approximately 100-200 meters before the junction at average deceleration around 0.2-0.4 m/sec<2>(with a maximum of about 0.5 m/sec<2>). By decreasing the speed of vehicle 200 in segment 1, system 100 may provide the driver and/or passengers of vehicle 200 with confidence that system 100 is aware of the approaching intersection.

[0327] In segment 2, system 100 may apply strong braking. In this segment, system 100 may receive information based on the analysis of image data that, for example, a traffic light is red and/or there is a stopped vehicle at the intersection in the line in which vehicle 200 is traveling. System 100 may cause vehicle 200 to experience significant speed reduction in this zone, while the distance to the stop line or stopped vehicle is still large. Doing so may provide the driver and/or passengers of vehicle 200 with a comfortable feeling that system 100 will have more than enough space to complete the braking maneuver. This zone may be approximately 30-100 meters away from the stop line, and the average deceleration may be about 1.5-2.5 m/sec<2>(with a maximum of about 3.5 m/sec<2>).

[0328] In segment 3, system 100 may make a moderate braking adjustment. For example, in this segment, system 100 may adjust the speed of vehicle 200 according to the remaining distance to the stop line or stopped vehicle and based on the current speed and deceleration of vehicle 200. By making this adjustment, system 100 may provide the driver and/or passengers of vehicle 200 with an indication that system 100 is releasing most of the braking power while the speed is still being reduced but slower than before. This zone may be approximately 5-30 meters away from the stop line or stopped vehicle, and the average deceleration may be about 0.5-1.5 m/sec<2>(with a maximum of about 2 m/sec<2>).

[0329] In segment 4, system 100 may make a small braking adjustment. In this segment, system 100 may close the remaining distance to the stop line or stopped vehicle at a very low speed gradually bringing it to zero. In this zone, the driver and/or passengers may feel that vehicle 200 is slowly sliding into its position (e.g., at the stop line or behind another vehicle). This zone may constitute approximately the last 5-7 meters to the stop line or stopped vehicle, and the average deceleration may be about 0.3-0.5 m/sec<2>(with a maximum of about 1 m/sec<2>).

[0330] FIG. 33 is an exemplary block diagram of memory 140 and/or 150, which may store instructions for performing one or more operations consistent with disclosed embodiments. As illustrated in FIG. 33 , memory 140 may store one or more modules for performing the trigger detection and responses described herein. For example, memory 140 may store a trigger identification module 3300 and an action response module 3310.

[0331] Trigger identification module 3300 may store instructions which, when executed by processing unit 110, may detect the presence and/or existence of a trigger for stopping vehicle 200. For example, trigger identification module 3300 may process the plurality of images received from at least one image capture device 122-124 to detect the presence of a trigger. As discussed above, this may include identifying object 3220, traffic light 3230, and/or traffic sign 3240.

[0332] Action response module 3310 may store instructions which, when executed by processing unit 110, may respond to the presence of a trigger. For example, action response module 3310 may execute control to decelerate and/or stop vehicle 200. For

Tesla EX1002 Page 1134 IPR2025-00944

example, action response module 3310 may execute a braking profile to decelerate and/or stop vehicle 200. The action response module 3310 may determining to execute one of a plurality of braking profiles based on a number of braking profiles based on, for example, environmental factors, the type of trigger detected, and the velocity of vehicle 200. To navigate vehicle 200 to a stop based on the detected trigger, processing unit 110 may transmit electronic signals to one or more of throttling system 220, braking system 230, and/or steering system 240 of vehicle 200. Further, one or more actuators may control throttling system 220, braking system 230, and/or steering system 240. For example, processing unit 110 may transmit electronic signals that cause system 100 to physically depress the brake by a predetermined amount.

[0333] Fig. 34 illustrates a process 3400 for navigating vehicle 200, consistent with disclosed embodiments. Process 3000 may identify a trigger (e.g., an object, vehicle, traffic light, traffic sign, etc.) for stopping vehicle 200 and execute a braking profile in response to the detection of the trigger. Process 3400 may detect various types of triggers determine which of a plurality of braking profiles to execute.

[0334] The braking profile may be based on the type of a detected trigger and/or other environmental conditions related to the detected trigger (e.g., a distance to an object). For example, in some embodiments, the braking profile may include a plurality of segments (e.g., three segments). In a three segment braking profile, a first segment of the braking profile may be associated with a first deceleration rate, a second segment of the braking profile may include a second deceleration rate less than the first deceleration rate, and a third segment of the braking profile may include causing a level of braking to be decreased as a target stopping location is approached, as determined based on an analysis of a plurality of images acquired by one or more of image capture devices 122-126. For example, in some embodiments, the first segment may cause vehicle 200 to progressively increase deceleration, the second segment may cause vehicle 200 to decelerate at a constant rate, and the third segment may cause vehicle 200 to progressively decrease its deceleration rate to zero (or near zero).

[0335] At step 3410, process 3400 may acquire, using at least one of image capture devices 122-126, a plurality of images of an area in the vicinity of vehicle 200. For example, processing unit 110 may receive the plurality of images may through data interface 128. Processing unit 110 may be configured to determine the presence of one or more objects and/or signs located on or near roadway 3210 based on the acquired images. For example, processing unit 110 may be configured to determine the presence of object 3220, traffic light 3230, and/or traffic sign 3240. Processing unit 110 may be further configured to read the content of traffic sign 3240. For example, processing unit 110 may be configured to differentiate between a speed limit sign and a stop sign. Further, processing unit 110 may be configured to process images to determine the speed limit based on traffic sign 3240.

[0336] At step 3420, process 300 may include a trigger for stopping vehicle 200. For example, processing unit 110 may identify a trigger based on the presence of object 3220 within the vicinity of vehicle 200. Additionally or alternatively, processing unit 110 may identify a trigger based on a current signal status of traffic light 3230. According to some embodiments, processing unit 110 may determine that traffic light 3230 constitutes a trigger if it is not green. According to some embodiments, processing unit 110 may determine that traffic sign 3240 constitutes a trigger if it is a stop sign, for example. In some embodiments, traffic sign 3240 may constitute a trigger if it is a speed limit sign and vehicle 200 is exceeding the posted speed limit. Further in some

IPR2025-00944
Tesla EX1002 Page 1135

embodiments, processing unit 110 may cause system 100 to provide an audible announcement identifying the trigger and/or indicating that vehicle 200 is braking or about to begin braking.

[0337] After processing unit 110 identifies the trigger, at step 3430, processing unit 110 may cause vehicle 200 to slow down and/or stop according to a braking profile. In some embodiments, the braking profile may comprise a number of profile segments. The different profile segments of a braking profile may be defined by a time period and/or a vehicle speed. For example, processing unit 110 may execute controls of vehicle 200 in accordance with a first profile segment for a predetermined period of time or until the speed of vehicle 200 reaches a target value. The target value may be a predetermined speed or it may be a percentage of the speed that vehicle 200 was traveling prior to execution of the braking profile. To implement the braking profile, processing unit 110 may transmit electronic signals to one or more of throttling system 220, braking system 230, and/or steering system 240 of vehicle 200. Further, one or more actuators may control throttling system 220, braking system 230, and/or steering system 240.

[0338] As discussed above, according to some embodiments, a profile segment of the braking profile may correspond to a deceleration rate. For example, a braking profile may include multiple segments, and the first segment may be associated with a first deceleration rate. The braking profile may further include a second segment associated with a second deceleration rate. The second deceleration rate may be less than first deceleration rate.

[0339] Further, as discussed above, according to some embodiments, braking profile may include a third profile segment in which a level of braking is decreased as a function of the distance between vehicle 200 and underlying factors that give rise to the trigger (e.g., object 3220, traffic light 3230, traffic sign 3240). For example, the braking profile may include decreasing the level of braking as vehicle 200 approaches target stopping location 3250. Further, reaching target stopping location 3250 may be accompanied by a release in braking pressure of vehicle 200.

[0340] The foregoing description has been presented for purposes of illustration. It is not exhaustive and is not limited to the precise forms or embodiments disclosed. Modifications and adaptations will be apparent to those skilled in the art from consideration of the specification and practice of the disclosed embodiments. Additionally, although aspects of the disclosed embodiments are described as being stored in memory, one skilled in the art will appreciate that these aspects can also be stored on other types of computer readable media, such as secondary storage devices, for example, hard disks or CD ROM, or other forms of RAM or ROM, USB media, DVD, Blu-ray, or other optical drive media.

[0341] Computer programs based on the written description and disclosed methods are within the skill of an experienced developer. The various programs or program modules can be created using any of the techniques known to one skilled in the art or can be designed in connection with existing software. For example, program sections or program modules can be designed in or by means of .Net Framework, Net Compact Framework (and related languages, such as Visual Basic, C, etc.), Java, C++, Objective-C, HTML, HTML/AJAX combinations, XML, or HTML with included Java applets.

IPR2025-00944
Tesla EX1002 Page 1136

Espacenet

Claims: EP3092599 (A1) — 2016-11-16

SYSTEMS AND METHODS FOR MIMICKING A LEADING VEHICLE

Claims of EP3092599 (B1)

A high quality text as facsimile in your desired language may be available amongst the following family members:

CN106030609 (A) US2015151725 (A1) WO2015083009 (A1) CN110920609 (A) CN111027420 (A) US2015151742 (A1) US2015151751 (A1) US2015151753 (A1) US2015153735 (A1) US2015367850 (A1) US2016052514 (A1) US2016167650 (A1) US2016229410 (A1) US2016347322 (A1) US2019248372 (A1) US2020094839 (A1) US2020108834 (A1) US2020331477 (A1) more

- Original claims
- Claims tree

The EPO does not accept any responsibility for the accuracy of data and information originating from other authorities than the EPO; in particular, the EPO does not guarantee that they are complete, up-to-date or fit for specific purposes.

1. A driver assist navigation system (100) for a primary vehicle (200a), the system comprising:
at least one image capture device (122, 124, 126) configured to acquire a plurality of images of an area in a vicinity of the primary vehicle;
a data interface (128); and
at least one processing device (110) configured to:
receive (1810) the plurality of images via the data interface (128);
determine (1820), from at least some of the plurality of images, a first lane constraint (810) on a first side (802) of the primary vehicle (200a);
determine (1830), from at least some of the plurality of images, a second lane constraint (820) on a second side (812) of the primary vehicle (200a) opposite to the first side (802) of the primary vehicle, wherein the first and second lane constraints define a lane within which the primary vehicle (200a) travels; and
send control signals to the control systems (220, 230, 240) of said primary vehicle for causing (1840) the primary vehicle (200a) to travel within the first and second lane constraints (810, 820);
wherein:
the first lane constraint includes a line marked on a road surface, an edge of a road surface, markers on a road surface, or is estimated based on a determined midpoint of a road surface width;
the second lane constraint includes a line marked on a road surface, an edge of a road surface, markers on a road surface, or is estimated based on a determined midpoint of

a road surface width; and
characterized in that
the at least one processing device (110) is further configured to:
locate (1850) in the plurality of images a leading vehicle (200b);
determine (1860), based on the plurality of images, at least one action taken by the leading vehicle (200b); and
send control signals to the control systems (220, 230, 240) for causing (1870) the primary vehicle (200a) to mimic the at least one action of the leading vehicle (200b);
wherein the at least one action taken by the leading vehicle (200b) includes performing a maneuver such that a first offset distance (c3) on a side of the leading vehicle (200b) adjacent to the first lane constraint (810) is different from a second offset distance (c4) on a side of the leading vehicle (200b) adjacent to the second lane constraint (820), and
wherein mimicking the at least one action includes the at least one processing device causing the primary vehicle to perform a maneuver such that a third offset distance (c1) on a side of the primary vehicle (200a) adjacent to the first lane constraint (810) is different from a fourth offset distance (c2) on a side of the primary vehicle (200a) adjacent to the second lane constraint (820), wherein the third offset distance (c1) equals to the first offset distance (c3) and the fourth offset distance (c2) equals to the second offset distance (c4).

2. The system of claim 1, wherein the at least one action taken by the leading vehicle (200b) further includes a turn, and mimicking the at least one action respectively includes the at least one processing device (110) causing the primary vehicle (200a) to execute a similar turn beginning at a location at which the leading vehicle (200b) initiated its turn.

3. The system of claim 1, wherein the at least one action taken by the leading vehicle further includes an acceleration and wherein mimicking the at least one action respectively includes the at least one processing device causing the primary vehicle to accelerate.

4. The system of claim 1, wherein the at least one action taken by the leading vehicle further includes braking wherein mimicking the at least one action respectively includes the at least one processing device causing the primary vehicle to brake.

5. The system of claim 1, wherein the at least one processing device (110) is further configured to:
determine (1925), based on the plurality of images or based on map data and position information, whether the leading vehicle (200b) is turning at an intersection; and
if the leading vehicle (200b) is determined to be turning at an intersection, decline (1930) to cause the primary vehicle to perform a turn similar to the turn of the leading vehicle (200b).

6. The system of claim 1, wherein the at least one processing device (110) is further configured to:
determine (2025), based on the plurality of images or based on map data and position information, whether the leading vehicle (200b) is turning at an intersection;
determine (2035) whether the leading vehicle (200b) is changing lanes within the intersection;
mimic the turn of the leading vehicle (200b) by causing the primary vehicle (200a) to execute a similar turn beginning at a location at which the leading vehicle (200b)

IPR2025-00944
Tesla EX1002 Page 1139

initiated its turn, if the leading vehicle (200b) is determined to be turning at an intersection and not changing lanes within the intersection; and
decline (2050) to cause the primary vehicle to perform a turn similar to the turn of the leading vehicle (200b) if the leading vehicle (200b) is determined to be turning at an intersection, but changing lanes within the intersection.

7. The system of claim 1, wherein:
the leading vehicle is a first leading vehicle;
at least one processing device (110) is further configured to: locate (2130) a second leading vehicle in the plurality of images; and
cause the primary vehicle (200a) to mimic the at least one action of the first leading vehicle if, based on the plurality of images, the first leading vehicle is determined to be continuing on a path having less of a turning radius than a path followed by the second leading vehicle.

8. The system of claim 1, wherein the at least one processing device (110) is further configured to:
determine (2230) position information for the leading vehicle (200b) and compare (2240) the position information to predetermined map data;
create (2250) a navigation history for the leading vehicle (200b) by tracking movements of the leading vehicle (200b) relative to the predetermined map data; and
determine (2260) whether to cause the primary vehicle (200a) to mimic the at least one action of the leading vehicle (200b) based on the navigation history of the leading vehicle (200b).

9. The system of claim 8, wherein the at least one processing device (110) is configured to:
cause the primary vehicle (200a) to perform a turn similar to a turn of the leading vehicle (200b), even where no turn is detected based on the plurality of images, when the navigation history indicates that the leading vehicle (200b) has successfully navigated at least one prior turn.

10. A primary vehicle for following a leading vehicle, the primary vehicle comprising the system of any preceding claim.

11. A method for navigating a primary vehicle (200a), the method comprising:
acquiring (1810), using at least one image capture device (122, 124, 126), a plurality of images of an area in a vicinity of the primary vehicle (200a);
determining (1820), from at least some of the plurality of images, a first lane constraint (810) on a first side (802) of the primary vehicle (200a);
determining (1830), from at least some of the plurality of images, a second lane constraint (820) on a second side (812) of the primary vehicle (200a) opposite to the first side (802) of the primary vehicle (200a), wherein the first and second lane constraints (810, 820) define a lane within which the primary vehicle (200a) travels; and
sending control signals to the control systems (220, 230, 240) of said primary vehicle for causing (1840) the primary vehicle (200a) to travel within the first and second lane constraints (810, 820); wherein:
the first lane constraint includes a line marked on a road surface, an edge of a road surface, markers on a road surface, or is estimated based on a determined midpoint of a road surface width;
the second lane constraint includes a line marked on a road surface, an edge of a road surface, markers on a road surface, or is estimated based on a determined midpoint of

Tesla EX1002 Page 1140 IPR2025-00944

a road surface width, and
characterized in that the method further comprises:
locating (1850) in the plurality of images a leading vehicle (200b);
determining (1860), based on the plurality of images, at least one action taken by the leading vehicle (200b); and
sending control signals to the control systems (220, 230, 240) for causing (1870) the primary vehicle (200a) to mimic the at least one action of the leading vehicle (200b);
wherein the at least one action taken by the leading vehicle (200b) includes performing a maneuver such that a first offset distance (c3) on a side of the leading vehicle (200b) adjacent to the first lane constraint (810) is different from a second offset distance (c4) on a side of the leading vehicle (200b) adjacent to the second lane constraint (820), and
wherein mimicking the at least one action includes the at least one processing device causing the primary vehicle to perform a maneuver such that a third offset distance (c1) on a side of the primary vehicle (200a) adjacent to the first lane constraint (810) is different from a fourth offset distance (c2) on a side of the primary vehicle (200a) adjacent to the second lane constraint (820), wherein the third offset distance (c1) equals to the first offset distance (c3) and the fourth offset distance (c2) equals to the second offset distance (c4).

12. The method of claim 11, further comprising:
determining (1925), based on the plurality of images or based on map data and position information, whether the leading vehicle (200b) is turning at an intersection; and
if the leading vehicle (200b) is determined to be turning at an intersection, declining (1930) to mimic the turn of the leading vehicle (200b).

13. The method of claim 11, further comprising:
determining (2025), based on the plurality of images or based on map data and position information, whether the leading vehicle (200b) is turning at an intersection;
determining (2035) whether the leading vehicle (200b) is changing lanes within the intersection;
mimicking the turn of the leading vehicle (200b) by causing the primary vehicle (200a) to execute a similar turn beginning at a location at which the leading vehicle (200b) initiated its turn, if the leading vehicle (200b) is determined to be turning at an intersection and not changing lanes within the intersection; and
declining (2050) to cause the primary vehicle to perform a turn similar to the turn of the leading vehicle (200b) if the leading vehicle (200b) is determined to be turning at an intersection, but changing lanes within the intersection.

14. The method of claim 11, wherein the leading vehicle is a first leading vehicle, the method further comprising:
locating (2130) a second leading vehicle in the plurality of images; and
causing the primary vehicle (200a) to mimic the at least one action of first leading vehicle if, based on the plurality of images, the first leading vehicle is determined to be continuing on a path having less of a turning radius than a path followed by the second leading vehicle.

15. The method of claim 11, further comprising to:
determining (2230) position information for the leading vehicle (200b) and compare (2240) the position information to predetermined map data;
creating (2250) a navigation history for the leading vehicle (200b) by tracking

movements of the leading vehicle (200b) relative to the predetermined map data; and determining (2260) whether to cause the primary vehicle (200a) to mimic the at least one action of the leading vehicle (200b) based on the navigation history of the leading vehicle (200b); and

optionally causing the primary vehicle (200a) to mimic a turn of the leading vehicle (200b), even where no turn is detected based on the plurality of images, when the navigation history indicates that the leading vehicle (200b) has successfully navigated at least one prior turn.

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Veröffentlichungsnummer:
(11) Publication number: **EP 3 092 599 A0**
(11) Numéro de publication:

Internationale Anmeldung veröffentlicht durch die Weltorganisation für geistiges Eigentum unter der Nummer:

WO 2015/083009 (Art. 153(3) EPÜ).

International application published by the World Intellectual Property Organization under number:

WO 2015/083009 (Art. 153(3) EPC).

Demande internationale publiée par l'Organisation Mondiale de la Propriété Intellectuelle sous le numéro:

WO 2015/083009 (art. 153(3) CBE).

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 3 092 599 B1

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention of the grant of the patent:
06.03.2019 Bulletin 2019/10

(21) Application number: **14843184.4**

(22) Date of filing: **04.12.2014**

(51) Int Cl.:
G06K 9/00 *(2006.01)* ***B60K 31/00*** *(2006.01)*

(86) International application number:
PCT/IB2014/003151

(87) International publication number:
WO 2015/083009 (11.06.2015 Gazette 2015/23)

(54) **SYSTEMS AND METHODS FOR MIMICKING A LEADING VEHICLE**

SYSTEME UND VERFAHREN ZUM SIMULIEREN EINES FÜHRUNGSFAHRZEUGS

SYSTEMES ET PROCEDES POUR IMITER UN VEHICULE DE TETE

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR

(30) Priority: **04.12.2013 US 201361911490 P**
14.05.2014 US 201461993084 P
14.05.2014 US 201461993111 P
23.06.2014 US 201462015524 P
09.07.2014 US 201462022221 P
21.08.2014 US 201462040224 P
21.08.2014 US 201462040269 P

(43) Date of publication of application:
16.11.2016 Bulletin 2016/46

(73) Proprietor: **Mobileye Vision Technologies Ltd.**
91450 Jerusalem (IL)

(72) Inventors:
- **CLARKE, Anna**
  9083600 Har Adar (IL)
- **BAGON, Eyal**
  Jerusalem (IL)
- **NIR, Gali**
  Tel Aviv (IL)

(74) Representative: **Finnegan Europe LLP**
1 London Bridge
London SE1 9BG (GB)

(56) References cited:
US-A1- 2011 190 972 US-B1- 8 504 233

- **XIN LIU ET AL: "Vision-based long-distance lane perception and front vehicle location for full autonomous vehicles on highway roads", JOURNAL OF CENTRAL SOUTH UNIVERSITY OF TECHNOLOGY ; SCIENCE & TECHNOLOGY OF MINING AND METALLURGY, CENTRAL SOUTH UNIVERSITY, HEIDELBERG, vol. 19, no. 5, 27 April 2012 (2012-04-27) , pages 1454-1465, XP035048188, ISSN: 1993-0666, DOI: 10.1007/S11771-012-1162-7**
- **NASSER KEHTARNAVAZ ET AL: "A Transportable Neural-Network Approach to Autonomous Vehicle Following", IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, IEEE SERVICE CENTER, PISCATAWAY, NJ, US, vol. 47, no. 2, 1 May 1998 (1998-05-01), XP011063694, ISSN: 0018-9545**
- **NASSER KEHTARNAVAZ ET AL: "VISUAL CONTROL OF AN AUTONOMOUS VEHICLE (BART)-THE VEHICLE-FOLLOWING PROBLEM", IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, IEEE SERVICE CENTER, PISCATAWAY, NJ, US, vol. 40, no. 3, 1 August 1991 (1991-08-01) , pages 654-662, XP000266040, ISSN: 0018-9545, DOI: 10.1109/25.97520**

Note: Within nine months of the publication of the mention of the grant of the European patent in the European Patent Bulletin, any person may give notice to the European Patent Office of opposition to that patent, in accordance with the Implementing Regulations. Notice of opposition shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

EP 3 092 599 B1

Printed by Jouve, 75001 PARIS (FR)

Description

BACKGROUND

I. Technical Field

[0001] The present disclosure relates generally to autonomous vehicle navigation and, more specifically, to systems and methods that use cameras to provide autonomous vehicle navigation features.

II. Background Information

[0002] As technology continues to advance, the goal of a fully autonomous vehicle that is capable of navigating on roadways is on the horizon. Primarily, an autonomous vehicle may be able to identify its environment and navigate without input from a human operator. Autonomous vehicles may also take into account a variety of factors and make appropriate decisions based on those factors to safely and accurately reach an intended destination. For example, various objects-such as other vehicles and pedestrians-are encountered when a vehicle typically travels a roadway. Autonomous driving systems may recognize these objects in a vehicle's environment and take appropriate and timely action to avoid collisions. Additionally, autonomous driving systems may identify other indicators-such as traffic signals, traffic signs, and lane markings-that regulate vehicle movement (e.g., when the vehicle must stop and may go, a speed at which the vehicle must not exceed, where the vehicle must be positioned on the roadway, etc.). Autonomous driving systems may need to determine when a vehicle should change lanes, turn at intersections, change roadways, etc. As is evident from these examples, many factors may need to be addressed in order to provide an autonomous vehicle that is capable of navigating safely and accurately.

[0003] "Vision-based Long-distance Lane Perception and Front Vehicle Location for Full Autonomous Vehicles on Highway Roads" by Xin Liu, et al is directed to a vision-based long-distance lane perception and front location method for decision making by full autonomous vehicles on highway roads.

[0004] "A Transportable Neural-Network Approach to Autonomous Vehicle Following" by Nasser Kehtarnavaz, et al is directed to utilizing neural networks to control the speed and steering of a vehicle for the purpose of performing autonomous vehicle following, while following a leading vehicle.

[0005] US 2011/0190972 A1 is directed to systems for detecting the presence of stationary and non-stationary objects in the vicinity of a travelling vehicle, and controlling vehicle operational parameters in response to presence of such objects.

[0006] US 8,504,233 B1 is directed to a method for controlling a vehicle in an autonomous mode, which includes obtaining lane information that provides an estimated location of a lane on a road on which the vehicle is travelling and, if the lane information has become unavailable or unreliable, using sensor(s) to monitor neighboring vehicle(s) and controlling the vehicle to maintain at least a predetermined minimum distance between the vehicle and the neighboring vehicle(s).

[0007] JP2012-168592 is directed to a method for environment recognition of a vehicle. It comprises recognizing an arrow signal of a traffic light which is located far from the vehicle. It can also comprise performing a preceding vehicle detection and means to detect lanes detection or the road surface.

SUMMARY

[0008] Embodiments consistent with the present disclosure provide systems and methods for autonomous vehicle navigation. The disclosed embodiments may use cameras to provide autonomous vehicle navigation features. For example, consistent with the disclosed embodiments, the disclosed systems may include one, two, or more cameras that monitor the environment of a vehicle and cause a navigational response based on an analysis of images captured by one or more of the cameras.

[0009] In accordance with aspects of the invention there are provided a driver assist navigation system, a vehicle, and a method for navigating a vehicle as set out in claims 1, 10, and 11 respectively.

[0010] Consistent with a disclosed embodiment, a driver assist navigation system is provided for a vehicle. The system may include at least one image capture device configured to acquire a plurality of images of an area in a vicinity of the vehicle; a data interface; and at least one processing device. The at least one processing device may be configured to: receive the plurality of images via the data interface; determine from the plurality of images a first lane constraint on a first side of the vehicle; determine from the plurality of images a second lane constraint on a second side of the vehicle opposite to the first side of the vehicle, wherein the first and second lane constraints define a lane within which the vehicle travels and wherein a first distance corresponds to a distance between the first side of the vehicle and the first lane constraint and a second distance corresponds to a distance between the second side of the vehicle and the second lane constraint; determine, based on the plurality of images, whether a lane offset condition exists on the first side of the vehicle; if a lane offset condition exists on the first side of the vehicle, cause the vehicle to travel within the first and second lane constraints such that the first distance is greater than the second distance; determine, based on the plurality of images, whether a lane offset condition exists on the second side of the vehicle; and if a lane offset condition exists on the second side of the vehicle, cause the vehicle to travel within the first and second lane constraints such that the first distance is less than the second distance.

[0011] Consistent with the disclosed embodiment, the

IPR2025-00944
Tesla EX1002 Page 1145

at least one processing device may be further configured to cause the first distance to be substantially the same as the second distance if, based on the plurality of images, a lane offset condition is determined to exist on both the first and second sides of the vehicle. The at least one processing device may be configured to determine, based on the plurality of images, whether a lane offset condition is substantially non-existent, and if the lane offset condition is substantially non-existent, cause the vehicle to travel within the first and second lane constraints such that the first distance is substantially equal to the second distance. If the driver assist navigation system determines that a lane offset condition exists, the at least one processing device may be configured to cause the first distance to be at least 1.25 times greater than the second distance. Alternatively, the at least one processing device may be further configured to cause the first distance to be at least 2 times greater than the second distance. If the driver assist navigation system determines that the lane offset condition includes a curved road, the at least one processing device may be configured to cause the vehicle to travel within the first and second lane constraints such that the first distance is toward an outside of the curve relative to the vehicle, and the second distance is toward an inside of the curve relative to the vehicle. If the driver assist navigation system determines that the lane offset condition includes a presence of one or more objects on one side of the lane, the at least one processing device may be configured to cause the vehicle to travel within the first and second lane constraints such that the first distance, which is greater than the second distance, is provided toward the one or more objects. The one or more objects may include parked cars, a wall, and/or pedestrians, and may have a height of greater than 10 cm from a road surface. At least one of the first and second lane constraints may include a line marked on a road surface, an edge of a road surface, or a determined midpoint of a road surface width. Determination of whether a lane offset condition exists on the first side of the vehicle may be performed based on a position of a target object on the first side of the vehicle. The causing of the vehicle to travel within the first and second lane constraints such that the first distance is greater than the second distance may be overridden based on input from a driver of the vehicle.

[0012] Consistent with another disclosed embodiment, a vehicle may include a first vehicle side; a second vehicle side opposite the first vehicle side; at least one image capture device configured to acquire a plurality of images of an area in a vicinity of the vehicle; a data interface; and at least one processing device. The at least one processing device may be configured to: receive the plurality of images via the data interface; determine from the plurality of images a first lane constraint on the first vehicle side; determine from the plurality of images a second lane constraint on the second vehicle side, wherein the first and second lane constraints define a lane within which the vehicle travels and wherein a first distance corresponds to a distance between the first vehicle side and the first lane constraint and a second distance corresponds to a distance between the second vehicle side and the second lane constraint; determine, based on the plurality of images, whether a lane offset condition exists on the first vehicle side; if a lane offset condition exists on the first vehicle side, cause the vehicle to travel within the first and second lane constraints such that the first distance is greater than the second distance; determine, based on the plurality of images, whether a lane offset condition exists on the second vehicle side; and if a lane offset condition exists on the second vehicle side, cause the vehicle to travel within the first and second lane constraints such that the first distance is less than the second distance.

[0013] Consistent with the disclosed embodiment, the at least one processing device may be further configured to cause the first distance to be substantially the same as the second distance if, based on the plurality of images, a lane offset condition is determined to exist on both the first and second vehicle sides. The at least one processing device may be configured to determine, based on the plurality of images, whether a lane offset condition is substantially non-existent, and if the lane offset condition is substantially non-existent, cause the vehicle to travel within the first and second lane constraints such that the first distance is substantially equal to the second distance. The lane offset condition may include a curved road, and the at least one processing device may be configured to cause the vehicle to travel within the first and second lane constraints such that the first distance is toward an outside of the curve relative to the vehicle, and the second distance is toward an inside of the curve relative to the vehicle. The lane offset condition may include a presence of one or more objects on one side of the lane, and the at least one processing device may be configured to cause the vehicle to travel within the first and second lane constraints such that the first distance, which is greater than the second distance, is provided toward the one or more objects. At least one of the first and second lane constraints may include an edge of a road surface and/or a determined midpoint of a road surface width. Determination of whether a lane offset condition exists on the first side of the vehicle may be performed based on a position of a target object on the first side of the vehicle. The causing of the vehicle to travel within the first and second lane constraints such that the first distance is greater than the second distance may be overridden based on input from a driver of the vehicle.

[0014] Consistent with another disclosed embodiment, a method is provided for navigating a vehicle. The method may include acquiring, using at least one image capture device, a plurality of images of an area in the vicinity of the vehicle; determining from the plurality of images a first lane constraint on a first side of the vehicle; determining from the plurality of images a second lane constraint on a second side of the vehicle opposite to the

IPR2025-00944
Tesla EX1002 Page 1146

first side of the vehicle, wherein the first and second lane constraints define a lane within which the vehicle travels and wherein a first distance corresponds to a distance between the first side of the vehicle and the first lane constraint and a second distance corresponds to a distance between the second side of the vehicle and the second lane constraint; determining, based on the plurality of images, whether a lane offset condition exists on the first side of the vehicle; if a lane offset condition exists on the first side of the vehicle, causing the vehicle to travel within the first and second lane constraints such that the first distance is greater than the second distance; determining, based on the plurality of images, whether a lane offset condition exists on the second side of the vehicle; and if a lane offset condition exists on the second side of the vehicle, causing the vehicle to travel within the first and second lane constraints such that the first distance is less than the second distance.

[0015] Consistent with a disclosed embodiment, a driver assist navigation system is provided for a vehicle. The system may include at least one image capture device configured to acquire a plurality of images of an area in a vicinity of the vehicle; a data interface; and at least one processing device. The at least processing device may be configured to: receive the plurality of images via the data interface; determine from the plurality of images a current lane of travel from among a plurality of available travel lanes; and cause the vehicle to change lanes if the current lane of travel is not the same as a predetermined default travel lane.

[0016] Consistent with the disclosed embodiment, the predetermined default travel lane may be a right-most lane among the plurality of available travel lanes. The at least one processing device may be configured to activate a turn signal prior to causing the vehicle to change lanes. The at least one processing device may be configured to cause an audible announcement prior to causing the vehicle to change lanes. The at least one processing device may be configured to cause an audible indicator prior to causing the vehicle to change lanes. The predetermined default travel lane may be determined based on the plurality of images of the area in the vicinity of the vehicle, based on data received from a global positioning system, and/or based on an input received from a user.

[0017] Consistent with another disclosed embodiment, a vehicle may include a body; at least one image capture device configured to acquire a plurality of images of an area in a vicinity of the vehicle; a data interface; and at least one processing device. The at least one processing device may be configured to: receive the plurality of images via the data interface; determine from the plurality of images a current lane of travel from among a plurality of available travel lanes; and cause the vehicle to change lanes if the current lane of travel is not the same as a predetermined default travel lane.

[0018] Consistent with the disclosed embodiment, the predetermined default travel lane may be a right-most lane among the plurality of available travel lanes. The at least one processing device may be configured to activate a turn signal prior to causing the vehicle to change lanes. The at least one processing device may be configured to cause an audible announcement prior to causing the vehicle to change lanes. The at least one processing device may be configured to cause an audible indicator prior to causing the vehicle to change lanes. The predetermined default travel lane may be determined based on the plurality of images of the area in the vicinity of the vehicle, based on data received from a global positioning system, and/or based on an input received from a user.

[0019] Consistent with another disclosed embodiment, a method is provided for navigating a vehicle. The method may include acquiring, using at least one image capture device, a plurality of images of an area in a vicinity of the vehicle; determining from the plurality of images a current lane of travel from among a plurality of available travel lanes; and causing the vehicle to change lanes if the current lane of travel is not the same as a predetermined default travel lane.

[0020] Consistent with the disclosed embodiment, the predetermined default travel lane may be a right-most lane among the plurality of available travel lanes. A turn signal may be activated prior to causing the vehicle to change lanes. The predetermined default travel lane may be determined based on the plurality of images of the area in the vicinity of the vehicle.

[0021] Consistent with a disclosed embodiment, a driver assist navigation system is provided for a vehicle. The system may include at least one image capture device configured to acquire a plurality of images of an area in a vicinity of the vehicle; a data interface; and at least one processing device. The at least one processing device may be configured to: receive the plurality of images via the data interface; recognize a curve to be navigated based on map data and vehicle position information; determine an initial target velocity for the vehicle based on at least one characteristic of the curve as reflected in the map data; adjust a velocity of the vehicle to the initial target velocity; determine, based on the plurality of images, one or more observed characteristics of the curve; determine an updated target velocity based on the one or more observed characteristics of the curve; and adjust the velocity of the vehicle to the updated target velocity.

[0022] Consistent with the disclosed embodiment, the at least one characteristic of the curve may include an expected radius of curvature for the curve as determined from the map data. The initial target velocity may be determined based on the expected radius of curvature for the curve and a lateral acceleration limit. The lateral acceleration limit may be between about 0.2 m/sec^2 and 3.0 m/sec^2. Alternatively, the lateral acceleration limit may be between about 1.5 m/sec^2 and 3.0 m/sec^2. The lateral acceleration limit may be adjusted based on a road width determined based on the plurality of images. The one or more observed characteristics of the curve may

IPR2025-00944
Tesla EX1002 Page 1147

include an actual radius of curvature for the curve as determined through analysis of the plurality of images. The updated target velocity may be determined based on the actual radius of curvature for the curve and a lateral acceleration limit. The one or more observed characteristics of the curve may include a rate of change in an actual radius of curvature for the curve as determined through analysis of the plurality of images. The one or more observed characteristics of the curve may include a degree of banking associated with the curve as determined through analysis of the plurality of images. The one or more observed characteristics of the curve may include information, determined through analysis of the plurality of images, relating to the curve conveyed by one or more road signs. The initial target velocity may also be based on a focal range of the at least one image capture device. The initial target velocity may also be based on whether a traffic light is determined to be located along the curve. The initial target velocity may also be based on whether a traffic light is determined to be located within a predetermined distance after the curve.

[0023] Consistent with another disclosed embodiment, a vehicle may include a body; at least one image capture device configured to acquire a plurality of images of an area in a vicinity of the vehicle; a data interface; and at least one processing device. The at least one processing device may be configured to: receive the plurality of images via the data interface; recognize a curve to be navigated based on map data and vehicle position information; determine an initial target velocity for the vehicle based on at least one characteristic of the curve as reflected in the map data; adjust a velocity of the vehicle to the initial target velocity; determine, based on the plurality of images, one or more observed characteristics of the curve; determine an updated target velocity based on the one or more observed characteristics of the curve; and adjust the velocity of the vehicle to the updated target velocity.

[0024] Consistent with the disclosed embodiment, the at least one characteristic of the curve may include an expected radius of curvature for the curve as determined from the map data. The initial target velocity may be determined based on the expected radius of curvature for the curve and a lateral acceleration limit. The lateral acceleration limit may be adjusted based on a road width determined based on the plurality of images. The one or more observed characteristics of the curve may include an actual radius of curvature for the curve as determined through analysis of the plurality of images. The updated target velocity may be determined based on the actual radius of curvature for the curve and a lateral acceleration limit. The one or more observed characteristics of the curve may include a rate of change in an actual radius of curvature for the curve as determined through analysis of the plurality of images. The one or more observed characteristics of the curve may include a degree of banking associated with the curve as determined through analysis of the plurality of images. The one or more observed characteristics of the curve may include information, determined through analysis of the plurality of images, relating to the curve conveyed by one or more road signs. The initial target velocity may also be based on a focal range of the at least one image capture device. The initial target velocity may also be based on whether a traffic light is determined to be located along the curve. The initial target velocity may also be based on whether a traffic light is determined to be located within a predetermined distance after the curve.

[0025] Consistent with another disclosed embodiment, a method is provided for navigating a vehicle. The method may include acquiring, using at least one image capture device, a plurality of images of an area in a vicinity of the vehicle; recognizing a curve to be navigated based on map data and vehicle position information; determining an initial target velocity for the vehicle based on at least one characteristic of the curve as reflected in the map data; adjusting a velocity of the vehicle to the initial target velocity; determining, based on the plurality of images, one or more observed characteristics of the curve; determining an updated target velocity based on the one or more observed characteristics of the curve; and adjusting the velocity of the vehicle to the updated target velocity.

[0026] Consistent with a disclosed embodiment, a driver assist navigation system is provided for a primary vehicle. The system may include at least one image capture device configured to acquire a plurality of images of an area in a vicinity of the primary vehicle; a data interface; and at least one processing device. The at least one processing device may be configured to: receive the plurality of images via the data interface; determine, from at least some of the plurality of images, a first lane constraint on a first side of the primary vehicle; determine, from at least some of the plurality of images, a second lane constraint on a second side of the primary vehicle opposite to the first side of the primary vehicle, wherein the first and second lane constraints define a lane within which the primary vehicle travels; cause the primary vehicle to travel within the first and second lane constraints; locate in the plurality of images a leading vehicle; determine, based on the plurality of images, at least one action taken by the leading vehicle; and cause the primary vehicle to mimic the at least one action of the leading vehicle.

[0027] Consistent with the disclosed embodiment, at least one of the first and second lane constraints may include a line marked on a road surface. At least one of the first and second lane constraints may include an edge of a road surface. At least one of the first and second lane constraints may include a determined midpoint of a road surface width. The at least one action taken by the leading vehicle may include a turn, and the at least one processing device may cause the primary vehicle to execute a similar turn beginning at a location at which the leading vehicle initiated its turn. The at least one action taken by the leading vehicle may include an acceleration. The at least one action taken by the leading vehicle may include braking. The at least one action taken by the lead-

IPR2025-00944
Tesla EX1002 Page 1148

ing vehicle may include providing a first offset distance on a side of the leading vehicle adjacent to the first lane constraint that is different from a second offset distance on a side of the leading vehicle adjacent to the second lane constraint. The at least one processing device may be further configured to determine, based on the plurality of images or based on map data and position information, whether the leading vehicle is turning at an intersection, and if the leading vehicle is determined to be turning at an intersection, decline to mimic the turn of the leading vehicle. The at least one processing device may be configured to determine, based on the plurality of images or based on map data and position information, whether the leading vehicle is turning at an intersection, determine whether the leading vehicle is changing lanes within the intersection, mimic the turn of the leading vehicle if the leading vehicle is determined to be turning at an intersection and not changing lanes within the intersection, and decline to mimic the turn of the leading vehicle if the leading vehicle is determined to be turning at an intersection, but changing lanes within the intersection. The at least one processing device may be configured to acquire a first leading vehicle based on the plurality of images, acquire a second leading vehicle based on the plurality of images, and cause the primary vehicle to mimic the first leading vehicle if, based on the plurality of images, the first leading vehicle is determined to be continuing on a path having less of a turning radius than a path followed by the second leading vehicle. The at least one processing device may be configured to determine position information for the leading vehicle and compare the position information to predetermined map data, create a navigation history for the leading vehicle by tracking movements of the leading vehicle relative to the predetermined map data, and determine whether to cause the primary vehicle to mimic the at least one action of the leading vehicle based on the navigation history of the leading vehicle. The at least one processing device may be configured to cause the primary vehicle to mimic a turn of the leading vehicle, even where no turn is detected based on the plurality of images, when the navigation history indicates that the leading vehicle has successfully navigated at least one prior turn.

[0028] Consistent with another disclosed embodiment, a primary vehicle may include a body; at least one image capture device configured to acquire a plurality of images of an area in a vicinity of the primary vehicle; a data interface; and at least one processing device. The at least one processing device may be configured to: receive the plurality of images via the data interface; determine, from at least some of the plurality of images, a first lane constraint on a first side of the primary vehicle; determine, from at least some of the plurality of images, a second lane constraint on a second side of the primary vehicle opposite to the first side of the primary vehicle, wherein the first and second lane constraints define a lane within which the primary vehicle travels; cause the primary vehicle to travel within the first and second lane constraints; locate in the plurality of images a leading vehicle; determine, based on the plurality of images, at least one action taken by the leading vehicle; and cause the primary vehicle to mimic the at least one action of the leading vehicle.

[0029] Consistent with the disclosed embodiment, at least one of the first and second lane constraints may include a line marked on a road surface. At least one of the first and second lane constraints may include an edge of a road surface. At least one of the first and second lane constraints may include a determined midpoint of a road surface width. The at least one action taken by the leading vehicle may include a turn, and the at least one processing device may cause the primary vehicle to execute a similar turn beginning at a location at which the leading vehicle initiated its turn. The at least one action taken by the leading vehicle may include an acceleration. The at least one action taken by the leading vehicle may include braking. The at least one action taken by the leading vehicle may include providing a first offset distance on a side of the leading vehicle adjacent to the first lane constraint that is different from a second offset distance on a side of the leading vehicle adjacent to the second lane constraint. The at least one processing device may be further configured to determine, based on the plurality of images or based on map data and position information, whether the leading vehicle is turning at an intersection, and if the leading vehicle is determined to be turning at an intersection, decline to mimic the turn of the leading vehicle. The at least one processing device may be configured to determine, based on the plurality of images or based on map data and position information, whether the leading vehicle is turning at an intersection, determine whether the leading vehicle is changing lanes within the intersection, mimic the turn of the leading vehicle if the leading vehicle is determined to be turning at an intersection and not changing lanes within the intersection, and decline to mimic the turn of the leading vehicle if the leading vehicle is determined to be turning at an intersection, but changing lanes within the intersection. The at least one processing device may be configured to acquire a first leading vehicle based on the plurality of images, acquire a second leading vehicle based on the plurality of images, and cause the primary vehicle to mimic the first leading vehicle if, based on the plurality of images, the first leading vehicle is determined to be continuing on a path having less of a turning radius than a path followed by the second leading vehicle. The at least one processing device may be configured to determine position information for the leading vehicle and compare the position information to predetermined map data, create a navigation history for the leading vehicle by tracking movements of the leading vehicle relative to the predetermined map data, and determine whether to cause the primary vehicle to mimic the at least one action of the leading vehicle based on the navigation history of the leading vehicle. The at least one processing device may be configured to cause the primary vehicle to mimic a

IPR2025-00944
Tesla EX1002 Page 1149

turn of the leading vehicle, even where no turn is detected based on the plurality of images, when the navigation history indicates that the leading vehicle has successfully navigated at least one prior turn.

[0030] Consistent with another disclosed embodiment, a method is provided for navigating a primary vehicle. The method ma include acquiring, using at least one image capture device, a plurality of images of an area in a vicinity of the primary vehicle; determining, from at least some of the plurality of images, a first lane constraint on a first side of the primary vehicle; determining, from at least some of the plurality of images, a second lane constraint on a second side of the primary vehicle opposite to the first side of the primary vehicle, wherein the first and second lane constraints define a lane within which the primary vehicle travels; causing the primary vehicle to travel within the first and second lane constraints; locating in the plurality of images a leading vehicle; determining, based on the plurality of images, at least one action taken by the leading vehicle; and causing the primary vehicle to mimic the at least one action of the leading vehicle.

[0031] Consistent with a disclosed embodiment, a driver assist navigation system is provided for a user vehicle. They system may include at least one image capture device configured to acquire a plurality of images of an area in a vicinity of the user vehicle; a data interface; and at least one processing device. The at least one processing device may be configured to: receive the plurality of images via the data interface; determine from the plurality of images a first lane constraint on a first side of the user vehicle; determine from the plurality of images a second lane constraint on a second side of the user vehicle opposite to the first side of the user vehicle, wherein the first and second lane constraints define a lane within which the user vehicle travels; acquire, based on the plurality of images, a target vehicle; enable the user vehicle to pass the target vehicle if the target vehicle is determined to be in a lane different from the lane in which the user vehicle is traveling; monitor a position of the target vehicle based on the plurality of images; and cause the user vehicle to abort the pass before completion of the pass, if the target vehicle is determined to be entering the lane in which the user vehicle is traveling.

[0032] Consistent with the disclosed embodiment, causing the user vehicle to abort the pass may include causing an application of brakes in the user vehicle. Causing the user vehicle to abort the pass may include causing an audible announcement. Causing the user vehicle to abort the pass may include causing the user vehicle to stop accelerating. Monitoring the position of the target vehicle may include estimating a speed of the target vehicle. At least one of the first and second lane constraints may include a line marked on a road surface, an edge of a road surface, and/or a determined midpoint of a road surface width.

[0033] Consistent with another disclosed embodiment, a user vehicle may include at least one image capture device configured to acquire a plurality of images of an area in a vicinity of the user vehicle; a data interface; and at least one processing device. The at least one processing device may be configured to: receive the plurality of images via the data interface; determine from the plurality of images a first lane constraint on a first side of the user vehicle; determine from the plurality of images a second lane constraint on a second side of the user vehicle opposite to the first side of the user vehicle, wherein the first and second lane constraints define a lane within which the user vehicle travels; acquire, based on the plurality of images, a target vehicle; enable the user vehicle to pass the target vehicle if the target vehicle is determined to be in a lane different from the lane in which the user vehicle is traveling; monitor a position of the target vehicle based on the plurality of images; and cause the user vehicle to abort the pass before completion of the pass, if the target vehicle is determined to be entering the lane in which the user vehicle is traveling.

[0034] Consistent with the disclosed embodiment, causing the user vehicle to abort the pass may include causing an application of brakes in the user vehicle. Causing the user vehicle to abort the pass may include causing an audible announcement. Causing the user vehicle to abort the pass may include causing the user vehicle to stop accelerating. Monitoring the position of the target vehicle may include estimating a speed of the target vehicle. At least one of the first and second lane constraints may include a line marked on a road surface, an edge of a road surface, and/or a determined midpoint of a road surface width.

[0035] Consistent with another disclosed embodiment, a method is provided for navigating a user vehicle. The method may include acquiring, using at least one image capture device, a plurality of images of an area in a vicinity of the user vehicle; determining from the plurality of images a first lane constraint on a first side of the user vehicle; determining from the plurality of images a second lane constraint on a second side of the user vehicle opposite to the first side of the user vehicle, wherein the first and second lane constraints define a lane within which the user vehicle travels; acquiring, based on the plurality of images, a target vehicle; enabling the user vehicle to pass the target vehicle if the target vehicle is determined to be in a lane different from the lane in which the user vehicle is traveling; monitoring a position of the target vehicle based on the plurality of images; and causing the user vehicle to abort the pass before completion of the pass, if the target vehicle is determined to be entering the lane in which the user vehicle is traveling.

[0036] Consistent with the disclosed embodiment, causing the user vehicle to abort the pass may include causing an application of brakes in the user vehicle. Causing the user vehicle to abort the pass may include causing the user vehicle to stop accelerating. At least one of the first and second lane constraints may include a line marked on a road surface, an edge of a road surface, and/or a determined midpoint of a road surface

IPR2025-00944
Tesla EX1002 Page 1150

width.

[0037] Consistent with a disclosed embodiment, a driver assist navigation system is provided for a user vehicle. The system may include at least one image capture device configured to acquire a plurality of images of an area in a vicinity of the user vehicle; a data interface; and at least one processing device. The at least one processing device may be configured to: receive the plurality of images via the data interface; determine from the plurality of images a first lane constraint on a first side of the user vehicle; determine from the plurality of images a second lane constraint on a second side of the user vehicle opposite to the first side of the user vehicle, wherein the first and second lane constraints define a lane within which the user vehicle travels; determine, based on the plurality of images, whether an encroaching vehicle is approaching from a side of the user vehicle; and cause the user vehicle to maintain a current velocity and to travel within the first and second lane constraints such that a first distance, on the side of the user vehicle that the encroaching vehicle is approaching from, is greater than a second distance.

[0038] Consistent with the disclosed embodiment, the at least one processing device may be further configured to cause the user vehicle to take evasive action if analysis of the plurality of images indicates that the encroaching vehicle has crossed at least one collision threshold. The evasive action may include changing lanes, altering course, and/or braking. The at least one collision threshold may include a minimum predetermined distance to maintain between the user vehicle and the encroaching vehicle. The at least one collision threshold may include a time to collision. The at least one processing device may be configured to cause the user vehicle to return to a center of the lane, within which the user vehicle originally traveled, after completion of the evasive action. The at least one image capture device may have a field of view of at least 140 degrees. The at least one processing device may be configured to cause the user vehicle to resume a preset speed after braking in response to an encroaching vehicle. The at least one processing device may be configured to cause the user vehicle to return to a lane location within the first and second lane constraints such that the first distance between the user vehicle and the first lane constraint is substantially equal to the second distance between the user vehicle and the second lane constraint after altering course in response to an encroaching vehicle.

[0039] Consistent with another disclosed embodiment, a user vehicle may include at least one image capture device configured to acquire a plurality of images of an area in a vicinity of the user vehicle; a data interface; and at least one processing device. The at least one processing device may be configured to: receive the plurality of images via the data interface; determine from the plurality of images a first lane constraint on a first side of the user vehicle; determine from the plurality of images a second lane constraint on a second side of the user vehicle opposite to the first side of the user vehicle, wherein the first and second lane constraints define a lane within which the user vehicle travels; determine, based on the plurality of images, whether an encroaching vehicle is approaching from a side of the user vehicle; and cause the user vehicle to maintain a current velocity and to travel within the first and second lane constraints such that a first distance, on the side of the user vehicle that the encroaching vehicle is approaching from, is greater than a second distance.

[0040] Consistent with the disclosed embodiment, the at least one processing device may be further configured to cause the user vehicle to take evasive action if analysis of the plurality of images indicates that the encroaching vehicle has crossed at least one collision threshold. The evasive action may include changing lanes, altering course, and/or braking. The at least one collision threshold may include a minimum predetermined distance to maintain between the user vehicle and the encroaching vehicle. The at least one collision threshold may include a time to collision. The at least one processing device may be configured to cause the user vehicle to return to a center of the lane, within which the user vehicle originally traveled, after completion of the evasive action. The at least one image capture device may have a field of view of at least 140 degrees. The at least one processing device may be configured to cause the user vehicle to resume a preset speed after braking in response to an encroaching vehicle. The at least one processing device may be configured to cause the user vehicle to return to a lane location within the first and second lane constraints such that the first distance between the user vehicle and the first lane constraint is substantially equal to the second distance between the user vehicle and the second lane constraint after altering course in response to an encroaching vehicle.

[0041] Consistent with another disclosed embodiment, a method for navigating a user vehicle is provided. The method may include acquiring, using at least one image capture device, a plurality of images of an area in a vicinity of the user vehicle; determining from the plurality of images a first lane constraint on a first side of the user vehicle; determining from the plurality of images a second lane constraint on a second side of the user vehicle opposite to the first side of the user vehicle, wherein the first and second lane constraints define a lane within which the user vehicle travels; determining, based on the plurality of images, whether an encroaching vehicle is approaching from a side of the user vehicle; and causing the user vehicle to maintain a current velocity and to travel within the first and second lane constraints such that a first distance, on the side of the user vehicle that the encroaching vehicle is approaching from, is greater than a second distance.

[0042] Consistent with a disclosed embodiment, a driver assist navigation system is provided for a primary vehicle. They system may include at least one image capture device configured to acquire a plurality of images of

IPR2025-00944
Tesla EX1002 Page 1151

an area in a vicinity of the primary vehicle; a data interface; and at least one processing device. The at least processing device may be configured to: receive the plurality of images via the data interface; identify a target object within the plurality of images; monitor, via the plurality of images, a motion of the target object and a distance between the primary vehicle and the target object; determine an indicator of an intercept time between the primary vehicle and the target object based on the monitored motion and the distance between the primary vehicle and the target object; and cause a response in the primary vehicle based on a comparison of the intercept time to a plurality of predetermined intercept thresholds.

[0043] Consistent with the disclosed embodiment, the target object may be a target vehicle traveling in a lane different from the primary vehicle. The intercept time may be based on an indicator of speed of the tracked object, as determined based on the plurality of images. The intercept time may be based on an indicator of an acceleration of the tracked object, as determined based on the plurality of images. The at least one processing device may be further configured to determine whether the target object is traveling in the same lane as the primary vehicle, and forego the response in the primary vehicle if the target object is determined to be traveling in a lane different from the primary vehicle. Determination of whether the target object is traveling in the same lane as the primary vehicle may be based on the plurality of images. Determination of whether the target object is traveling in the same lane as the primary vehicle may be based on at least one of predetermined map data or yaw rate of the primary vehicle. The response may be based on an averaged motion profile associated with the target object if the indicator of intercept time indicates that the target object is outside a first intercept threshold relative to the vehicle. The response may be based on actual motion of the target object if the indicator of intercept time indicates that the target object is between the first intercept threshold and a second intercept threshold relative to the vehicle. The response may include an emergency avoidance action if the indicator of intercept time indicates that the target object is within the second intercept threshold. The emergency avoidance action may include changing lanes and/or emergency braking.

[0044] Consistent with another disclosed embodiment, a primary vehicle may include a body; at least one image capture device configured to acquire a plurality of images of an area in a vicinity of the primary vehicle; a data interface; and at least one processing device. The at least one processing device may be configured to: receive the plurality of images via the data interface; identify a target object within the plurality of images; monitor, via the plurality of images, a motion of the target object and a distance between the primary vehicle and the target object; determine an indicator of an intercept time between the primary vehicle and the target object based on the monitored motion and the distance between the primary vehicle and the target object; and cause a response in the primary vehicle based on a comparison of the intercept time to a plurality of predetermined intercept thresholds.

[0045] Consistent with the disclosed embodiment, the target object may be a target vehicle traveling in a lane different from the primary vehicle. The intercept time may be based on an indicator of speed of the tracked object, as determined based on the plurality of images. The intercept time may be based on an indicator of an acceleration of the tracked object, as determined based on the plurality of images. The at least one processing device may be further configured to determine whether the target object is traveling in the same lane as the primary vehicle, and forego the response in the primary vehicle if the target object is determined to be traveling in a lane different from the primary vehicle. Determination of whether the target object is traveling in the same lane as the primary vehicle may be based on the plurality of images. Determination of whether the target object is traveling in the same lane as the primary vehicle may be based on at least one of predetermined map data or yaw rate of the primary vehicle. The response may be based on an averaged motion profile associated with the target object if the indicator of intercept time indicates that the target object is outside a first intercept threshold relative to the vehicle. The response may be based on actual motion of the target object if the indicator of intercept time indicates that the target object is between the first intercept threshold and a second intercept threshold relative to the vehicle. The response may include an emergency avoidance action if the indicator of intercept time indicates that the target object is within the second intercept threshold. The emergency avoidance action may include changing lanes and/or emergency braking.

[0046] Consistent with another disclosed embodiment, a method for navigating a primary vehicle may include acquiring, via at least one image capture device, a plurality of images of an area in a vicinity of the primary vehicle; identifying a target object within the plurality of images; monitoring, based on the plurality of images, a motion of the target object and a distance between the primary vehicle and the target object; determining an indicator of an intercept time between the primary vehicle and the target object based on the monitored motion and the distance between the primary vehicle and the target object; and causing a response in the primary vehicle based on a comparison of the intercept time to a plurality of predetermined intercept thresholds.

[0047] Consistent with a disclosed embodiment, a driver assist navigation system is provided for a vehicle. The system may include at least one image capture device configured to acquire a plurality of images of an area in a vicinity of the vehicle; a data interface; and at least one processing device. The at least one processing device may be configured to: receive the plurality of images via the data interface; identify, based on analysis of the plurality of images, a trigger for stopping the vehicle; and based on the identified trigger, cause the vehicle to stop according to a braking profile including a first segment

IPR2025-00944
Tesla EX1002 Page 1152

associated with a first deceleration rate, a second segment which includes a second deceleration rate less than the first deceleration rate, and a third segment in which a level of braking is decreased as a target stopping location is approached, as determined based on the analysis of the plurality of images.

[0048] Consistent with the disclosed embodiment, reaching of the target stopping location may be accompanied by a release in braking pressure. The trigger may include an object. The object may include another vehicle in front of the vehicle. The trigger may include a traffic light signal, a stop sign, and/or a speed limit sign. The at least one processing device may be further configured to provide an audible announcement identifying the trigger. The at least one processing device may be configured to cause an audible announcement indicating that the vehicle is braking.

[0049] Consistent with another disclosed embodiment, a vehicle may include a body; at least one image capture device configured to acquire a plurality of images of an area in a vicinity of the vehicle; a data interface; and at least one processing device. The at least one processing device may be configured to: receive the plurality of images via the data interface; identify, based on analysis of the plurality of images, a trigger for stopping the vehicle; and based on the identified trigger, cause the vehicle to stop according to a braking profile including a first segment associated with a first deceleration rate, a second segment which includes a second deceleration rate less than the first deceleration rate, and a third segment in which a level of braking is decreased as a target stopping location is approached, as determined based on the analysis of the plurality of images.

[0050] Consistent with the disclosed embodiment, reaching of the target stopping location may be accompanied by a release in braking pressure. The trigger may include an object. The object may include another vehicle in front of the vehicle. The trigger may include a traffic light signal, a stop sign, and/or a speed limit sign. The at least one processing device may be further configured to provide an audible announcement identifying the trigger. The at least one processing device may be configured to cause an audible announcement indicating that the vehicle is braking.

[0051] Consistent with another disclosed embodiment, a method is provided for navigating a vehicle. The method may include acquiring, via at least one image capture plurality of images of an area in a vicinity of the vehicle; identifying, based on analysis of the plurality of images, a trigger for stopping the vehicle; and based on the identified trigger, causing the vehicle to stop according to a braking profile including a first segment associated with a first deceleration rate, a second segment which includes a second deceleration rate less than the first deceleration rate, and a third segment in which a level of braking is decreased as a target stopping location is approached, as determined based on the analysis of the plurality of images.

[0052] Consistent with other disclosed embodiments, non-transitory computer-readable storage media may store program instructions, which are executed by at least one processing device and perform any of the methods described herein.

[0053] The foregoing general description and the following detailed description are exemplary and explanatory.

BRIEF DESCRIPTION OF THE DRAWINGS

[0054] The accompanying drawings, which are incorporated in and constitute a part of this disclosure, illustrate various disclosed embodiments. In the drawings:

FIG. 1 is a diagrammatic representation of an exemplary system consistent with the disclosed embodiments.

FIG. 2A is a diagrammatic side view representation of an exemplary vehicle including a system consistent with the disclosed embodiments.

FIG. 2B is a diagrammatic top view representation of the vehicle and system shown in FIG. 2A consistent with the disclosed embodiments.

FIG. 2C is a diagrammatic top view representation of another embodiment of a vehicle including a system consistent with the disclosed embodiments.

FIG. 2D is a diagrammatic top view representation of yet another embodiment of a vehicle including a system consistent with the disclosed embodiments.

FIG. 2E is a diagrammatic top view representation of yet another embodiment of a vehicle including a system consistent with the disclosed embodiments.

FIG. 2F is a diagrammatic representation of exemplary vehicle control systems consistent with the disclosed embodiments.

FIG. 3A is a diagrammatic representation of an interior of a vehicle including a rearview mirror and a user interface for a vehicle imaging system consistent with the disclosed embodiments.

FIG. 3B is an illustration of an example of a camera mount that is configured to be positioned behind a rearview mirror and against a vehicle windshield consistent with the disclosed embodiments.

FIG. 3C is an illustration of the camera mount shown in FIG. 3B from a different perspective consistent with the disclosed embodiments.

FIG. 3D is an illustration of an example of a camera mount that is configured to be positioned behind a rearview mirror and against a vehicle windshield consistent with the disclosed embodiments.

FIG. 4 is an exemplary block diagram of a memory configured to store instructions for performing one or more operations consistent with the disclosed embodiments.

FIG. 5A is a flowchart showing an exemplary process for causing one or more navigational responses based on monocular image analysis consistent with

IPR2025-00944
Tesla EX1002 Page 1153

disclosed embodiments.

FIG. 5B is a flowchart showing an exemplary process for detecting one or more vehicles and/or pedestrians in a set of images consistent with the disclosed embodiments.

FIG. 5C is a flowchart showing an exemplary process for detecting road marks and/or lane geometry information in a set of images consistent with the disclosed embodiments.

FIG. 5D is a flowchart showing an exemplary process for detecting traffic lights in a set of images consistent with the disclosed embodiments.

FIG. 5E is a flowchart showing an exemplary process for causing one or more navigational responses based on a vehicle path consistent with the disclosed embodiments.

FIG. 5F is a flowchart showing an exemplary process for determining whether a leading vehicle is changing lanes consistent with the disclosed embodiments.

FIG. 6 is a flowchart showing an exemplary process for causing one or more navigational responses based on stereo image analysis consistent with the disclosed embodiments.

FIG. 7 is a flowchart showing an exemplary process for causing one or more navigational responses based on an analysis of three sets of images consistent with the disclosed embodiments.

FIG. 8A is a diagrammatic representation of an exemplary vehicle traveling within a first and a second lane constraint in which a lane offset condition exists on the first vehicle side consistent with the disclosed embodiments.

FIG. 8B is a diagrammatic representation of an exemplary vehicle traveling within a first and a second lane constraint in which a lane offset condition exists on the second vehicle side consistent with the disclosed embodiments.

FIG. 8C is a diagrammatic representation of an exemplary vehicle traveling within a first and a second lane constraint on a curved road consistent with the disclosed embodiments.

FIG. 8D is a diagrammatic representation of an exemplary vehicle traveling within a first and a second lane constraint in which a lane offset condition exists on both the first vehicle side and on the second vehicle side consistent with the disclosed embodiments.

FIG. 9 is a diagrammatic representation of the memory of an exemplary navigation system consistent with the disclosed embodiments.

FIG. 10 is a flowchart of an exemplary method for navigating a vehicle consistent with the disclosed embodiments.

FIG 11 is a diagrammatic representation of an exemplary vehicle including a navigation system traveling within a default lane consistent with the disclosed embodiments.

FIG 12A is a diagrammatic representation of the memory of an exemplary navigation system consistent with the disclosed embodiments.

FIG 12B is a flowchart of an exemplary process for navigating a vehicle consistent with the disclosed embodiments.

FIGS. 13A and 13B are diagrammatic representations of an exemplary vehicle approaching, and navigating, a curve with one or more characteristics consistent with the disclosed embodiments.

FIG. 14 is an exemplary block diagram of a memory configured to store instructions for performing one or more operations consistent with the disclosed embodiments.

FIG. 15 is a flow chart showing an exemplary process for controlling the velocity of a vehicle based on a detected curve and observed characteristics of the curve consistent with disclosed embodiments.

FIGS. 16A-16D are diagrammatic representations of a primary vehicle mimicking one or more actions of a leading vehicle consistent with the disclosed embodiments.

FIG. 17 is an exemplary block diagram of a memory configured to store instructions for performing one or more operations consistent with the disclosed embodiments.

FIG. 18 is a flow chart showing an exemplary process for causing a primary vehicle to mimic one or more actions of a leading vehicle consistent with disclosed embodiments.

FIG. 19 is a flow chart showing an exemplary process for causing a primary vehicle to decline to mimic a turn of a leading vehicle consistent with disclosed embodiments.

FIG. 20 is a flow chart showing an exemplary process for causing a primary vehicle to mimic or decline a turn of a leading vehicle consistent with disclosed embodiments.

FIG. 21 is a flow chart showing another exemplary process for causing a primary vehicle to mimic one or more actions of a leading vehicle consistent with disclosed embodiments.

FIG. 22 is a flow chart showing an exemplary process for causing a primary vehicle to mimic one or more actions of a leading vehicle based on a navigation history consistent with disclosed embodiments.

FIG. 23 is a diagrammatic representation of an exemplary vehicle traveling within a first and a second lane constraint on a roadway along on which another vehicle is traveling consistent with the disclosed embodiments.

FIG. 24A is a diagrammatic representation of a memory storing instructions for navigating a vehicle among encroaching vehicles consistent with the disclosed embodiments.

FIG. 24B is a flowchart of an exemplary process for navigating a vehicle among encroaching vehicles consistent with the disclosed embodiments.

IPR2025-00944
Tesla EX1002 Page 1154

FIG. 25 is a diagrammatic representation of an exemplary vehicle traveling within a first and a second lane constraint on a roadway along on which another vehicle is traveling consistent with the disclosed embodiments.

FIG. 26 is a diagrammatic representation of a memory storing instructions for detecting and responding to traffic laterally encroaching on a vehicle consistent with the disclosed embodiments.

FIG. 27 is a flowchart of an exemplary process for detecting and responding to traffic laterally encroaching on a vehicle consistent with the disclosed embodiments.

FIGS. 28A and 28B are diagrammatic representations of causing a response in a primary vehicle consistent with the disclosed embodiments.

FIG. 29 is an exemplary block diagram of a memory configured to store instructions for performing one or more operations consistent with the disclosed embodiments.

FIG. 30 is a flow chart showing an exemplary process for causing a response in a primary vehicle consistent with disclosed embodiments.

FIG. 31 is a flow chart showing an exemplary process for declining to cause a response in a primary vehicle consistent with disclosed embodiments.

FIG. 32A is a diagrammatic representation of an exemplary vehicle in which an object is present in the vicinity of vehicle consistent with the disclosed embodiments.

FIG. 32B illustrates an exemplary braking profile, consistent with the disclosed embodiments.

FIG. 33 is a diagrammatic representation of the memory of an exemplary navigation system consistent with the disclosed embodiments.

FIG. 34 is a flowchart of an exemplary method for implementing a braking profile while navigating a vehicle consistent with the disclosed embodiments.

DETAILED DESCRIPTION

[0055] The following detailed description refers to the accompanying drawings. Wherever possible, the same reference numbers are used in the drawings and the following description to refer to the same or similar parts. While several illustrative embodiments are described herein, modifications, adaptations and other implementations are possible. For example, substitutions, additions or modifications may be made to the components illustrated in the drawings, and the illustrative methods described herein may be modified by substituting, reordering, removing, or adding steps to the disclosed methods. Accordingly, the following detailed description is not limited to the disclosed embodiments and examples. Instead, the proper scope is defined by the appended claims.

[0056] Disclosed embodiments provide systems and methods that use cameras to provide autonomous navigation features. In various embodiments, the system may include one, two or more cameras that monitor the environment of a vehicle. In one embodiment, the system may monitor and adjust the free space between a vehicle and the boundaries of the lane within which the vehicle is traveling. In another embodiment, the system may select a particular lane as a default lane for the vehicle to use while traveling. In another embodiment, the system may control the speed of the vehicle in different scenarios, such as while making a turn. In yet another embodiment, the system may mimic the actions of a leading vehicle. In yet another embodiment, the system may monitor a target vehicle and enable the vehicle to pass the target vehicle under certain conditions (e.g., if the target vehicle is traveling in a lane different from the lane within which the vehicle is traveling. In still yet another embodiment, the system may produce a natural feeling response to a laterally encroaching vehicle, such as a vehicle attempting to move into the lane within which the vehicle is traveling.

[0057] FIG. 1 is a block diagram representation of a system 100 consistent with the exemplary disclosed embodiments. System 100 may include various components depending on the requirements of a particular implementation. In some embodiments, system 100 may include a processing unit 110, an image acquisition unit 120, a position sensor 130, one or more memory units 140, 150, a map database 160, and a user interface 170. Processing unit 110 may include one or more processing devices. In some embodiments, processing unit 110 may include an applications processor 180, an image processor 190, or any other suitable processing device. Similarly, image acquisition unit 120 may include any number of image acquisition devices and components depending on the requirements of a particular application. In some embodiments, image acquisition unit 120 may include one or more image capture devices (e.g., cameras), such as image capture device 122, image capture device 124, and image capture device 126. System 100 may also include a data interface 128 communicatively connecting processing device 110 to image acquisition device 120. For example, data interface 128 may include any wired and/or wireless link or links for transmitting image data acquired by image accusation device 120 to processing unit 110.

[0058] Both applications processor 180 and image processor 190 may include various types of processing devices. For example, either or both of applications processor 180 and image processor 190 may include a microprocessor, preprocessors (such as an image preprocessor), graphics processors, a central processing unit (CPU), support circuits, digital signal processors, integrated circuits, memory, or any other types of devices suitable for running applications and for image processing and analysis. In some embodiments, applications processor 180 and/or image processor 190 may include any type of single or multi-core processor, mobile device microcontroller, central processing unit, etc. Various

IPR2025-00944
Tesla EX1002 Page 1155

processing devices may be used, including, for example, processors available from manufacturers such as Intel®, AMD®, etc. and may include various architectures (e.g., x86 processor, ARM®, etc.).

[0059] In some embodiments, applications processor 180 and/or image processor 190 may include any of the EyeQ series of processor chips available from Mobileye®. These processor designs each include multiple processing units with local memory and instruction sets. Such processors may include video inputs for receiving image data from multiple image sensors and may also include video out capabilities. In one example, the EyeQ2® uses 90nm-micron technology operating at 332Mhz. The EyeQ2® architecture consists of two floating point, hyper-thread 32-bit RISC CPUs (MIPS32® 34K® cores), five Vision Computing Engines (VCE), three Vector Microcode Processors (VMP®), Denali 64-bit Mobile DDR Controller, 128-bit internal Sonics Interconnect, dual 16-bit Video input and 18-bit Video output controllers, 16 channels DMA and several peripherals. The MIPS34K CPU manages the five VCEs, three VMP™ and the DMA, the second MIPS34K CPU and the multi-channel DMA as well as the other peripherals. The five VCEs, three VMP® and the MIPS34K CPU can perform intensive vision computations required by multi-function bundle applications. In another example, the EyeQ3®, which is a third generation processor and is six times more powerful that the EyeQ2®, may be used in the disclosed embodiments.

[0060] While FIG. 1 depicts two separate processing devices included in processing unit 110, more or fewer processing devices may be used. For example, in some embodiments, a single processing device may be used to accomplish the tasks of applications processor 180 and image processor 190. In other embodiments, these tasks may be performed by more than two processing devices.

[0061] Processing unit 110 may comprise various types of devices. For example, processing unit 110 may include various devices, such as a controller, an image preprocessor, a central processing unit (CPU), support circuits, digital signal processors, integrated circuits, memory, or any other types of devices for image processing and analysis. The image preprocessor may include a video processor for capturing, digitizing and processing the imagery from the image sensors. The CPU may comprise any number of microcontrollers or microprocessors. The support circuits may be any number of circuits generally well known in the art, including cache, power supply, clock and input-output circuits. The memory may store software that, when executed by the processor, controls the operation of the system. The memory may include databases and image processing software. The memory may comprise any number of random access memories, read only memories, flash memories, disk drives, optical storage, tape storage, removable storage and other types of storage. In one instance, the memory may be separate from the processing unit 110. In another instance, the memory may be integrated into the processing unit 110.

[0062] Each memory 140, 150 may include software instructions that when executed by a processor (e.g., applications processor 180 and/or image processor 190), may control operation of various aspects of system 100. These memory units may include various databases and image processing software. The memory units may include random access memory, read only memory, flash memory, disk drives, optical storage, tape storage, removable storage and/or any other types of storage. In some embodiments, memory units 140, 150 may be separate from the applications processor 180 and/or image processor 190. In other embodiments, these memory units may be integrated into applications processor 180 and/or image processor 190.

[0063] Position sensor 130 may include any type of device suitable for determining a location associated with at least one component of system 100. In some embodiments, position sensor 130 may include a GPS receiver. Such receivers can determine a user position and velocity by processing signals broadcasted by global positioning system satellites. Position information from position sensor 130 may be made available to applications processor 180 and/or image processor 190.

[0064] User interface 170 may include any device suitable for providing information to or for receiving inputs from one or more users of system 100. In some embodiments, user interface 170 may include user input devices, including, for example, a touchscreen, microphone, keyboard, pointer devices, track wheels, cameras, knobs, buttons, etc. With such input devices, a user may be able to provide information inputs or commands to system 100 by typing instructions or information, providing voice commands, selecting menu options on a screen using buttons, pointers, or eye-tracking capabilities, or through any other suitable techniques for communicating information to system 100.

[0065] User interface 170 may be equipped with one or more processing devices configured to provide and receive information to or from a user and process that information for use by, for example, applications processor 180. In some embodiments, such processing devices may execute instructions for recognizing and tracking eye movements, receiving and interpreting voice commands, recognizing and interpreting touches and/or gestures made on a touchscreen, responding to keyboard entries or menu selections, etc. In some embodiments, user interface 170 may include a display, speaker, tactile device, and/or any other devices for providing output information to a user.

[0066] Map database 160 may include any type of database for storing map data useful to system 100. In some embodiments, map database 160 may include data relating to the position, in a reference coordinate system, of various items, including roads, water features, geographic features, businesses, points of interest, restaurants, gas stations, etc. Map database 160 may store not

IPR2025-00944
Tesla EX1002 Page 1156

only the locations of such items, but also descriptors relating to those items, including, for example, names associated with any of the stored features. In some embodiments, map database 160 may be physically located with other components of system 100. Alternatively or additionally, map database 160 or a portion thereof may be located remotely with respect to other components of system 100 (e.g., processing unit 110). In such embodiments, information from map database 160 may be downloaded over a wired or wireless data connection to a network (e.g., over a cellular network and/or the Internet, etc.).

[0067] Image capture devices 122, 124, and 126 may each include any type of device suitable for capturing at least one image from an environment. Moreover, any number of image capture devices may be used to acquire images for input to the image processor. Some embodiments may include only a single image capture device, while other embodiments may include two, three, or even four or more image capture devices. Image capture devices 122,124, and 126 will be further described with reference to FIGS. 2B-2E, below.

[0068] System 100, or various components thereof, may be incorporated into various different platforms. In some embodiments, system 100 may be included on a vehicle 200, as shown in FIG. 2A. For example, vehicle 200 may be equipped with a processing unit 110 and any of the other components of system 100, as described above relative to FIG. 1. While in some embodiments vehicle 200 may be equipped with only a single image capture device (e.g., camera), in other embodiments, such as those discussed in connection with FIGS. 2B-2E, multiple image capture devices may be used. For example, either of image capture devices 122 and 124 of vehicle 200, as shown in FIG. 2A, may be part of an ADAS (Advanced Driver Assistance Systems) imaging set.

[0069] The image capture devices included on vehicle 200 as part of the image acquisition unit 120 may be positioned at any suitable location. In some embodiments, as shown in FIGS. 2A-2E and 3A-3C, image capture device 122 may be located in the vicinity of the rearview mirror. This position may provide a line of sight similar to that of the driver of vehicle 200, which may aid in determining what is and is not visible to the driver. Image capture device 122 may be positioned at any location near the rearview mirror, but placing image capture device 122 on the driver side of the mirror may further aid in obtaining images representative of the driver's field of view and/or line of sight.

[0070] Other locations for the image capture devices of image acquisition unit 120 may also be used. For example, image capture device 124 may be located on or in a bumper of vehicle 200. Such a location may be especially suitable for image capture devices having a wide field of view. The line of sight of bumper-located image capture devices can be different from that of the driver and, therefore, the bumper image capture device and driver may not always see the same objects. The image capture devices (e.g., image capture devices 122, 124, and 126) may also be located in other locations. For example, the image capture devices may be located on or in one or both of the side mirrors of vehicle 200, on the roof of vehicle 200, on the hood of vehicle 200, on the trunk of vehicle 200, on the sides of vehicle 200, mounted on, positioned behind, or positioned in front of any of the windows of vehicle 200, and mounted in or near light figures on the front and/or back of vehicle 200, etc.

[0071] In addition to image capture devices, vehicle 200 may include various other components of system 100. For example, processing unit 110 may be included on vehicle 200 either integrated with or separate from an engine control unit (ECU) of the vehicle. Vehicle 200 may also be equipped with a position sensor 130, such as a GPS receiver and may also include a map database 160 and memory units 140 and 150.

[0072] FIG. 2A is a diagrammatic side view representation of an exemplary vehicle imaging system consistent with the disclosed embodiments. FIG. 2B is a diagrammatic top view illustration of the embodiment shown in FIG. 2A. As illustrated in FIG. 2B, the disclosed embodiments may include a vehicle 200 including in its body a system 100 with a first image capture device 122 positioned in the vicinity of the rearview mirror and/or near the driver of vehicle 200, a second image capture device 124 positioned on or in a bumper region (e.g., one of bumper regions 210) of vehicle 200, and a processing unit 110.

[0073] As illustrated in FIG. 2C, image capture devices 122 and 124 may both be positioned in the vicinity of the rearview mirror and/or near the driver of vehicle 200. Additionally, while two image capture devices 122 and 124 are shown in FIGS. 2B and 2C, it should be understood that other embodiments may include more than two image capture devices. For example, in the embodiments shown in FIGS. 2D and 2E, first, second, and third image capture devices 122, 124, and 126, are included in the system 100 of vehicle 200.

[0074] As illustrated in FIG. 2D, image capture device 122 may be positioned in the vicinity of the rearview mirror and/or near the driver of vehicle 200, and image capture devices 124 and 126 may be positioned on or in a bumper region (e.g., one of bumper regions 210) of vehicle 200. And as shown in FIG. 2E, image capture devices 122, 124, and 126 may be positioned in the vicinity of the rearview mirror and/or near the driver seat of vehicle 200. The disclosed embodiments are not limited to any particular number and configuration of the image capture devices, and the image capture devices may be positioned in any appropriate location within and/or on vehicle 200.

[0075] It is to be understood that the disclosed embodiments are not limited to vehicles and could be applied in other contexts. It is also to be understood that disclosed embodiments are not limited to a particular type of vehicle 200 and may be applicable to all types of vehicles includ-

IPR2025-00944
Tesla EX1002 Page 1157

ing automobiles, trucks, trailers, and other types of vehicles.

[0076] The first image capture device 122 may include any suitable type of image capture device. Image capture device 122 may include an optical axis. In one instance, the image capture device 122 may include an Aptina M9V024 WVGA sensor with a global shutter. In other embodiments, image capture device 122 may provide a resolution of 1280x960 pixels and may include a rolling shutter. Image capture device 122 may include various optical elements. In some embodiments one or more lenses may be included, for example, to provide a desired focal length and field of view for the image capture device. In some embodiments, image capture device 122 may be associated with a 6mm lens or a 12mm lens. In some embodiments, image capture device 122 may be configured to capture images having a desired field-of-view (FOV) 202, as illustrated in FIG. 2D. For example, image capture device 122 may be configured to have a regular FOV, such as within a range of 40 degrees to 56 degrees, including a 46 degree FOV, 50 degree FOV, 52 degree FOV, or greater. Alternatively, image capture device 122 may be configured to have a narrow FOV in the range of 23 to 40 degrees, such as a 28 degree FOV or 36 degree FOV. In addition, image capture device 122 may be configured to have a wide FOV in the range of 100 to 180 degrees. In some embodiments, image capture device 122 may include a wide angle bumper camera or one with up to a 180 degree FOV.

[0077] The first image capture device 122 may acquire a plurality of first images relative to a scene associated with the vehicle 200. Each of the plurality of first images may be acquired as a series of image scan lines, which may be captured using a rolling shutter. Each scan line may include a plurality of pixels.

[0078] The first image capture device 122 may have a scan rate associated with acquisition of each of the first series of image scan lines. The scan rate may refer to a rate at which an image sensor can acquire image data associated with each pixel included in a particular scan line.

[0079] Image capture devices 122, 124, and 126 may contain any suitable type and number of image sensors, including CCD sensors or CMOS sensors, for example. In one embodiment, a CMOS image sensor may be employed along with a rolling shutter, such that each pixel in a row is read one at a time, and scanning of the rows proceeds on a row-by-row basis until an entire image frame has been captured. In some embodiments, the rows may be captured sequentially from top to bottom relative to the frame.

[0080] The use of a rolling shutter may result in pixels in different rows being exposed and captured at different times, which may cause skew and other image artifacts in the captured image frame. On the other hand, when the image capture device 122 is configured to operate with a global or synchronous shutter, all of the pixels may be exposed for the same amount of time and during a common exposure period. As a result, the image data in a frame collected from a system employing a global shutter represents a snapshot of the entire FOV (such as FOV 202) at a particular time. In contrast, in a rolling shutter application, each row in a frame is exposed and data is capture at different times. Thus, moving objects may appear distorted in an image capture device having a rolling shutter. This phenomenon will be described in greater detail below.

[0081] The second image capture device 124 and the third image capturing device 126 may be any type of image capture device. Like the first image capture device 122, each of image capture devices 124 and 126 may include an optical axis. In one embodiment, each of image capture devices 124 and 126 may include an Aptina M9V024 WVGA sensor with a global shutter. Alternatively, each of image capture devices 124 and 126 may include a rolling shutter. Like image capture device 122, image capture devices 124 and 126 may be configured to include various lenses and optical elements. In some embodiments, lenses associated with image capture devices 124 and 126 may provide FOVs (such as FOVs 204 and 206) that are the same as, or narrower than, a FOV (such as FOV 202) associated with image capture device 122. For example, image capture devices 124 and 126 may have FOVs of 40 degrees, 30 degrees, 26 degrees, 23 degrees, 20 degrees, or less.

[0082] Image capture devices 124 and 126 may acquire a plurality of second and third images relative to a scene associated with the vehicle 200. Each of the plurality of second and third images may be acquired as a second and third series of image scan lines, which may be captured using a rolling shutter. Each scan line or row may have a plurality of pixels. Image capture devices 124 and 126 may have second and third scan rates associated with acquisition of each of image scan lines included in the second and third series.

[0083] Each image capture device 122, 124, and 126 may be positioned at any suitable position and orientation relative to vehicle 200. The relative positioning of the image capture devices 122, 124, and 126 may be selected to aid in fusing together the information acquired from the image capture devices. For example, in some embodiments, a FOV (such as FOV 204) associated with image capture device 124 may overlap partially or fully with a FOV (such as FOV 202) associated with image capture device 122 and a FOV (such as FOV 206) associated with image capture device 126.

[0084] Image capture devices 122, 124, and 126 may be located on vehicle 200 at any suitable relative heights. In one instance, there may be a height difference between the image capture devices 122, 124, and 126, which may provide sufficient parallax information to enable stereo analysis. For example, as shown in FIG. 2A, the two image capture devices 122 and 124 are at different heights. There may also be a lateral displacement difference between image capture devices 122, 124, and 126, giving additional parallax information for stereo anal-

IPR2025-00944
Tesla EX1002 Page 1158

ysis by processing unit 110, for example. The difference in the lateral displacement may be denoted by d_x, as shown in FIGS. 2C and 2D. In some embodiments, fore or aft displacement (e.g., range displacement) may exist between image capture devices 122, 124, and 126. For example, image capture device 122 may be located 0.5 to 2 meters or more behind image capture device 124 and/or image capture device 126. This type of displacement may enable one of the image capture devices to cover potential blind spots of the other image capture device(s).

[0085] Image capture devices 122 may have any suitable resolution capability (e.g., number of pixels associated with the image sensor), and the resolution of the image sensor(s) associated with the image capture device 122 may be higher, lower, or the same as the resolution of the image sensor(s) associated with image capture devices 124 and 126. In some embodiments, the image sensor(s) associated with image capture device 122 and/or image capture devices 124 and 126 may have a resolution of 640 x 480, 1024 x 768, 1280 x 960, or any other suitable resolution.

[0086] The frame rate (e.g., the rate at which an image capture device acquires a set of pixel data of one image frame before moving on to capture pixel data associated with the next image frame) may be controllable. The frame rate associated with image capture device 122 may be higher, lower, or the same as the frame rate associated with image capture devices 124 and 126. The frame rate associated with image capture devices 122, 124, and 126 may depend on a variety of factors that may affect the timing of the frame rate. For example, one or more of image capture devices 122, 124, and 126 may include a selectable pixel delay period imposed before or after acquisition of image data associated with one or more pixels of an image sensor in image capture device 122, 124, and/or 126. Generally, image data corresponding to each pixel may be acquired according to a clock rate for the device (e.g., one pixel per clock cycle). Additionally, in embodiments including a rolling shutter, one or more of image capture devices 122, 124, and 126 may include a selectable horizontal blanking period imposed before or after acquisition of image data associated with a row of pixels of an image sensor in image capture device 122, 124, and/or 126. Further, one or more of image capture devices 122, 124, and/or 126 may include a selectable vertical blanking period imposed before or after acquisition of image data associated with an image frame of image capture device 122, 124, and 126.

[0087] These timing controls may enable synchronization of frame rates associated with image capture devices 122,124, and 126, even where the line scan rates of each are different. Additionally, as will be discussed in greater detail below, these selectable timing controls, among other factors (e.g., image sensor resolution, maximum line scan rates, etc.) may enable synchronization of image capture from an area where the FOV of image capture device 122 overlaps with one or more FOVs of image capture devices 124 and 126, even where the field of view of image capture device 122 is different from the FOVs of image capture devices 124 and 126.

[0088] Frame rate timing in image capture device 122,124, and 126 may depend on the resolution of the associated image sensors. For example, assuming similar line scan rates for both devices, if one device includes an image sensor having a resolution of 640 x 480 and another device includes an image sensor with a resolution of 1280 x 960, then more time will be required to acquire a frame of image data from the sensor having the higher resolution.

[0089] Another factor that may affect the timing of image data acquisition in image capture devices 122, 124, and 126 is the maximum line scan rate. For example, acquisition of a row of image data from an image sensor included in image capture device 122, 124, and 126 will require some minimum amount of time. Assuming no pixel delay periods are added, this minimum amount of time for acquisition of a row of image data will be related to the maximum line scan rate for a particular device. Devices that offer higher maximum line scan rates have the potential to provide higher frame rates than devices with lower maximum line scan rates. In some embodiments, one or more of image capture devices 124 and 126 may have a maximum line scan rate that is higher than a maximum line scan rate associated with image capture device 122. In some embodiments, the maximum line scan rate of image capture device 124 and/or 126 may be 1.25, 1.5, 1.75, or 2 times or more than a maximum line scan rate of image capture device 122.

[0090] In another embodiment, image capture devices 122, 124, and 126 may have the same maximum line scan rate, but image capture device 122 may be operated at a scan rate less than or equal to its maximum scan rate. The system may be configured such that one or more of image capture devices 124 and 126 operate at a line scan rate that is equal to the line scan rate of image capture device 122. In other instances, the system may be configured such that the line scan rate of image capture device 124 and/or image capture device 126 may be 1.25, 1.5, 1.75, or 2 times or more than the line scan rate of image capture device 122.

[0091] In some embodiments, image capture devices 122, 124, and 126 may be asymmetric. That is, they may include cameras having different fields of view (FOV) and focal lengths. The fields of view of image capture devices 122, 124, and 126 may include any desired area relative to an environment of vehicle 200, for example. In some embodiments, one or more of image capture devices 122, 124, and 126 may be configured to acquire image data from an environment in front of vehicle 200, behind vehicle 200, to the sides of vehicle 200, or combinations thereof.

[0092] Further, the focal length associated with each image capture device 122,124, and/or 126 may be selectable (e.g., by inclusion of appropriate lenses etc.) such that each device acquires images of objects at a

IPR2025-00944
Tesla EX1002 Page 1159

desired distance range relative to vehicle 200. For example, in some embodiments image capture devices 122, 124, and 126 may acquire images of close-up objects within a few meters from the vehicle. Image capture devices 122, 124, and 126 may also be configured to acquire images of objects at ranges more distant from the vehicle (e.g., 25 m, 50 m, 100 m, 150 m, or more). Further, the focal lengths of image capture devices 122, 124, and 126 may be selected such that one image capture device (e.g., image capture device 122) can acquire images of objects relatively close to the vehicle (e.g., within 10 m or within 20 m) while the other image capture devices (e.g., image capture devices 124 and 126) can acquire images of more distant objects (e.g., greater than 20 m, 50 m, 100 m, 150 m, etc.) from vehicle 200.

[0093] According to some embodiments, the FOV of one or more image capture devices 122, 124, and 126 may have a wide angle. For example, it may be advantageous to have a FOV of 140 degrees, especially for image capture devices 122, 124, and 126 that may be used to capture images of the area in the vicinity of vehicle 200. For example, image capture device 122 may be used to capture images of the area to the right or left of vehicle 200 and, in such embodiments, it may be desirable for image capture device 122 to have a wide FOV (e.g., at least 140 degrees).

[0094] The field of view associated with each of image capture devices 122, 124, and 126 may depend on the respective focal lengths. For example, as the focal length increases, the corresponding field of view decreases.

[0095] Image capture devices 122, 124, and 126 may be configured to have any suitable fields of view. In one particular example, image capture device 122 may have a horizontal FOV of 46 degrees, image capture device 124 may have a horizontal FOV of 23 degrees, and image capture device 126 may have a horizontal FOV in between 23 and 46 degrees. In another instance, image capture device 122 may have a horizontal FOV of 52 degrees, image capture device 124 may have a horizontal FOV of 26 degrees, and image capture device 126 may have a horizontal FOV in between 26 and 52 degrees. In some embodiments, a ratio of the FOV of image capture device 122 to the FOVs of image capture device 124 and/or image capture device 126 may vary from 1.5 to 2.0. In other embodiments, this ratio may vary between 1.25 and 2.25.

[0096] System 100 may be configured so that a field of view of image capture device 122 overlaps, at least partially or fully, with a field of view of image capture device 124 and/or image capture device 126. In some embodiments, system 100 may be configured such that the fields of view of image capture devices 124 and 126, for example, fall within (e.g., are narrower than) and share a common center with the field of view of image capture device 122. In other embodiments, the image capture devices 122, 124, and 126 may capture adjacent FOVs or may have partial overlap in their FOVs. In some embodiments, the fields of view of image capture devices 122, 124, and 126 may be aligned such that a center of the narrower FOV image capture devices 124 and/or 126 may be located in a lower half of the field of view of the wider FOV device 122.

[0097] FIG. 2F is a diagrammatic representation of exemplary vehicle control systems, consistent with the disclosed embodiments. As indicated in FIG. 2F, vehicle 200 may include throttling system 220, braking system 230, and steering system 240. System 100 may provide inputs (e.g., control signals) to one or more of throttling system 220, braking system 230, and steering system 240 over one or more data links (e.g., any wired and/or wireless link or links for transmitting data). For example, based on analysis of images acquired by image capture devices 122, 124, and/or 126, system 100 may provide control signals to one or more of throttling system 220, braking system 230, and steering system 240 to navigate vehicle 200 (e.g., by causing an acceleration, a turn, a lane shift, etc.). Further, system 100 may receive inputs from one or more of throttling system 220, braking system 230, and steering system 24 indicating operating conditions of vehicle 200 (e.g., speed, whether vehicle 200 is braking and/or turning, etc.). Further details are provided in connection with FIGS. 4-7, below.

[0098] As shown in FIG. 3A, vehicle 200 may also include a user interface 170 for interacting with a driver or a passenger of vehicle 200. For example, user interface 170 in a vehicle application may include a touch screen 320, knobs 330, buttons 340, and a microphone 350. A driver or passenger of vehicle 200 may also use handles (e.g., located on or near the steering column of vehicle 200 including, for example, turn signal handles), buttons (e.g., located on the steering wheel of vehicle 200), and the like, to interact with system 100. In some embodiments, microphone 350 may be positioned adjacent to a rearview mirror 310. Similarly, in some embodiments, image capture device 122 may be located near rearview mirror 310. In some embodiments, user interface 170 may also include one or more speakers 360 (e.g., speakers of a vehicle audio system). For example, system 100 may provide various notifications (e.g., alerts) via speakers 360.

[0099] FIGS. 3B-3D are illustrations of an exemplary camera mount 370 configured to be positioned behind a rearview mirror (e.g., rearview mirror 310) and against a vehicle windshield, consistent with disclosed embodiments. As shown in FIG. 3B, camera mount 370 may include image capture devices 122, 124, and 126. Image capture devices 124 and 126 may be positioned behind a glare shield 380, which may be flush against the vehicle windshield and include a composition of film and/or anti-reflective materials. For example, glare shield 380 may be positioned such that it aligns against a vehicle windshield having a matching slope. In some embodiments, each of image capture devices 122, 124, and 126 may be positioned behind glare shield 380, as depicted, for example, in FIG. 3D. The disclosed embodiments are not limited to any particular configuration of image cap-

IPR2025-00944
Tesla EX1002 Page 1160

ture devices 122, 124, and 126, camera mount 370, and glare shield 380. FIG. 3C is an illustration of camera mount 370 shown in FIG. 3B from a front perspective.

[0100] As will be appreciated by a person skilled in the art having the benefit of this disclosure, numerous variations and/or modifications may be made to the foregoing disclosed embodiments. For example, not all components are essential for the operation of system 100. Further, any component may be located in any appropriate part of system 100 and the components may be rearranged into a variety of configurations while providing the functionality of the disclosed embodiments. Therefore, the foregoing configurations are examples and, regardless of the configurations discussed above, system 100 can provide a wide range of functionality to analyze the surroundings of vehicle 200 and navigate vehicle 200 in response to the analysis.

[0101] As discussed below in further detail and consistent with various disclosed embodiments, system 100 may provide a variety of features related to autonomous driving and/or driver assist technology. For example, system 100 may analyze image data, position data (e.g., GPS location information), map data, speed data, and/or data from sensors included in vehicle 200. System 100 may collect the data for analysis from, for example, image acquisition unit 120, position sensor 130, and other sensors. Further, system 100 may analyze the collected data to determine whether or not vehicle 200 should take a certain action, and then automatically take the determined action without human intervention. For example, when vehicle 200 navigates without human intervention, system 100 may automatically control the braking, acceleration, and/or steering of vehicle 200 (e.g., by sending control signals to one or more of throttling system 220, braking system 230, and steering system 240). Further, system 100 may analyze the collected data and issue warnings and/or alerts to vehicle occupants based on the analysis of the collected data.

[0102] Further, consistent with disclosed embodiments, the functionality provided by system 100 may cause vehicle 200 to take different actions to navigate vehicle 200 within a lane and/or relative to other vehicles and/or objects. For example, system 100 may adjust the positioning of vehicle 200 relative to a lane within which vehicle 200 is traveling and/or relative to objects positioned near vehicle 200, select a particular lane for vehicle 200 to use while traveling, and take action in response to an encroaching vehicle, such as a vehicle attempting to move into the lane within which vehicle 200 is traveling. Additionally, system 100 may control the speed of vehicle 200 in different scenarios, such as when vehicle 200 is making a turn. System 100 may further cause vehicle 200 to mimic the actions of a leading vehicle or monitor a target vehicle and navigate vehicle 200 so that it passes the target vehicle. Additional details regarding the various embodiments that are provided by system 100 are provided below.

Forward-Facing Multi-Imaging System

[0103] As discussed above, system 100 may provide drive assist functionality that uses a multi-camera system. The multi-camera system may use one or more cameras facing in the forward direction of a vehicle. In other embodiments, the multi-camera system may include one or more cameras facing to the side of a vehicle or to the rear of the vehicle. In one embodiment, for example, system 100 may use a two-camera imaging system, where a first camera and a second camera (e.g., image capture devices 122 and 124) may be positioned at the front and/or the sides of a vehicle (e.g., vehicle 200). The first camera may have a field of view that is greater than, less than, or partially overlapping with, the field of view of the second camera. In addition, the first camera may be connected to a first image processor to perform monocular image analysis of images provided by the first camera, and the second camera may be connected to a second image processor to perform monocular image analysis of images provided by the second camera. The outputs (e.g., processed information) of the first and second image processors may be combined. In some embodiments, the second image processor may receive images from both the first camera and second camera to perform stereo analysis. In another embodiment, system 100 may use a three-camera imaging system where each of the cameras has a different field of view. Such a system may, therefore, make decisions based on information derived from objects located at varying distances both forward and to the sides of the vehicle. References to monocular image analysis may refer to instances where image analysis is performed based on images captured from a single point of view (e.g., from a single camera). Stereo image analysis may refer to instances where image analysis is performed based on two or more images captured with one or more variations of an image capture parameter. For example, captured images suitable for performing stereo image analysis may include images captured: from two or more different positions, from different fields of view, using different focal lengths, along with parallax information, etc.

[0104] For example, in one embodiment, system 100 may implement a three camera configuration using image capture devices 122-126. In such a configuration, image capture device 122 may provide a narrow field of view (e.g., 34 degrees, or other values selected from a range of about 20 to 45 degrees, etc.), image capture device 124 may provide a wide field of view (e.g., 150 degrees or other values selected from a range of about 100 to about 180 degrees), and image capture device 126 may provide an intermediate field of view (e.g., 46 degrees or other values selected from a range of about 35 to about 60 degrees). In some embodiments, image capture device 126 may act as a main or primary camera. Image capture devices 122-126 may be positioned behind rearview mirror 310 and positioned substantially side-by-side (e.g., 6 cm apart). Further, in some embod-

IPR2025-00944
Tesla EX1002 Page 1161

iments, as discussed above, one or more of image capture devices 122-126 may be mounted behind glare shield 380 that is flush with the windshield of vehicle 200. Such shielding may act to minimize the impact of any reflections from inside the car on image capture devices 122-126.

[0105] In another embodiment, as discussed above in connection with FIGS. 3B and 3C, the wide field of view camera (e.g., image capture device 124 in the above example) may be mounted lower than the narrow and main field of view cameras (e.g., image devices 122 and 126 in the above example). This configuration may provide a free line of sight from the wide field of view camera. To reduce reflections, the cameras may be mounted close to the windshield of vehicle 200, and may include polarizers on the cameras to damp reflected light.

[0106] A three camera system may provide certain performance characteristics. For example, some embodiments may include an ability to validate the detection of objects by one camera based on detection results from another camera. In the three camera configuration discussed above, processing unit 110 may include, for example, three processing devices (e.g., three EyeQ series of processor chips, as discussed above), with each processing device dedicated to processing images captured by one or more of image capture devices 122-126.

[0107] In a three camera system, a first processing device may receive images from both the main camera and the narrow field of view camera, and perform vision processing of the narrow FOV camera to, for example, detect other vehicles, pedestrians, lane marks, traffic signs, traffic lights, and other road objects. Further, the first processing device may calculate a disparity of pixels between the images from the main camera and the narrow camera and create a 3D reconstruction of the environment of vehicle 200. The first processing device may then combine the 3D reconstruction with 3D map data or with 3D information calculated based on information from another camera.

[0108] The second processing device may receive images from main camera and perform vision processing to detect other vehicles, pedestrians, lane marks, traffic signs, traffic lights, and other road objects. Additionally, the second processing device may calculate a camera displacement and, based on the displacement, calculate a disparity of pixels between successive images and create a 3D reconstruction of the scene (e.g., a structure from motion). The second processing device may send the structure from motion based 3D reconstruction to the first processing device to be combined with the stereo 3D images.

[0109] The third processing device may receive images from the wide FOV camera and process the images to detect vehicles, pedestrians, lane marks, traffic signs, traffic lights, and other road objects. The third processing device may further execute additional processing instructions to analyze images to identify objects moving in the image, such as vehicles changing lanes, pedestrians, etc.

[0110] In some embodiments, having streams of image-based information captured and processed independently may provide an opportunity for providing redundancy in the system. Such redundancy may include, for example, using a first image capture device and the images processed from that device to validate and/or supplement information obtained by capturing and processing image information from at least a second image capture device.

[0111] In some embodiments, system 100 may use two image capture devices (e.g., image capture devices 122 and 124) in providing navigation assistance for vehicle 200 and use a third image capture device (e.g., image capture device 126) to provide redundancy and validate the analysis of data received from the other two image capture devices. For example, in such a configuration, image capture devices 122 and 124 may provide images for stereo analysis by system 100 for navigating vehicle 200, while image capture device 126 may provide images for monocular analysis by system 100 to provide redundancy and validation of information obtained based on images captured from image capture device 122 and/or image capture device 124. That is, image capture device 126 (and a corresponding processing device) may be considered to provide a redundant sub-system for providing a check on the analysis derived from image capture devices 122 and 124 (e.g., to provide an automatic emergency braking (AEB) system).

[0112] One of skill in the art will recognize that the above camera configurations, camera placements, number of cameras, camera locations, etc., are examples only. These components and others described relative to the overall system may be assembled and used in a variety of different configurations without departing from the scope of the disclosed embodiments. Further details regarding usage of a multi-camera system to provide driver assist and/or autonomous vehicle functionality follow below.

[0113] FIG. 4 is an exemplary functional block diagram of memory 140 and/or 150, which may be stored/programmed with instructions for performing one or more operations consistent with the disclosed embodiments. Although the following refers to memory 140, one of skill in the art will recognize that instructions may be stored in memory 140 and/or 150.

[0114] As shown in FIG. 4, memory 140 may store a monocular image analysis module 402, a stereo image analysis module 404, a velocity and acceleration module 406, and a navigational response module 408. The disclosed embodiments are not limited to any particular configuration of memory 140. Further, application processor 180 and/or image processor 190 may execute the instructions stored in any of modules 402-408 included in memory 140. One of skill in the art will understand that references in the following discussions to processing unit 110 may refer to application processor 180 and image processor 190 individually or collectively. Accordingly,

IPR2025-00944
Tesla EX1002 Page 1162

steps of any of the following processes may be performed by one or more processing devices.

[0115] In one embodiment, monocular image analysis module 402 may store instructions (such as computer vision software) which, when executed by processing unit 110, performs monocular image analysis of a set of images acquired by one of image capture devices 122, 124, and 126. In some embodiments, processing unit 110 may combine information from a set of images with additional sensory information (e.g., information from radar) to perform the monocular image analysis. As described in connection with FIGS. 5A-5D below, monocular image analysis module 402 may include instructions for detecting a set of features within the set of images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, hazardous objects, and any other feature associated with an environment of a vehicle. Based on the analysis, system 100 (e.g., via processing unit 110) may cause one or more navigational responses in vehicle 200, such as a turn, a lane shift, a change in acceleration, and the like, as discussed below in connection with navigational response module 408.

[0116] In one embodiment, stereo image analysis module 404 may store instructions (such as computer vision software) which, when executed by processing unit 110, performs stereo image analysis of first and second sets of images acquired by a combination of image capture devices selected from any of image capture devices 122, 124, and 126. In some embodiments, processing unit 110 may combine information from the first and second sets of images with additional sensory information (e.g., information from radar) to perform the stereo image analysis. For example, stereo image analysis module 404 may include instructions for performing stereo image analysis based on a first set of images acquired by image capture device 124 and a second set of images acquired by image capture device 126. As described in connection with FIG. 6 below, stereo image analysis module 404 may include instructions for detecting a set of features within the first and second sets of images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, hazardous objects, and the like. Based on the analysis, processing unit 110 may cause one or more navigational responses in vehicle 200, such as a turn, a lane shift, a change in acceleration, and the like, as discussed below in connection with navigational response module 408.

[0117] In one embodiment, velocity and acceleration module 406 may store software configured to analyze data received from one or more computing and electromechanical devices in vehicle 200 that are configured to cause a change in velocity and/or acceleration of vehicle 200. For example, processing unit 110 may execute instructions associated with velocity and acceleration module 406 to calculate a target speed for vehicle 200 based on data derived from execution of monocular image analysis module 402 and/or stereo image analysis module 404. Such data may include, for example, a target position, velocity, and/or acceleration, the position and/or speed of vehicle 200 relative to a nearby vehicle, pedestrian, or road object, position information for vehicle 200 relative to lane markings of the road, and the like. In addition, processing unit 110 may calculate a target speed for vehicle 200 based on sensory input (e.g., information from radar) and input from other systems of vehicle 200, such as throttling system 220, braking system 230, and/or steering system 240 of vehicle 200. Based on the calculated target speed, processing unit 110 may transmit electronic signals to throttling system 220, braking system 230, and/or steering system 240 of vehicle 200 to trigger a change in velocity and/or acceleration by, for example, physically depressing the brake or easing up off the accelerator of vehicle 200.

[0118] In one embodiment, navigational response module 408 may store software executable by processing unit 110 to determine a desired navigational response based on data derived from execution of monocular image analysis module 402 and/or stereo image analysis module 404. Such data may include position and speed information associated with nearby vehicles, pedestrians, and road objects, target position information for vehicle 200, and the like. Additionally, in some embodiments, the navigational response may be based (partially or fully) on map data, a predetermined position of vehicle 200, and/or a relative velocity or a relative acceleration between vehicle 200 and one or more objects detected from execution of monocular image analysis module 402 and/or stereo image analysis module 404. Navigational response module 408 may also determine a desired navigational response based on sensory input (e.g., information from radar) and inputs from other systems of vehicle 200, such as throttling system 220, braking system 230, and steering system 240 of vehicle 200. Based on the desired navigational response, processing unit 110 may transmit electronic signals to throttling system 220, braking system 230, and steering system 240 of vehicle 200 to trigger a desired navigational response by, for example, turning the steering wheel of vehicle 200 to achieve a rotation of a predetermined angle. In some embodiments, processing unit 110 may use the output of navigational response module 408 (e.g., the desired navigational response) as an input to execution of velocity and acceleration module 406 for calculating a change in speed of vehicle 200.

[0119] FIG. 5A is a flowchart showing an exemplary process 500A for causing one or more navigational responses based on monocular image analysis, consistent with disclosed embodiments. At step 510, processing unit 110 may receive a plurality of images via data interface 128 between processing unit 110 and image acquisition unit 120. For instance, a camera included in image acquisition unit 120 (such as image capture device 122 having field of view 202) may capture a plurality of images of an area forward of vehicle 200 (or to the sides or rear of a vehicle, for example) and transmit them over a data connection (e.g., digital, wired, USB, wireless, Bluetooth,

IPR2025-00944
Tesla EX1002 Page 1163

etc.) to processing unit 110. Processing unit 110 may execute monocular image analysis module 402 to analyze the plurality of images at step 520, as described in further detail in connection with FIGS. 5B-5D below. By performing the analysis, processing unit 110 may detect a set of features within the set of images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, and the like.

[0120] Processing unit 110 may also execute monocular image analysis module 402 to detect various road hazards at step 520, such as, for example, parts of a truck tire, fallen road signs, loose cargo, small animals, and the like. Road hazards may vary in structure, shape, size, and color, which may make detection of such hazards more challenging. In some embodiments, processing unit 110 may execute monocular image analysis module 402 to perform multi-frame analysis on the plurality of images to detect road hazards. For example, processing unit 110 may estimate camera motion between consecutive image frames and calculate the disparities in pixels between the frames to construct a 3D-map of the road. Processing unit 110 may then use the 3D-map to detect the road surface, as well as hazards existing above the road surface.

[0121] At step 530, processing unit 110 may execute navigational response module 408 to cause one or more navigational responses in vehicle 200 based on the analysis performed at step 520 and the techniques as described above in connection with FIG. 4. Navigational responses may include, for example, a turn, a lane shift, a change in acceleration, and the like. In some embodiments, processing unit 110 may use data derived from execution of velocity and acceleration module 406 to cause the one or more navigational responses. Additionally, multiple navigational responses may occur simultaneously, in sequence, or any combination thereof. For instance, processing unit 110 may cause vehicle 200 to shift one lane over and then accelerate by, for example, sequentially transmitting control signals to steering system 240 and throttling system 220 of vehicle 200. Alternatively, processing unit 110 may cause vehicle 200 to brake while at the same time shifting lanes by, for example, simultaneously transmitting control signals to braking system 230 and steering system 240 of vehicle 200.

[0122] FIG. 5B is a flowchart showing an exemplary process 500B for detecting one or more vehicles and/or pedestrians in a set of images, consistent with disclosed embodiments. Processing unit 110 may execute monocular image analysis module 402 to implement process 500B. At step 540, processing unit 110 may determine a set of candidate objects representing possible vehicles and/or pedestrians. For example, processing unit 110 may scan one or more images, compare the images to one or more predetermined patterns, and identify within each image possible locations that may contain objects of interest (e.g., vehicles, pedestrians, or portions thereof). The predetermined patterns may be designed in such a way to achieve a high rate of "false hits" and a low rate of "misses." For example, processing unit 110 may use a low threshold of similarity to predetermined patterns for identifying candidate objects as possible vehicles or pedestrians. Doing so may allow processing unit 110 to reduce the probability of missing (e.g., not identifying) a candidate object representing a vehicle or pedestrian.

[0123] At step 542, processing unit 110 may filter the set of candidate objects to exclude certain candidates (e.g., irrelevant or less relevant objects) based on classification criteria. Such criteria may be derived from various properties associated with object types stored in a database (e.g., a database stored in memory 140). Properties may include object shape, dimensions, texture, position (e.g., relative to vehicle 200), and the like. Thus, processing unit 110 may use one or more sets of criteria to reject false candidates from the set of candidate objects.

[0124] At step 544, processing unit 110 may analyze multiple frames of images to determine whether objects in the set of candidate objects represent vehicles and/or pedestrians. For example, processing unit 110 may track a detected candidate object across consecutive frames and accumulate frame-by-frame data associated with the detected object (e.g., size, position relative to vehicle 200, etc.). Additionally, processing unit 110 may estimate parameters for the detected object and compare the object's frame-by-frame position data to a predicted position.

[0125] At step 546, processing unit 110 may construct a set of measurements for the detected objects. Such measurements may include, for example, position, velocity, and acceleration values (relative to vehicle 200) associated with the detected objects. In some embodiments, processing unit 110 may construct the measurements based on estimation techniques using a series of time-based observations such as Kalman filters or linear quadratic estimation (LQE), and/or based on available modeling data for different object types (e.g., cars, trucks, pedestrians, bicycles, road signs, etc.). The Kalman filters may be based on a measurement of an object's scale, where the scale measurement is proportional to a time to collision (e.g., the amount of time for vehicle 200 to reach the object). Thus, by performing steps 540-546, processing unit 110 may identify vehicles and pedestrians appearing within the set of captured images and derive information (e.g., position, speed, size) associated with the vehicles and pedestrians. Based on the identification and the derived information, processing unit 110 may cause one or more navigational responses in vehicle 200, as described in connection with FIG. 5A, above.

[0126] At step 548, processing unit 110 may perform an optical flow analysis of one or more images to reduce the probabilities of detecting a "false hit" and missing a candidate object that represents a vehicle or pedestrian. The optical flow analysis may refer to, for example, analyzing motion patterns relative to vehicle 200 in the one or more images associated with other vehicles and pedestrians, and that are distinct from road surface motion.

IPR2025-00944
Tesla EX1002 Page 1164

Processing unit 110 may calculate the motion of candidate objects by observing the different positions of the objects across multiple image frames, which are captured at different times. Processing unit 110 may use the position and time values as inputs into mathematical models for calculating the motion of the candidate objects. Thus, optical flow analysis may provide another method of detecting vehicles and pedestrians that are nearby vehicle 200. Processing unit 110 may perform optical flow analysis in combination with steps 540-546 to provide redundancy for detecting vehicles and pedestrians and increase the reliability of system 100.

[0127] FIG. 5C is a flowchart showing an exemplary process 500C for detecting road marks and/or lane geometry information in a set of images, consistent with disclosed embodiments. Processing unit 110 may execute monocular image analysis module 402 to implement process 500C. At step 550, processing unit 110 may detect a set of objects by scanning one or more images. To detect segments of lane markings, lane geometry information, and other pertinent road marks, processing unit 110 may filter the set of objects to exclude those determined to be irrelevant (e.g., minor potholes, small rocks, etc.). At step 552, processing unit 110 may group together the segments detected in step 550 belonging to the same road mark or lane mark. Based on the grouping, processing unit 110 may develop a model to represent the detected segments, such as a mathematical model.

[0128] At step 554, processing unit 110 may construct a set of measurements associated with the detected segments. In some embodiments, processing unit 110 may create a projection of the detected segments from the image plane onto the real-world plane. The projection may be characterized using a 3rd-degree polynomial having coefficients corresponding to physical properties such as the position, slope, curvature, and curvature derivative of the detected road. In generating the projection, processing unit 110 may take into account changes in the road surface, as well as pitch and roll rates associated with vehicle 200. In addition, processing unit 110 may model the road elevation by analyzing position and motion cues present on the road surface. Further, processing unit 110 may estimate the pitch and roll rates associated with vehicle 200 by tracking a set of feature points in the one or more images.

[0129] At step 556, processing unit 110 may perform multi-frame analysis by, for example, tracking the detected segments across consecutive image frames and accumulating frame-by-frame data associated with detected segments. As processing unit 110 performs multi-frame analysis, the set of measurements constructed at step 554 may become more reliable and associated with an increasingly higher confidence level. Thus, by performing steps 550-556, processing unit 110 may identify road marks appearing within the set of captured images and derive lane geometry information. Based on the identification and the derived information, processing unit 110 may cause one or more navigational responses in vehicle 200, as described in connection with FIG. 5A, above.

[0130] At step 558, processing unit 110 may consider additional sources of information to further develop a safety model for vehicle 200 in the context of its surroundings. Processing unit 110 may use the safety model to define a context in which system 100 may execute autonomous control of vehicle 200 in a safe manner. To develop the safety model, in some embodiments, processing unit 110 may consider the position and motion of other vehicles, the detected road edges and barriers, and/or general road shape descriptions extracted from map data (such as data from map database 160). By considering additional sources of information, processing unit 110 may provide redundancy for detecting road marks and lane geometry and increase the reliability of system 100.

[0131] FIG. 5D is a flowchart showing an exemplary process 500D for detecting traffic lights in a set of images, consistent with disclosed embodiments. Processing unit 110 may execute monocular image analysis module 402 to implement process 500D. At step 560, processing unit 110 may scan the set of images and identify objects appearing at locations in the images likely to contain traffic lights. For example, processing unit 110 may filter the identified objects to construct a set of candidate objects, excluding those objects unlikely to correspond to traffic lights. The filtering may be done based on various properties associated with traffic lights, such as shape, dimensions, texture, position (e.g., relative to vehicle 200), and the like. Such properties may be based on multiple examples of traffic lights and traffic control signals and stored in a database. In some embodiments, processing unit 110 may perform multi-frame analysis on the set of candidate objects reflecting possible traffic lights. For example, processing unit 110 may track the candidate objects across consecutive image frames, estimate the real-world position of the candidate objects, and filter out those objects that are moving (which are unlikely to be traffic lights). In some embodiments, processing unit 110 may perform color analysis on the candidate objects and identify the relative position of the detected colors appearing inside possible traffic lights.

[0132] At step 562, processing unit 110 may analyze the geometry of a junction. The analysis may be based on any combination of: (i) the number of lanes detected on either side of vehicle 200, (ii) markings (such as arrow marks) detected on the road, and (iii) descriptions of the junction extracted from map data (such as data from map database 160). Processing unit 110 may conduct the analysis using information derived from execution of monocular analysis module 402. In addition, Processing unit 110 may determine a correspondence between the traffic lights detected at step 560 and the lanes appearing near vehicle 200.

[0133] As vehicle 200 approaches the junction, at step 564, processing unit 110 may update the confidence level associated with the analyzed junction geometry and the detected traffic lights. For instance, the number of traffic

IPR2025-00944
Tesla EX1002 Page 1165

lights estimated to appear at the junction as compared with the number actually appearing at the junction may impact the confidence level. Thus, based on the confidence level, processing unit 110 may delegate control to the driver of vehicle 200 in order to improve safety conditions. By performing steps 560-564, processing unit 110 may identify traffic lights appearing within the set of captured images and analyze junction geometry information. Based on the identification and the analysis, processing unit 110 may cause one or more navigational responses in vehicle 200, as described in connection with FIG. 5A, above.

[0134] FIG. 5E is a flowchart showing an exemplary process 500E for causing one or more navigational responses in vehicle 200 based on a vehicle path, consistent with the disclosed embodiments. At step 570, processing unit 110 may construct an initial vehicle path associated with vehicle 200. The vehicle path may be represented using a set of points expressed in coordinates (x, z), and the distance d_i between two points in the set of points may fall in the range of 1 to 5 meters. In one embodiment, processing unit 110 may construct the initial vehicle path using two polynomials, such as left and right road polynomials. Processing unit 110 may calculate the geometric midpoint between the two polynomials and offset each point included in the resultant vehicle path by a predetermined offset (e.g., a smart lane offset), if any (an offset of zero may correspond to travel in the middle of a lane). The offset may be in a direction perpendicular to a segment between any two points in the vehicle path. In another embodiment, processing unit 110 may use one polynomial and an estimated lane width to offset each point of the vehicle path by half the estimated lane width plus a predetermined offset (e.g., a smart lane offset).

[0135] At step 572, processing unit 110 may update the vehicle path constructed at step 570. Processing unit 110 may reconstruct the vehicle path constructed at step 570 using a higher resolution, such that the distance d_k between two points in the set of points representing the vehicle path is less than the distance d_i described above. For example, the distance d_k may fall in the range of 0.1 to 0.3 meters. Processing unit 110 may reconstruct the vehicle path using a parabolic spline algorithm, which may yield a cumulative distance vector S corresponding to the total length of the vehicle path (i.e., based on the set of points representing the vehicle path).

[0136] At step 574, processing unit 110 may determine a look-ahead point (expressed in coordinates as (x_l, z_l)) based on the updated vehicle path constructed at step 572. Processing unit 110 may extract the look-ahead point from the cumulative distance vector S, and the look-ahead point may be associated with a look-ahead distance and look-ahead time. The look-ahead distance, which may have a lower bound ranging from 10 to 20 meters, may be calculated as the product of the speed of vehicle 200 and the look-ahead time. For example, as the speed of vehicle 200 decreases, the look-ahead distance may also decrease (e.g., until it reaches the lower bound). The look-ahead time, which may range from 0.5 to 1.5 seconds, may be inversely proportional to the gain of one or more control loops associated with causing a navigational response in vehicle 200, such as the heading error tracking control loop. For example, the gain of the heading error tracking control loop may depend on the bandwidth of a yaw rate loop, a steering actuator loop, car lateral dynamics, and the like. Thus, the higher the gain of the heading error tracking control loop, the lower the look-ahead time.

[0137] At step 576, processing unit 110 may determine a heading error and yaw rate command based on the look-ahead point determined at step 574. Processing unit 110 may determine the heading error by calculating the arctangent of the look-ahead point, e.g., arctan (x_l/z_l). Processing unit 110 may determine the yaw rate command as the product of the heading error and a high-level control gain. The high-level control gain may be equal to: (2 / look-ahead time), if the look-ahead distance is not at the lower bound. Otherwise, the high-level control gain may be equal to: (2 * speed of vehicle 200 / look-ahead distance).

[0138] FIG. 5F is a flowchart showing an exemplary process 500F for determining whether a leading vehicle is changing lanes, consistent with the disclosed embodiments. At step 580, processing unit 110 may determine navigation information associated with a leading vehicle (e.g., a vehicle traveling ahead of vehicle 200). For example, processing unit 110 may determine the position, velocity (e.g., direction and speed), and/or acceleration of the leading vehicle, using the techniques described in connection with FIGS. 5A and 5B, above. Processing unit 110 may also determine one or more road polynomials, a look-ahead point (associated with vehicle 200), and/or a snail trail (e.g., a set of points describing a path taken by the leading vehicle), using the techniques described in connection with FIG. 5E, above.

[0139] At step 582, processing unit 110 may analyze the navigation information determined at step 580. In one embodiment, processing unit 110 may calculate the distance between a snail trail and a road polynomial (e.g., along the trail). If the variance of this distance along the trail exceeds a predetermined threshold (for example, 0.1 to 0.2 meters on a straight road, 0.3 to 0.4 meters on a moderately curvy road, and 0.5 to 0.6 meters on a road with sharp curves), processing unit 110 may determine that the leading vehicle is likely changing lanes. In the case where multiple vehicles are detected traveling ahead of vehicle 200, processing unit 110 may compare the snail trails associated with each vehicle. Based on the comparison, processing unit 110 may determine that a vehicle whose snail trail does not match with the snail trails of the other vehicles is likely changing lanes. Processing unit 110 may additionally compare the curvature of the snail trail (associated with the leading vehicle) with the expected curvature of the road segment in which the leading vehicle is traveling. The expected

IPR2025-00944
Tesla EX1002 Page 1166

curvature may be extracted from map data (e.g., data from map database 160), from road polynomials, from other vehicles' snail trails, from prior knowledge about the road, and the like. If the difference in curvature of the snail trail and the expected curvature of the road segment exceeds a predetermined threshold, processing unit 110 may determine that the leading vehicle is likely changing lanes.

[0140] In another embodiment, processing unit 110 may compare the leading vehicle's instantaneous position with the look-ahead point (associated with vehicle 200) over a specific period of time (e.g., 0.5 to 1.5 seconds). If the distance between the leading vehicle's instantaneous position and the look-ahead point varies during the specific period of time, and the cumulative sum of variation exceeds a predetermined threshold (for example, 0.3 to 0.4 meters on a straight road, 0.7 to 0.8 meters on a moderately curvy road, and 1.3 to 1.7 meters on a road with sharp curves), processing unit 110 may determine that the leading vehicle is likely changing lanes. In another embodiment, processing unit 110 may analyze the geometry of the snail trail by comparing the lateral distance traveled along the trail with the expected curvature of the snail trail. The expected radius of curvature may be determined according to the calculation: $(\delta_x^2 + \delta_x^2) / 2 / (\delta_x)$, where δ_x represents the lateral distance traveled and δ_z represents the longitudinal distance traveled. If the difference between the lateral distance traveled and the expected curvature exceeds a predetermined threshold (e.g., 500 to 700 meters), processing unit 110 may determine that the leading vehicle is likely changing lanes. In another embodiment, processing unit 110 may analyze the position of the leading vehicle. If the position of the leading vehicle obscures a road polynomial (e.g., the leading vehicle is overlaid on top of the road polynomial), then processing unit 110 may determine that the leading vehicle is likely changing lanes. In the case where the position of the leading vehicle is such that, another vehicle is detected ahead of the leading vehicle and the snail trails of the two vehicles are not parallel, processing unit 110 may determine that the (closer) leading vehicle is likely changing lanes.

[0141] At step 584, processing unit 110 may determine whether or not leading vehicle 200 is changing lanes based on the analysis performed at step 582. For example, processing unit 110 may make the determination based on a weighted average of the individual analyses performed at step 582. Under such a scheme, for example, a decision by processing unit 110 that the leading vehicle is likely changing lanes based on a particular type of analysis may be assigned a value of "1" (and "0" to represent a determination that the leading vehicle is not likely changing lanes). Different analyses performed at step 582 may be assigned different weights, and the disclosed embodiments are not limited to any particular combination of analyses and weights.

[0142] FIG. 6 is a flowchart showing an exemplary process 600 for causing one or more navigational responses based on stereo image analysis, consistent with disclosed embodiments. At step 610, processing unit 110 may receive a first and second plurality of images via data interface 128. For example, cameras included in image acquisition unit 120 (such as image capture devices 122 and 124 having fields of view 202 and 204) may capture a first and second plurality of images of an area forward of vehicle 200 and transmit them over a digital connection (e.g., USB, wireless, Bluetooth, etc.) to processing unit 110. In some embodiments, processing unit 110 may receive the first and second plurality of images via two or more data interfaces. The disclosed embodiments are not limited to any particular data interface configurations or protocols.

[0143] At step 620, processing unit 110 may execute stereo image analysis module 404 to perform stereo image analysis of the first and second plurality of images to create a 3D map of the road in front of the vehicle and detect features within the images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, road hazards, and the like. Stereo image analysis may be performed in a manner similar to the steps described in connection with FIGS. 5A-5D, above. For example, processing unit 110 may execute stereo image analysis module 404 to detect candidate objects (e.g., vehicles, pedestrians, road marks, traffic lights, road hazards, etc.) within the first and second plurality of images, filter out a subset of the candidate objects based on various criteria, and perform multi-frame analysis, construct measurements, and determine a confidence level for the remaining candidate objects. In performing the steps above, processing unit 110 may consider information from both the first and second plurality of images, rather than information from one set of images alone. For example, processing unit 110 may analyze the differences in pixel-level data (or other data subsets from among the two streams of captured images) for a candidate object appearing in both the first and second plurality of images. As another example, processing unit 110 may estimate a position and/or velocity of a candidate object (e.g., relative to vehicle 200) by observing that the object appears in one of the plurality of images but not the other or relative to other differences that may exist relative to objects appearing if the two image streams. For example, position, velocity, and/or acceleration relative to vehicle 200 may be determined based on trajectories, positions, movement characteristics, etc. of features associated with an object appearing in one or both of the image streams.

[0144] At step 630, processing unit 110 may execute navigational response module 408 to cause one or more navigational responses in vehicle 200 based on the analysis performed at step 620 and the techniques as described above in connection with FIG. 4. Navigational responses may include, for example, a turn, a lane shift, a change in acceleration, a change in velocity, braking, and the like. In some embodiments, processing unit 110 may use data derived from execution of velocity and ac-

IPR2025-00944
Tesla EX1002 Page 1167

celeration module 406 to cause the one or more navigational responses. Additionally, multiple navigational responses may occur simultaneously, in sequence, or any combination thereof.

[0145] FIG. 7 is a flowchart showing an exemplary process 700 for causing one or more navigational responses based on an analysis of three sets of images, consistent with disclosed embodiments. At step 710, processing unit 110 may receive a first, second, and third plurality of images via data interface 128. For instance, cameras included in image acquisition unit 120 (such as image capture devices 122, 124, and 126 having fields of view 202, 204, and 206) may capture a first, second, and third plurality of images of an area forward and/or to the side of vehicle 200 and transmit them over a digital connection (e.g., USB, wireless, Bluetooth, etc.) to processing unit 110. In some embodiments, processing unit 110 may receive the first, second, and third plurality of images via three or more data interfaces. For example, each of image capture devices 122, 124, 126 may have an associated data interface for communicating data to processing unit 110. The disclosed embodiments are not limited to any particular data interface configurations or protocols.

[0146] At step 720, processing unit 110 may analyze the first, second, and third plurality of images to detect features within the images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, road hazards, and the like. The analysis may be performed in a manner similar to the steps described in connection with FIGS. 5A-5D and 6, above. For instance, processing unit 110 may perform monocular image analysis (e.g., via execution of monocular image analysis module 402 and based on the steps described in connection with FIGS. 5A-5D, above) on each of the first, second, and third plurality of images. Alternatively, processing unit 110 may perform stereo image analysis (e.g., via execution of stereo image analysis module 404 and based on the steps described in connection with FIG. 6, above) on the first and second plurality of images, the second and third plurality of images, and/or the first and third plurality of images. The processed information corresponding to the analysis of the first, second, and/or third plurality of images may be combined. In some embodiments, processing unit 110 may perform a combination of monocular and stereo image analyses. For example, processing unit 110 may perform monocular image analysis (e.g., via execution of monocular image analysis module 402) on the first plurality of images and stereo image analysis (e.g., via execution of stereo image analysis module 404) on the second and third plurality of images. The configuration of image capture devices 122, 124, and 126-including their respective locations and fields of view 202, 204, and 206-may influence the types of analyses conducted on the first, second, and third plurality of images. The disclosed embodiments are not limited to a particular configuration of image capture devices 122, 124, and 126, or the types of analyses conducted on the first, second, and third plurality of images.

[0147] In some embodiments, processing unit 110 may perform testing on system 100 based on the images acquired and analyzed at steps 710 and 720. Such testing may provide an indicator of the overall performance of system 100 for certain configurations of image capture devices 122, 124, and 126. For example, processing unit 110 may determine the proportion of "false hits" (e.g., cases where system 100 incorrectly determined the presence of a vehicle or pedestrian) and "misses."

[0148] At step 730, processing unit 110 may cause one or more navigational responses in vehicle 200 based on information derived from two of the first, second, and third plurality of images. Selection of two of the first, second, and third plurality of images may depend on various factors, such as, for example, the number, types, and sizes of objects detected in each of the plurality of images. Processing unit 110 may also make the selection based on image quality and resolution, the effective field of view reflected in the images, the number of captured frames, the extent to which one or more objects of interest actually appear in the frames (e.g., the percentage of frames in which an object appears, the proportion of the object that appears in each such frame, etc.), and the like.

[0149] In some embodiments, processing unit 110 may select information derived from two of the first, second, and third plurality of images by determining the extent to which information derived from one image source is consistent with information derived from other image sources. For example, processing unit 110 may combine the processed information derived from each of image capture devices 122, 124, and 126 (whether by monocular analysis, stereo analysis, or any combination of the two) and determine visual indicators (e.g., lane markings, a detected vehicle and its location and/or path, a detected traffic light, etc.) that are consistent across the images captured from each of image capture devices 122, 124, and 126. Processing unit 110 may also exclude information that is inconsistent across the captured images (e.g., a vehicle changing lanes, a lane model indicating a vehicle that is too close to vehicle 200, etc.). Thus, processing unit 110 may select information derived from two of the first, second, and third plurality of images based on the determinations of consistent and inconsistent information.

[0150] Navigational responses may include, for example, a turn, a lane shift, a change in acceleration, and the like. Processing unit 110 may cause the one or more navigational responses based on the analysis performed at step 720 and the techniques as described above in connection with FIG. 4. Processing unit 110 may also use data derived from execution of velocity and acceleration module 406 to cause the one or more navigational responses. In some embodiments, processing unit 110 may cause the one or more navigational responses based on a relative position, relative velocity, and/or relative acceleration between vehicle 200 and an object detected within any of the first, second, and third plurality

IPR2025-00944
Tesla EX1002 Page 1168

of images. Multiple navigational responses may occur simultaneously, in sequence, or any combination thereof.

Smart Lane Offset

[0151] System 100 may provide driver assist functionality that monitors and adjusts the free space between vehicle 200 and boundaries (e.g., lane boundaries) within which vehicle 200 is traveling. A lane may refer to a designated or intended travel path of a vehicle and may have marked (e.g., lines on a road) or unmarked boundaries (e.g., an edge of a road, a road barrier, guard rail, parked vehicles, etc.). For instance, as a default, system 100 may maximize or increase the free space between vehicle 200 and the current lane boundaries. Further, under certain conditions, system 100 may offset vehicle 200 toward one side of the current lane of travel without leaving the lane. For example, where vehicle 200 is traveling around a curve, system 100 may decrease the offset towards the inside of the curve (e.g., adjust the position of vehicle 200 so that it is closer to the inside of the curve). As another example, where obstacles-such as parked cars, pedestrians, or cyclists-are present on one side of the lane, system 100 may increase the offset on the side of vehicle 200 where the obstacles are located (e.g., adjust the position of vehicle 200 so that it is further from the obstacles).

[0152] FIG. 8A illustrates vehicle 200 traveling on a roadway 800 in which the disclosed systems and methods for identifying lane constraints and operating vehicle 200 within the lane constraints may be used. As shown, vehicle 200 may have a first vehicle side 802, which may be a first distance 805 from a first lane constraint 810. Similarly, vehicle 200 may have a second vehicle side 812 opposite from first vehicle side 802, and second vehicle side 812 may be a second distance 815 from second lane constraint 820. In this manner, first lane constraint 810 and second lane constraint 820 may define a lane within which vehicle 200 may travel.

[0153] Processing unit 110 may be configured to determine first lane constraint 810 and second lane constraint 820 based on a plurality of images acquired by image capture device 122-126 that processing unit 110 may receive via data interface 128. According to some embodiments, first lane constraint 810 and/or second lane constraint 820 may be identified by visible lane boundaries, such as dashed or solid lines marked on a road surface. Additionally or alternatively, first lane constraint 810 and/or second lane constraint 820 may include an edge of a road surface or a barrier. Additionally or alternatively, first lane constraint 810 and/or second lane constrain 820 may include markers (e.g., Botts' dots). According to some embodiments, processing unit 110 may determine first lane constraint 810 and/or second lane constraint 820 by identifying a midpoint of a road surface width. According to some embodiments, if processing unit 110 identifies only first lane constraint 810, processing unit 110 may estimate second lane constraint 820, such as based on an estimated lane width or road width. Processing unit 110 may identify lane constraints in this manner when, for example, lines designating road lanes are not painted or otherwise labeled.

[0154] Detection of first lane constraint 810 and/or second lane constraint 820 may include processing unit 110 determining their 3D models in camera coordinate system. For example, the 3D models of first lane constraint 810 and/or second lane constraint 820 may be described by a third-degree polynomial. In addition to 3D modeling of first lane constraint 810 and/or second lane constraint 820, processing unit 110 may perform multi-frame estimation of host motion parameters, such as the speed, yaw and pitch rates, and acceleration of vehicle 200. Optionally, processing unit 110 may detect static and moving vehicles and their position, heading, speed, and acceleration, all relative to vehicle 200. Processing unit 110 may further determine a road elevation model to transform all of the information acquired from the plurality of images into 3D space.

[0155] Generally, as a default condition, vehicle 200 may travel relatively centered within first and second lane constraints 810 and/or 820. However, in some circumstances, environmental factors may make this undesirable or unfeasible, such as when objects or structures are present on one side of the road. Thus, there may be circumstances in which it may be desirable or advantageous for vehicle 200 to be closer to one lane constraint or the other. Processing unit 110 may also be configured to determine whether such circumstances, called lane offset conditions, exist based on the plurality of images.

[0156] To determine whether lane offset conditions exist, processing unit 110 may be configured to determine the presence of one or more objects 825 in the vicinity of vehicle 200. For example object 825 may comprise another vehicle, such as a car, truck or motorcycle traveling on roadway 800. Processing unit 110 may determine, for each object 825, an offset profile. The offset profile may include a determination of whether in its current and predicted position object 825 will be within a predefined range of vehicle 200. If processing unit 110 determines that object 825 is or will be within a predefined range of vehicle 200, processing unit 110 may determine whether there is enough space within first lane constraint 810 and second lane constraint 820, for vehicle 200 to bypass object 825. If there is not enough space, processing unit 110 may execute a lane change. If there is enough space, processing unit 110 may determine whether there is enough time to bypass object 825 before the distance between vehicle 200 and object 825 closes. If there is enough time, processing unit 110 may initiate and or activate the offset maneuver to bypass object 825. The offset maneuver may be executed by processing unit 110 so that the movement of vehicle 200 is smooth and comfortable for the driver. For example, the offset slew rate may be approximately 0.15 to 0.75 m/sec. The offset maneuver may be executed so that the maximum amplitude of the offset maneuver will occur when the gap be-

IPR2025-00944
Tesla EX1002 Page 1169

tween vehicle 200 and object 825 closes. If there is not enough time, processing unit 110 may initiate braking (e.g., by transmitting electronic signals to braking system 230) and then initiate offset maneuver.

[0157] After each offset profile is set for each object 825, processing unit may combine offset profiles such that the maximum lateral distance is kept from each object 825 that vehicle 200 is bypassing. That is, offset maneuvers may be modified and/or executed such that the lateral distance between vehicle 200 and each object 825 that vehicle 200 is bypassing is maximized.

[0158] Additionally or alternatively, object 825 may be a parked or stationary vehicle, a wall, or a person, such as pedestrian or cycling traffic. Processing unit 110 may determine a lane offset condition exists based on certain characteristics of object 825. For example, processing unit 110 may determine object 825 constitutes a lane offset condition if the height of object 825 exceeds a predetermined threshold, such as 10 cm from a road surface. In this manner, processing unit 110 may be configured to filter out objects that would not present a lane offset condition, such as small pieces of debris.

[0159] Additionally or alternatively, processing unit 110 may be configured to determine whether a lane offset condition exists on first vehicle side 802 based on a position of a target object 830 on vehicle first side 802. Target object 830 may include one or more objects that constitute a lane offset condition. Additionally or alternatively, target object 830 may not constitute a lane offset condition. For example, target object 830 may be a street sign or other object a sufficient distance from the travel lane or occurring with sufficient infrequency to not constitute a lane offset condition. Processing unit 110 may be configured to conclude a lane offset condition does not exist if such a target object 830 is detected.

[0160] Processing unit 110 may be configured to cause vehicle 200 to travel such that first distance 805 and second distance 810 are substantially different to avoid or stay farther away from object 825. For example, if a lane offset condition exists on first vehicle side 802, processing unit 110 may be configured to cause vehicle 200 to travel within first lane constraint 810 and second lane constraint 820 such that first distance 805 is greater than second distance 815. This is illustrated, for example, in FIG. 8A. Alternatively, if a lane offset condition exists on second vehicle side 812, processing unit 110 may create or execute an offset profile that is configured to cause vehicle 200 to travel within first lane constraint 810 and second lane constraint 820 such that first distance 805 is less than second distance 815. This is illustrated, for example, in FIG 8B.

[0161] As another example, a parked vehicle may be located on one side of vehicle 200 and a pedestrian may be located on the other side of vehicle 200. In such a situation, system 100 may determine that is desirable to reduce the distance between vehicle 200 and the parked vehicle, in order to provide more space between vehicle 200 and the pedestrian. Accordingly, processing unit 110 may cause vehicle 200 to travel such that the distance between the parked vehicle is reduced.

[0162] The identity of first vehicle side 802 and second vehicle side 812 need not be static, such that at different points in time, processing unit 110 may identify first vehicle side 802 as the driver's side or as the passenger's side of vehicle 200. For example, the identity of first vehicle side 802 may change depending on the presence of one or more objects 825 on that side of vehicle 200, such that if one lane offset condition exists based on object 825, it exists on first vehicle side 802. In such circumstances, processing unit 110 may create or execute an offset profile that is configured to cause vehicle 200 to travel within first lane constraint 810 and second lane constraint 820 such that first distance 805, which is greater than second distance 815, is provided toward object 825.

[0163] According to some embodiments, when a lane offset condition exists on first vehicle side 802, processing unit 110 may create an offset profile configured to cause vehicle 200 to travel within first lane constraint 810 and second lane constraint 820 such that first distance 805 is at least 1.25 times greater than second distance 815. Additionally or alternatively, when a lane offset condition exists on first vehicle side 802, processing unit 110 may create an offset profile that is configured to cause vehicle 200 to travel within first lane constraint 810 and second lane constraint 820 such that first distance 805 is at least 2 times greater than second distance 815.

[0164] In some embodiments, as shown in FIG. 8C, processing unit 110 may determine a lane offset condition exists on both first vehicle side 802 and on second vehicle side 812. Processing unit 110 may then create an offset profile for this offset condition that is configured to cause vehicle 200 to travel within first lane constraint 810 and second lane constraint 820 such that first distance 805 is substantially the same as second distance 815.

[0165] Processing unit 110 may be configured so that the offset profiles for moving objects 825 as well as stationary objects 825 are prioritized. For example, processing unit 110 may prioritize offset profiles to keep a predefined distance between vehicle 200 and moving objects 825 and/or static objects 825. Processing unit 200 may prioritize these offset conditions over those offset conditions based on certain lane characteristics, such as a curve. The last priority lane offset condition may include mimicking a lead vehicle.

[0166] Processing unit 110 may factor offset profiles or offset conditions into a desired look-ahead point, which may be located at the middle of the lane defined by first lane constraint 810 and second lane constraint 820. Processing unit 110 may compute and implement a yaw rate command based on the look-ahead point and the offset profiles. The yaw rate command may be used to compute a torque command based on the yaw rate command and the yaw rate of vehicle 200 measured by using a gyroscope. Processing unit 110 may transmit the

IPR2025-00944
Tesla EX1002 Page 1170

torque command to a steering system (e.g., steering system 240) of vehicle 200.

[0167] According to some embodiments, input from a driver of vehicle 200, such as via user interface 170, may override processing unit 110. For example, a user input may override the control of processing unit 110 to cause vehicle 200 to travel within first lane constraint 810 and second lane constraint 820 such that first distance 805 is greater than second distance 815.

[0168] Additionally or alternatively, processing unit 110 may be configured to determine, based on the plurality of images captured by image capture device 122, whether a lane offset condition is substantially non-existent. This may include circumstances in which one or more objects 825 below a certain size threshold are detected, one or more objects 825 detected are outside of a predetermined distance from the lane constraint, and/or when no objects 825 are detected. Under such circumstances, processing unit 110 may be configured to cause vehicle 200 to travel within first lane constraint 810 and second lane constraint 820 such that first distance 805 is equal to or substantially equal to second distance 815.

[0169] A lane offset condition may include a curved road, as illustrated in FIG. 8D. Processing unit 110 may be configured to cause vehicle 200 to travel within first lane constraint 810 and second lane constraint 820 such that first distance 805 is toward the outside of the curved road relative to vehicle 200 and second distance 815 is toward an inside of the curved road relative to vehicle 200. For example, first distance 805 may be less than second distance 815.

[0170] FIG. 9 is an exemplary block diagram of memory 140 and/or 150, which may stare instructions for performing one or more operations consistent with disclosed embodiments. As illustrated in FIG. 9, memory 140 may store one or more modules for performing the lane offset condition detection and responses described herein. For example, memory 140 may store a lane constraint module 910 and a lane offset module 920. Application processor 180 and/or image processor 190 may execute the instructions stored in any of modules 910 and 920 included in memory 140. One of skill in the art will understand that references in the following discussions to processing unit 110 may refer to application processor 180 and image processor 190 individually or collectively. Accordingly, steps of any of the following processes may be performed by one or more processing devices.

[0171] Lane constraint module 910 may store instructions which, when executed by processing unit 110, may detect and define first lane constraint 810 and second lane constraint 820. For example, lane offset module 910 may process the plurality of images received from at least one image capture device 122-124 to detect first lane constraint 810 and second lane constraint 820. As discussed above, this may include identifying painted lane lines and/or measuring a midpoint of a road surface.

[0172] Lane offset module 920 may store instructions which, when executed by processing unit 110, may detect the presence of a lane offset condition and/or identify whether lane offset condition is present on first vehicle side 802 and/or second vehicle side 812. For example, lane offset module 920 may process the plurality of images to detect the presence of object(s) 825 on first vehicle side 802 and/or second vehicle side 812. Lane offset module 920 may process the plurality of images to detect a curve in the road on which vehicle 200 is traveling. Additionally or alternatively, lane offset module 920 may receive information from another module or other system indicative of the presence of object(s) 825 and/or a curved lane. Lane offset module 920 may execute control to change the position of vehicle 200 to first lane constraint 810 and/or second lane constraint 820. For example, if lane offset module 920 determines a lane offset condition only on first vehicle side 802, lane offset module 920 may move vehicle 200 closer to second lane constraint 820. On the other hand, if lane offset module 920 determines a lane offset condition only on first vehicle side 812, lane offset module 920 may move vehicle 200 closer to first lane constraint 810.

[0173] Fig. 10 illustrates a process 1000 for navigating vehicle 200, consistent with disclosed embodiments. Process 1000 may identify lane constraints that define a lane of travel for vehicle travel. Process 1000 include causing vehicle 200 to travel closer to one or the other lane constraint, such as first lane constraint 810 or second lane constraint 820, in response to detecting a lane offset condition. In this manner, process 1000 may be used to drive vehicle 200 within lane constraints as well as account for environmental factors that may make it more desirable to shift vehicle 200 within first lane constraint 810 and second lane constraint 820.

[0174] At step 1010, process 1000 may include acquiring, using at least one image capture device 122, 124, and/or 126, a plurality of images of an area in the vicinity of vehicle 200. For example, processing unit 110 may receive the plurality of images may through data interface 128. For example, processing unit 110 may be configured to determine the presence of one or more objects 825 on the side of the road. For example, object 825 may be a parked or stationary vehicle, a wall, or a person, such as pedestrian or cycling traffic. Processing unit 110 may determine if a lane offset condition exists based on certain characteristics of object 825. For example, processing unit 110 may determine that object 825 constitutes a lane offset condition if the height of object 825 exceeds a predetermined threshold, such as 10 cm from a road surface. In this manner, processing unit 110 may be configured to filter out objects that would not present a lane offset condition, such as small pieces of debris.

[0175] Additionally or alternatively, processing unit 110 may be configured to determine whether a lane offset condition exists on first vehicle side 802 based on a position of a target object 830 on vehicle first side 802. Target object 830 may be one or more objects that constitute a lane offset condition. Additionally or alternatively, target object 830 may not constitute a lane offset condition. For

IPR2025-00944
Tesla EX1002 Page 1171

example, target object 830 may be a road barrier or a street sign. Processing unit 110 may be configured to conclude a lane offset condition does not exist if such a target object 830 is detected.

[0176] At step 1020, process 1000 may determine from the plurality of images a first lane constraint on first vehicle side 802 and a second lane constraint on second vehicle side 812. For example, processing unit 110 may be configured to determine first lane constraint 810 and second lane constraint 820 based on the plurality of images received via data interface 128.

[0177] At step 1030, process 1000 may determine whether a lane offset condition exists on first vehicle side 802. If a lane offset condition does exist on first vehicle side 802, at step 1040, process 1000 may include causing vehicle 200 to travel within first lane constraint 810 and second lane constraint 820 such that second distance 815 is less than first distance 805. Processing unit 110 may be configured to cause vehicle 200 to travel such that first distance 805 and second distance 810 are substantially different to avoid or stay farther away from object 825. For example, if a lane offset condition exists on first vehicle side 802, processing unit 110 may be configured to cause vehicle 200 to travel within first lane constraint 810 and second lane constraint 820 such that first distance 805 is greater than second distance 815. This is illustrated, for example, in FIG. 8A. Alternatively, if a lane offset condition exists on second vehicle side 812, processing unit 110 may be configured to cause vehicle 200 to travel within first lane constraint 810 and second lane constraint 820 such that first distance 805 is less than second distance 815. This is illustrated, for example, in FIG 8B. For example, to navigate vehicle 200 according to process 1000, processing unit 110 may transmit electronic signals to one or more of throttling system 220, braking system 230, and/or steering system 240 of vehicle 200.

[0178] If a lane offset condition does not exist on first vehicle side 802, in step 1050, method may include determining whether lane offset condition exists on second vehicle side 812. If a lane offset condition exists on second vehicle side 812, the method may include causing vehicle 200 to travel between first lane constraint 810 and second lane constraint 820 such that first distance 805 is less than second distance 815.

Navigating a Vehicle to a Default Lane

[0179] System 100 may provide driver assist functionality that monitors the location of vehicle 200 in a current lane of travel and moves vehicle 200 into a predetermined default travel lane if the current lane of travel is not the predetermined default travel lane. The predetermined default travel lane may be defined by a user input and/or as a result of processing images received from at least one image capture device 122-126.

[0180] System 100 may also provide notification of an impending lane change of vehicle 200 to the predetermined default travel lane to the driver and passengers of vehicle 200 and/or other drivers. For example, system 100 may activate a turn signal positioned at or near one of bumper regions 210. Additionally or alternatively, system 100 may sound an audible notification, such as through speakers 360. Vehicle 200 may use system 100 to select a particular lane as predetermined default travel lane when vehicle is traveling. When the current lane within which vehicle 200 is traveling-e.g., the current lane of travel-differs from the predetermined default travel lane, system 100 may cause vehicle 200 to make a lane change, e.g., into the predetermined default travel lane.

[0181] FIG. 11 illustrates vehicle 200 traveling in a current lane of travel 1110 of roadway 800. According to some embodiments, vehicle 200 may have a predetermined default travel lane 1120. For example, predetermined default travel lane 1120 may be the right-most lane of roadway 800. Navigation system 100 may be configured to determine whether vehicle 200 is traveling in predetermined default travel lane 1120 and, if not, cause vehicle 200 to change to predetermined default travel lane 1120. For example, processing unit 110 may be configured to compare current lane of travel 1110 with predetermined default travel lane 1120 and cause vehicle 200 to return to predetermined default travel lane 1120.

[0182] According to some embodiments, processing unit 110 may detect an object 1130 in predetermined default travel lane 1120 based on a plurality of images acquired by one or more of image capture devices 122-126. Processing unit 110 may detect the position and speed of vehicle 200 relative to object 1130 by analyzing the images using techniques described in connection with FIGS. 5A-5D, above. According to some embodiments, processing unit 110 may determine whether vehicle 200 should pass object 1130. This may be desirable if, for example, the speed and/or acceleration of vehicle 200 relative to object 1130 exceeds a certain threshold.

[0183] FIG. 12A is an exemplary block diagram of memory 140, which may store instructions for performing one or more operations consistent with disclosed embodiments. As illustrated in FIG. 12A, memory 140 may store one or more modules for performing the default travel lane and current travel lane detection and responses described herein. For example, memory 140 may store a current lane detection module 1210, a default lane detection module 1220, and an action module 1230.

[0184] Current lane detection module 1210 may store instructions which, when executed by processing unit 110, may detect and define current lane of travel 1110. For example, processing unit 110 may execute current lane detection module 1212 to process the plurality of images received from at least one image capture device 122-124 and detect current lane of travel 1110. This may include detecting the relative position of vehicle 200 with respect to the edge of roadway 800, a midpoint of roadway 800, and/or lane markers, such as painted lane lines. Processing unit 110 may detect the position of vehicle

IPR2025-00944
Tesla EX1002 Page 1172

200 relative to roadway 800 by analyzing the images using techniques described in connection with FIGS. 5A-5D, above.

[0185] Default lane detection module 1220 may store instructions which, when executed by processing unit 110, may identify predetermined default travel lane 1120 and compare default lane detection module 1220 to current lane of travel 1110 identified by current lane detection module 1210. Default lane detection module 1220 may identify predetermined default travel lane 1120 based upon a user input and/or the plurality of images received from at least one image capture device 122-124. For example, a user may select a default lane through voice commands received by speakers 360 or by making a selection on a menu displayed on touch screen 320.

[0186] In some embodiments, predetermined default travel lane 1120 may be the right-most travel lane, or it may be the lane that has the more desirable features. For example, default lane detection module 1220 may include instructions to select the lane with less traffic, the lane in which vehicles are traveling closest to speed of vehicle 200, and/or a lane closest to an approaching exit or street to which vehicle 200 will turn. In some embodiments, predetermined default travel lane 1120 may be determined based on data received from position sensor 130, such as a global positioning system.

[0187] Action module 1230 may store instructions which, when executed by processing unit 110, may cause a response to the identification of current lane of travel 1110 and predetermined default travel lane 1120. For example, processing unit 110 may execute action module 1230 to cause vehicle 200 to change lanes if current lane of travel 1110 is not the same as predetermined default travel lane 1120. Additionally or alternatively, action module 1230 may initiate a notification of the lane change, such as activating turn signal 210 and or causing an audible announcement.

[0188] Action module 1230 may include instructions that cause processing unit 110 to determine whether vehicle 200 should move out of predetermined default travel lane 1120. For example, action module 1230 may determine that object 1130 in predetermined default travel lane 1120 should be bypassed by vehicle 200.

[0189] Action module 1230 may further include instructions that cause processing unit 110 to determine whether it is safe to make a lane change. For example, prior to executing a lane change, processing unit 110 may evaluate determine one or more characteristics of roadway 800 based on, for example, analysis of the plurality of images conducted using techniques described in connection with FIGS. 5A-5D, above. If the analysis indicates conditions are not safe to change lanes, processing unit 110 may determine that vehicle 200 should stay in the current lane.

[0190] Action module 1230 may include additional instructions for manipulating vehicle 200 to maneuver out of and return to predetermined default travel lane 1120, such as for bypassing object 1130. When executing these instructions, processing unit 110 may estimate a location of predetermined default travel lane 1120 using estimated motion of vehicle 200. This may be advantageous in circumstances in which predetermined default travel lane 1120 does not remain detectable by image capture devices 122-124.

[0191] Processing unit 110 may be configured to execute an offset profile that offsets vehicle 200 from predetermined default travel lane 1120 by a full lane width, which may be estimated based on the distance between detected lane markers. The offset maneuver defined by offset profile may be executed by processing unit 110 so that the movement of vehicle 200 is smooth and comfortable for the driver. For example, the offset slew rate may be approximately 0.15 to 0.75 m/sec. During the offset maneuver, processing unit 110 may execute instructions from action module 1230 that switch to following the center of current lane of travel 1110. These instructions may include gradually transitioning from following the predicted lane with offset to following current lane of travel 1110. A gradual transition may be accomplished by defining desired look-ahead point so that it is not allowed to change in time more than the recommended slew rate, which may be around 0.25 m/sec.

[0192] The instructions for transitioning vehicle 200 from predetermined default travel lane 1120 may be used to transition vehicle 200 back into predetermined default travel lane. For example, processing unit 110 may be configured to execute an offset profile that offsets vehicle 200 from current lane of travel 1110 by a full lane width, which may be estimated based on the distance between detected lane markers. The offset maneuver defined by offset profile may be executed by processing unit 110 so that the movement of vehicle 200 is smooth and comfortable for the driver. For example, the offset slew rate may be approximately 0.15 to 0.75 m/sec. During the offset maneuver, processing unit 110 may execute instructions from action module 1230 that switch to following the center of predetermined default travel lane 1120. These instructions may include gradually transitioning from following the predicted lane with offset to following predetermined default travel lane 1120. A gradual transition may be accomplished by defining desired look-ahead point so that it is not allowed to change in time more than the recommended slew rate, which may be around 0.25 m/sec.

[0193] At one or more of the steps for transitioning vehicle 200 between two lanes, processing unit 110 may compute and implement a yaw rate command based on the look-ahead point and the offset profiles. The yaw rate command may be used to compute a torque command based on the yaw rate command and the yaw rate of vehicle 200 measured using a gyroscope. Processing unit 110 may transmit a torque command to a steering system (e.g., steering system 240) of vehicle 200.

[0194] FIG. 12B illustrates a process 1250 for navigating vehicle 200 to a default lane consistent with disclosed embodiments. According to some embodiments, proc-

IPR2025-00944
Tesla EX1002 Page 1173

ess 1250 may be implemented by one or more components of navigation system 100, such as at least one processing unit 110. Process 1250 may identify current lane of travel 1110 of vehicle 200 and compare it with predetermined default travel lane 1120 to determine whether vehicle 200 is traveling within predetermined default travel lane 1120. If vehicle 200 is not traveling in predetermined default travel lane 1120, processing unit 110 may cause vehicle 200 to change lanes so that it navigates to default travel lane 1120.

[0195] At step 1260, one or more of image capture devices 122-124 may acquire a plurality of images of an area in a vicinity of vehicle 200. Processing unit 110 may receive the plurality of images via data interface 128.

[0196] At step 1270, processing unit 110 may execute current lane detection module 1210 detect and define current lane of travel 1110. For example, processing unit 110 may process the plurality of images to, for example, detect the relative position of vehicle 200 with respect to the edge of roadway 100, a midpoint of roadway 100, and/or painted lane markers. Processing unit 110 may detect the position of vehicle 200 relative to roadway 800 by analyzing the images using techniques described in connection with FIGS. 5A-5D, above.

[0197] At step 1280, processing unit 110 may execute default lane detection module 1220 to identify predetermined default travel lane 1120 and compare default lane detection module 1220 to current lane of travel 1110 identified by current lane detection module 1210. Default lane detection module 1220 may identify predetermined default travel lane 1120 based upon a user input and/or the plurality of images received from at least one image capture device 122-124.

[0198] For example, as discussed above, processing unit 110 may determine predetermined default travel lane 1120 based on any number of factors. Predetermined default travel lane 1120 may be the right-most lane among the plurality of travel lanes 1110 and 1120. Additionally or alternatively, predetermined default travel lane 1120 may be determined based on an input received from a user via user interface 170, by processing the plurality of images, and/or based on data received from position sensor 130, such as a global positioning system. Predetermined default travel lane 1120 may be determined once, once per vehicle trip, or on regular intervals. For example, predetermined default travel lane 1120 may be determined dynamically in response to conditions at a given location.

[0199] If current lane 1110 is not the same as predetermined default travel lane 1120, processing unit 110 may execute action module 1230 to cause vehicle 100 to navigate to default travel lane 1120. For example, processing unit 110 may transmit electronic signals to one or more of throttling system 220, braking system 230, and/or steering system 240 of vehicle 200 to trigger the desired response. For example, processing unit 110 may cause steering system 240 to turn the steering wheel of vehicle 200 to achieve a rotation of a predetermined angle.

[0200] Further, prior to executing a lane change, processing unit 110 may determine whether it is safe (e.g., there are no vehicles or objects in the way) to change lanes. As discussed above, processing unit 110 may evaluate one or more characteristics of roadway 800 based on, for example, analysis of the plurality of images conducted using techniques described in connection with FIGS. 5A-5D, above. If conditions are not safe to change lanes, processing unit 110 may determine that vehicle 200 should stay in the current lane.

[0201] In some embodiments, execution of action module 1230 may initiate a notification of the lane change. For example, processing unit 110 may cause a signal to activate a turn signal and/or cause an audible announcement via speakers 360. Accordingly, such notification may occur prior to vehicle 200 changing lanes. In other embodiments, execution of action module 1230 may initiate a notification to a driver of vehicle 200 that vehicle 200 is not traveling in a default lane without causing the lane change to occur. Accordingly, in such an embodiment, the driver of vehicle 200 may determine whether to manually take action and steer vehicle 200 to change lanes.

[0202] In some embodiments, system 100 may notify the user of vehicle 200 and/or other drivers in the vicinity of vehicle 200 that vehicle 200 will be changing lanes if vehicle 200 will be moved to predetermined default travel lane 1120. For example, processing unit 110 may activate a turn signal of vehicle 200 prior to causing vehicle 200 to change lanes. Additionally or alternatively, an audible announcement to notify other drivers in the vicinity may be caused by processing unit 110. The audible announcement may be transmitted via, for example, Bluetooth, to receivers included in other nearby vehicles. Additionally or alternatively, at least one processing unit 110 may cause an audible indicator to notify the driver of vehicle 200 prior to causing vehicle 200 to switch lanes. In this manner, driver of vehicle 200 can anticipate the control system 100 will exert on vehicle 200.

Controlling Velocity of a Turning Vehicle

[0203] System 100 may provide driver assist functionality that controls the velocity (i.e., speed and/or direction) of vehicle 200 in different scenarios, such as while making a turn (e.g., through a curve). For example, system 100 may be configured to use a combination of map, position, and/or visual data to control the velocity of vehicle 200 when turning. In particular, system 100 may control the velocity of vehicle 200 as it approaches a curve, while navigating a curve, and/or as it exits a curve. System 100 may consider different factors to control the velocity depending on the position of vehicle 200 relative to the curve at a given time. For example, system 100 may initially cause vehicle 200 to slow down in response to learning that a curve is located a certain distance ahead of vehicle 200. While vehicle 200 turns through

IPR2025-00944
Tesla EX1002 Page 1174

the curve, however, system 100 may cause vehicle 200 to accelerate and/or charge direction based on an analysis of the characteristics of the curve.

[0204] FIGS. 13A and 13B are diagrammatic representations of a vehicle (e.g., vehicle 200) approaching, and navigating, a curve with one or more characteristics on road 1300, consistent with the disclosed embodiments. As illustrated in FIG. 13A, visual indicators, such as road sign 1302 located near road 1300, may inform vehicle 200 that a curve lies ahead. In response to the warning provided by road sign 1302, a driver of vehicle 200 and/or a system 100 providing driver assist functionality of vehicle 200 may adjust the velocity of vehicle 200 (e.g., by slowing down and/or steering) to safely navigate the curve.

[0205] As vehicle 200 approaches the curve, system 100 may detect the lane width 1306 and the curve radius 1308 by using image capture devices 122 and/or 124 (which may be cameras) of vehicle 200. Based on the curve radius 1308, the lane width 1306, and other characteristics associated with the curve detected using image capture devices 122 and/or 124, system 100 may adjust the velocity of vehicle 200. Specifically, system 100 may adjust the velocity as vehicle 200 approaches the curve and/or while vehicle 200 navigates the curve. System 100 may also adjust the velocity as vehicle 200 exits the curve, in response to, for example, traffic light 1304 appearing after the curve. Further detail regarding adjustments to the velocity of vehicle 200 and detection of the lane width 1306, the radius of curvature or curve radius 1308, and characteristics associated with the curve are described in connection with FIGS. 14 and 15, below.

[0206] In some embodiments, system 100 may recognize a curve to be navigated based on map data and/or vehicle position information (e.g., GPS data). System 100 may determine an initial target velocity for the vehicle based on one or more characteristics of the curve as reflected in the map data. System 100 may adjust a velocity of the vehicle to the initial target velocity and determine, based on one or more images acquired by one or more of image capture devices 122-126, one or more observed characteristics of the curve. System 100 may determine an updated target velocity based on the one or more observed characteristics of the curve and adjust the velocity of the vehicle to the updated target velocity.

[0207] FIG. 14 is an exemplary block diagram of a memory configured to store instructions for performing one or more operations consistent with the disclosed embodiments. As indicated in FIG. 14, memory 140 may store a curve recognition module 1402, a characteristic observation module 1404, and a velocity module 1406. The disclosed embodiments are not limited to any particular configuration of memory 140. Further, processing unit 110 may execute the instructions stored in any of modules 1402-1406 included in memory 140.

[0208] In one embodiment, curve recognition module 1402 may store software instructions (such as computer vision software) which, when executed by processing unit 110, detects a curve by, for example, analyzing one or more sets of images and/or using map and/or position data (e.g., data stored in map database 160). The image sets may be acquired by one or more of image capture devices 122, 124, and 126. Based on the analysis, system 100 (e.g., via processing unit 110) may cause a change in velocity of vehicle 200. For example, processing unit 110 may cause vehicle 200 to reduce its speed by a predetermined amount and/or adjust its direction by a particular angle if processing unit 110 detects the presence of a curve located within a minimum threshold distance ahead of vehicle 200.

[0209] In one embodiment, characteristic observation module 1404 may store software instructions (such as computer vision software) which, when executed by processing unit 110, observes characteristics associated with a curve detected by system 100 (e.g., via execution of curve recognition module 1402). Processing unit 110 may observe characteristics associated with a detected curve by analyzing one or more sets of images acquired by one or more of image capture devices 122, 124, and 126 and/or using map and/or position data (e.g., data stored in map database 160). The analysis may yield information associated with the curve, such as curve radius 1308, lane width 1306, a degree of curvature, a rate of change in curvature, a degree of banking, a length or arc length of the curve, and the like. These characteristics may include estimated values (e.g., curve characteristics calculated using preexisting data and/or mathematical models based on such data) and actual values (e.g., curve characteristics calculated by analyzing captured images of the curve). Based on the analysis, processing unit 110 may cause a change in velocity of vehicle 200. For example, processing unit 110 may cause vehicle 200 to reduce its speed and/or adjust its direction by a particular angle in view of a high rate of change in curvature associated with a curve located 25 meters ahead of vehicle 200.

[0210] In one embodiment, velocity module 1406 may store software instructions configured to analyze data from one or more computing and electromechanical devices configured to determine a target velocity and cause a change in velocity of vehicle 200 to the target velocity. For example, processing unit 110 may execute velocity module 1406 to calculate a target velocity for vehicle 200 based on data derived from execution of curve recognition module 1402 and characteristic observation module 1404. Such data may include, for example, an initial target position and initial velocity, an updated target position and updated velocity, a final target position and final velocity, and the like. In addition, processing unit 110 may calculate a velocity for vehicle 200 based on input from other systems of vehicle 200, such as a throttling system 220, braking system 230, and/or steering system 240 of vehicle 200. Based on the calculated target velocity, processing unit 110 may transmit electronic signals (e.g., via a controller area network bus ("CAN bus")) to throttling

IPR2025-00944
Tesla EX1002 Page 1175

system 220, braking system 230, and/or steering system 240 of vehicle 200 to trigger a change in the speed and/or direction of vehicle 200 by, for example, physically depressing the brake, easing up off the accelerator of vehicle 200, or steering vehicle 200 in a particular direction (e.g., at a particular angle). An increase in speed may be associated with a corresponding acceleration (e.g., between 0.2 m/sec^2 and 3.0 m/sec^2). Conversely, a decrease in speed may be associated with a corresponding deceleration. In addition, the acceleration or deceleration may be based on the road type (e.g., highway, city street, country road, etc.), the presence of any speed constraints nearby (e.g., sharp curves, traffic lights), as well as a road width or lane width 1306. For purposes of this disclosure, deceleration may refer to an acceleration having a negative value.

[0211] FIG. 15 is a flow chart showing an exemplary process 1500 for controlling the velocity of a vehicle based on a detected curve and observed characteristics of the curve and/or map data regarding the curve consistent with disclosed embodiments. At step 1510, processing unit 110 may receive a plurality of images via data interface 128 between processing unit 110 and image acquisition unit 120. For instance, a camera included in image acquisition unit 120 (such as image capture device 122) may capture a plurality of images and transmit them over a digital connection (e.g., USB, wireless, Bluetooth, etc.) to processing unit 110. In some embodiments, processing unit 110 may receive more than one plurality of images via a plurality of data interfaces. For example, processing unit 110 may receive a plurality of images from each of image capture devices 122, 124, 126, each of which may have an associated data interface for communicating data to processing unit 110. The disclosed embodiments are not limited to any particular data interface configurations or protocols.

[0212] At step 1520, processing unit 110 may execute curve recognition module 1502 to detect a curve located ahead of vehicle 200 based on, for example, map and/or position data (e.g., data stored in map database 160). Such data may indicate the location of a curve that is present on a given road. The data may indicate the absolute position of a particular curve (e.g., via GPS coordinates) or the relative position of the curve (e.g., by describing the curve relative to vehicle 200, road sign 1302, traffic light 1304, and/or other descriptive landmarks). In some embodiments, processing unit 110 may execute curve recognition module 1502 to detect the curve by analyzing the plurality of images. The analysis may be conducted using techniques described in connection with FIGS. 5A-5D and 6, above.

[0213] At step 1530, processing unit 110 may execute velocity module 1406 to determine an initial target velocity for vehicle 200 based on, for example, data derived from execution of curve recognition module 1402. The initial target velocity may reflect a speed and a direction for vehicle 200 to safely enter an upcoming curve and may be determined using techniques described in connection with FIG. 14, above. For example, velocity module 1406 may determine the initial target velocity based on known or estimated characteristics associated with the upcoming curve, such as the curve radius 1308 and the degree of banking. Velocity module 1406 may calculate the initial target velocity in accordance with mathematical models and/or formulae and based on various constraints. In some embodiments, for example, velocity module 1406 may determine the initial target velocity based on the presence of traffic lights (such as traffic light 1304) before, during, or after a curve, weather conditions (e.g., light rain, heavy rain, windy, clear, etc.), road conditions (e.g., paved road, unpaved road, potholes, etc.), posted speed limits, the presence of vehicles or pedestrians nearby, etc. The initial target velocity may also be based on a lateral acceleration limit associated with vehicle 200. In some embodiments, the lateral acceleration limit may be in the range of 0.2 m/sec^2 and 3.0 m/sec^2 and may be adjusted depending on the curve characteristics, the road width, or lane width 1306. Additionally, the initial target velocity may be based on a focal range associated with one or more of cameras, such as image capture devices 122 and 124.

[0214] For example, velocity module 1406 may determine the initial target velocity based on the radius of curvature of the road ahead of vehicle 200. Velocity module 1406 may determine the radius of curvature at various distances from vehicle 200 and store the information in a vector R (1:n). In one embodiment, velocity module 1406 may determine the radius of curvature at: (1) the current position of vehicle 200, (2) a distance D_{end} from vehicle 200, where D_{end} represents the distance ahead of vehicle 200 at which knowledge of the road characteristics allows system 100 to control navigational responses in vehicle 200 such that the driver may feel comfortable, and (3) at incremental distances from vehicle 200, such as every 1 meter, etc. The value of distance D_{end} may depend on the current speed of vehicle 200, V_{ego}. For example, the greater the value of V_{ego} (e.g., the faster vehicle 200 is traveling), the greater the value of D_{end} (e.g., the longer the distance may be taken into account to provide system 100 with sufficient time to provide a navigational response, such as comfortably reducing the speed of vehicle 200 based on an approaching sharp turn or curve). In one embodiment, D_{end} may be calculated according to the following equation: $D_{end} = (V_{ego}^2/2)/(0.7/40*V_{ego}+0.5)$. Thus, the vector $D(1{:}n) = [1{:}1{:}n]$, where $n = \text{floor}(D_{end})$.

[0215] In some embodiments, velocity module 1406 may determine the initial target velocity based on a lateral acceleration constraint, a_{max}, e.g., the maximum allowed lateral acceleration, a_{max} may fall in the range of 1.5 to 3.5 meters / sec^2. Based on a_{max} and the radius of curvature vector R (discussed above), velocity module 1406 may limit the speed of vehicle 200 (e.g., at different points or distances from the current position of vehicle 200) to a speed v_{lim}, calculated as: $v_{lim}(1{:}n)=\text{sqrt}(a_{max}*R(1{:}n))$. In addition, a deceleration constraint that follows from

IPR2025-00944
Tesla EX1002 Page 1176

the determination of v_{lim} may be calculated as: dec_{lim} = min{{$(v_{lim}(1:n)^2 - V_{ego}^2)/2$), max{$d_{min}$,$D(1:n)-2*V_{ego}$}}, where d_{min} is equal to 40 meters. The value of d_{min} may represent a point at which deceleration feels comfortable for the driver of vehicle 200.

[0216] In some embodiments, velocity module 1406 may determine the initial target velocity based on a curvature slew rate constraint, $slew_{max}$, which may be based on the curvature rate of change. $slew_{max}$ may fall in the range of 0.5 to 1.5 meters / sec^3. Velocity module 1406 may determine v_{lim} according the following equation: $v_{lim}(1:n) = (slew_{max}/\text{abs}\{\text{diff}\{1/R(1:n)\}\})^{1/3}$. Further, velocity module 1406 may determine dec_{lim} according to the equation: dec_{lim} = min{{$(v_{lim}(1:n)^2 - V_{ego}^2)/2$), max {$d_{min}$, $D(1:n) - 2*V_{ego}$}}, where d_{min} is equal to 40 meters.

[0217] In some embodiments, velocity module 1406 may determine a comfort constraint such that the driver of vehicle 200 is comfortable while vehicle 200 is undergoing various navigational responses, such as making a turn or navigating a curve. Velocity module 1406 may determine v_{lim} according to the following equation: $v_{lim}(1:n) = 4/500*R(1:n) + 22.5$. Further, velocity module 1406 may determine dec_{lim} according to the equation: dec_{lim} = min{{$(v_{lim}(1:n)^2 - V_{ego}^2)/2$), max {$d_{min}$, $D(1:n) - 2*V_{ego}$}}, where d_{min} is equal to 40 meters.

[0218] Velocity module 1406 may merge the three constraints determined above-the lateral acceleration constraint, the curvature slew rate constraint, and comfort constraint-by using the smallest value of dec_{lim} calculated in association with each constraint. After merging the constraints, velocity module 1406 may arrive at the initial target velocity by determining how much to adjust the speed of vehicle 200, denoted v_{com}, according to the following equation: $v_{com} = v_{com(prev)} + dec_{lim}*\Delta t$, where $v_{com(prev)}$ represents the prior speed of vehicle 200 (e.g., a previously determined speed, the speed of vehicle 200 prior to any adjustment by velocity module 1406, etc.), and Δt represents the time between the determination of v_{com} and $v_{com(prev)}$.

[0219] At step 1540, processing unit 110 may execute velocity module 1406 to adjust a velocity of vehicle 200 to the initial target velocity determined at step 1530. Based on the current velocity of vehicle 200 and the initial target velocity determined at step 1530, processing unit 110 may transmit electronic signals to throttling system 220, braking system 230, and/or steering system 240 of vehicle 200 to change the velocity of vehicle 200 to the initial target velocity. Further, one or more actuators may control throttling system 220, braking system 230, and/or steering system 240. For example, processing unit 110 may transmit electronic signals that cause system 100 to physically depress the brake by a predetermined amount or ease partially off the accelerator of vehicle 200. Further, processing unit 110 may transmit electronic signals that cause system 100 to steer vehicle 200 in a particular direction. In some embodiments, the magnitude of acceleration (or de-acceleration) may be based on the difference between the current velocity of vehicle 200 and the initial target velocity. Where the difference is large (e.g., the current velocity is 10 km/hr greater than the initial target velocity), for example, system 100 may change the current velocity by braking in a manner that maximizes deceleration and results in vehicle 200 achieving the initial target velocity in the shortest amount of time safely possible. Alternatively, system 100 may apply the brakes in a manner that minimizes deceleration and results in vehicle 200 achieving the initial target velocity gradually. More generally, system 100 may change the current velocity to the initial target velocity according to any particular braking profile (e.g., a braking profile calling for a high level of braking for the first two seconds, and a low level of braking for the subsequent three seconds). The disclosed embodiments are not limited to any particular braking profile or manner of braking.

[0220] At step 1550, processing unit 110 may execute characteristic observation module 1404 to determine one or more characteristics of a curve located ahead of vehicle 200 based on, for example, analysis of the plurality of images conducted using techniques described in connection with FIGS. 5A-5D, above. The characteristics may include, for example, estimated and/or actual values associated with curve radius 1308, a road width or lane width 1306, a degree of curvature, a rate of change in curvature, a degree of banking, and the like. The characteristics may also include, for example, information conveyed by one or more road signs (such as road sign 1302). Furthermore, the characteristics may include the length or arc length of the detected curve, which may be a function of the speed of vehicle 200 as it approaches the curve. Processing unit 110 may execute characteristic observation module 1404 to calculate values associated with these characteristics (e.g., curve radius 1308, the degree of curvature, the rate of change in curvature, the arc length of the curve, etc.) based on mathematical models and/or formulae. In some embodiments, processing unit 110 may execute characteristic observation module 1404 to construct a mathematical model expressing the curvature of the curve as a function of distance (e.g., distance traveled). Such a model may be subject to constraints such as a maximum lateral acceleration constraint (e.g., in the range of 1.5 to 3.0 m/sec^2) and/or a maximum lateral acceleration derivative constraint (e.g., in the range of 0.8 to 1.2 m/sec^2). These constraints may yield a maximum speed for vehicle 200 as a function of distance traveled.

[0221] At step 1560, processing unit 110 may execute velocity module 1406 to calculate an updated target velocity based on the characteristics determined at step 1550. The updated target velocity may reflect an updated speed and/or direction for vehicle 200 to safely perform any combination of: (i) entering an upcoming curve, (ii) navigating through a curve, or (iii) exiting a curve. The updated target velocity may be determined using techniques described in connection with FIG. 14 and step 1530, above. For example, velocity module 1406 may determine the updated target velocity based on charac-

IPR2025-00944
Tesla EX1002 Page 1177

teristics associated with a curve and mathematical models and/or formulae.

[0222] At step 1570, processing unit 110 may execute velocity module 1406 to adjust the velocity of vehicle 200 to the updated target velocity determined at step 1560. This adjustment may be accomplished using techniques described in connection with step 1540, above. For example, based on the current velocity of vehicle 200 and the updated target velocity determined at step 1560, processing unit 110 may transmit electronic signals (e.g., via a CAN bus) to throttling system 220, braking system 230, and/or steering system 240 of vehicle 200 to change the velocity of vehicle 200 to the updated target velocity. At this point, the current velocity of vehicle 200 is not necessarily the same as the initial target velocity determined at step 1530. For example, processing unit 110 may have determined the updated target velocity at step 1560 while in the process of adjusting the velocity of vehicle 200 to the initial target velocity. This may be true in the case where the difference in the velocity of vehicle 200 and the initial target velocity is large (e.g., greater than 10 km/hr) and the associated acceleration is small.

[0223] In some embodiments, system 100 may make regular updates to the target velocity based on continued observations of characteristics of a curve (e.g., based on image data depicting portions of the curve that vehicle 200 is approaching). Additionally, after system 100 has navigated vehicle 200 through a curve, system 100 may cause vehicle 200 to accelerate to a new target velocity suitable for traveling a segment of the road that does not have curves, as determined by analyzing map data, positional data, and/or image data acquired by system 100.

Mimicking A Leading Vehicle

[0224] System 100 may provide driver assist functionality that causes vehicle 200 to mimic (or decline to mimic) a leading vehicle in different scenarios, such as when the leading vehicle switches lanes, accelerates, or makes a turn. For example, system 100 may be configured to detect the leading vehicle by analyzing a plurality of images and determine one or more actions taken by the leading vehicle. In some scenarios, such as when the leading vehicle is turning at an intersection, system 100 may be configured to cause vehicle 200 to decline to mimic the turn. In other scenarios, such as when the leading vehicle is turning at an intersection without changing lanes within the intersection, system 100 may be configured to cause vehicle 200 to mimic the turn. System 100 may also be configured to cause vehicle 200 to mimic one or more actions of the leading vehicle based on a navigation history of the leading vehicle.

[0225] FIGS. 16A and 16B are diagrammatic representations of a primary vehicle 200a mimicking one or more actions of a leading vehicle 200b on road 1600a, consistent with the disclosed embodiments. As illustrated in FIGS. 16A and 16B, primary vehicle 200a may be trailing leading vehicle 200b while traveling in the same lane of road 1600a. The lane may have a left edge 1620, a right edge 1630, and a midpoint 1610, with w_1 and w_2 indicating the left and right halves of the lane, respectively. Primary vehicle 200a and leading vehicle 200b may be positioned in the lane such that neither vehicle is centered on the midpoint 1610. As shown in FIG. 16A, for example, primary vehicle 200a may be centered to the right of midpoint 1610 such that the distance c_1 between primary vehicle 200a and left edge 1620 is greater than the distance c_2 between primary vehicle 200a and right edge 1630. In contrast, leading vehicle 200b may be centered to the left of midpoint 1610 such that the distance c_3 between leading vehicle 200b and left edge 1620 is less than the distance c_4 between leading vehicle 200b and right edge 1630. Different configurations of the positions of primary vehicle 200a and leading vehicle 200b are possible. For example, primary vehicle 200a and leading vehicle 200b may be positioned anywhere on road 1600a. Further, although FIG. 16A depicts leading vehicle 200b as having a camera system, the disclosed embodiments are not limited to a configuration in which leading vehicle 200b includes a camera system.

[0226] System 100 of primary vehicle 200a may cause primary vehicle 200a to mimic one or more actions of leading vehicle 200b. As shown in FIG. 16B, for example, primary vehicle 200a may mimic a lane shift performed by leading vehicle 200b. After mimicking the lane shift, distance c_1 may be equal to distance c_3, and distance c_2 may be equal to distance c_4. In some embodiments, primary vehicle 200a may mimic a lane shift performed by leading vehicle 200b in a way such that distances c_1 and c_3, and c_2 and c_4, are unequal. For example, in the case where primary vehicle 200a mimics a leftward lane shift of primary vehicle 200b and is positioned (after the lane shift) to the left of primary vehicle 200a, c_1 will be less than c_3 and c_2 will be greater than c_4. Alternatively, where primary vehicle 200a mimics a leftward lane shift of primary vehicle 200b and is positioned to the right of primary vehicle 200a, c_1 will be greater than c_3 and c_2 will be less than c_4.

[0227] FIGS. 16C and 16D are diagrammatic representations of a primary vehicle 200a mimicking one or more actions of a leading vehicle 200b on road 1600b, consistent with the disclosed embodiments. As illustrated in FIG. 16C, primary vehicle 200a may be traveling behind leading vehicle 200b, which may make a left turn at an intersection. System 100 of primary vehicle 200a may cause primary vehicle 200a to mimic the left turn of leading vehicle 200b, as shown in FIG. 16D. System 100 may cause primary vehicle 200a to mimic the left turn based on various factors, such as whether leading vehicle 200b changes lanes while turning within the intersection, the navigation history associated with leading vehicle 200b, the relative difference in speed between primary vehicle 200a and leading vehicle 200b, and the like. Additionally, primary vehicle 200a may mimic leading vehicle 200b under various conditions, such as based on whether leading vehicle 200b is traveling along the same route or

IPR2025-00944
Tesla EX1002 Page 1178

the same portion of a route that primary vehicle 200a is traveling along. Further detail regarding scenarios in which system 100 may cause primary vehicle 200a to mimic one or more actions of leading vehicle 200b are described in connection with FIGS. 17-22, below.

[0228] In some embodiments, system 100 may be configured to determine whether to mimic a particular movement of a leading vehicle. For example, system 100 may determine that certain movements of the leading vehicle do not need to be mimicked because the movements do not affect the

[0229] course of the leading vehicle. For example, system 100 may determine that small changes in movements, such as moving within a particular lane, do not need to be mimicked. In other embodiments, system 100 may implement a smoothing operation to filter out small movements of a leading vehicle. The smoothing operation may cause the primary vehicle to implement more significant and/or important movements (e.g., changing lanes) of a leading vehicle while filtering out smaller movements not affecting the overall course of the leading vehicle.

[0230] FIG. 17 is an exemplary block diagram of a memory configured to store instructions for performing one or more operations consistent with the disclosed embodiments. As shown in FIG. 17, memory 140 may store a lane constraint module 1702, an action detection module 1704, and an action response module 1706. The disclosed embodiments are not limited to any particular configuration of memory 140. Further, processing unit 110 may execute the instructions stored in any of modules 1702-1706 included in memory 140.

[0231] In one embodiment, lane constraint module 1702 may store software instructions (such as computer vision software) which, when executed by processing unit 110, determine one or more constraints associated with a lane that primary vehicle 200a is traveling in. The lane constraints may include a midpoint 1610, a left edge 1620, and a right edge 1630 of the lane. The lane constraints may also include one or more lines or other symbols marked on the surface of road 1600a or 1600b. Processing unit 110 may execute lane constraint module 1702 to determine the constraints by, for example, analyzing one or more sets of images and/or using map and/or position data (e.g., data stored in map database 160). The image sets may be acquired by one or more of image capture devices 122, 124, and 126. Based on the analysis, system 100 (e.g., via processing unit 110) may cause primary vehicle 200a to travel within a lane defined by left edge 1620 and right edge 1630. In some embodiments, system 100 may cause primary vehicle 200a to move away from left edge 1620 or right edge 1630 if one of distance c_1 or c_2 is less than a predetermined threshold.

[0232] In one embodiment, action detection module 1704 may store software instructions (such as computer vision software) which, when executed by processing unit 110, detects one or more actions taken by leading vehicle 200b traveling in front of primary vehicle 200a. Leading vehicle 200b may be traveling in the same or different lane as primary vehicle 200a. Processing unit 110 may execute action detection module 1704 to detect one or more actions taken by leading vehicle 200b by, for example, analyzing one or more sets of images acquired by one or more of image capture devices 122, 124, and 126. Based on the analysis, processing unit 110 may determine, for example, that leading vehicle 200b has shifted lanes, made a turn, accelerated, decelerated, applied its brakes, and/or the like. As another example, leading vehicle 200b may perform a maneuver such that a first offset distance on a side of leading vehicle 200b adjacent to a first lane constraint is different from a second offset distance on a side of leading vehicle 200b adjacent to a second lane constraint. For example, processing unit 110 may perform the analysis based on the techniques described in connection with FIGS. 5A-5D, above.

[0233] In one embodiment, action response module 1706 may store software instructions (such as computer vision software) which, when executed by processing unit 110, determines one or more actions for primary vehicle 200a to take based on one or more actions taken by leading vehicle 200b detected by system 100 (e.g., via execution of action detection module 1704). For example, processing unit 110 may execute action response module 1706 to determine whether to mimic one or more actions of leading vehicle 200b. This determination may be based on the nature of leading vehicle 200b's actions (e.g., turn, lane shift, etc.), information associated with primary vehicle 200a (e.g., speed, distance c_1 and c_2 from the lane edges, etc.), road and environmental conditions (e.g., potholes, heavy rain or wind, etc.), and the like. In the case where processing unit 110 determines to mimic one or more actions of leading vehicle 200b, processing unit 110 may accomplish this by, for example, transmitting electronic signals (e.g., via a CAN bus) to throttling system 220, braking system 230, and/or steering system 240 of primary vehicle 200a to trigger a turn, lane shift, change in speed, and/or change in direction of primary vehicle 200a. Processing unit 110 may use the techniques described in connection with FIGS. 4-7, above, to cause primary vehicle 200a to mimic the one or more actions. To mimic the one or more actions, system 100 may provide control signals to one or more of throttling system 220, braking system 230, and steering system 240 to navigate vehicle 200 (e.g., by causing an acceleration, a turn, a lane shift, etc.). Further, one or more actuators may control throttling system 220, braking system 230, and/or steering system 240.

[0234] FIG. 18 is a flow chart showing an exemplary process for causing primary vehicle 200a to mimic one or more actions of leading vehicle 200b consistent with disclosed embodiments. At step 1810, processing unit 110 may receive a plurality of images via data interface 128 between processing unit 110 and image acquisition unit 120. For instance, a camera included in image ac-

IPR2025-00944
Tesla EX1002 Page 1179

quisition unit 120 (such as image capture device 122) may capture a plurality of images and transmit them over a digital connection (e.g., USB, wireless, Bluetooth, etc.) to processing unit 110. In some embodiments, processing unit 110 may receive more than one plurality of images via a plurality of data interfaces. For example, processing unit 110 may receive a plurality of images from each of image capture devices 122, 124, 126, each of which may have an associated data interface for communicating data to processing unit 110. The disclosed embodiments are not limited to any particular data interface configurations or protocols.

[0235] At steps 1820 and 1830, processing unit 110 may execute lane constraint module 1702 to determine constraints for each side of primary vehicle 200a associated with a lane that primary vehicle 200a is traveling in. Processing unit 110 may determine the constraints by analyzing the images received at step 1810. Thus, at step 1820, processing unit 110 may determine a first lane constraint on a first side of primary vehicle 200a (e.g., the distance between the left side of primary vehicle 200a and left edge 1620 of the lane, c_1) and, at step 1830, a second lane constraint on a second side opposite the first side (e.g., the distance between the right side of primary vehicle 200a and right edge 1620 of the lane, c_2). In some embodiments, processing unit 110 may determine the constraints by using map and/or position data (e.g., data stored in map database 160) indicating the position of primary vehicle 200a relative to midpoint 1610, left edge 1620, and right edge 1630 of the lane. Processing unit 110 may determine the constraints based on both an analysis of the images received at step 1810 and map and/or position data; doing so may increase the confidence level associated with the constraints.

[0236] At step 1840, processing unit 110 may execute lane constraint module 1702 to cause primary vehicle 200a to travel within the lane constraints determined at steps 1820 and 1830. For example, the left side of primary vehicle 200a may be positioned to the left of left edge 1620, or the right side of primary vehicle 200a may be positioned to the right of right edge 1630. In such a scenario, processing unit 110 may transmit electronic signals (e.g., via a CAN bus) to throttling system 220, braking system 230, and/or steering system 240 of primary vehicle 200a to cause primary vehicle 200a to adjust its position and travel within the lane constraints.

[0237] At step 1850, processing unit 110 may execute action detection module 1704 to locate a leading vehicle 200b within the images received at step 1810 by analyzing the images. Processing unit 110 may detect the presence of vehicles in the images using the techniques described in connection with FIGS. 5A and 5B, above. For each detected vehicle, processing unit 110 may construct a multi-frame model of the detected vehicle's position, velocity (e.g., speed and direction), and acceleration relative to primary vehicle 200a. Processing unit 110 may use the techniques described in connection with FIGS. 5A and 5B, above, to construct such a model. In addition, processing unit 110 may determine a lane of travel associated with each detected vehicle, based on, for example, lane or other road markings and lane constraints determined at steps 1820 and 1830. Processing unit 110 may determine leading vehicle 200b to be the detected vehicle that is closest to, and is traveling in the same lane as, primary vehicle 200a. At step 1860, processing unit 110 may execute action detection module 1704 to determine one or more actions taken by the leading vehicle 200b detected within the images at step 1850. For example, processing unit 110 may determine that leading vehicle 200b has shifted lanes, made a turn, accelerated, decelerated, and/or the like. In addition, processing unit 110 may analyze the images to determine characteristics associated with leading vehicle 200b, such as its speed and position (e.g., relative to primary vehicle 200a). These determinations may be based on the techniques described in connection with FIGS. 5A and 5B, above. In some embodiments, processing unit 110 may analyze the images to determine characteristics associated with an action performed by leading vehicle 200b. For example, where leading vehicle 200b makes a turn, processing unit 110 may determine position information for leading vehicle 200b right before it begins the turn, position information for leading vehicle 200b as it completes the turn, a turning radius, and a speed profile describing changes in speed (e.g., over time) for leading vehicle 200b as it makes the turn.

[0238] At step 1870, processing unit 110 may execute action response module 1706 to cause primary vehicle 200a to mimic the one or more actions of leading vehicle 200b determined at step 1850. For example, processing unit 110 may transmit electronic signals (e.g., via a CAN bus) to throttling system 220, braking system 230, and/or steering system 240 of primary vehicle 200a to trigger a turn, lane shift, change in speed, and/or change in direction of primary vehicle 200a. In addition, processing unit 110 may construct a set of path points to guide changes in position and speed for primary vehicle 200a. The path points may be based on lane marks (e.g., midpoint 1610, left edge 1620, right edge 1630, etc.) represented by a polynomial (e.g., of the third-degree) and target coordinates associated with a target vehicle, such as leading vehicle 200b. Processing unit 110 may adjust the coefficients of the polynomial so that the path points pass through the target coordinates. Processing unit 110 may also offset the path points (e.g., relative to left edge 1620 and right edge 1630) based on a predetermined offset amount. As processing unit 110 adjusts the polynomial coefficients and offsets the path points, processing unit 110 may also calculate local curvature information for each of the resulting segments formed between the path points.

[0239] In some embodiments, processing unit 110 (e.g., via execution of action response module 1706) may cause primary vehicle 200a to mimic leading vehicle 200b within certain position, speed, and/or acceleration constraints associated with the two vehicles. For exam-

IPR2025-00944
Tesla EX1002 Page 1180

ple, the constraints may require a predetermined minimum distance (e.g., a safety distance) between primary vehicle 200a and leading vehicle 200b or a predetermined maximum difference in speed and/or acceleration between the vehicles. Thus, if mimicking an action performed by leading vehicle 200b would violate one or more constraints (e.g., mimicking an acceleration would result in primary vehicle 200a being too close to leading vehicle 200b), processing unit 110 may cause primary vehicle 200a to decline to mimic the one or more actions of leading vehicle 200b. In addition, even while mimicking an action performed by leading vehicle 200b (e.g., a lane shift), processing unit 110 may cause primary vehicle 200a to follow leading vehicle 200b's speed profile and maintain a predetermined distance or a safety distance from leading vehicle 200b. Processing unit 110 may accomplish this by causing primary vehicle 200a to: (i) continuously close the distance gap from leading vehicle 200b when the gap is greater than the predetermined minimum distance, (ii) continuously close the gap between the speeds of primary vehicle 200a and leading vehicle 200b, and (iii) match the acceleration of leading vehicle 200b to the extent it is consistent with predetermined constraints. To control the extent to which of these actions should be performed, processing unit 110 may associate each action with weights in real-time based on the current and predicted distance to leading vehicle 200b. Thus, for example, as the distance to leading vehicle 200b decreases more weight may be put on decelerating rather than matching the acceleration of leading vehicle 200b.

[0240] As another example, in the case where primary vehicle 200a is approaching leading vehicle 200b at a speed greater than the speed of leading vehicle 200b, processing unit 110 may cause primary vehicle 200a to gradually slow down such that, when the distance between the two vehicles equals a predetermined safety distance such as a minimum following distance, the speed of the two vehicles is the same. In another example, where a leading vehicle 200b abruptly appears in front of primary vehicle 200a, causing the distance between the two vehicles to be less than the minimum following distance, processing unit 110 may cause primary vehicle 200a to gradually slow down to avoid approaching leading vehicle 200b any closer while at the same increasing the distance between the vehicles until it reaches the minimum following distance.

[0241] FIG. 19 is a flow chart showing an exemplary process 1900 for causing primary vehicle 200a to decline to mimic a turn of leading vehicle 200b, consistent with disclosed embodiments. At step 1910, processing unit 110 may receive a plurality of images via data interface 128 between processing unit 110 and image acquisition unit 120, as described in connection with step 1810 of FIG. 18, above. At step 1920, processing unit 110 may execute action detection module 1704 to locate a leading vehicle 200b within the images received at step 1910 by analyzing the images, as described in connection with step 1850 of FIG. 18, above. At step 1925, processing unit 110 may execute action detection module 1704 to determine whether leading vehicle 200b is turning at an intersection, using the techniques described in connection with step 1860 of FIG. 18, above. If leading vehicle 200b is determined to be turning at the intersection (step 1925, yes), processing unit 110 may cause primary vehicle 200a to decline to mimic the turn of leading vehicle 200b at step 1930. Otherwise (step 1925, no), process 1900 concludes.

[0242] FIG. 20 is a flow chart showing an exemplary process 2000 for causing primary vehicle 200a to mimic or decline a turn of leading vehicle 200b, consistent with disclosed embodiments. At step 2010, processing unit 110 may receive a plurality of images via data interface 128 between processing unit 110 and image acquisition unit 120, as described in connection with step 1810 of FIG. 18, above. At step 2020, processing unit 110 may execute action detection module 1704 to locate a leading vehicle 200b within the images received at step 2010 by analyzing the images, as described in connection with step 1850 of FIG. 18, above. At step 2025, processing unit 110 may execute action detection module 1704 to determine whether leading vehicle 200b is turning at an intersection, using the techniques described in connection with step 1860 of FIG. 18, above. If leading vehicle 200b is determined to be turning at the intersection (step 2025, yes), at step 2035, processing unit 110 may execute action detection module 1704 to determine whether leading vehicle 200b is changing lanes within the intersection, using the techniques described in connection with step 1860 of FIG. 18, above. Otherwise (step 2025, no), process 2000 concludes. If leading vehicle 200b is determined to be changing lanes within the intersection (step 2035, yes), processing unit 110 may cause primary vehicle 200a to decline to mimic the turn of leading vehicle 200b at step 2050. Otherwise (step 2035, no), processing unit 110 may cause primary vehicle 200a to mimic the turn of leading vehicle 200b at step 2040.

[0243] FIG. 21 is a flow chart showing another exemplary process 2100 for causing primary vehicle 200a to mimic one or more actions of a first leading vehicle, consistent with disclosed embodiments. At step 2110, processing unit 110 may receive a plurality of images via data interface 128 between processing unit 110 and image acquisition unit 120, as described in connection with step 1810 of FIG. 18, above. At steps 2120 and 2130, processing unit 110 may execute action detection module 1704 to locate a first and second leading vehicle within the images received at step 2110 by analyzing the images, as described in connection with step 1850 of FIG. 18, above. At step 2135, processing unit 110 may execute action detection module 1704 to determine whether the first leading vehicle path has a lower turning radius than the second leading vehicle path, using the techniques described in connection with step 1860 of FIG. 18, above. If the first leading vehicle path is determined to have a lower turning radius than the second leading

IPR2025-00944
Tesla EX1002 Page 1181

vehicle path (step 2135, yes), processing unit 110 may cause primary vehicle 200a to mimic the turn of the first leading vehicle at step 2140. Otherwise (step 2135, no), process 2100 concludes.

[0244] FIG. 22 is a flow chart showing an exemplary process 2200 for causing primary vehicle 200a to mimic one or more actions of leading vehicle 200b based on a navigation history, consistent with disclosed embodiments. At step 2210, processing unit 110 may receive a plurality of images via data interface 128 between processing unit 110 and image acquisition unit 120, as described in connection with step 1810 of FIG. 18, above. At step 2220, processing unit 110 may execute action detection module 1704 to locate a leading vehicle 200b within the images received at step 2210 by analyzing the images, as described in connection with step 1850 of FIG. 18, above.

[0245] At step 2230, processing unit 110 may determine position information for leading vehicle 200b by analyzing the images received at step 2210. Position information for leading vehicle 200b may indicate the position of leading vehicle 200b relative to primary vehicle 200a. Processing unit 110 may analyze the images to derive position information for leading vehicle 200b and one or more actions taken by leading vehicle 200b by using the techniques described in connection with FIGS. 5A-5D, above.

[0246] At step 2240, processing unit 110 may compare the position information for leading vehicle 200b determined at step 2230 to predetermined map data (e.g., data stored in map database 160). For example, by comparing the position information to predetermined map data, processing unit 110 may derive position coordinates for leading vehicle 200b. In addition, processing unit 110 may use the predetermined map data as an indicator of a confidence level associated with the position information determined at step 2230.

[0247] At step 2250, processing unit 110 may create a navigation history associated with leading vehicle 200b by, for example, tracking the actions taken by leading vehicle 200b. Processing unit 110 may associate the actions taken by leading vehicle 200b (e.g., determined from analysis of the images at step 2230) with predetermined map data from step 2240. Thus, the navigation history may indicate that leading vehicle 200b performed certain actions (e.g., a turn, a lane shift, etc.) at certain position coordinates.

[0248] At step 2260, processing unit 110 may execute action response module 1706 to determine whether to cause primary vehicle 200a to mimic one or more actions of leading vehicle 200b based on the navigation history created at step 2250 and using the techniques described in connection with step 1870 of FIG. 18, above. For example, processing unit 110 (e.g., via action response module 1706) may cause primary vehicle 200a to mimic a turn of leading vehicle 200b, even where no turn is detected from analysis of the images (e.g., at step 2230), when the navigation history indicates that leading vehicle 200b has successfully navigated at least one prior turn. More generally, processing unit 110 may cause primary vehicle 200a to mimic an action taken by leading vehicle 200b in cases where the navigation history indicates the leading vehicle 200b had successfully performed such an action.

Navigating a Vehicle to Pass Another Vehicle

[0249] System 100 may provide driver assist functionality that monitors the vicinity around vehicle 200 and aborts lane changes in the event that the lane change is deemed to be at risk of causing a collision, such as with another vehicle (e.g., a target vehicle). For example, system 100 may enable vehicle 200 to complete a pass of a target vehicle if the target vehicle is determined to be in a different lane than vehicle 200. If, before completion of a pass, system 100 determines that the target vehicle is entering the lane in which vehicle 200 is traveling, system 100 may cause vehicle 200 to abort the pass. For example, aborting the pass may include causing vehicle 200 to brake or causing vehicle 200 to stop accelerating.

[0250] As illustrated in FIG. 23, vehicle 200 may travel on roadway 2300. Vehicle 200 may be equipped with system 100 and may implement any of the disclosed embodiments for identifying lane constraints and operating vehicle 200 within the lane constraints. System 100 may also facilitate lane changes made by vehicle 200. For example, system 100 may determine whether circumstances allow vehicle 200 to safely change lanes and, if the circumstances change during the lane change, cause vehicle 200 to abort lane change.

[0251] As shown in FIG. 23 and as previously discussed with respect to FIG. 8, vehicle 200 may include first vehicle side 802, which may be first distance 805 from a first lane constraint 2302. Similarly, vehicle 200 may include second vehicle side 812 opposite from first vehicle side 802, and second vehicle side 812 may be second distance 815 from a second lane constraint 2304. In this manner, first lane constraint 810 and second lane constraint 2304 may define a lane 2330 within which vehicle 200 may travel.

[0252] System 100 may determine first lane constraint 2302 and second lane constraint 2304 based on a plurality of images acquired by one or more of image capture devices 122-126. According to some embodiments, first lane constraint 2302 and/or second lane constraint 2304 may be identified by visible lane boundaries, including lines marked on a road surface. Additionally or alternatively, first lane constraint 2302 and/or second lane constraint 2304 may include an edge of a road surface. According to some embodiments, system 100 may determine first lane constraint 2302 and/or second lane constraint 2304 by identifying a midpoint of a road surface width 2320 (or any other suitable road feature that vehicle 200 may use as a navigation reference). System 100 may identify lane constraints in this manner when, for example, lines designating road lanes are not painted or

IPR2025-00944
Tesla EX1002 Page 1182

otherwise labeled.

[0253] When passing a target vehicle 2310, the lane change of vehicle 200 may be complicated by a change in the trajectory of target vehicle 2330. Thus, system 100 may identify target vehicle 2310 based on an analysis of a plurality of images acquired by one or more of image capture devices 122-126. System 100 may also identify second lane 2340 in which target vehicle 2310 is traveling. Further, system 100 may determine whether second lane 2340 is different than lane 2330 in which vehicle 200 is traveling. If second lane 2340 is different than lane 2330, system 100 may be configured to enable vehicle 200 to pass target vehicle 2310.

[0254] System 100 may also monitor target vehicle 2310 based on the plurality of images. For example, system 100 may monitor a position of target vehicle 2310, such as monitoring its position relative to vehicle 200, relative to first lane constraint 2302 and/or second lane constraint 2304, an absolute position, or relative to another reference point. Additionally or alternatively, monitoring a position of target vehicle 2310 may include estimating a speed of target vehicle 2310.

[0255] If system 100 determines that target vehicle 2310 is entering lane 2330 in which vehicle 200 is traveling, system 100 may cause vehicle 200 to abort the pass of target vehicle 2310. For example, system 100 may cause vehicle 200 to stop accelerating or brake. Additionally or alternatively, system 100 may issue an audible announcement, such as to notify driver of vehicle 200 that the lane change is being aborted.

[0256] FIG. 24A is an exemplary block diagram of memory 140, which may store instructions for performing one or more operations consistent with disclosed embodiments. As illustrated in FIG. 24A, memory 140 may store one or more modules for identification of the target vehicle and responses described herein. For example, memory 140 may store a lane constraint module 2400, a target vehicle acquisition module 2410, and an action module 2415.

[0257] In one embodiment, lane constraint module 2400 may store instructions which, when executed by processing unit 110, may detect and define first lane constraint 2302 and second lane constraint 2304. For example, processing unit 110 may execute lane offset module 2400 to process the plurality of images received from at least one image capture device 122-124 and detect first lane constraint 2302 and second lane constraint 2304. As discussed above, this may include identifying painted lane lines or markers and/or measuring a midpoint of a road surface. Processing unit 110 may perform the analysis based on the techniques described in connection with FIGS. 5A-5D, above.

[0258] In some embodiments, target acquisition module 2410 may store instructions which, when executed by processing unit 110, may detect the presence of target vehicle 2310 and monitor the position of target vehicle 2310. For example, target acquisition module 2410 may process the plurality of images to detect and monitor target vehicle 2310. Additionally or alternatively, when executing target acquisition module 2410, processing unit 110 may receive information from another module or other system indicative of the presence and/or location of target vehicle 2310. Processing unit 110 may perform the analysis based on the techniques described in connection with FIGS. 5A-5D, above.

[0259] Processing unit 110 may be configured to determine the presence of one or more objects such as target vehicle 2310 in the vicinity of vehicle 200. For example target vehicle 2310 may comprise another vehicle, such as a car, truck or motorcycle traveling near vehicle 200. Processing unit 110 may determine, for each target vehicle 2310, an offset profile. The offset profile may include a determination of whether in its current and predicted position, target vehicle 2310 will be within a predefined range of vehicle 200. If processing unit 110 determines that target vehicle 2310 is or will be within a predefined range of vehicle 200, processing unit may determine whether there is enough space within first lane constraint 810 and second lane constraint 820, for vehicle 200 to bypass target vehicle 2310. If there is not enough space, processing unit 110 may execute a lane change. If there is enough space, processing unit 110 may determine whether there is enough time to bypass target vehicle 2310 before the distance between vehicle 200 and target vehicle 2310 closes. If there is enough time, processing unit 110 may initiate and or activate the offset maneuver to bypass target vehicle 2310. The offset maneuver may be executed by processing unit 110 so that the movement of vehicle 200 is smooth and comfortable for the driver. For example, the offset slew rate may be approximately 0.15 to 0.75 m/sec. The offset maneuver may be executed so that the maximum amplitude of the offset maneuver will occur when the gap between vehicle 200 and object 825 closes. If there is not enough time, processing unit 110 may initiate braking and then initiate an offset maneuver.

[0260] Further, processing unit 110 may execute action module 2415 to abort the plan of vehicle 200 to pass target vehicle 2310. If target vehicle 2310 is not in lane 2330 of vehicle 200, instructions included in action module 1230 may enable vehicle 200 to pass target vehicle 2330. However, if execution of target vehicle acquisition module 2410 indicates that target vehicle 2310 is changing lanes into lane 2330 of vehicle 200, processing unit 110 may execute instructions included in action module 2415 to abort the lane change of vehicle 200. For example, processing unit 110 may transmit electronic signals to throttling system 220, braking system 230, and/or steering system 240 of vehicle 200 to abort the lane change. Additionally or alternatively, processing unit may cause an audible notification to sound.

[0261] FIG. 24B illustrates a flowchart of an exemplary process 2420 for navigating a vehicle among encroaching vehicles, consistent with disclosed embodiments. Process 2420 may identify lane constraints that define a lane of travel for a vehicle (e.g., vehicle 200) to travel,

IPR2025-00944
Tesla EX1002 Page 1183

identify and monitor a target vehicle (e.g., vehicle 2310), and enable the vehicle to pass the target vehicle or abort the pass if it is determined that the target vehicle is entering a lane in which the vehicle is traveling.

[0262] At step 2430, at least one of image capture device 122, 124, and/or 126 may acquire a plurality of images of an area in the vicinity of vehicle 200. For example, a camera included in image acquisition unit 120 (such as image capture device 122) may capture a plurality of images and transmit them over a digital connection (e.g., USB, wireless, Bluetooth, etc.) to processing unit 110. In some embodiments, processing unit 110 may receive more than one plurality of images via a plurality of data interfaces. For example, processing unit 110 may receive a plurality of images from each of image capture devices 122, 124, 126, each of which may have an associated data interface for communicating data to processing unit 110. The disclosed embodiments are not limited to any particular data interface configurations or protocols.

[0263] Next, at step 2440, processing unit 110 may determine a first lane constraint on first vehicle side 802 and a second lane constraint on second vehicle side 812 using the plurality of images received via data interface 128. For example, as part of implementing step 2440, processing unit 110 may execute instructions of lane constraint module 910. Further, as part of determining the first and second lane constraints, processing unit 110 may use one or more of the processes discussed above in connection with FIGS. 5A-5D.

[0264] At step 2450, processing unit 110 may acquire target vehicle 2310. To acquire target vehicle 2310, processing unit 110 may execute target vehicle acquisition module 2410. For example, processing unit 110 may execute instructions to analyze acquired images to identify second lane 2340 in which target vehicle 2310 is traveling and determine whether target vehicle 2310 is traveling in lane 2330 in which vehicle 200 is traveling. This determination may be based on the techniques described in connection with FIGS. 5A-5D, above.

[0265] At step 2455, processing unit 110 may determine whether second lane 2340 in which target vehicle 2310 is traveling is different than lane 2330 in which vehicle 200 is traveling.

[0266] At step 2460, processing unit 110 may enable vehicle 200 to pass target vehicle 2310 if second lane 2340 is different than lane 2330 in which vehicle 200 is traveling. To enable vehicle 200 to pass target vehicle 2310, processing unit may execute action module 2415.

[0267] Next, at step 2465, processing unit 110 may execute target acquisition module 2410 to monitor target vehicle 2310. This may include estimating a velocity of target vehicle 2310 and determining the location of target vehicle 2310, such as with respect to vehicle 200.

[0268] Step 2470 may include determining whether target vehicle 2310 has entered lane 2330. If processing unit 110 determines that target vehicle 2310 has not entered lane 2330, at step 2480, processing unit 110 may allow vehicle 200 to complete the pass.

[0269] At step 2490, if processing unit 110 determines that target vehicle 2310 has entered lane 2330 (or is entering lane 2330 or is otherwise on a trajectory to bring vehicle 2310 into lane 2330), processing unit 110 may execute action module 2415 to abort the pass. This may include, for example, issuing an audible notification to driver of vehicle 200 via speakers 360. Additionally or alternatively, step 2490 may include executing instructions included in action module 2415 to cause processing unit 110 to transmit electronic signals to throttling system 220, braking system 230, and/or steering system 240 of vehicle 200 to abort lane change.

Navigating a Vehicle to Avoid Encroaching Traffic

[0270] System 100 may provide driver assist functionality that monitors the vicinity around vehicle 200 and responds to the presence of another vehicle encroaching upon vehicle 200. Monitoring the vicinity around vehicle 200 may include monitoring a collision threshold. For example a collision threshold may include a time to collision or a minimum predetermined distance between vehicle 200 and other traffic. System 100 may further cause vehicle 200 to take evasive action to avoid encroaching traffic. For example, evasive action may include braking, changing lanes, or otherwise altering the course of vehicle 200.

[0271] As illustrated in FIG. 25, vehicle 200 may travel on roadway 2500. Vehicle 200 may include system 100 to detect and respond to encroaching traffic. For example, vehicle 200 may determine that another vehicle is entering into the lane in which vehicle 200 is traveling or otherwise determine that another vehicle is crossing a collision threshold. Vehicle 200 may react to an encroaching vehicle, such as by altering course, braking, or accelerating.

[0272] As discussed with respect to FIG. 8 and as shown in FIG. 25, vehicle 200 may include first vehicle side 802, which may be first distance 805 from a first lane constraint 2502. Similarly, vehicle 200 may include second vehicle side 812 opposite from first vehicle side 802, and second vehicle side 812 may be second distance 815 from a second lane constraint 2504. In this manner, first lane constraint 2502 and second lane constraint 2504 may define a lane 2505 within which vehicle 200 may travel.

[0273] System 100 may determine first lane constraint 2502 and second lane constraint 2504 based on a plurality of images acquired by one or more of image capture devices 122-126. According to some embodiments, first lane constraint 2502 and/or second lane constraint 2504 may be identified by visible lane boundaries, such a line marked on a road surface. Additionally or alternatively, first lane constraint 2502 and/or second lane constraint 2504 may include an edge of a road surface. According to some embodiments, system 100 may determine first lane constraint 2502 and/or second lane constraint 2502 by identifying a midpoint 2520 of a road surface width (or

IPR2025-00944
Tesla EX1002 Page 1184

any other suitable road feature vehicle 200 may use as a navigation reference). System 100 may identify lane constraints in this manner when, for example, lines designating road lanes are not painted or otherwise labeled.

[0274] Detection of first lane constraint 2502 and/or second lane constraint 2504 may include processing unit 110 determining their 3D models in a camera coordinate system. For example, the 3D models of first lane constraint 2502 and/or second lane constraint 2504 may be described by a third-degree polynomial. In addition to 3D modeling of first lane constraint 2502 and/or second lane constraint 2504, processing unit 110 may estimate motion parameters, such as the speed, yaw and pitch rates, and acceleration of vehicle 200. Optionally, processing unit may detect static and moving vehicles and their position, heading, speed, and acceleration, all relative to vehicle 200. Processing unit 110 may determine a road elevation model to transform the information acquired from the plurality of images into 3D space.

[0275] The driver of vehicle 200 cannot assume that other traffic traveling in adjacent or near-adjacent lanes will only change lanes if it is safe for that traffic to shift into lane 2505 in which vehicle 200 is driving. Thus, it may be desirable to consider and monitor traffic traveling in nearby lanes to detect lateral encroachment into lane 2505 of vehicle 200, so that vehicle 200 may react to lateral encroachment to avoid a collision. System 100 may identify an encroaching vehicle 2510 based on the plurality of images and determine that encroaching vehicle 2510 is approaching vehicle 200. System 100 may then cause an action to account for the encroaching vehicle. For example, system 100 may cause vehicle 200 to maintain a current velocity and travel within first lane constraint 2502 and second lane constraint 2504 such that a first distance 2530, which is on first vehicle side 802 on which encroaching vehicle 2510 is approaching, is greater than a second distance 2540.

[0276] Processing unit 110 may be configured to determine the presence of one or more encroaching vehicles 2510 in the vicinity of vehicle 200. Processing unit 110 may determine, for each encroaching vehicle 2510 an offset profile. The offset profile may include a determination of whether in its current and predicted position, encroaching vehicle 2510 will be within a predefined range of vehicle 200. If processing unit 110 determines that encroaching vehicle 2510 is or will be within a predefined range of vehicle 200, processing unit may determine whether there is enough space within first lane constraint 2502 and second lane constraint 2504, for vehicle 200 to bypass encroaching vehicle 2510. If there is not enough space, processing unit 110 may execute a lane change. If there is enough space, processing unit 110 may determine whether there is enough time to bypass encroaching vehicle 2510 before the distance between vehicle 200 and encroaching vehicle 2510 closes.

[0277] For example, if there is enough time, processing unit 110 may initiate and or activate the offset maneuver to bypass encroaching vehicle 2510. The offset maneuver may be executed by processing unit 110 so that the movement of vehicle 200 is smooth and comfortable for the driver. For example, the offset slew rate may be approximately 0.15 to 0.75 m/sec. The offset maneuver may be executed so that the maximum amplitude of the offset maneuver will occur when the gap between vehicle 200 and encroaching vehicle 2510 closes. If there is not enough time, processing unit 110 may initiate braking (e.g., via transmitting electronic signals to braking system 230) and then initiate offset maneuver (e.g., by transmitting electronic signals to throttling system 220 and/or steering system 240).

[0278] System 100 may determine, based on a plurality of images acquired by one or more of image capture devices 122-126, that encroaching vehicle 2510 has crossed at least one collision threshold. For example, the at least one collision threshold may include a minimum predetermined distance 2550 between vehicle 200 and encroaching vehicle 2510. For example, encroaching vehicle 2510, as illustrated in FIG. 25, is shown exceeding minimum predetermined distance 2550. Additionally or alternatively, the at least one collision threshold may include a time to collision of vehicle 200 and encroaching vehicle 2510. The time-to-collision threshold may be determined based on the distance between vehicle 200 and encroaching vehicle 2510, the velocity of vehicle 200 and/or encroaching vehicle 2510, the lateral velocity of vehicle 200 and/or encroaching vehicle 2510, and/or the acceleration of vehicle 200 and/or encroaching vehicle 2510. A collision threshold based on a time to collision may be set to allow enough time for system 100 to effectively conduct evasive action to avoid collision.

[0279] If system 100 determines that at least one collision threshold has been crossed, system 100 may cause vehicle 200 to take evasive action. For example, this may include causing other subsystems of vehicle 200 to operate. Thus, system 100 may operate to cause vehicle 200 to increase speed by communicating with throttling system 220, braking system 230, and/or steering system 240 to change the speed and/or direction of vehicle 200.

[0280] The evasive action may include causing vehicle 200 to change lanes. This may be a desirable response if, for example, vehicle 200 is traveling on a multilane highway and system 100 determines that changing lanes to move away from encroaching vehicle 2510 will avoid a collision. Evasive action may include vehicle 200 altering its course. Additionally or alternatively, evasive action may include altering course of vehicle 200. For example, to take evasive action, processing unit 110 may transmit electronic signals to throttling system 220, braking system 230, and/or steering system 240.

[0281] System 100 may also be configured to cause vehicle 200 to take an action after the evasive action has been completed. For example, system 100 may cause vehicle 200 to return to midpoint 2520 of lane 2505 in which vehicle 200 is traveling. Additionally or alternatively, system 100 may cause vehicle 200 to resume a preset

IPR2025-00944
Tesla EX1002 Page 1185

speed after braking in response to encroaching vehicle 2510.

[0282] FIG. 26 is an exemplary block diagram of memory 140, which may store instructions for detecting and responding to traffic laterally encroaching on a vehicle consistent with disclosed embodiments. As illustrated in FIG. 26, memory 140 may store a lane constraint module 2610, an encroaching vehicle acquisition module 2620, and an action module 2630.

[0283] In one embodiment, lane constraint module 2610 may store instructions which, when executed by processing unit 110, may detect and define first lane constraint 2502 and second lane constraint 2504. For example, processing unit 110 may execute lane offset constraint module 2610 to process the plurality of images received from at least one image capture device 122-126 and detect first lane constraint 2502 and second lane constraint 2504. As discussed above, this may include identifying painted lane lines and/or measuring a midpoint of a road surface. Processing unit 110 may perform the analysis based on the techniques described in connection with FIGS. 5A-5D, above.

[0284] In one embodiment, encroaching vehicle acquisition module 2620 may store instructions which, when executed by processing unit 110, may detect the presence of encroaching vehicle 2510 and monitor encroaching vehicle 2510. For example, encroaching vehicle acquisition module 2620 may include instructions for determining a relative location of encroaching vehicle 2510 to vehicle 200. In some embodiments, this may include monitoring the velocity of encroaching vehicle 200. According to some embodiments, encroaching vehicle acquisition module 2620 may determine whether encroaching vehicle 2510 has exceeded a collision threshold. A collision threshold may be, for example, a minimum predetermined distance between vehicle 200 and encroaching vehicle 2510, If encroaching vehicle 2510 comes closer to vehicle 200 than the minimum predetermined distance, processing unit 110 may determine that encroaching vehicle 2510 is indeed encroaching and execute instructions to avoid collision with encroaching vehicle 2510.

[0285] Additionally or alternatively, a collision threshold may include a time to collision. The time-to-collision threshold may be determined based on the distance between vehicle 200 and encroaching vehicle 2510, the velocity of vehicle 200 and/or encroaching vehicle 2510, the lateral velocity of vehicle 200 and/or encroaching vehicle 2510, and/or the acceleration of vehicle 200 and/or encroaching vehicle 2510. A collision threshold based on a time to collision may be set to allow enough time for processing unit 110 and system 100 to effectively conduct evasive action to avoid collision.

[0286] Further, processing unit 110 may execute action module 2630 to respond to the detection of encroaching vehicle 2510 and/or the crossing of a collision threshold. For example, action module 2630 may include instructions to cause vehicle 200 to maintain a current velocity and to travel within first lane constraint 2502 and second lane constraint 2504 such that first distance 2530 is greater than second distance 2540. The first distance 2530 may be on the same side of vehicle 200 as encroaching vehicle 2510, so that vehicle 200 is farther away from encroaching vehicle 2510 while still within first lane constraint 2502 and second lane constraint 2504. If processing unit 110 determines that a collision threshold has been crossed, processing unit 110 may execute action module 2630 to cause vehicle 200 to take evasive action, such as altering course, braking, and/or changing lanes. Action module 2630 may also include instructions that cause vehicle 200 to resume course after completing evasive action. For example, processing unit 110 may cause vehicle 200 to resume a preset speed after braking or return to the center of lane 2505.

[0287] FIG. 27 illustrates a flowchart of an exemplary process 2720 detecting and responding to traffic laterally encroaching on a vehicle, consistent with disclosed embodiments. Process 2720 may identify lane constraints that define a lane of travel for vehicle travel, determine whether an encroaching vehicle 2510 is approaching, and cause vehicle 200 to maintain current velocity and to travel within first lane constraint 2501 and second lane constraint 2504 such that first distance 2530, on the side of vehicle 200 that encroaching vehicle 2510 is approaching, is greater than second distance 2540.

[0288] At step 2730, at least one image capture device 122,124, and/or 126 may acquire a plurality of images of an area in the vicinity of vehicle 200. For example, a camera included in image acquisition unit 120 (such as image capture device 122) may capture a plurality of images and transmit them over a digital connection (e.g., USB, wireless, Bluetooth, etc.) to processing unit 110. In some embodiments, processing unit 110 may receive more than one plurality of images via a plurality of data interfaces. For example, processing unit 110 may receive a plurality of images from each of image capture devices 122, 124, 126, each of which may have an associated data interface for communicating data to processing unit 110. The disclosed embodiments are not limited to any particular data interface configurations or protocols.

[0289] Next, at step 2740, processing unit 110 may determine a first lane constraint on first vehicle side 802 and a second lane constraint on second vehicle side 812 using the plurality of images received via data interface 128. For example, as part of implementing step 2740, processing unit 110 may execute instructions of lane constraint module 2610. Further, as part of determining the first and second lane constraints, processing unit 110 may use one or more of the processes discussed above in connection with FIGS. 5A-5D.

[0290] At step 2750, processing unit 110 may determine whether encroaching vehicle 2510 is approaching. To make this determination, processing unit 110 may execute encroaching vehicle acquisition module 2620. Encroaching vehicle acquisition module 2620 may include instructions to process images to detect other traffic. This

IPR2025-00944
Tesla EX1002 Page 1186

may include identifying vehicles within a certain range of vehicle 200, such as those vehicles that are adjacent to vehicle 200. These adjacent vehicles may be encroaching vehicles 2510.

[0291] Once encroaching vehicle 2510 is identified, processing unit may determine more information regarding vehicle 2510. For example, at step 2555, processing unit 110 may determine whether target vehicle 2330 has exceeded a collision threshold. A collision threshold may be, for example, a minimum predetermined distance between vehicle 200 and encroaching vehicle 2510. If encroaching vehicle 2510 comes closer to vehicle 200 than the minimum predetermined distance, processing unit 110 may determine that encroaching vehicle 2510 is indeed encroaching and execute instructions to avoid collision with encroaching vehicle 2510.

[0292] Additionally or alternatively, a collision threshold may be a time to collision. The time-to-collision threshold may be determined based on the distance between vehicle 200 and encroaching vehicle 2510, the velocity of vehicle 200 and/or encroaching vehicle 2510, the lateral velocity of vehicle 200 and/or encroaching vehicle 2510, and/or the acceleration of vehicle 200 and/or encroaching vehicle 2510. A collision threshold based on a time to collision may be set to allow enough time for processing unit 110 and system 100 to effectively conduct evasive action to avoid collision.

[0293] At step 2760, processing unit 110 may cause vehicle 200 to maintain current velocity and to travel within first lane constraint 2502 and second lane constraint 2504. To cause vehicle 200 to operate in this manner, processing unit 110 may execute action module 2630. Additionally or alternatively, step 2760 may include executing instructions on action module 1230 to cause processing unit 110 to transmit electronic signals to the accelerator 2610, brakes 2620, and/or steering system of vehicle 200 to cause vehicle 200 to maintain current velocity and to travel within first and second lane constraints 810 and 820.

[0294] If, at step 2755, processing unit 110 determined that a collision threshold has been exceeded, the processing unit 110 may take evasive action. The evasive action may include causing vehicle 200 to change lanes. This may be a desirable response if, for example, vehicle 200 is traveling on a multilane highway and system 100 determines that changing lanes to move away from encroaching vehicle 2510 will avoid a collision. Evasive action may include vehicle 200 altering its course. Additionally or alternatively, evasive action may include altering course of vehicle 200. For example, processing unit 110 may undertake evasive action by transmitting electronic signals to throttling system 220, braking system 230, and/or steering system 240 of vehicle 200. After the evasive action is completed, at step 2780, processing unit 110 may control vehicle 200 to resume its prior navigation (e.g., to resume navigation to a desired location).

Multi-Threshold Reaction Zone for Vehicle Navigation

[0295] System 100 may provide driver assist functionality that causes a response in vehicle 200, such as braking, accelerating, switching lanes, turning, and/or other navigational responses. For example, if system 100 detects the presence of another vehicle ahead of vehicle 200 that is traveling at a rate of speed slower than vehicle 200, system 100 may cause vehicle 200 to reduce its speed by braking. System 100 may apply a braking profile that accounts for the immediacy of the risk (e.g., the risk of collision) to provide the driver with a natural driving sensation.

[0296] System 100 may cause various different responses depending on the situation and a determined risk of collision with another vehicle, for example. System 100 may take no action in response to changes in speed or direction of a target vehicle ahead if the target vehicle is sufficiently far ahead and the determined risk of collision is low. Where the target vehicle may be closer or when the risk of collision rises above a predetermined level, system 100 may cause a response, such as braking, etc. When the risk of collision is even higher (e.g., where system 100 determines that a collision would be imminent without evasive action), system 100 may cause vehicle 100 to take an evasive action, such as maximum or near maximum braking, change of direction, etc.

[0297] FIGS. 28A and 28B are diagrammatic representations of a primary vehicle 200a on a road 2800, consistent with disclosed embodiments. In some instances, road 2800 may be divided into two lanes, lane 2810 and lane 2820. A target object, such as target vehicle 200b, may be traveling ahead of primary vehicle 200a in the same lane (e.g., lane 2820 as shown in FIGS. 28A and 28B) or in a different lane. As shown in the example depicted in FIG. 28A, primary vehicle 200a and target vehicle 200b may be separated by a distance d_1. At a given time after the scene depicted in FIG. 28A, primary vehicle 200a and target vehicle 200b may be separated by a distance d_2 as shown in FIG. 28B, where d_2 is less than d_1. For example, primary vehicle 200a may be traveling at a rate of speed greater than target vehicle 200b, causing the distance between the vehicles to decrease from d_1 to d_2.

[0298] As another example, primary vehicle 200a and target vehicle 200b may be traveling at the same rate of speed until target vehicle 200b reduces its speed by braking, or primary vehicle 200a increases its speed by accelerating, causing the distance between the vehicles to decrease from d_1 to d_2. As another example, target vehicle 200b may encroach from a side of primary vehicle 200a. In some circumstances, if target vehicle 200b is ahead or a sufficient distance away from primary vehicle 200a (e.g., five meters or more, 10 meters or more, etc.) and moves toward primary vehicle 200a, system 100 may not take any action. However, if target vehicle 200b continues to move closer to primary vehicle 200a, then system 100 may cause primary vehicle 200a to brake,

IPR2025-00944
Tesla EX1002 Page 1187

change speed, and/or change lanes. For example, if target vehicle 200b is within a predetermined threshold (e.g., within five meters), primary vehicle 200a may then take action. Different scenarios involving the relative speeds and positions of primary vehicle 200a and target vehicle 200b are possible, and the disclosed embodiments are not limited to any particular scenario.

[0299] In some embodiments, the multi-threshold reaction zone may provide a natural feeling driving experience to the user. For example, as discussed above, if the risk of collision is low, then system 100 may disregard movements by target vehicle 200b. As a result, primary vehicle 200a may not respond to movements or actions of target vehicle 200b when there is not a risk of a collision (e.g., primary vehicle 200a would not need to mimic every motion of target vehicle 200b when it's 100 meters ahead). However, if target vehicle 200b is close to primary vehicle 200a and/or if system 100 determines that the risk of collision is high, then a responsive maneuver may be taken such that the user feels safer by increasing space between primary vehicle 200a and target vehicle 200b. If system 100 determines that a crash may be imminent, then system 100 may take a more dramatic action without causing the user to be surprised, because the action may be necessary to avoid a collision.

[0300] Accordingly, in some embodiments, system 100 may associate the different reaction zones with different degrees of responsive actions. For example, in some embodiments, system 100 may take no action when target vehicle 200b is within a first threshold (e.g., when target vehicle 200b is far away from primary vehicle 200a), system 100 may take a medium action with target vehicle 200b is within a second threshold (e.g., when target vehicle 200b is becoming closer to primary vehicle 200a), and system 100 may take a more dramatic action when target vehicle 200 is within a third threshold (e.g., when target vehicle 200b is sufficiently close to primary vehicle 200a that there is a risk of a collision). In a first reaction zone associated with the first threshold, system 100 may take no action when the estimated distance-to-target associated with target vehicle 200b is greater than twice a predetermined safety distance. In some embodiments, the first reaction zone may apply where the estimated distance-to-target associated with target vehicle 200b falls in the range of 1.5 to 2.5 times the predetermined safety distance.

[0301] The estimated distance-to-target may be the current distance between primary vehicle 200a and target vehicle 200b plus the expected change of this distance over a prediction time (e.g., in the range of 0.1 to 1.0 seconds, depending on factors associated with system 100 such as velocity tracking control loops, actuators, and vehicle dynamics). The expected change of distance may be calculated as: $(v_{target}-v_{primary})*t_{prediction} + ((a_{target}-a_{primary})*t_{prediction}^{1/2})$, where v_{target} and $v_{primary}$ correspond to the speed of target vehicle 200b and primary vehicle 200a, respectively, a_{target} and $a_{primary}$ correspond to the acceleration of target vehicle 200b and primary vehicle 200a, respectively, and $t_{prediction}$ corresponds to the prediction time. Choosing a prediction time approximating the overall reaction time of system 100 may account for the fact that desired navigational responses for primary vehicle 200a may not be implemented immediately, and thereby may provide the driver and/or passengers of primary vehicle 200a with a smooth, natural driving experience.

[0302] The predetermined safety distance may be defined as: max $\{d_{static}, d_{dynamic}\}$. d_{static}, which may fall in the range of 2 to 10 meters, may represent a desired safety distance between primary vehicle 200a and another vehicle (such as target vehicle 200b) while the vehicles are stopping and/or moving at a low rate of speed. $d_{dynamic}$ may represent a desired safety distance when the vehicles are moving at a speed greater than a low rate of speed, and may be calculated as: $t_{safety}*\min(v_{target}, v_{primary})$, where v_{target} and $v_{primary}$ correspond to the speed of target vehicle 200b and primary vehicle 200a, respectively. t_{safety} may fall in the range of 0.5 to 2.5 seconds and may be adjusted by the driver of primary vehicle 200a via user interface 170 according to the driver's preferences. Thus, the driver may be able to control the distance that system 100 maintains between primary vehicle 200a and other vehicles.

[0303] In a second reaction zone associated with the second threshold, system 100 may place target vehicle 200b in a safety zone and apply different weights to different maneuvers performed by target vehicle 200b. For example, as target vehicle 200b may approach a lower boundary of the safety zone, e.g., the estimated distance-to-target associated with target vehicle 200b may approach the predetermined safety distance (discussed above in connection with the first reaction zone), system 100 may assign more weight to deceleration maneuvers (e.g., mimicked more) and less weight to acceleration maneuvers (e.g., mimicked less). In some embodiments, the lower boundary may fall in the range of 0.5 to 1.5 times the predetermined safety distance, depending on driver preferences. Conversely, as target vehicle 200b approaches an upper boundary of the safety zone, e.g., the estimated distance-to-target associated with target vehicle 200b may approach twice the predetermined safety distance, system 100 may assign less weight to deceleration maneuvers (e.g., mimicked less) and more weight to acceleration maneuvers (e.g., mimicked more). In some embodiments, the lower boundary may fall in the range of 1.5 to 2.5 times the predetermined safety distance, depending on driver preferences.

[0304] In a third reaction zone associated with the third threshold, system 100 may place target vehicle 200b in a danger zone when the estimated distance-to-target associated with target vehicle 200b is less than the predetermined safety distance. In this scenario, system 100 may cause primary vehicle 200a to mimic (e.g., immediately) any deceleration maneuvers (while disregarding any acceleration maneuvers) performed by target vehicle

IPR2025-00944
Tesla EX1002 Page 1188

200b. In some embodiments, the first reaction zone may apply where the estimated distance-to-target associated with target vehicle 200b falls in the range of 0.5 to 1.5 times the predetermined safety distance, depending on driver preferences.

[0305] In some embodiments, system 100 may cause a response in primary vehicle 200a, such as braking. As indicated in FIG. 28B, for example, the distance between the primary vehicle 200a and target vehicle 200b may decrease from d_1 to d_2. Depending on the distance d_2 and, by extension, the time before collision, system 100 may cause primary vehicle 200a to reduce its speed so as to avoid a collision with target vehicle 200b. In addition, system 100 may apply one or more braking profiles when causing primary vehicle 200a to reduce its speed, also based on, for example, the time before collision. Thus, if the time before collision (and/or the distance d_2) exceeds a first predetermined threshold, system 100 may cause primary vehicle 200a to brake in a gradual manner, change lanes, and/or change speed, whereas if the time before collision (and/or the distance d_2) is below a second predetermined threshold, system 100 may cause primary vehicle 200a to brake, change lanes, and/or change speed in a more rapid manner.

[0306] FIG. 29 is an exemplary block diagram of a memory configured to store instructions for performing one or more operations, consistent with the disclosed embodiments. As shown in FIG. 29, memory 140 may store a target monitoring module 2902, an intercept time module 2904, and an action response module 2906. The disclosed embodiments are not limited to any particular configuration of memory 140. Further, processing unit 110 may execute the instructions stored in any of modules 2902-2906 included in memory 140.

[0307] In one embodiment, target monitoring module 2902 may store software instructions (such as computer vision software) which, when executed by processing unit 110, identifies and monitors a target object (such as target vehicle 200b) appearing in a plurality of images representing an environment in the vicinity of primary vehicle 200. For purposes of this disclosure, the target object may also be referred to as target vehicle 200b, although the target object may be a pedestrian, road hazard (e.g., large debris), and the like. The plurality of images may be acquired by one or more of image capture devices 122, 124, and 126. Processing unit 110 may execute target monitoring module 2902 to identify and monitor target vehicle 200b by, for example, analyzing the plurality of images using the techniques described in connection with FIGS. 5A-5D, above. Thus, target monitoring module 2902 may track different variables, such as the motion (e.g., direction of travel and trajectory), speed, and acceleration associated with target vehicle 200b, as well as the distance between primary vehicle 200a and the target vehicle 200b. Such information may be tracked continuously or at predetermined intervals.

[0308] In one embodiment, intercept time module 2904 may store software instructions which, when executed by processing unit 110, determines an indicator of an intercept time between primary vehicle 200a and target vehicle 200b. The intercept time may refer to an amount of time before primary vehicle 200a makes contact (e.g., collides) with target vehicle 200b. Processing unit 110 may execute intercept time module 2904 to determine an indicator of the intercept time based on an analysis of images acquired by one or more of image capture devices 122, 124, and 126. For example, intercept time module 2904 may make the determination using a series of time-based observations, e.g., of the position, speed, and/or acceleration of target vehicle 200b (relative to primary vehicle 200a), such as Kalman filters or linear quadratic estimation (LQE). The Kalman filters may be based on a measurement of the scale of target vehicle 200b, where the scale measurement is proportional to the intercept time.

[0309] In one embodiment, action response module 2906 may store software instructions (such as computer vision software) which, when executed by processing unit 110, causes one or more responses in primary vehicle 200a. The responses may be based on the identification and monitoring of target vehicle 200b (e.g., performed via execution of target monitoring module 2902) and/or the indicator of an intercept time between primary vehicle 200a and target vehicle 200b (e.g., determined via execution of intercept time module 2904). Processing unit 110 may cause one or more responses in primary vehicle 200a by, for example, transmitting electronic signals (e.g., via a CAN bus) to throttling system 220, braking system 230, and/or steering system 240 of primary vehicle 200a to brake, accelerate, or trigger a turn, lane shift, and/or change in direction of primary vehicle 200a. Processing unit 110 may use extended Kalman Filters to estimate data associated with target vehicle 200b and primary vehicle 200a, such as: speed, acceleration, turn rate, pitch angle, pitch rate, position and heading of target vehicle 200b relative to primary vehicle 200a, the size of target vehicle 200b, road geometry, and the like. For example, processing unit 110 may receive measurements from components on primary vehicle 200a, such as an odometer, an inertial measurement unit (IMU), image processor 190, and the like. Thus, processing unit 110 may calculate data associated with target vehicle 200b and/or primary vehicle 200a using state equations derived from kinematic equations of motion, rigidity constraints, and motion assumptions, where inputs to the equations may include measurements from components on primary vehicle 200a. Processing unit 110 may also use the techniques described in connection with FIGS. 4-7, above, to cause one or more responses in primary vehicle 200a.

[0310] FIG. 30 is a flow chart showing an exemplary process 3000 for causing a response in a primary vehicle, consistent with disclosed embodiments. At step 3010, processing unit 110 may receive a plurality of images via data interface 128 between processing unit 110 and image acquisition unit 120. For instance, one or more cam-

IPR2025-00944
Tesla EX1002 Page 1189

eras included in image acquisition unit 120 may capture a plurality of images of an area forward of primary vehicle 200a (or to the sides or rear of a vehicle, for example) and transmit them over a data connection (e.g., digital, wired, USB, wireless, Bluetooth, etc.) to processing unit 110.

[0311] At step 3020, processing unit 110 may execute target monitoring module 2902 to identify a target object (such as target vehicle 200b) within the plurality of images. For example, target monitoring module 2902 may detect the presence of vehicles in the images using the techniques described in connection with FIGS. SA and 5B, above. Multiple vehicles (and/or objects) may be detected in the same lane or in a different lane as primary vehicle 200a, and at varying distances from primary vehicle 200a, based on image data, position data (e.g., GPS location information), map data, the yaw rate of primary vehicle 200a, lane or other road markings, speed data, and/or data from sensors included in primary vehicle 200a. In a scenario where multiple vehicles are detected, target monitoring module 2902 may determine target vehicle 200b to be the vehicle traveling in the same lane as primary vehicle 200a and/or located closest to primary vehicle 200a.

[0312] At step 3030, processing unit 110 may execute target monitoring module 2902 to monitor target vehicle 200b identified at step 3020. For example, target monitoring module 2902 may track information associated with target vehicle 200b, such as the motion (e.g., direction of travel and trajectory), speed, and acceleration, as well as the distance between primary vehicle 200a and the target vehicle 200b. Target monitoring module 2902 may track such information continuously, or at predetermined intervals, by analyzing the plurality of images using the techniques described in connection with FIGS. 5A-5D, above. At step 3040, processing unit 110 may execute intercept time module 2904 to determine an indicator of an intercept time between primary vehicle 200a and target vehicle 200b. Intercept time module 2904 may make the determination by tracking the position, speed, and/or acceleration of target vehicle 200b (relative to primary vehicle 200a) over one or more periods of time. Intercept time module 2904 may use the techniques described in connection with FIGS. 5A and 5B, above to track such information.

[0313] At step 3050, processing unit 110 may execute action response module 2906 to cause one or more responses in primary vehicle 200a. For example, processing unit 110 may transmit electronic signals (e.g., via a CAN bus) to throttling system 220, braking system 230, and/or steering system 240 of primary vehicle 200a to brake, accelerate, or trigger a turn, lane shift, and/or change in direction of primary vehicle 200a. Further, one or more actuators may control throttling system 220, braking system 230, and/or steering system 240. For example, processing unit 110 may transmit electronic signals that cause system 100 to physically depress the brake by a predetermined amount or ease partially off the accelerator of vehicle 200. Further, processing unit 110 may transmit electronic signals that cause system 100 to steer vehicle 200 in a particular direction. Such responses may be based on the monitoring of target vehicle 200b (performed at step 3030) and/or the indicator of an intercept time between primary vehicle 200a and target vehicle 200b (determined at step 3040). In some embodiments, the intercept time may be compared with one or more predetermined thresholds, e.g., a first and second intercept threshold. For example, the response may include an emergency avoidance action if the indicator of intercept time indicates that a target object is within a particular intercept threshold (e.g., within the second intercept threshold). The emergency avoidance action may include changing lanes and/or emergency braking.

[0314] The thresholds may be expressed as units of time. Thus, for example, if the indicator of intercept time indicates that target vehicle 200b is outside of the first threshold (e.g., the time before collision between primary vehicle 200a and target vehicle 200b exceeds the first threshold), then action response module 2906 may cause a response in primary vehicle 200a based on an averaged motion profile associated with target vehicle 200b. The averaged motion profile associated with target vehicle 200b may represent an averaged position and/or averaged speed of target vehicle 200b for a particular threshold. In some embodiments, if target vehicle 200b is a certain distance and/or traveling a certain speed, system 100 may elect to take no action. For example, if target vehicle 200b is ahead or a sufficient distance away from primary vehicle 200a (e.g., five meters, 10 meters, etc.), system 100 may not take any action unless target vehicle 200b becomes closer to primary vehicle 200.

[0315] If the indicator of intercept time indicates that target vehicle 200b is between the first and second thresholds (e.g., the time before collision between primary vehicle 200a and target vehicle 200b falls in between the first and second thresholds), then action response module 2906 may cause a response in primary vehicle 200a based on the actual motion associated with target vehicle 200b. The actual motion may represent a position and/or speed of target vehicle 200b (e.g., not averaged or delayed) If the indicator of intercept time indicates that target vehicle 200b is within the second threshold (e.g., the time before collision between primary vehicle 200a and target vehicle 200b is below the second threshold), then action response module 2906 may cause an emergency action in primary vehicle 200a (such as a lane change or emergency braking).

[0316] FIG. 31 is a flow chart showing an exemplary process 3100 for declining to cause a response in a primary vehicle, consistent with disclosed embodiments. At step 3110, processing unit 110 may receive a plurality of images via data interface 128 between processing unit 110 and image acquisition unit 120, as described in connection with step 3010 of FIG. 30, above.

[0317] At step 3120, processing unit 110 may execute

IPR2025-00944
Tesla EX1002 Page 1190

target monitoring module 2902 to locate a target object (e.g., target vehicle 200b) within the images received at step 3010 by analyzing the images, as described in connection with step 3020 of FIG. 30, above. At step 3125, processing unit 110 may execute target monitoring module 2902 to determine whether the target object (e.g., target vehicle 200b) is traveling in a lane different from primary vehicle 200a, using the techniques described in connection with step 3030 of FIG. 30, above. If target vehicle 200b is determined to be traveling in a lane different from primary vehicle 200a (step 3125, yes), processing unit 110 may decline to cause a response in primary vehicle 200a at step 3130. Alternatively, if system 100 determines that target vehicle 200b does not constitute an imminent risk of collision (e.g., target vehicle 200b a nonintersecting trajectory with primary vehicle 200a or a potential time to collision beyond a threshold such that the target object does not pose an imminent risk at the present time), processing unit 110 may decline to cause a response in primary vehicle 200a at step 3130. Otherwise (step 3125, no), process 3100 may conclude and no response occurs.

Stopping an Autonomously Driven Vehicle Using a Multi-Segment Braking Profile

[0318] System 100 may provide driver assist functionality that monitors and responds to environmental conditions to control the braking of vehicle 200. For instance, system 100 may detect, based on a plurality of images, an object that triggers the stopping of vehicle 200. For example, this object may include another vehicle in the vicinity of vehicle 200, a traffic light, and/or a traffic sign. System 100 may use a braking profile to slow and/or stop vehicle 200. According to some embodiments, the braking profile may include multiple segments, where deceleration is executed at different stages. For example, the braking profile may cause system 100 to stop vehicle 200 at a decreasing rate of deceleration.

[0319] By providing a braking profile that includes multiple segments, system 100 may cause vehicle 200 to operate in a manner that feels natural to a driver and/or passengers. For example, when a driver stops a vehicle at a traffic light, the brakes are typically not applied at a constant rate from the moment they are applied until the moment when the vehicle stops. Operating brakes in such a manner result in an abrupt stop at the end. Instead, the driver may apply the brakes gently at first, then with increasing pressure, and then as the vehicle reaches its target stopping location, the driver may decrease the level of braking until it reaches zero (or near zero) at the moment the car stops. Operating the vehicle in this manner may provide a natural stopping progression, i.e., the driver may cause-a progressive deceleration at first so the vehicle does not experience an abrupt initial deceleration and then at the end, so that vehicle does not come to an abrupt stop.

[0320] System 100 may provide autonomous vehicle navigation that may implement braking according to an approach that feels natural to the driver and/or passengers of vehicle 200. For example, in some embodiments, system 100 may cause one or more of image capture devices 122-126 to acquire a plurality of images of an area in a vicinity of vehicle 200. Processing device 110 may receive the receive the plurality of images via data interface 128 and identify, based on analysis of the plurality of images, a trigger (e.g., an object, a traffic light, and/or a traffic sign) for stopping vehicle 200. Based on the identified trigger, system 100 may cause vehicle 200 to stop according to a braking profile. In some embodiments, the braking profile may include a plurality of segments (e.g., including a first segment associated with a first deceleration rate, a second segment which includes a second deceleration rate less than the first deceleration rate, and a third segment in which a level of braking is decreased as a target stopping location is approached, as determined based on the analysis of the plurality of images.

[0321] FIG. 32A illustrates vehicle 200 traveling on a roadway 3210 that includes an object 3220. System 100 may identify a trigger (e.g., object 3220) for stopping vehicle 200 in respect to detection of the trigger. For example, system 100 may first slow vehicle 200 after the trigger is detected and bring vehicle 200 to a stop before reaching object 3220.

[0322] Processing unit 110 may be configured to determine a trigger for stopping vehicle 200 based on a plurality of images acquired by one or more of image capture devices 122-126. Processing unit 110 may receive the images from image capture devices 122-126 via data interface 128. According to some embodiments, the trigger may include object 3220. For example, object 3220 may be another vehicle on roadway 2810. According to some embodiments, processing unit 110 may identify a trigger if object 3220 is within a certain vicinity of vehicle 200. For example, identification of the trigger may depend on a relative distance between object 3220 and vehicle 200. Further, this identification may optionally consider the relative velocity and/or relative acceleration between object 3220 and vehicle 200.

[0323] Additionally or alternatively, the trigger may include a traffic sign, such as traffic sign 3240 shown in FIG. 32A. For example, traffic sign 3240 may be a stop sign and/or a speed limit sign. According to some embodiments, processing unit 110 may process the plurality of images to determine the content of traffic sign 3240 as part of determining whether a trigger exists. For example, processing unit 110 may determine that a trigger is present based on a speed limit specified by traffic sign 3240 if vehicle 200 is exceeding the speed limit. As another example, processing unit 110 may determine a trigger exists based on a traffic light 2830 if the yellow and/or red light is on.

[0324] Based on the plurality of images, processing unit 110 may determine a target stopping location 3250 at which vehicle 200 will stop once a braking profile is

IPR2025-00944
Tesla EX1002 Page 1191

executed. Reaching target stopping location 3250 may be accompanied by a release in braking pressure of vehicle 200. Target stopping location 3250 may be based on the location of traffic light 3230 and/or traffic sign 3240 such that vehicle 200 complies with traffic light 3230 and/or traffic sign 3240. Alternatively, target stopping location 3250 may be based on object 3220 such that vehicle 200 stops before reaching object 3220. Accordingly, target stopping location 3250 may be based on the location of object 3220 to avoid a collision between vehicle 200 and object 3220. In some embodiments, target stopping location 3250 may be updated based on new information obtained from the plurality of images, such as newly identified targets, newly visible obstacles, newly visible signs, etc.

[0325] FIG. 32B illustrates an exemplary braking profile, consistent with disclosed embodiments. In the example shown in FIG. 32B, the braking profile includes four segments. However, the number of segments is exemplary and a braking profile consistent with the disclosed embodiments may include any appropriate number of segments (e.g., 2, 3, 4, 5, 6, etc., segments).

[0326] In segment 1, system 100 may prepare for a possible braking scenario. For example, in this zone, system 100 may receive information (e.g., map data and/or image data) about an upcoming intersection (e.g., an intersection with traffic lights). System 100 may gradually reduce the speed of vehicle 200 (e.g., the speed of vehicle 200 may be gradually reduced to about 70-50 km/hr, depending on the characteristics of the intersection). System 100 may cause the speed adjustment, for example, approximately 100-200 meters before the junction at average deceleration around 0.2-0.4 m/sec^2 (with a maximum of about 0.5 m/sec^2). By decreasing the speed of vehicle 200 in segment 1, system 100 may provide the driver and/or passengers of vehicle 200 with confidence that system 100 is aware of the approaching intersection.

[0327] In segment 2, system 100 may apply strong braking. In this segment, system 100 may receive information based on the analysis of image data that, for example, a traffic light is red and/or there is a stopped vehicle at the intersection in the line in which vehicle 200 is traveling. System 100 may cause vehicle 200 to experience significant speed reduction in this zone, while the distance to the stop line or stopped vehicle is still large. Doing so may provide the driver and/or passengers of vehicle 200 with a comfortable feeling that system 100 will have more than enough space to complete the braking maneuver. This zone may be approximately 30-100 meters away from the stop line, and the average deceleration may be about 1.5-2.5 m/sec^2 (with a maximum of about 3.5 m/sec^2).

[0328] In segment 3, system 100 may make a moderate braking adjustment. For example, in this segment, system 100 may adjust the speed of vehicle 200 according to the remaining distance to the stop line or stopped vehicle and based on the current speed and deceleration of vehicle 200. By making this adjustment, system 100 may provide the driver and/or passengers of vehicle 200 with an indication that system 100 is releasing most of the braking power while the speed is still being reduced but slower than before. This zone may be approximately 5-30 meters away from the stop line or stopped vehicle, and the average deceleration may be about 0.5-1.5 m/sec^2 (with a maximum of about 2 m/sec^2).

[0329] In segment 4, system 100 may make a small braking adjustment. In this segment, system 100 may close the remaining distance to the stop line or stopped vehicle at a very low speed gradually bringing it to zero. In this zone, the driver and/or passengers may feel that vehicle 200 is slowly sliding into its position (e.g., at the stop line or behind another vehicle). This zone may constitute approximately the last 5-7 meters to the stop line or stopped vehicle, and the average deceleration may be about 0.3-0.5 m/sec^2 (with a maximum of about 1 m/sec^2).

[0330] FIG. 33 is an exemplary block diagram of memory 140 and/or 150, which may store instructions for performing one or more operations consistent with disclosed embodiments. As illustrated in FIG. 33, memory 140 may store one or more modules for performing the trigger detection and responses described herein. For example, memory 140 may store a trigger identification module 3300 and an action response module 3310.

[0331] Trigger identification module 3300 may store instructions which, when executed by processing unit 110, may detect the presence and/or existence of a trigger for stopping vehicle 200. For example, trigger identification module 3300 may process the plurality of images received from at least one image capture device 122-124 to detect the presence of a trigger. As discussed above, this may include identifying object 3220, traffic light 3230, and/or traffic sign 3240.

[0332] Action response module 3310 may store instructions which, when executed by processing unit 110, may respond to the presence of a trigger. For example, action response module 3310 may execute control to decelerate and/or stop vehicle 200. For example, action response module 3310 may execute a braking profile to decelerate and/or stop vehicle 200. The action response module 3310 may determining to execute one of a plurality of braking profiles based on a number of braking profiles based on, for example, environmental factors, the type of trigger detected, and the velocity of vehicle 200. To navigate vehicle 200 to a stop based on the detected trigger, processing unit 110 may transmit electronic signals to one or more of throttling system 220, braking system 230, and/or steering system 240 of vehicle 200. Further, one or more actuators may control throttling system 220, braking system 230, and/or steering system 240. For example, processing unit 110 may transmit electronic signals that cause system 100 to physically depress the brake by a predetermined amount.

[0333] Fig. 34 illustrates a process 3400 for navigating vehicle 200, consistent with disclosed embodiments. Process 3000 may identify a trigger (e.g., an object, ve-

IPR2025-00944
Tesla EX1002 Page 1192

hicle, traffic light, traffic sign, etc.) for stopping vehicle 200 and execute a braking profile in response to the detection of the trigger. Process 3400 may detect various types of triggers determine which of a plurality of braking profiles to execute.

[0334] The braking profile may be based on the type of a detected trigger and/or other environmental conditions related to the detected trigger (e.g., a distance to an object). For example, in some embodiments, the braking profile may include a plurality of segments (e.g., three segments). In a three segment braking profile, a first segment of the braking profile may be associated with a first deceleration rate, a second segment of the braking profile may include a second deceleration rate less than the first deceleration rate, and a third segment of the braking profile may include causing a level of braking to be decreased as a target stopping location is approached, as determined based on an analysis of a plurality of images acquired by one or more of image capture devices 122-126. For example, in some embodiments, the first segment may cause vehicle 200 to progressively increase deceleration, the second segment may cause vehicle 200 to decelerate at a constant rate, and the third segment may cause vehicle 200 to progressively decrease its deceleration rate to zero (or near zero).

[0335] At step 3410, process 3400 may acquire, using at least one of image capture devices 122-126, a plurality of images of an area in the vicinity of vehicle 200. For example, processing unit 110 may receive the plurality of images may through data interface 128. Processing unit 110 may be configured to determine the presence of one or more objects and/or signs located on or near roadway 3210 based on the acquired images. For example, processing unit 110 may be configured to determine the presence of object 3220, traffic light 3230, and/or traffic sign 3240. Processing unit 110 may be further configured to read the content of traffic sign 3240. For example, processing unit 110 may be configured to differentiate between a speed limit sign and a stop sign. Further, processing unit 110 may be configured to process images to determine the speed limit based on traffic sign 3240.

[0336] At step 3420, process 300 may include a trigger for stopping vehicle 200. For example, processing unit 110 may identify a trigger based on the presence of object 3220 within the vicinity of vehicle 200. Additionally or alternatively, processing unit 110 may identify a trigger based on a current signal status of traffic light 3230. According to some embodiments, processing unit 110 may determine that traffic light 3230 constitutes a trigger if it is not green. According to some embodiments, processing unit 110 may determine that traffic sign 3240 constitutes a trigger if it is a stop sign, for example. In some embodiments, traffic sign 3240 may constitute a trigger if it is a speed limit sign and vehicle 200 is exceeding the posted speed limit. Further in some embodiments, processing unit 110 may cause system 100 to provide an audible announcement identifying the trigger and/or indicating that vehicle 200 is braking or about to begin braking.

[0337] After processing unit 110 identifies the trigger, at step 3430, processing unit 110 may cause vehicle 200 to slow down and/or stop according to a braking profile. In some embodiments, the braking profile may comprise a number of profile segments. The different profile segments of a braking profile may be defined by a time period and/or a vehicle speed. For example, processing unit 110 may execute controls of vehicle 200 in accordance with a first profile segment for a predetermined period of time or until the speed of vehicle 200 reaches a target value. The target value may be a predetermined speed or it may be a percentage of the speed that vehicle 200 was traveling prior to execution of the braking profile. To implement the braking profile, processing unit 110 may transmit electronic signals to one or more of throttling system 220, braking system 230, and/or steering system 240 of vehicle 200. Further, one or more actuators may control throttling system 220, braking system 230, and/or steering system 240.

[0338] As discussed above, according to some embodiments, a profile segment of the braking profile may correspond to a deceleration rate. For example, a braking profile may include multiple segments, and the first segment may be associated with a first deceleration rate. The braking profile may further include a second segment associated with a second deceleration rate. The second deceleration rate may be less than first deceleration rate.

[0339] Further, as discussed above, according to some embodiments, braking profile may include a third profile segment in which a level of braking is decreased as a function of the distance between vehicle 200 and underlying factors that give rise to the trigger (e.g., object 3220, traffic light 3230, traffic sign 3240). For example, the braking profile may include decreasing the level of braking as vehicle 200 approaches target stopping location 3250. Further, reaching target stopping location 3250 may be accompanied by a release in braking pressure of vehicle 200.

[0340] The foregoing description has been presented for purposes of illustration. It is not exhaustive and is not limited to the precise forms or embodiments disclosed. Modifications and adaptations will be apparent to those skilled in the art from consideration of the specification and practice of the disclosed embodiments. Additionally, although aspects of the disclosed embodiments are described as being stored in memory, one skilled in the art will appreciate that these aspects can also be stored on other types of computer readable media, such as secondary storage devices, for example, hard disks or CD ROM, or other forms of RAM or ROM, USB media, DVD, Blu-ray, or other optical drive media.

[0341] Computer programs based on the written description and disclosed methods are within the skill of an experienced developer. The various programs or program modules can be created using any of the techniques known to one skilled in the art or can be designed in

IPR2025-00944
Tesla EX1002 Page 1193

connection with existing software. For example, program sections or program modules can be designed in or by means of .Net Framework, Net Compact Framework (and related languages, such as Visual Basic, C, etc.), Java, C++, Objective-C, HTML, HTML/AJAX combinations, XML, or HTML with included Java applets.

Claims

1. A driver assist navigation system (100) for a primary vehicle (200a), the system comprising:

at least one image capture device (122, 124, 126) configured to acquire a plurality of images of an area in a vicinity of the primary vehicle;
a data interface (128); and
at least one processing device (110) configured to:

receive (1810) the plurality of images via the data interface (128);
determine (1820), from at least some of the plurality of images, a first lane constraint (810) on a first side (802) of the primary vehicle (200a);
determine (1830), from at least some of the plurality of images, a second lane constraint (820) on a second side (812) of the primary vehicle (200a) opposite to the first side (802) of the primary vehicle, wherein the first and second lane constraints define a lane within which the primary vehicle (200a) travels; and
send control signals to the control systems (220, 230, 240) of said primary vehicle for causing (1840) the primary vehicle (200a) to travel within the first and second lane constraints (810, 820);

wherein:

the first lane constraint includes a line marked on a road surface, an edge of a road surface, markers on a road surface, or is estimated based on a determined midpoint of a road surface width;
the second lane constraint includes a line marked on a road surface, an edge of a road surface, markers on a road surface, or is estimated based on a determined midpoint of a road surface width; and
characterized in that
the at least one processing device (110) is further configured to:

locate (1850) in the plurality of images a leading vehicle (200b);
determine (1860), based on the plurality of images, at least one action taken by the leading vehicle (200b); and
send control signals to the control systems (220, 230, 240) for causing (1870) the primary vehicle (200a) to mimic the at least one action of the leading vehicle (200b);

wherein the at least one action taken by the leading vehicle (200b) includes performing a maneuver such that a first offset distance (c_3) on a side of the leading vehicle (200b) adjacent to the first lane constraint (810) is different from a second offset distance (c_4) on a side of the leading vehicle (200b) adjacent to the second lane constraint (820), and
wherein mimicking the at least one action includes the at least one processing device causing the primary vehicle to perform a maneuver such that a third offset distance (c_1) on a side of the primary vehicle (200a) adjacent to the first lane constraint (810) is different from a fourth offset distance (c_2) on a side of the primary vehicle (200a) adjacent to the second lane constraint (820), wherein the third offset distance (c_1) equals to the first offset distance (c_3) and the fourth offset distance (c_2) equals to the second offset distance (c_4).

2. The system of claim 1, wherein the at least one action taken by the leading vehicle (200b) further includes a turn, and mimicking the at least one action respectively includes the at least one processing device (110) causing the primary vehicle (200a) to execute a similar turn beginning at a location at which the leading vehicle (200b) initiated its turn.

3. The system of claim 1, wherein the at least one action taken by the leading vehicle further includes an acceleration and wherein mimicking the at least one action respectively includes the at least one processing device causing the primary vehicle to accelerate.

4. The system of claim 1, wherein the at least one action taken by the leading vehicle further includes braking wherein mimicking the at least one action respectively includes the at least one processing device causing the primary vehicle to brake.

5. The system of claim 1, wherein the at least one processing device (110) is further configured to:

determine (1925), based on the plurality of images or based on map data and position infor-

IPR2025-00944
Tesla EX1002 Page 1194

mation, whether the leading vehicle (200b) is turning at an intersection; and
if the leading vehicle (200b) is determined to be turning at an intersection, decline (1930) to cause the primary vehicle to perform a turn similar to the turn of the leading vehicle (200b).

6. The system of claim 1, wherein the at least one processing device (110) is further configured to:

determine (2025), based on the plurality of images or based on map data and position information, whether the leading vehicle (200b) is turning at an intersection;
determine (2035) whether the leading vehicle (200b) is changing lanes within the intersection;
mimic the turn of the leading vehicle (200b) by causing the primary vehicle (200a) to execute a similar turn beginning at a location at which the leading vehicle (200b) initiated its turn, if the leading vehicle (200b) is determined to be turning at an intersection and not changing lanes within the intersection; and
decline (2050) to cause the primary vehicle to perform a turn similar to the turn of the leading vehicle (200b) if the leading vehicle (200b) is determined to be turning at an intersection, but changing lanes within the intersection.

7. The system of claim 1, wherein:

the leading vehicle is a first leading vehicle;
at least one processing device (110) is further configured to: locate (2130) a second leading vehicle in the plurality of images; and
cause the primary vehicle (200a) to mimic the at least one action of the first leading vehicle if, based on the plurality of images, the first leading vehicle is determined to be continuing on a path having less of a turning radius than a path followed by the second leading vehicle.

8. The system of claim 1, wherein the at least one processing device (110) is further configured to:

determine (2230) position information for the leading vehicle (200b) and compare (2240) the position information to predetermined map data;
create (2250) a navigation history for the leading vehicle (200b) by tracking movements of the leading vehicle (200b) relative to the predetermined map data; and
determine (2260) whether to cause the primary vehicle (200a) to mimic the at least one action of the leading vehicle (200b) based on the navigation history of the leading vehicle (200b).

9. The system of claim 8, wherein the at least one processing device (110) is configured to:

cause the primary vehicle (200a) to perform a turn similar to a turn of the leading vehicle (200b), even where no turn is detected based on the plurality of images, when the navigation history indicates that the leading vehicle (200b) has successfully navigated at least one prior turn.

10. A primary vehicle for following a leading vehicle, the primary vehicle comprising the system of any preceding claim.

11. A method for navigating a primary vehicle (200a), the method comprising:

acquiring (1810), using at least one image capture device (122, 124, 126), a plurality of images of an area in a vicinity of the primary vehicle (200a);
determining (1820), from at least some of the plurality of images, a first lane constraint (810) on a first side (802) of the primary vehicle (200a);
determining (1830), from at least some of the plurality of images, a second lane constraint (820) on a second side (812) of the primary vehicle (200a) opposite to the first side (802) of the primary vehicle (200a), wherein the first and second lane constraints (810, 820) define a lane within which the primary vehicle (200a) travels; and
sending control signals to the control systems (220, 230, 240) of said primary vehicle for causing (1840) the primary vehicle (200a) to travel within the first and second lane constraints (810, 820); wherein:
the first lane constraint includes a line marked on a road surface, an edge of a road surface, markers on a road surface, or is estimated based on a determined midpoint of a road surface width;
the second lane constraint includes a line marked on a road surface, an edge of a road surface, markers on a road surface, or is estimated based on a determined midpoint of a road surface width, and
characterized in that the method further comprises:

locating (1850) in the plurality of images a leading vehicle (200b);
determining (1860), based on the plurality of images, at least one action taken by the leading vehicle (200b); and
sending control signals to the control systems (220, 230, 240) for causing (1870) the primary vehicle (200a) to mimic the at least

IPR2025-00944
Tesla EX1002 Page 1195

one action of the leading vehicle (200b);

wherein the at least one action taken by the leading vehicle (200b) includes performing a maneuver such that a first offset distance (c_3) on a side of the leading vehicle (200b) adjacent to the first lane constraint (810) is different from a second offset distance (c_4) on a side of the leading vehicle (200b) adjacent to the second lane constraint (820), and wherein mimicking the at least one action includes the at least one processing device causing the primary vehicle to perform a maneuver such that a third offset distance (c_1) on a side of the primary vehicle (200a) adjacent to the first lane constraint (810) is different from a fourth offset distance (c_2) on a side of the primary vehicle (200a) adjacent to the second lane constraint (820), wherein the third offset distance (c_1) equals to the first offset distance (c_3) and the fourth offset distance (c_2) equals to the second offset distance (c_4).

12. The method of claim 11, further comprising:

determining (1925), based on the plurality of images or based on map data and position information, whether the leading vehicle (200b) is turning at an intersection; and
if the leading vehicle (200b) is determined to be turning at an intersection, declining (1930) to mimic the turn of the leading vehicle (200b).

13. The method of claim 11, further comprising:

determining (2025), based on the plurality of images or based on map data and position information, whether the leading vehicle (200b) is turning at an intersection;
determining (2035) whether the leading vehicle (200b) is changing lanes within the intersection;
mimicking the turn of the leading vehicle (200b) by causing the primary vehicle (200a) to execute a similar turn beginning at a location at which the leading vehicle (200b) initiated its turn, if the leading vehicle (200b) is determined to be turning at an intersection and not changing lanes within the intersection; and
declining (2050) to cause the primary vehicle to perform a turn similar to the turn of the leading vehicle (200b) if the leading vehicle (200b) is determined to be turning at an intersection, but changing lanes within the intersection.

14. The method of claim 11, wherein the leading vehicle is a first leading vehicle, the method further comprising:

locating (2130) a second leading vehicle in the plurality of images; and
causing the primary vehicle (200a) to mimic the at least one action of first leading vehicle if, based on the plurality of images, the first leading vehicle is determined to be continuing on a path having less of a turning radius than a path followed by the second leading vehicle.

15. The method of claim 11, further comprising to:

determining (2230) position information for the leading vehicle (200b) and compare (2240) the position information to predetermined map data;
creating (2250) a navigation history for the leading vehicle (200b) by tracking movements of the leading vehicle (200b) relative to the predetermined map data; and
determining (2260) whether to cause the primary vehicle (200a) to mimic the at least one action of the leading vehicle (200b) based on the navigation history of the leading vehicle (200b); and optionally causing the primary vehicle (200a) to mimic a turn of the leading vehicle (200b), even where no turn is detected based on the plurality of images, when the navigation history indicates that the leading vehicle (200b) has successfully navigated at least one prior turn.

Patentansprüche

1. Fahrerassistenz-Navigationssystem (100) für ein Primärfahrzeug (200a), wobei das System Folgendes umfasst:

mindestens eine Bilderfassungsvorrichtung (122, 124, 126), die dazu konfiguriert ist, eine Vielzahl von Bildern eines Bereiches in einer Nähe des Primärfahrzeugs zu erfassen;
eine Datenschnittstelle (128); und
mindestens eine Verarbeitungsvorrichtung (110), die dazu konfiguriert ist:

die Vielzahl von Bildern über die Datenschnittstelle (128) zu empfangen (1810);
von zumindest einigen der Vielzahl von Bildern eine erste Fahrspurbeschränkung (810) an einer ersten Seite (802) des Primärfahrzeugs (200a) festzustellen (1820);
von zumindest einigen der Vielzahl von Bildern eine zweite Fahrspurbeschränkung (820) an einer zweiten Seite (812) des Primärfahrzeugs (200a), die der ersten Seite

IPR2025-00944
Tesla EX1002 Page 1196

(802) des Primärfahrzeugs entgegengesetzt ist, festzustellen (1830), wobei die erste und zweite Fahrspurbeschränkung eine Fahrspur definieren, innerhalb welcher das Primärfahrzeug (200a) fährt; und

Steuersignale an die Steuersysteme (220, 230, 240) des Primärfahrzeugs zu senden, um zu veranlassen (1840), dass das Primärfahrzeug (200a) innerhalb der ersten und zweiten Fahrspurbeschränkung (810, 820) fährt;

wobei

die erste Fahrspurbeschränkung eine auf einer Fahrbahnoberfläche markierte Linie, einen Rand einer Fahrbahnoberfläche, Markierungen auf einer Fahrbahnoberfläche umfasst, oder basierend auf einem festgestellten Mittelpunkt einer Fahrbahnoberflächenbreite geschätzt wird;

die zweite Fahrspurbeschränkung eine auf einer Fahrbahnoberfläche markierte Linie, einen Rand einer Fahrbahnoberfläche, Markierungen auf einer Fahrbahnoberfläche umfasst, oder basierend auf einem festgestellten Mittelpunkt einer Fahrbahnoberflächenbreite geschätzt wird; und

dadurch gekennzeichnet, dass mindestens eine Verarbeitungsvorrichtung (110) ferner dazu konfiguriert ist:

in der Vielzahl von Bilder ein führendes Fahrzeug (200b) ausfindig zu machen (1850);

basierend auf der Vielzahl von Bildern, mindestens eine durch das führende Fahrzeug (200b) ausgeführte Handlung festzustellen (1860); und

Steuersignale an die Steuersysteme (220, 230, 240) des Primärfahrzeugs zu senden, um zu veranlassen (1870), dass das Primärfahrzeug (200a) mindestens eine Handlung des führenden Fahrzeugs (200b) nachahmt;

wobei die mindestens eine durch das führende Fahrzeug (200b) ausgeführte Handlung das Ausführen eines Manövers umfasst, so dass sich ein erster Versatzabstand (c_3) an einer Seite des führenden Fahrzeugs (200b) angrenzend an die erste Fahrspurbeschränkung (810) von einem zweiten Versatzabstand (c_4) an einer Seite des führenden Fahrzeugs (200b) angrenzend an die zweite Fahrspurbeschränkung (820) unterscheidet, und

wobei das Nachahmen der mindestens einen Handlung umfasst, dass die mindestens eine Verarbeitungsvorrichtung veranlasst, dass das Primärfahrzeug ein Manöver ausführt, so dass sich ein dritter Versatzabstand (c_1) an einer Seite des Primärfahrzeugs (200a) angrenzend an die erste Fahrspurbeschränkung (810) von einem vierten Versatzabstand (c_2) an einer Seite des Primärfahrzeugs (200a) angrenzend an die zweite Fahrspurbeschränkung (820) unterscheidet,

wobei der dritte Versatzabstand (c_1) gleich dem ersten Versatzabstand (c_3) ist und der vierte Versatzabstand (c_2) gleich dem zweiten Versatzabstand (c_4) ist.

2. System nach Anspruch 1, wobei die mindestens eine durch das führende Fahrzeug (200b) ausgeführte Handlung ferner ein Abbiegen umfasst, und das Nachahmen der mindestens einen Handlung jeweils umfasst, dass die mindestens eine Verarbeitungsvorrichtung (110) veranlasst, dass das Primärfahrzeug (200a) ein ähnliches Abbiegen beginnend an einer Stelle ausführt, an welcher das führende Fahrzeug (200b) sein Abbiegen begonnen hat.

3. System nach Anspruch 1, wobei die mindestens eine durch das führende Fahrzeug ausgeführte Handlung ferner eine Beschleunigung umfasst und wobei das Nachahmen der mindestens einen Handlung jeweils umfasst, dass die mindestens eine Verarbeitungsvorrichtung veranlasst, dass das Primärfahrzeug beschleunigt.

4. System nach Anspruch 1, wobei die mindestens eine durch das führende Fahrzeug ausgeführte Handlung ferner ein Bremsen umfasst, wobei das Nachahmen der mindestens einen Handlung jeweils umfasst, dass die mindestens eine Verarbeitungsvorrichtung veranlasst, dass das Primärfahrzeug bremst.

5. System nach Anspruch 1, wobei die mindestens eine Verarbeitungsvorrichtung (110) ferner dazu konfiguriert ist,

basierend auf der Vielzahl von Bildern oder basierend auf Kartendaten und Postitionsinformation festzustellen (1925), ob das führende Fahrzeug (200b) an einer Kreuzung abbiegt; und,

wenn festgestellt wird, dass das führende Fahrzeug (200b) an einer Kreuzung abbiegt, das Veranlassen, dass das Primärfahrzeug ein ähnliches Abbiegen wie das Abbiegen des führenden Fahrzeugs (200b) ausführt, abzulehnen (1930).

6. System nach Anspruch 1, wobei die mindestens eine Verarbeitungsvorrichtung (110) ferner dazu konfiguriert ist,

basierend auf der Vielzahl von Bildern oder basierend auf Kartendaten und Postitionsinformation festzustellen (2025), ob das führende Fahrzeug (200b) an einer Kreuzung abbiegt;

festzustellen (2035), ob das führende Fahrzeug

IPR2025-00944
Tesla EX1002 Page 1197

(200b) Fahrspuren innerhalb der Kreuzung wechselt;
das Abbiegen des führenden Fahrzeugs (200b) durch Veranlassen, dass das Primärfahrzeug (200a) ein ähnliches Abbiegen an einer Stelle ausführt, an welcher das führende Fahrzeug (200b) sein Abbiegen begonnen hat, nachzuahmen, wenn festgestellt wird, dass das führende Fahrzeug (200b) an einer Kreuzung abbiegt und keine Fahrspuren innerhalb der Kreuzung wechselt; und
ein Veranlassen, dass das Primärfahrzeug ein ähnliches Abbiegen wie das Abbiegen des führenden Fahrzeugs (200b) ausführt, abzulehnen (2050), wenn festgestellt wird, dass das führende Fahrzeug (200b) an einer Kreuzung abbiegt, jedoch Fahrspuren innerhalb der Kreuzung wechselt.

7. System nach Anspruch 1, wobei:

das führende Fahrzeug ein erstes führendes Fahrzeug ist;
mindestens eine Verarbeitungsvorrichtung (110) ferner dazu konfiguriert ist: ein zweites führendes Fahrzeug in der Vielzahl von Bildern ausfindig zu machen (2130); und
zu veranlassen, dass das Primärfahrzeug (200a) die mindestens eine Handlung des ersten führenden Fahrzeugs nachahmt, wenn, basierend auf der Vielzahl von Bildern, festgestellt wird, dass das erste führende Fahrzeug auf einer Bahn verbleibt, die einen geringeren Kurvenradius aufweist als eine vom zweiten führenden Fahrzeug gefolgte Bahn.

8. System nach Anspruch 1, wobei die mindestens eine Verarbeitungsvorrichtung (110) ferner dazu konfiguriert ist:

Positionsinformation für das führende Fahrzeug (200b) festzustellen (2230) und die Positionsinformation mit vorbestimmten Kartendaten zu vergleichen (2240);
einen Navigationsverlauf für das führende Fahrzeug (220b) durch Verfolgen von Bewegungen des führenden Fahrzeugs (200b) relativ zu den vorbestimmten Kartendaten zu erzeugen (2250); und
festzustellen (2260), ob veranlasst werden soll, dass das Primärfahrzeug (200a) die mindestens eine Handlung des führenden Fahrzeugs (200b) basierend auf dem Navigationsverlauf des führenden Fahrzeugs (200b) nachahmt.

9. System nach Anspruch 8, wobei die mindestens eine Verarbeitungsvorrichtung (110) dazu konfiguriert ist:

zu veranlassen, dass das Primärfahrzeug (200a) ein Abbiegen ausführt, das einem Abbiegen des führenden Fahrzeugs (200b) ähnlich ist, selbst dort, wo kein Abbiegen basierend auf der Vielzahl von Bildern erkannt wird, wenn der Navigationsverlauf angibt, dass das führende Fahrzeug (200b) mindestens ein vorhergehendes Abbiegen erfolgreich befahren hat.

10. Primärfahrzeug zum Folgen eines führenden Fahrzeugs, wobei das Primärfahrzeug das System nach einem der vorhergehenden Ansprüche umfasst.

11. Verfahren zum Steuern eines Primärfahrzeugs (200a), wobei das Verfahren Folgendes umfasst:

Erfassen (1810), unter Verwendung von mindestens einer Bilderfassungsvorrichtung (122, 124, 126), einer Vielzahl von Bildern eines Bereiches in einer Nähe des Primärfahrzeugs (200a);
Feststellen (1820), von zumindest einigen der Vielzahl von Bildern, einer ersten Fahrspurbeschränkung (810) an einer ersten Seite (802) des Primärfahrzeugs (200a);
Feststellen (1830), von zumindest einigen der Vielzahl von Bildern, einer zweiten Fahrspurbeschränkung (820) an einer zweiten Seite (812) des Primärfahrzeugs (200a), die der ersten Seite (802) des Primärfahrzeugs (200a) entgegengesetzt ist, wobei die erste und zweite Fahrspurbeschränkung (810, 820) eine Fahrspur definieren, innerhalb welcher das Primärfahrzeug (200a) fährt; und
Senden von Steuersignalen an die Steuersysteme (220, 230, 240) des Primärfahrzeugs, um zu veranlassen (1840), dass das Primärfahrzeug (200a) innerhalb der ersten und zweiten Fahrspurbeschränkung (810, 820) fährt;
wobei
die erste Fahrspurbeschränkung eine auf einer Fahrbahnoberfläche markierte Linie, einen Rand einer Fahrbahnoberfläche, Markierungen auf einer Fahrbahnoberfläche umfasst, oder basierend auf einem festgestellten Mittelpunkt einer Fahrbahnoberflächenbreite geschätzt wird;
die zweite Fahrspurbeschränkung eine auf einer Fahrbahnoberfläche markierte Linie, einen Rand einer Fahrbahnoberfläche, Markierungen auf einer Fahrbahnoberfläche umfasst, oder basierend auf einem festgestellten Mittelpunkt einer Fahrbahnoberflächenbreite geschätzt wird; und
dadurch gekennzeichnet, dass das Verfahren ferner Folgendes umfasst:

Ausfindigmachen (1850) eines führenden Fahrzeugs (200b) in der Vielzahl von Bildern;
Feststellen (1860), basierend auf der Viel-

IPR2025-00944
Tesla EX1002 Page 1198

zahl von Bildern, von mindestens einer durch das führende Fahrzeug (200b) ausgeführten Handlung; und

Senden von Steuersignalen an die Steuersysteme (220, 230, 240) des Primärfahrzeugs, um zu veranlassen (1870), dass das Primärfahrzeug (200a) die mindestens eine Handlung des führenden Fahrzeugs (200b) nachahmt;

wobei die mindestens eine durch das führende Fahrzeug (200b) ausgeführte Handlung das Ausführen eines Manövers umfasst, so dass sich ein erster Versatzabstand (c_3) an einer Seite des führenden Fahrzeugs (200b) angrenzend an die erste Fahrspurbeschränkung (810) von einem zweiten Versatzabstand (c_4) an einer Seite des führenden Fahrzeugs (200b) angrenzend an die zweite Fahrspurbeschränkung (820) unterscheidet, und

wobei das Nachahmen der mindestens einen Handlung umfasst, dass die mindestens eine Verarbeitungsvorrichtung veranlasst, dass das Primärfahrzeug ein Manöver ausführt, so dass sich ein dritter Versatzabstand (c_1) an einer Seite des Primärfahrzeugs (200a) angrenzend an die erste Fahrspurbeschränkung (810) von einem vierten Versatzabstand (c_2) an einer Seite des Primärfahrzeugs (200a) angrenzend an die zweite Fahrspurbeschränkung (820) unterscheidet,

wobei der dritte Versatzabstand (c_1) gleich dem ersten Versatzabstand (c_3) ist und der vierte Versatzabstand (c_2) gleich dem zweiten Versatzabstand (c_4) ist.

12. Verfahren nach Anspruch 11, ferner umfassend:

Feststellen (1925), basierend auf der Vielzahl von Bildern oder basierend auf Kartendaten und Postitionsinformation, ob das führende Fahrzeug (200b) an einer Kreuzung abbiegt; und,

wenn festgestellt wird, dass das führende Fahrzeug (200b) an einer Kreuzung abbiegt, Ablehnen (1930), das Abbiegen des führenden Fahrzeugs (200b) nachzuahmen.

13. Verfahren nach Anspruch 11, ferner umfassend:

Feststellen (2025), basierend auf der Vielzahl von Bildern oder basierend auf Kartendaten und Postitionsinformation, ob das führende Fahrzeug (200b) an einer Kreuzung abbiegt;

Feststellen (2035), ob das führende Fahrzeug (200b) Fahrspuren innerhalb der Kreuzung wechselt;

Nachahmen des Abbiegens des führenden Fahrzeugs (200b) durch Veranlassen, dass das Primärfahrzeug (200a) ein ähnliches Abbiegen an einer Stelle ausführt, an welcher das führende Fahrzeug (200b) sein Abbiegen begonnen hat, wenn festgestellt wird, dass das führende Fahrzeug (200b) an einer Kreuzung abbiegt und keine Fahrspuren innerhalb der Kreuzung wechselt; und

Ablehnen (2050) des Veranlassens, dass das Primärfahrzeug ein ähnliches Abbiegen wie das Abbiegen des führenden Fahrzeugs (200b) ausführt, wenn festgestellt wird, dass das führende Fahrzeug (200b) an einer Kreuzung abbiegt, jedoch Fahrspuren innerhalb der Kreuzung wechselt.

14. Verfahren nach Anspruch 11, wobei das führende Fahrzeug ein erstes führendes Fahrzeug ist, wobei das Verfahren ferner Folgendes umfasst:

Ausfindigmachen (2130) eines zweiten führenden Fahrzeugs in der Vielzahl von Bildern; und Veranlassen, dass das Primärfahrzeug (200a) die mindestens ein Handlung des ersten führenden Fahrzeugs nachahmt, wenn, basierend auf der Vielzahl von Bildern, festgestellt wird, dass das erste führende Fahrzeug auf einer Bahn verbleibt, die einen geringeren Kurvenradius aufweist als eine vom zweiten führenden Fahrzeug gefolgte Bahn.

15. Verfahren nach Anspruch 11, ferner umfassend:

Feststellen (2230) von Positionsinformation für das führende Fahrzeug (200b) und Vergleichen (2240) der Positionsinformation mit vorbestimmten Kartendaten;

Erzeugen (2250) eines Navigationsverlaufes für das führende Fahrzeug (200b) durch Verfolgen von Bewegungen des führenden Fahrzeugs (200b) relativ zu den vorbestimmten Kartendaten; und

Feststellen (2260), ob veranlasst werden soll, dass das Primärfahrzeug (200a) die mindestens eine Handlung des führenden Fahrzeugs (200b) basierend auf dem Navigationsverlauf des führenden Fahrzeugs (200b) nachahmt; und

optional Veranlassen, dass das Primärfahrzeug (200a) ein Abbiegen des führenden Fahrzeugs (200b) nachahmt, selbst dort, wo kein Abbiegen basierend auf der Vielzahl von Bildern erkannt wird, wenn der Navigationsverlauf angibt, dass das führende Fahrzeug (200b) mindestens ein vorhergehendes Abbiegen erfolgreich befahren hat.

IPR2025-00944
Tesla EX1002 Page 1199

Revendications

1. Système de navigation d'assistance au conducteur (100) pour un véhicule principal (200a), ledit système comprenant : au moins un dispositif de capture d'image (122, 124, 126) conçu pour acquérir une pluralité d'images d'une zone dans un voisinage du véhicule principal ; une interface de données (128) ; au moins un dispositif de traitement (110) conçu pour :

   recevoir (1810) la pluralité d'images par l'intermédiaire de l'interface de données (128) ;
   déterminer (1820), à partir d'au moins certaines images de la pluralité d'images, une première contrainte de voie (810) sur un premier côté (802) du véhicule principal (200a) ;
   déterminer (1830), à partir d'au moins certaines images de la pluralité d'images, une seconde contrainte de voie (820) sur un second côté (812) du véhicule principal (200a) opposé au premier côté (802) du véhicule principal, lesdites première et seconde contraintes de voie définissant une voie dans laquelle le véhicule principal (200a) circule ;
   et pour envoyer des signaux de commande aux systèmes de commande (220, 230, 240) dudit véhicule principal pour amener (1840) le véhicule principal (200a) à circuler dans les première et seconde contraintes de voie (810, 820) ;
   ladite première contrainte de voie comprenant une ligne marquée sur une surface de chaussée, un bord d'une surface de chaussée, des marqueurs sur une surface de chaussée ou étant estimée sur la base d'un point médian déterminé d'une largeur de surface de chaussée ;
   ladite seconde contrainte de voie comprenant une ligne marquée sur une surface de chaussée, un bord d'une surface de chaussée, des marqueurs sur une surface de chaussée, ou étant estimée sur la base d'un point médian déterminé d'une largeur de surface de chaussée ;
   et **caractérisé en ce que** le au moins un dispositif de traitement (110) est conçu en outre pour: localiser (1850) dans la pluralité d'images un véhicule de tête (200b) ; déterminer (1860), sur la base de la pluralité d'images, au moins une action prise par le véhicule de tête (200b) ;
   et envoyer des signaux de commande aux systèmes de commande (220, 230, 240) dudit véhicule principal pour amener (1870) le véhicule principal (200a) à imiter la au moins une action du véhicule de tête (200b) ;
   ladite au moins une action prise par le véhicule de tête (200b) comprenant la réalisation d'une manoeuvre de sorte qu'une première distance de décalage (c_3) sur un côté du véhicule de tête (200b) adjacent à la première contrainte de voie (810) soit différente d'une seconde distance de décalage (c_4) sur un côté du véhicule de tête (200b) adjacent à la seconde contrainte de voie (820), et ladite imitation de la au moins une action comprenant le au moins un dispositif de traitement amenant le véhicule principal à réaliser une manoeuvre de sorte qu'une troisième distance de décalage (c_1) sur un côté du véhicule principal (200a) adjacent à la première contrainte de voie (810) soit différente d'une quatrième distance de décalage (c_2) sur un côté du véhicule principal (200ba) adjacent à la seconde contrainte de voie (820), ladite troisième distance de décalage (c_1) étant égale à la première distance de décalage (c_3) et ladite quatrième distance de décalage (c_2) étant égale à la seconde distance de décalage (c_4).

2. Système selon la revendication 1, ladite au moins une action prise par le véhicule de tête (200b) comprenant en outre une prise de virage et ladite imitation de la au moins une action respectivement comprenant le au moins un dispositif de traitement (110) amenant le véhicule principal (200a) à exécuter une prise de virage semblable commençant au niveau d'un emplacement au niveau duquel le véhicule de tête (200b) a amorcé sa prise de virage.

3. Système selon la revendication 1, ladite au moins une action prise par le véhicule de tête comprenant en outre une accélération et ladite imitation de la au moins une action respectivement comprenant ledit au moins un dispositif de traitement amenant le véhicule principal à accélérer.

4. Système selon la revendication 1, ladite au moins une action prise par le véhicule de tête comprenant en outre le freinage, ladite imitation de la au moins une action respectivement comprenant le au moins un dispositif de traitement amenant le véhicule principal à freiner.

5. Système selon la revendication 1, ledit au moins un dispositif de traitement (110) étant conçu en outre pour déterminer (1925) sur la base de la pluralité d'images ou sur la base des informations de données cartographiques et de position, si le véhicule de tête (200b) tourne à une intersection ;
   et si le véhicule de tête (200b) est déterminé à tourner à une intersection, pour refuser (1930) d'amener le véhicule principal à effectuer une prise de virage semblable à la prise de virage du véhicule de tête (200b).

6. Système selon la revendication 1, ledit au moins un dispositif de traitement (110) étant conçu en outre pour déterminer (2025) sur la base de la pluralité d'images ou sur la base des informations de don-

IPR2025-00944
Tesla EX1002 Page 1200

nées cartographiques et de position, si le véhicule de tête (200b) tourne à une intersection ;
déterminer (2035) si le véhicule de tête (200b) change de voies dans l'intersection ;
imiter la prise de virage du véhicule de tête (200b) en amenant le véhicule principal (200a) à exécuter une prise de virage semblable commençant au niveau d'un emplacement au niveau duquel le véhicule de tête (200b) a amorcé sa prise de virage, si le véhicule de tête (200 b) est déterminé à tourner à une intersection et à ne pas changer de voie dans l'intersection ;
et refuser (2050) d'amener le véhicule principal à effectuer une prise de virage semblable à la prise de virage du véhicule de tête (200b) si le véhicule de tête (200b) est déterminé à tourner à une intersection mais changeant de voies dans l'intersection.

7. Système selon la revendication 1,
ledit véhicule de tête étant un premier véhicule de tête ;
au moins un dispositif de traitement (110) étant conçu en outre pour : localiser (2130) un second véhicule de tête dans la pluralité d'images ;
et pour amener le véhicule principal (200a) à imiter la au moins une action du premier véhicule de tête si, sur la base de la pluralité d'images, le premier véhicule de tête est déterminé à continuer sur un trajet présentant un rayon de braquage inférieur à celui d'un trajet suivi par le second véhicule de tête.

8. Système selon la revendication 1, ledit au moins un dispositif de traitement (110) étant en outre conçu pour :

déterminer (2230) des informations de position pour le véhicule de tête (200b) et comparer (2240) les informations de position aux données cartographiques prédéfinies ;
créer (2250) un historique de navigation pour le véhicule de tête (200b) en suivant les déplacements du véhicule de tête (200b) par rapport aux données cartographiques prédéfinies ;
et pour déterminer (2260) d'amener ou non le véhicule principal (200a) à imiter la au moins une action du véhicule de tête (200b) sur la base de l'historique de navigation du véhicule de tête (200b).

9. Système selon la revendication 8, ledit au moins un dispositif de traitement (110) étant conçu pour :

amener le véhicule principal (200a) à effectuer une prise de virage semblable à la prise de virage du véhicule de tête (200b), même lorsque aucun virage n'est détecté sur la base de la pluralité d'images, lorsque l'historique de navigation indique que le véhicule de tête (200b) présente une navigation réussie pour au moins une prise de virage précédente.

10. Véhicule principal destiné à suivre un véhicule de tête, ledit véhicule principal comprenant le système selon l'une quelconque des revendications précédentes.

11. Procédé permettant la navigation d'un véhicule principal (200a), ledit procédé comprenant : l'acquisition (1810), à l'aide d'au moins un dispositif de capture d'image (122, 124, 126), d'une pluralité d'images d'une zone dans un voisinage du véhicule principal (200a) ;
la détermination (1820), à partir d'au moins certaines images de la pluralité d'images, d'une première contrainte de voie (810) sur un premier côté (802) du véhicule principal (200a) ;
la détermination (1830), à partir d'au moins certaines images de la pluralité d'images, d'une seconde contrainte de voie (820) sur un second côté (812) du véhicule principal (200a) opposé au premier côté (802) du véhicule principal (200a), lesdites première et seconde contraintes de voie (810, 820) définissant une voie dans laquelle le véhicule principal (200a) circule ;
et l'envoi de signaux de commande aux systèmes de commande (220, 230, 240) dudit véhicule principal pour amener (1840) le véhicule principal (200a) à circuler dans les première et seconde contraintes de voie (810, 820) ;
ladite première contrainte de voie comprenant une ligne marquée sur une surface de chaussée, un bord d'une surface de chaussée, des marqueurs sur une surface de chaussée ou étant estimée sur la base d'un point médian déterminé d'une largeur de surface de chaussée ;
la seconde contrainte de voie comprenant une ligne marquée sur une surface de chaussée, un bord de surface de chaussée, des marqueurs sur une surface de chaussée, ou étant estimée sur la base d'un point médian déterminé de la largeur de surface de chaussée et **caractérisé en ce que** le procédé comprend en outre : la localisation (1850) dans la pluralité d'images d'un véhicule de tête (200b) ;
la détermination (1860), sur la base de la pluralité d'images, d'au moins une action prise par le véhicule de tête (200b) ;
et l'envoi de signaux de commande aux systèmes de commande (220, 230, 240) dudit véhicule principal pour amener (1870) le véhicule principal (200a) à imiter la au moins une action du véhicule de tête (200b) ;
ladite au moins une action prise par le véhicule de tête (200b) comprenant la réalisation d'une manoeuvre de sorte qu'une première distance de décalage (c_3) sur un côté du véhicule de tête (200b) adjacent à la première contrainte de voie (810) est différente

IPR2025-00944
Tesla EX1002 Page 1201

d'une seconde distance de décalage (c_4) sur un côté du véhicule de tête (200b) adjacent à la seconde contrainte de voie (820), et l'imitation de la au moins une action comprenant ledit au moins un dispositif de traitement amenant le véhicule principal à réaliser une manoeuvre de sorte qu'une troisième distance de décalage (c_1) sur un côté du véhicule principal (200a) adjacent à la première contrainte de voie (810) est différente d'une quatrième distance de décalage (c_2) sur un côté du véhicule principal (200a) adjacent à la seconde contrainte de voie (820), ladite troisième distance de décalage (c_1) étant égale à la première distance de décalage (c_3) et la quatrième distance de décalage (c_2) étant égale à la seconde distance de décalage (c_4).

12. Procédé selon la revendication 11, comprenant en outre :

la détermination (1925) sur la base de la pluralité d'images ou sur la base des informations de données cartographiques et de position, pour savoir si le véhicule de tête (200b) tourne à une intersection ;
et si le véhicule de tête (200 b) est déterminé à tourner à une intersection, le refus (1930) d'imiter la prise de virage du véhicule de tête (200b).

13. Procédé selon la revendication 11, comprenant en outre :

la détermination (2025) sur la base de la pluralité d'images ou sur la base des informations de données cartographiques et de position, pour savoir si le véhicule de tête (200b) tourne à une intersection ;
la détermination (2035) pour savoir si le véhicule de tête (200b) change de voies dans l'intersection ; l'imitation de la prise de virage du véhicule de tête (200b) en amenant le véhicule principal (200a) à exécute une prise de virage semblable commençant au niveau d'un emplacement au niveau duquel le véhicule de tête (200b) a amorcé sa prise de virage, si le véhicule de tête (200b) est déterminé à tourner à une intersection et à ne pas changer de voies dans l'intersection ;
et le refus (2050) à amener le véhicule principal à effectuer une prise de virage semblable à la prise de virage du véhicule de tête (200b), si le véhicule de tête (200b) est déterminé à tourner à une intersection, mais changeant de voie dans l'intersection.

14. Procédé selon la revendication 11, ledit véhicule de tête étant un premier véhicule de tête, ledit procédé comprenant en outre :

la localisation (2130) d'un second véhicule de tête dans la pluralité d'images ; et l'incitation du véhicule principal (200a) à imiter la au moins une action du premier véhicule de tête si, sur la base de la pluralité d'images, le premier véhicule de tête est déterminer à continuer sur un trajet présentant un rayon de braquage inférieur à celui du trajet suivi par le second véhicule de tête.

15. Procédé selon la revendication 11, comprenant en outre :

la détermination (2230) des informations de position pour le véhicule de tête (200b) et la comparaison (2240) des informations de position aux données cartographiques prédéfinies ;
la création (2250) d'un historique de navigation pour le véhicule de tête (200b) en suivant les déplacements du véhicule de tête (200b) par rapport aux données cartographiques prédéfinies ;
et la détermination (2260) pour amener ou non le véhicule principal (200a) à imiter la au moins une action du véhicule de tête (200b) sur la base de l'historique de navigation du véhicule de tête (200b) ;
et éventuellement l'incitation du véhicule principal (200a) à imiter une prise de virage du véhicule de tête (200b), même lorsque aucun virage n'est détecté sur la base de la pluralité d'images, lorsque l'historique de navigation indique que le véhicule de tête (200 b) présente une navigation réussie pour au moins une prise de virage précédente.

IPR2025-00944
Tesla EX1002 Page 1202

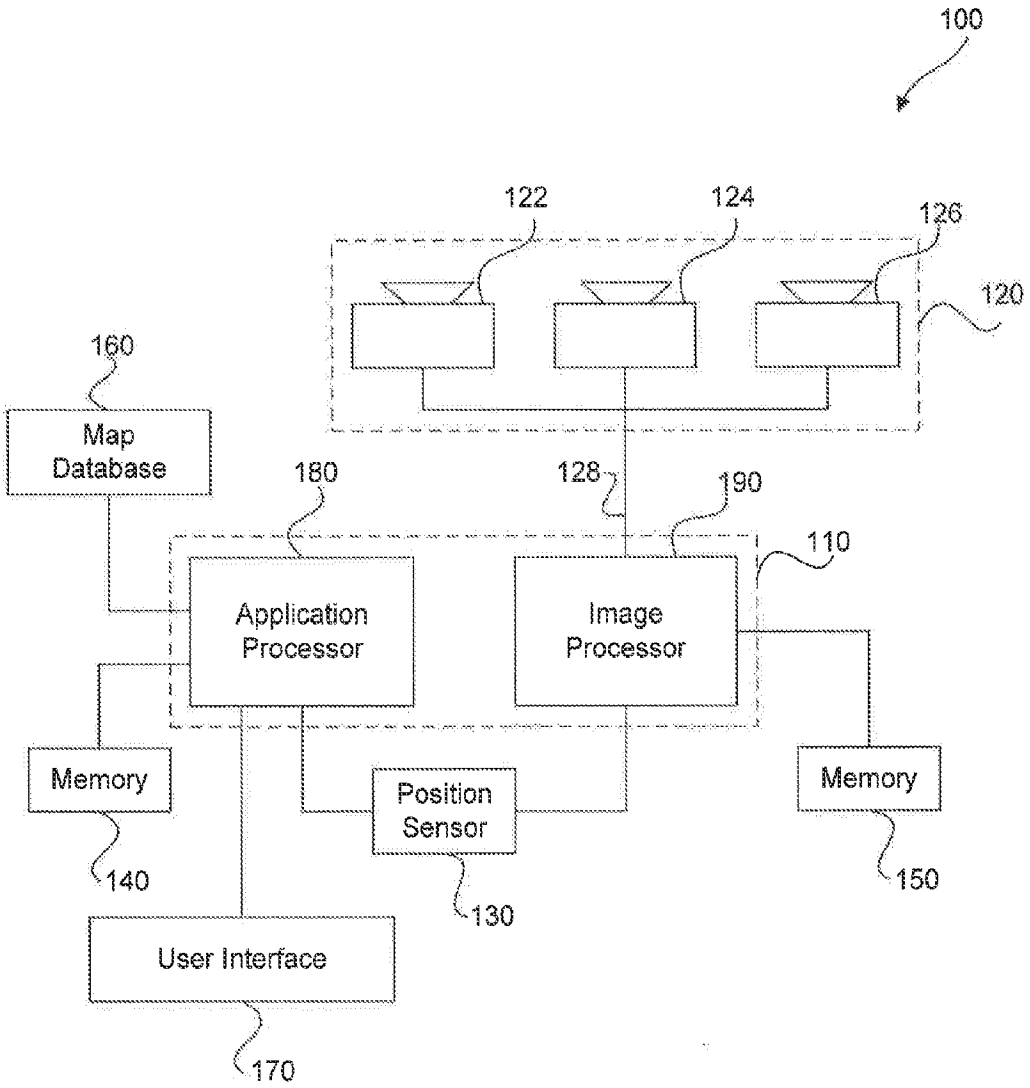


FIG. 1

IPR2025-00944
Tesla EX1002 Page 1203

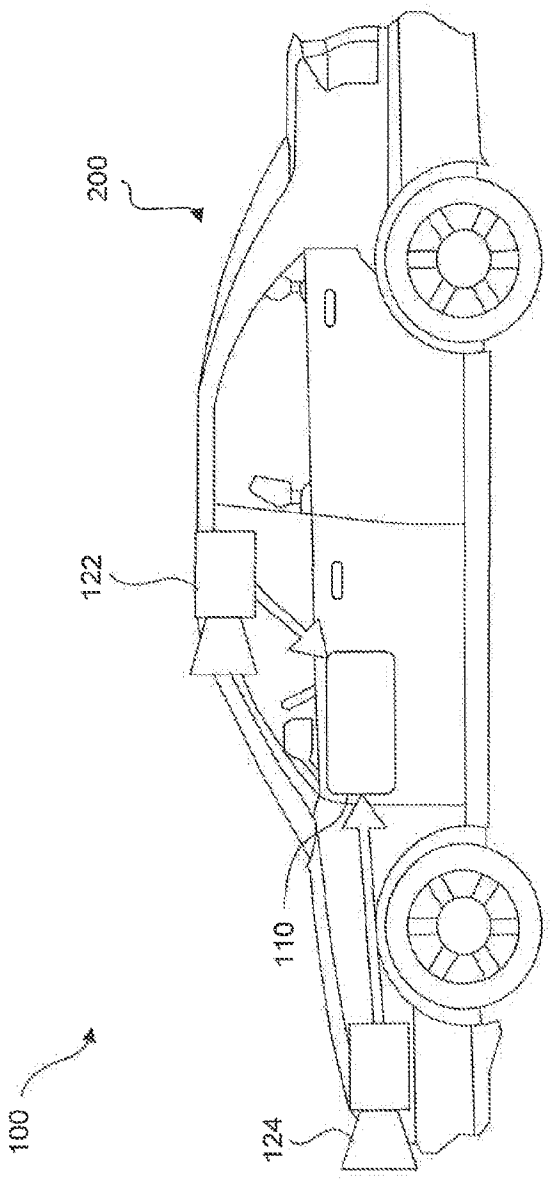


FIG. 2A

IPR2025-00944
Tesla EX1002 Page 1204

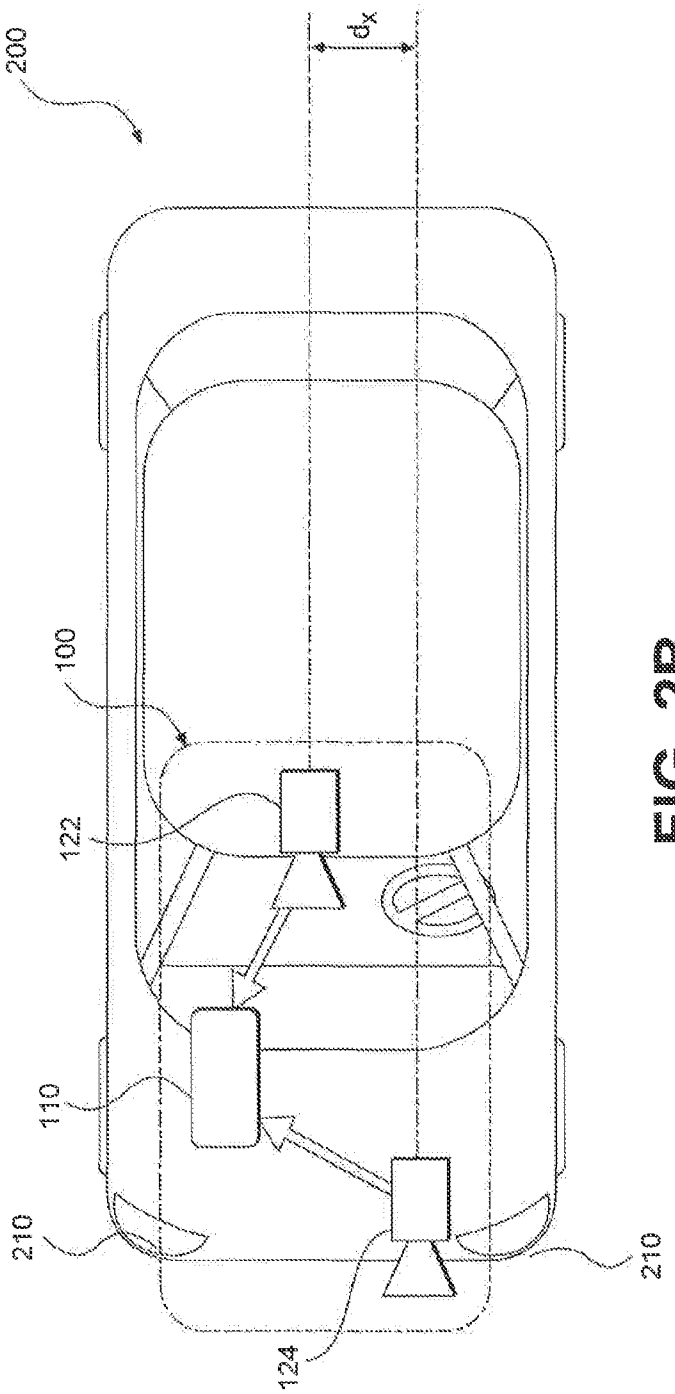


FIG. 2B

IPR2025-00944
Tesla EX1002 Page 1205

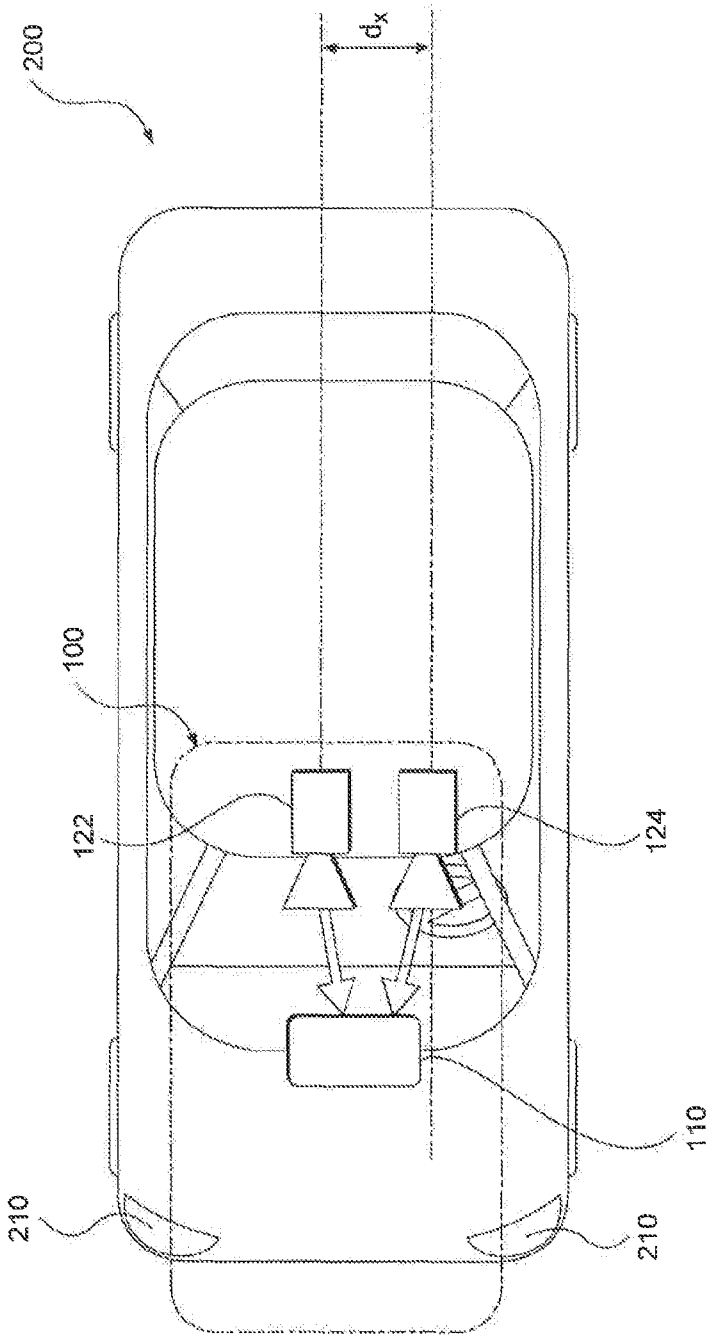


FIG. 2C

IPR2025-00944
Tesla EX1002 Page 1206

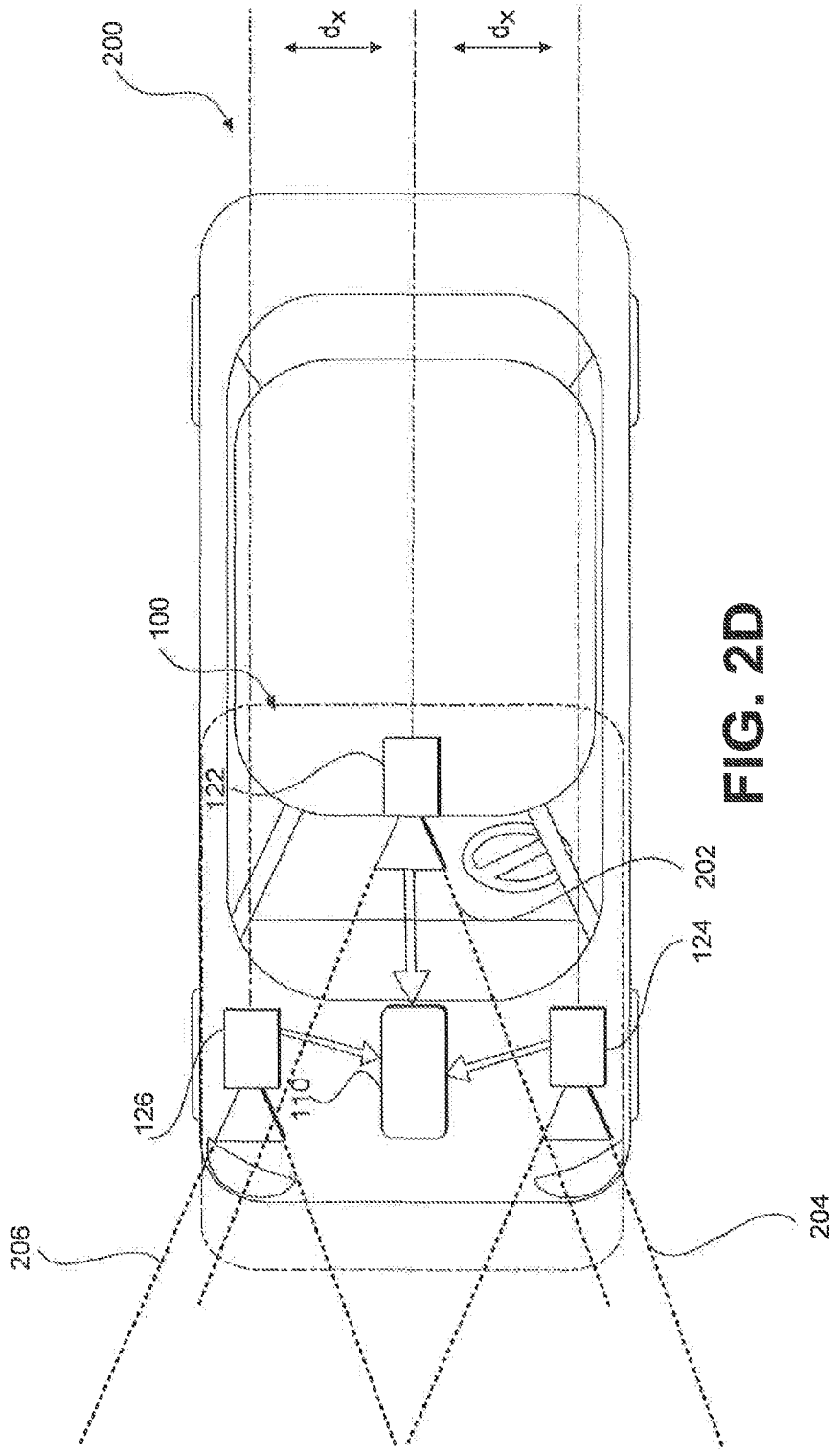


FIG. 2D

IPR2025-00944
Tesla EX1002 Page 1207

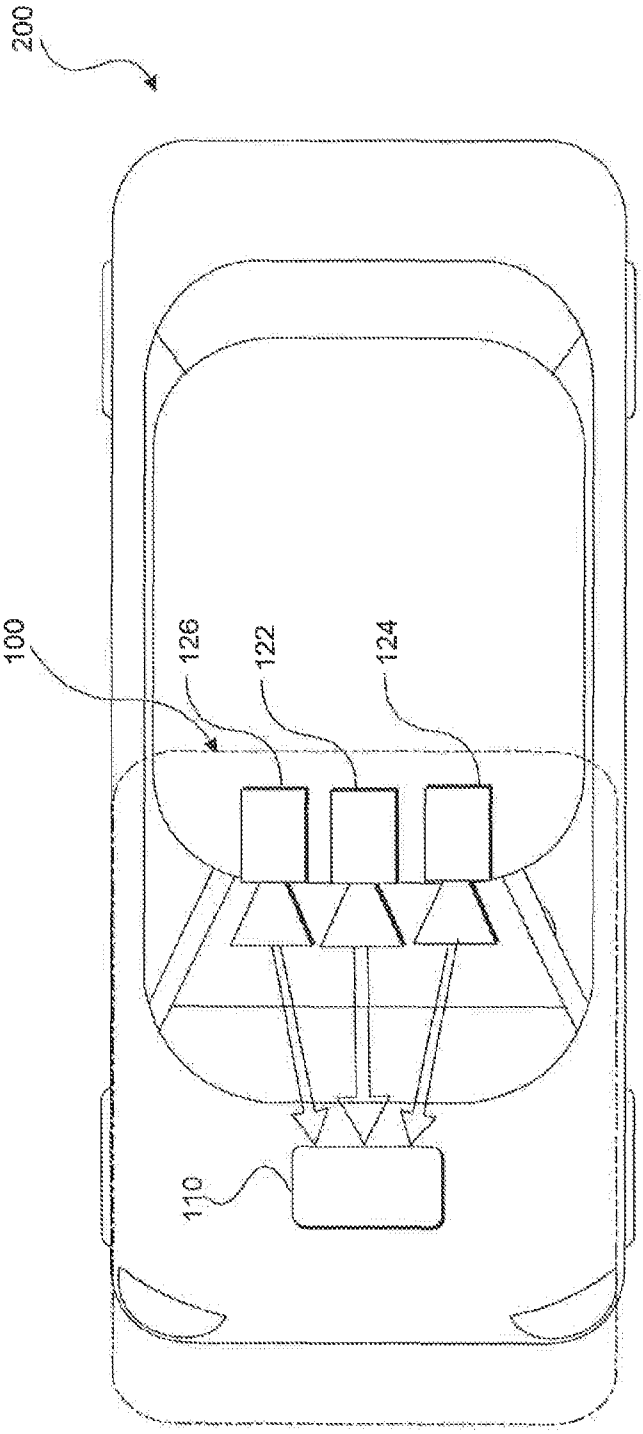


FIG. 2E

IPR2025-00944
Tesla EX1002 Page 1208

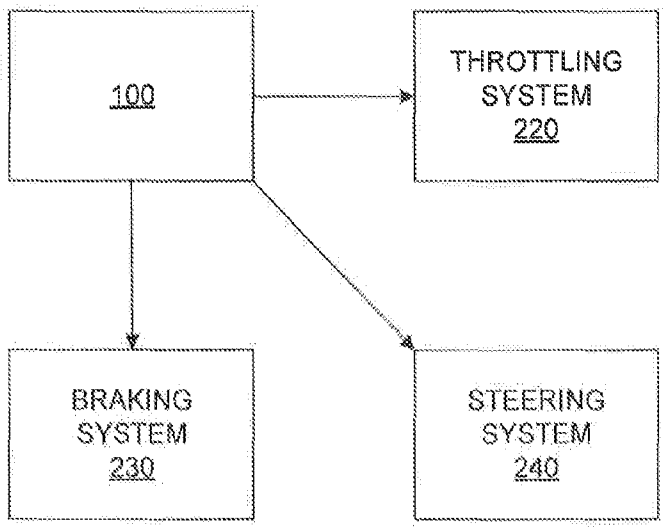


FIG. 2F

IPR2025-00944
Tesla EX1002 Page 1209

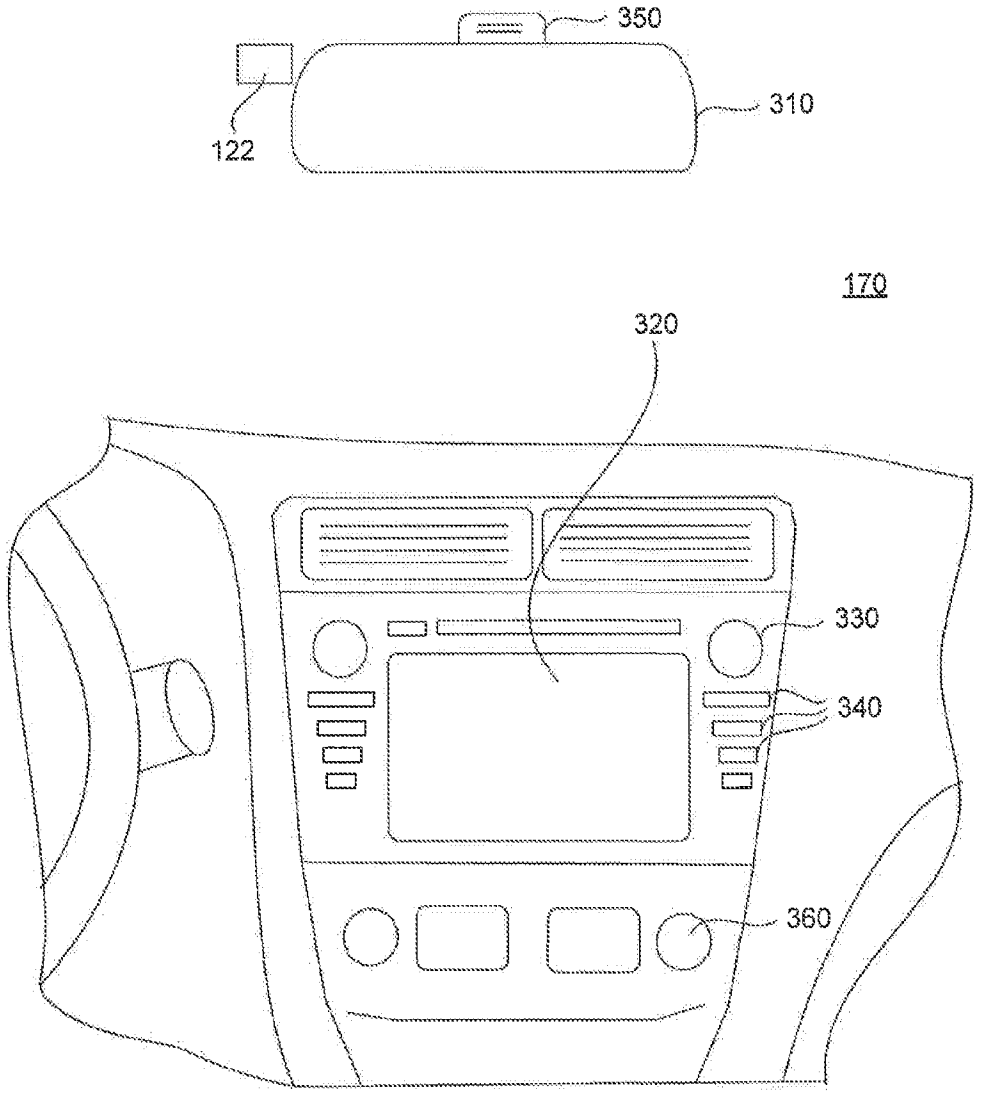


FIG. 3A

IPR2025-00944
Tesla EX1002 Page 1210

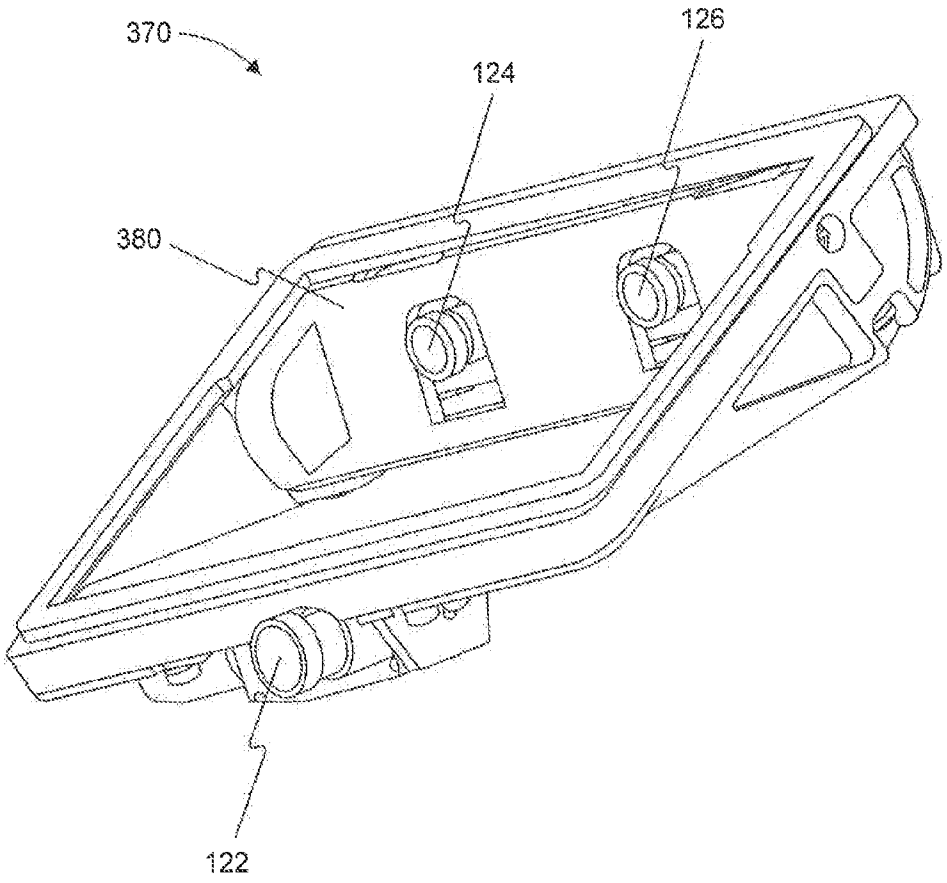


FIG. 3B

IPR2025-00944
Tesla EX1002 Page 1211

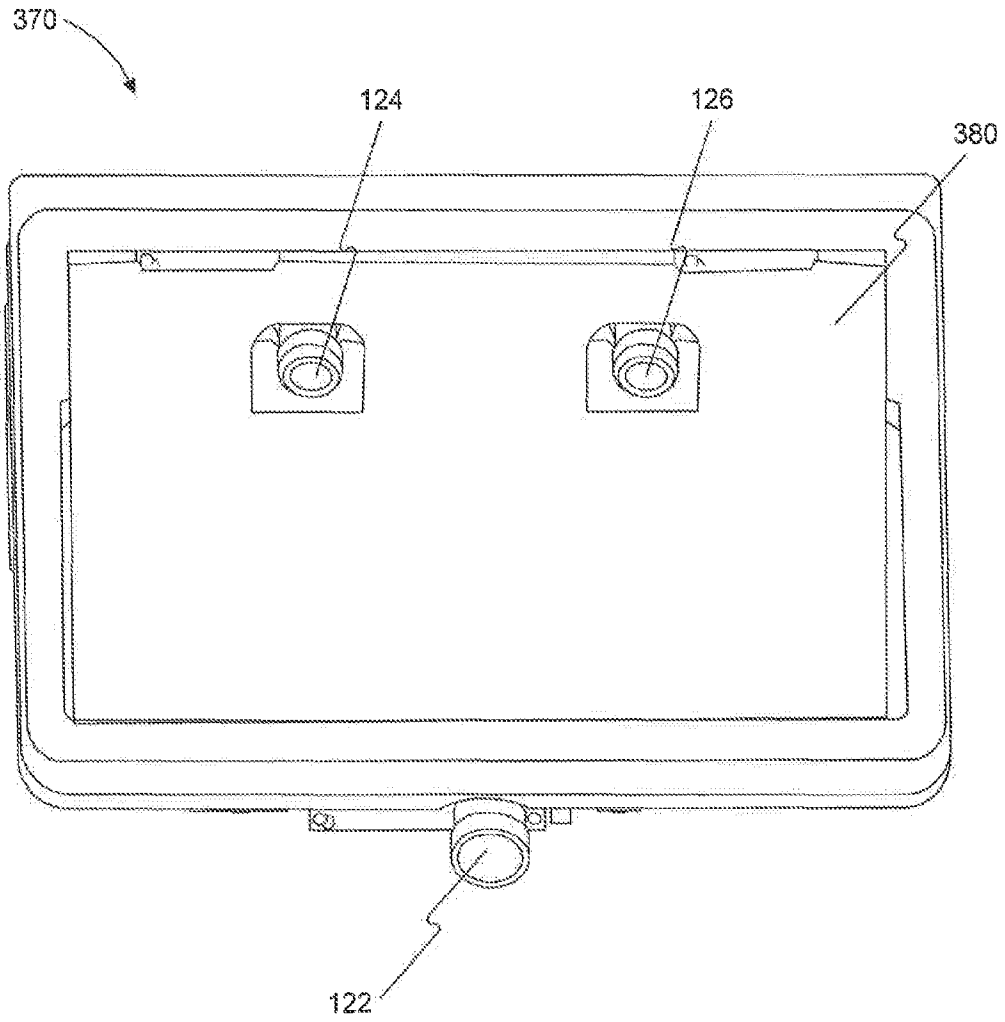


FIG. 3C

IPR2025-00944
Tesla EX1002 Page 1212

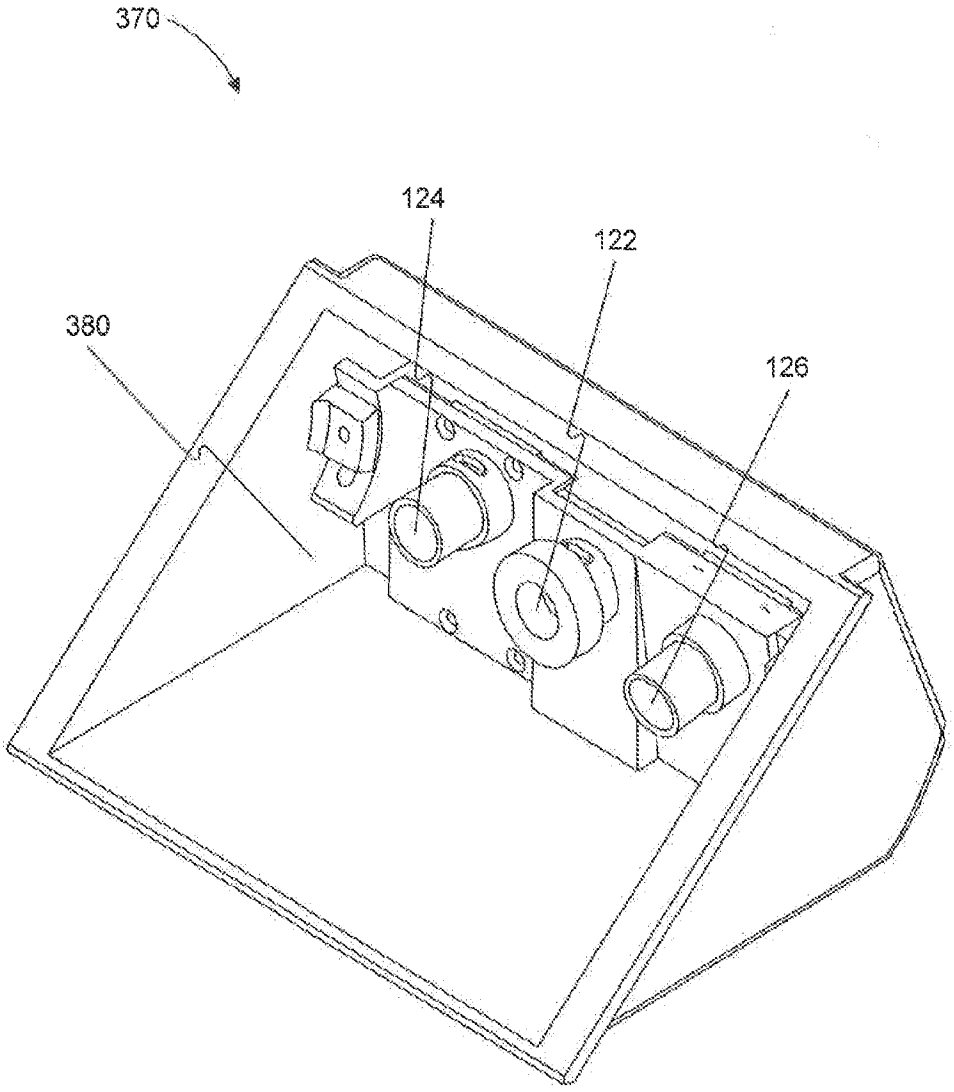


FIG. 3D

IPR2025-00944
Tesla EX1002 Page 1213

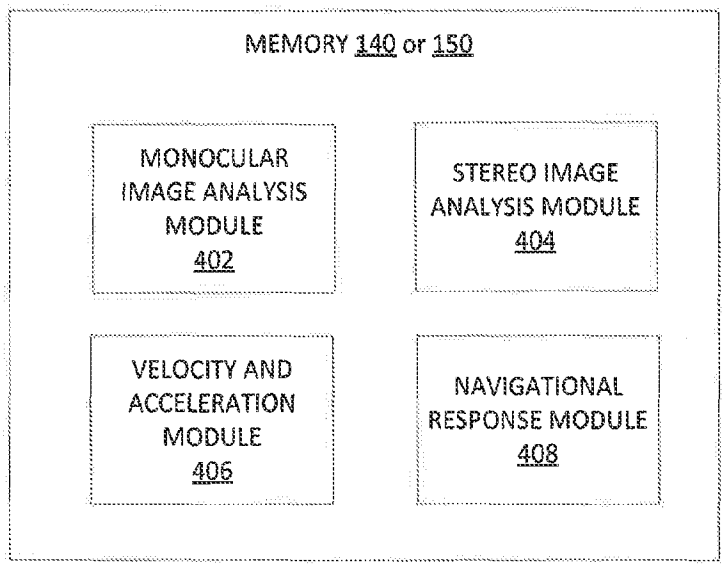


FIG. 4

IPR2025-00944
Tesla EX1002 Page 1214

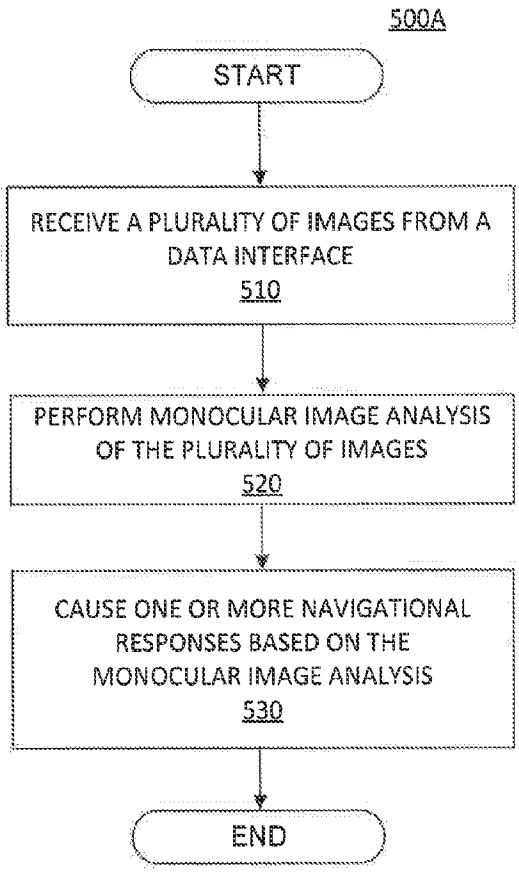


FIG. 5A

IPR2025-00944
Tesla EX1002 Page 1215

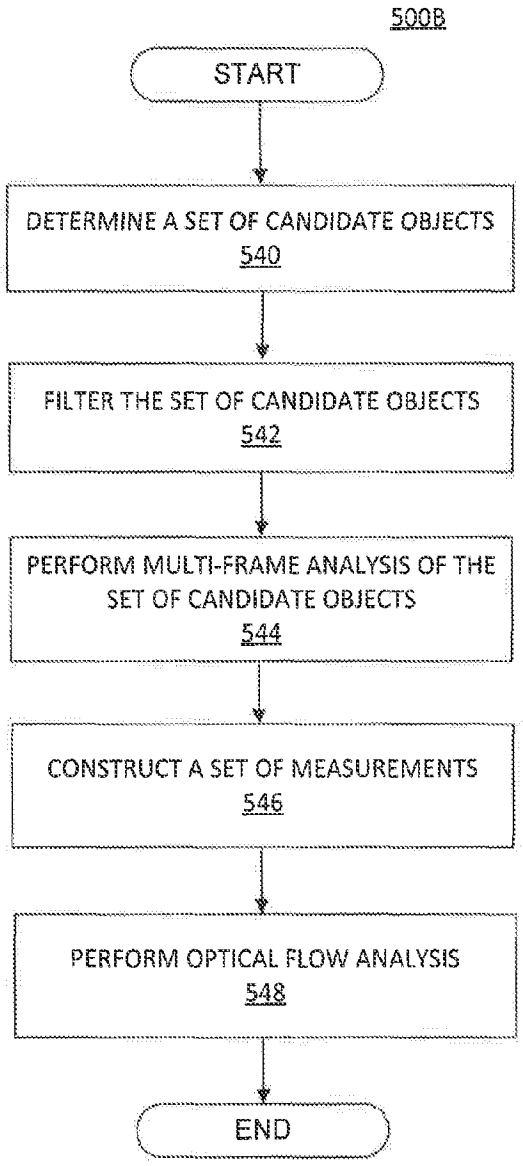


FIG. 5B

IPR2025-00944
Tesla EX1002 Page 1216

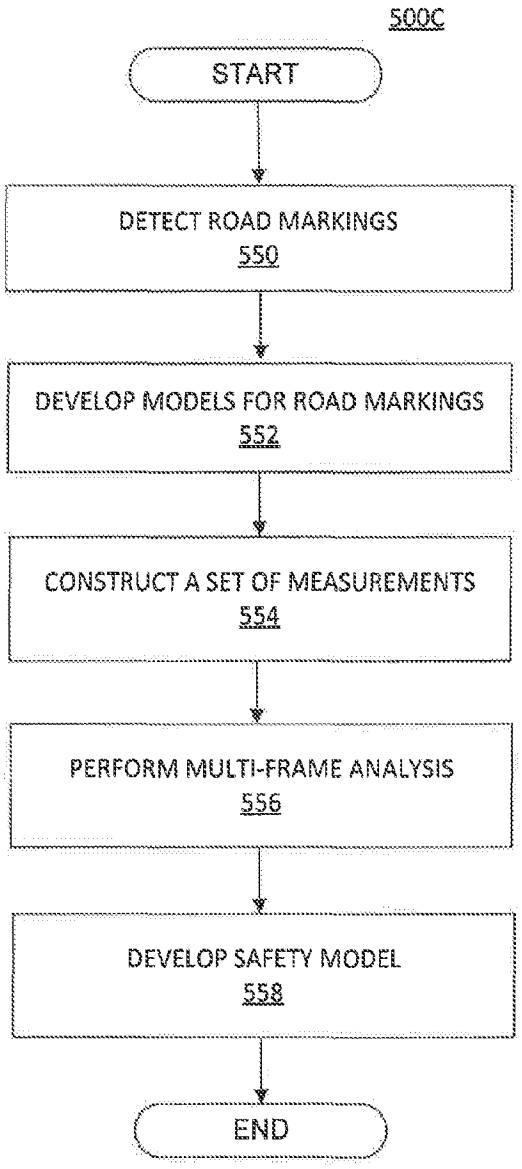


FIG. 5C

IPR2025-00944
Tesla EX1002 Page 1217

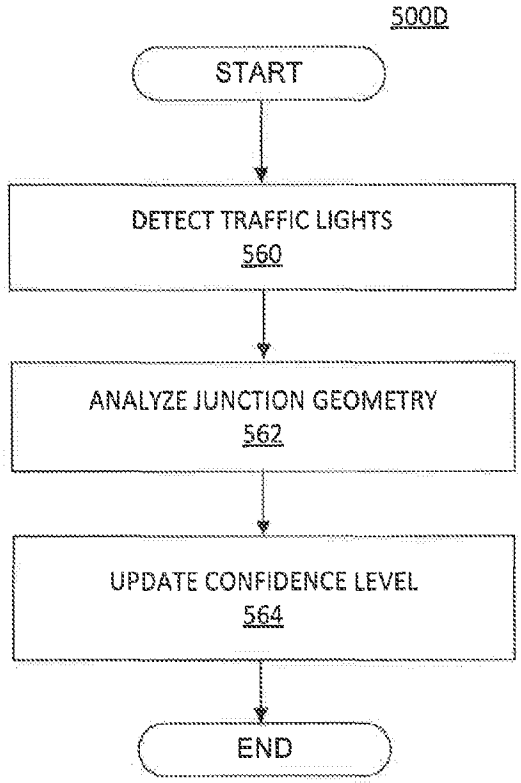


FIG. 5D

IPR2025-00944
Tesla EX1002 Page 1218

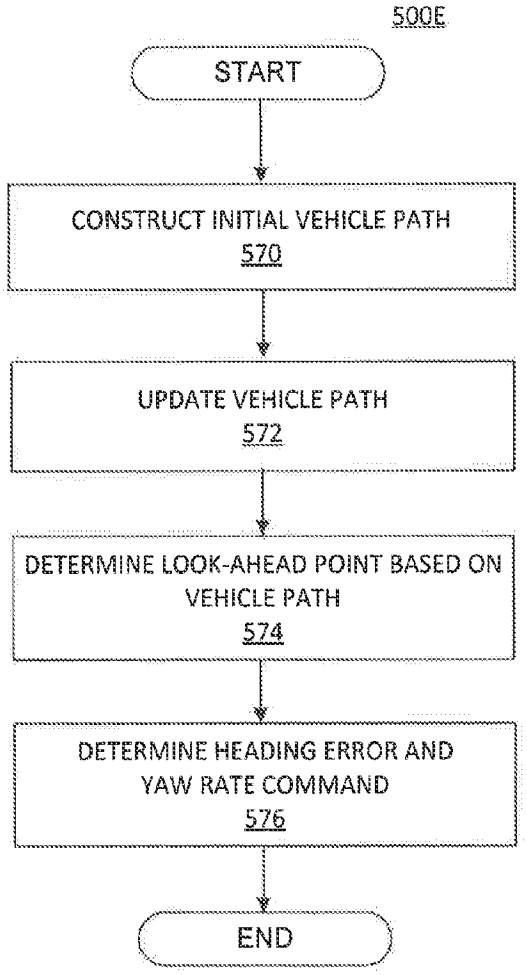


FIG. 5E

IPR2025-00944
Tesla EX1002 Page 1219

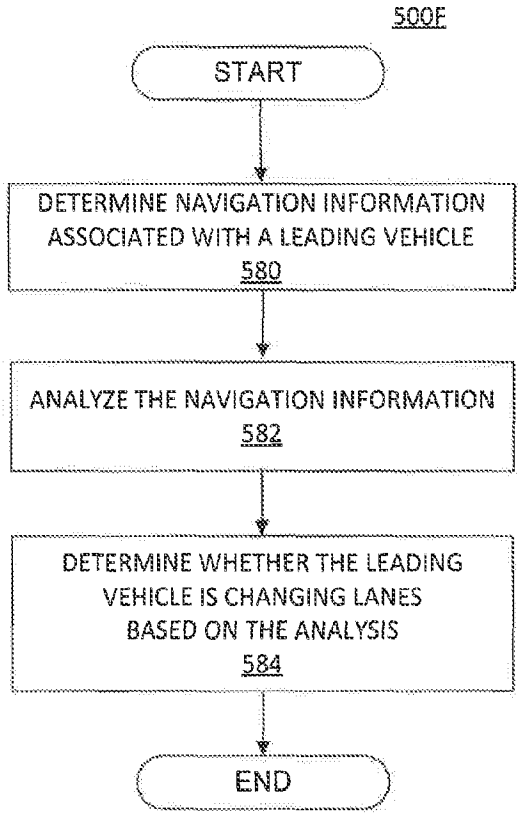


FIG. 5F

IPR2025-00944
Tesla EX1002 Page 1220

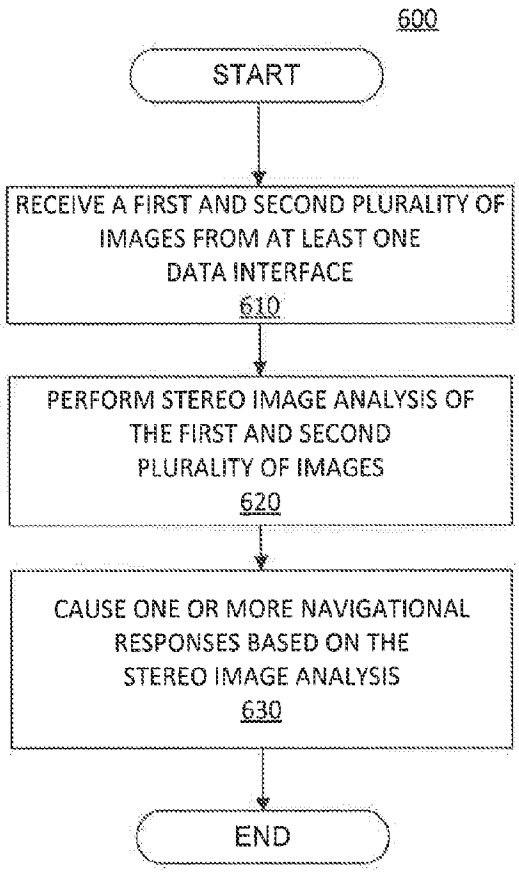


FIG. 6

IPR2025-00944
Tesla EX1002 Page 1221

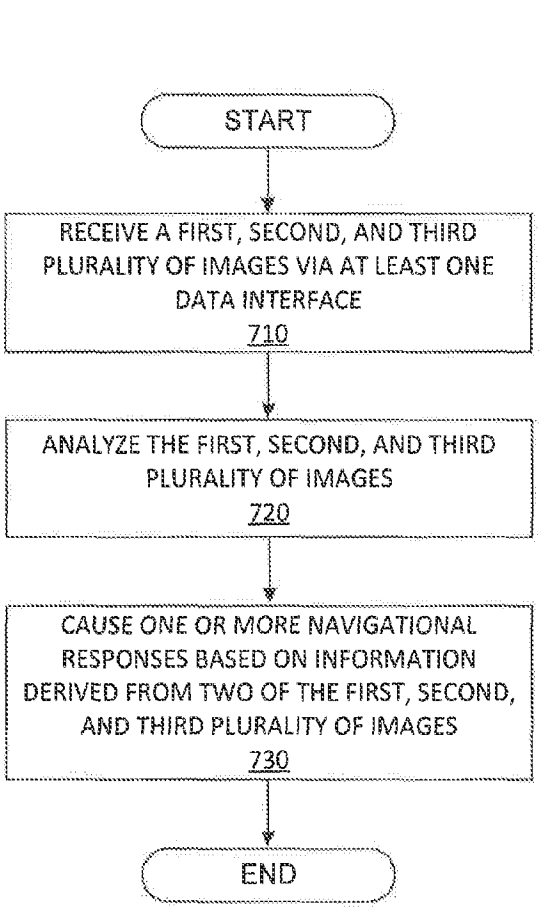


FIG. 7

IPR2025-00944
Tesla EX1002 Page 1222

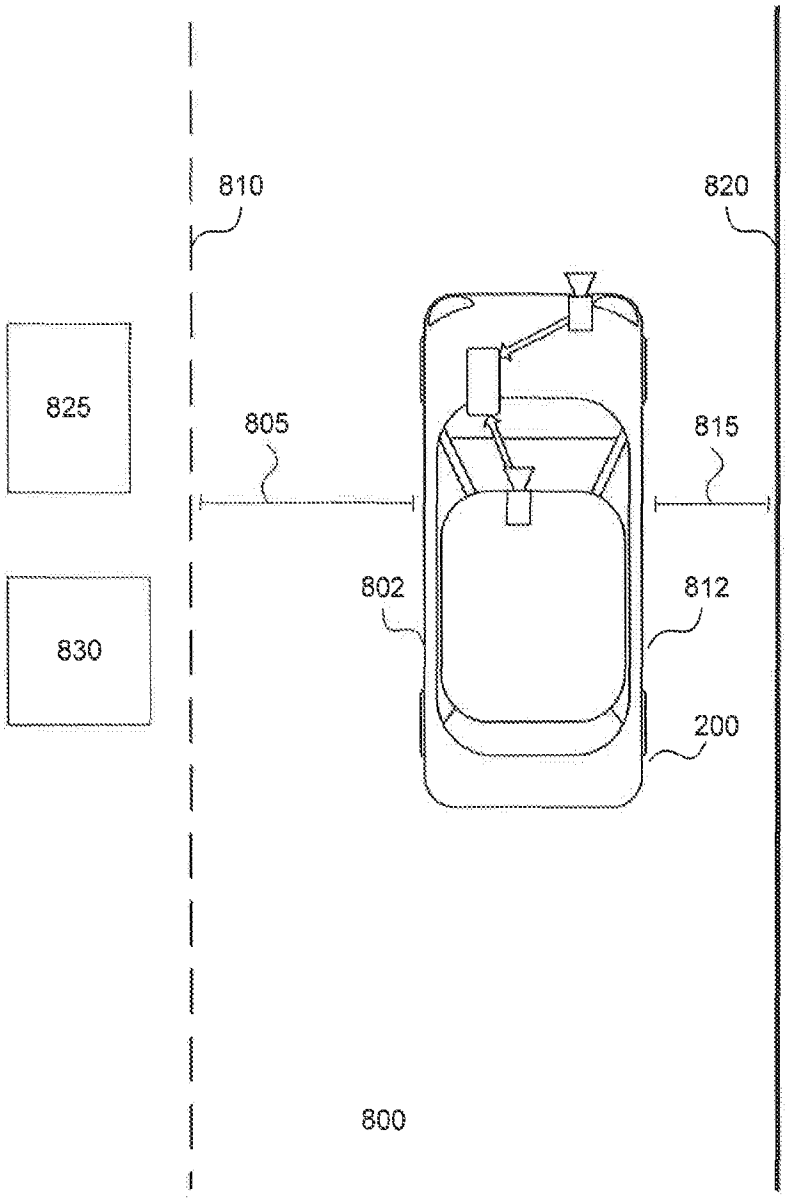


FIG. 8A

IPR2025-00944
Tesla EX1002 Page 1223

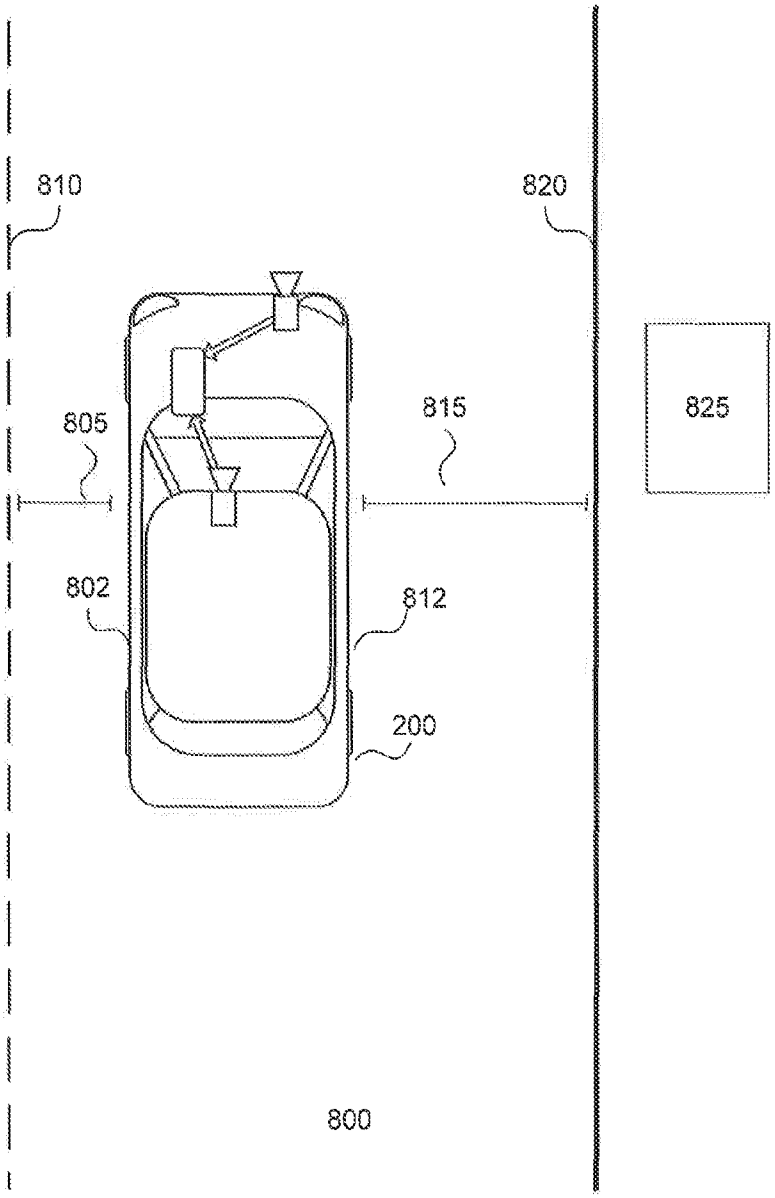


FIG. 8B

IPR2025-00944
Tesla EX1002 Page 1224

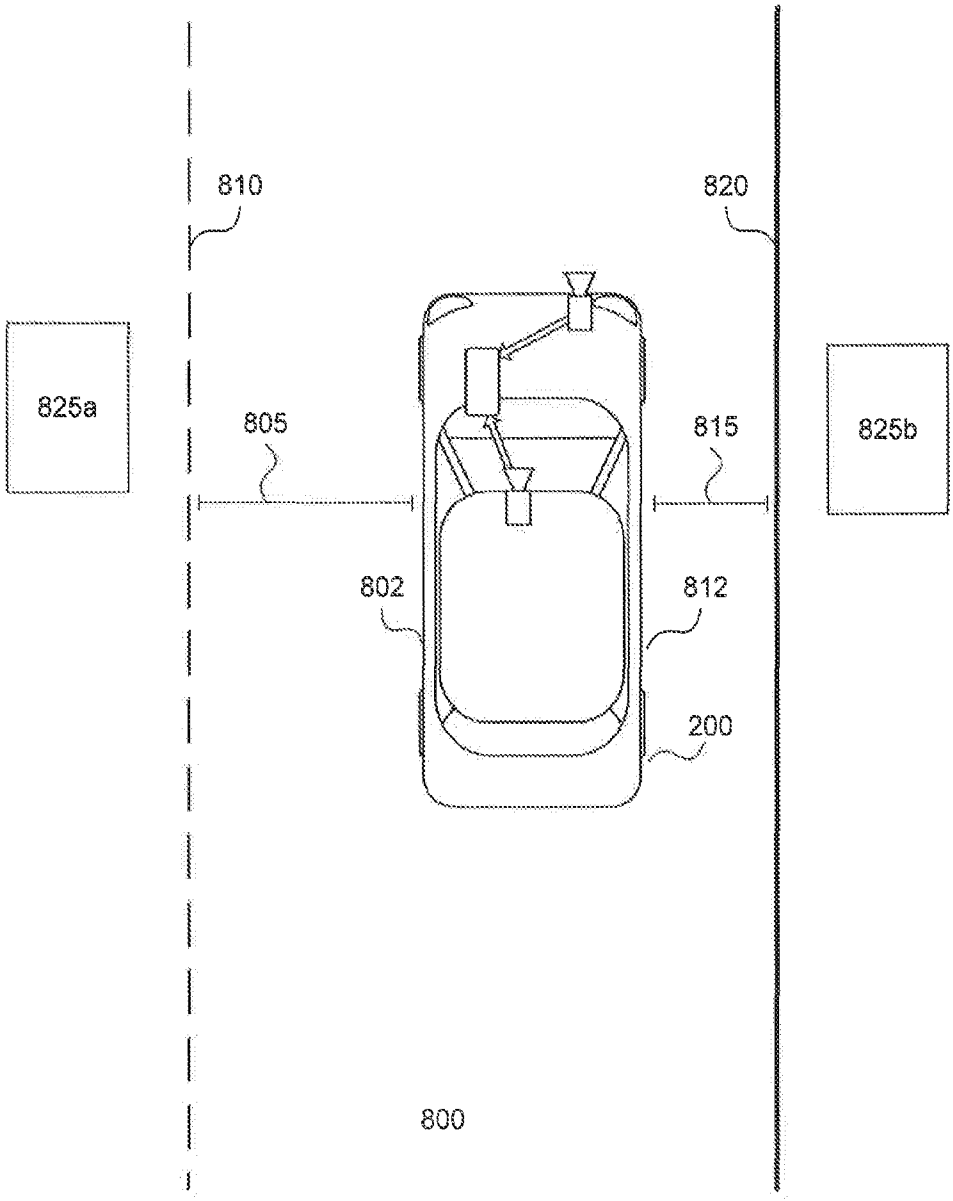


FIG. 8C

IPR2025-00944
Tesla EX1002 Page 1225

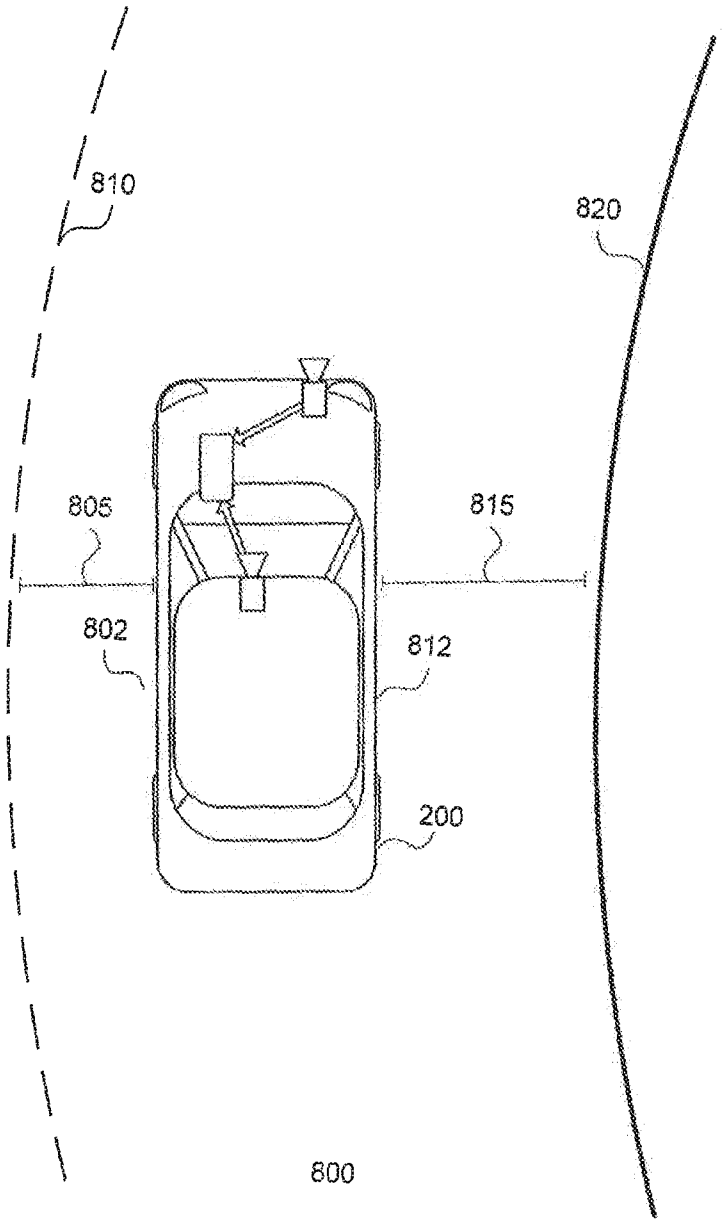


FIG. 8D

IPR2025-00944
Tesla EX1002 Page 1226

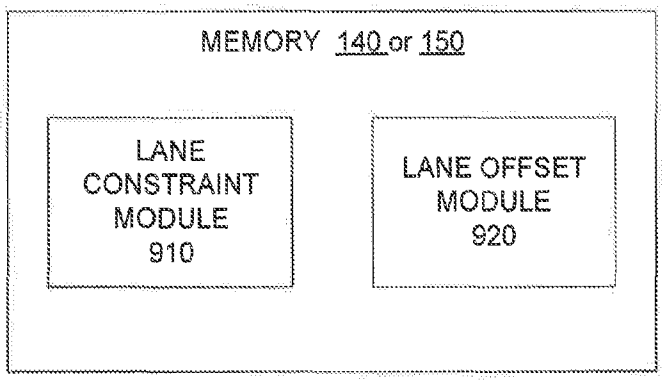


FIG. 9

IPR2025-00944
Tesla EX1002 Page 1227

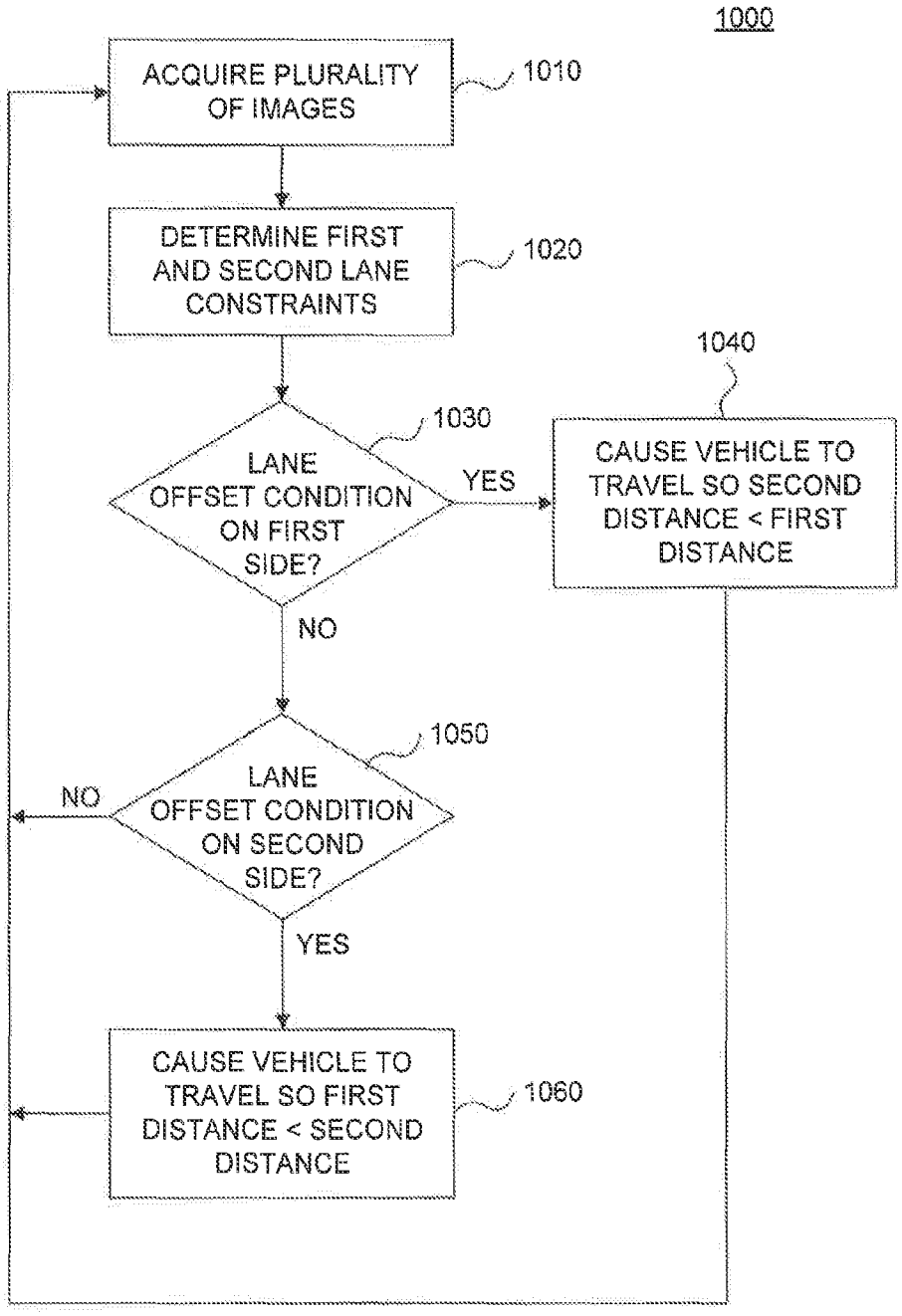


FIG. 10

IPR2025-00944
Tesla EX1002 Page 1228

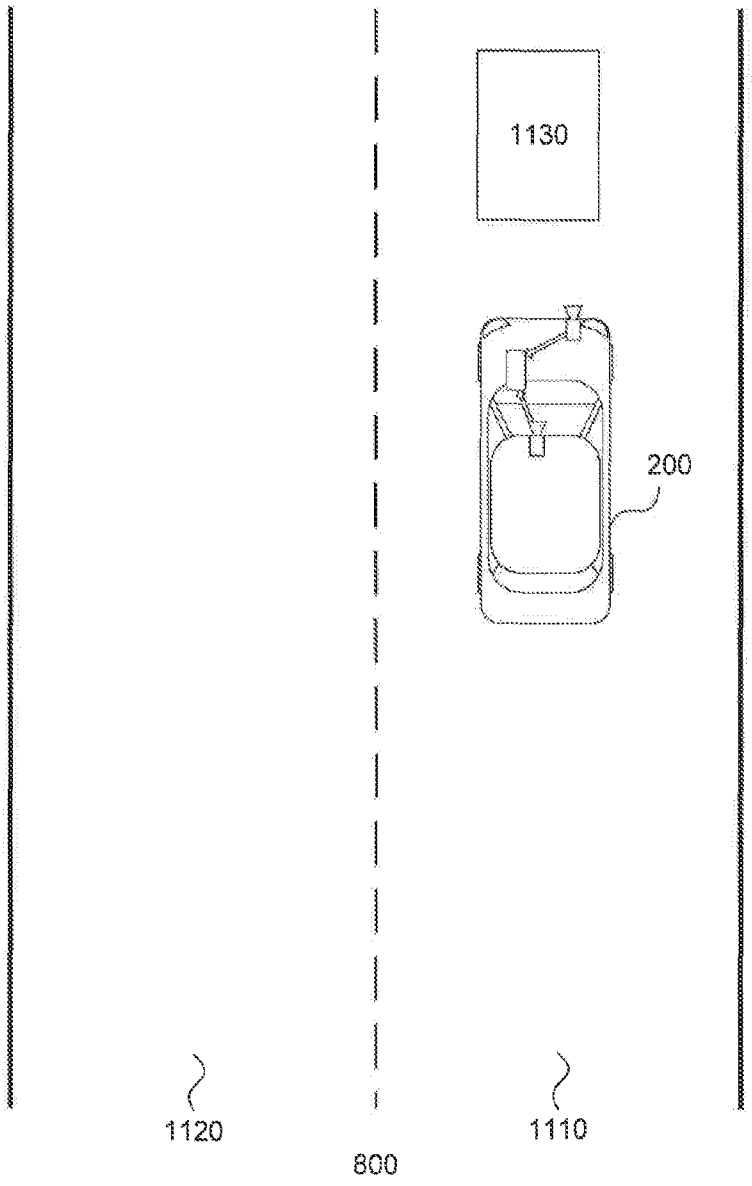


FIG. 11

IPR2025-00944
Tesla EX1002 Page 1229

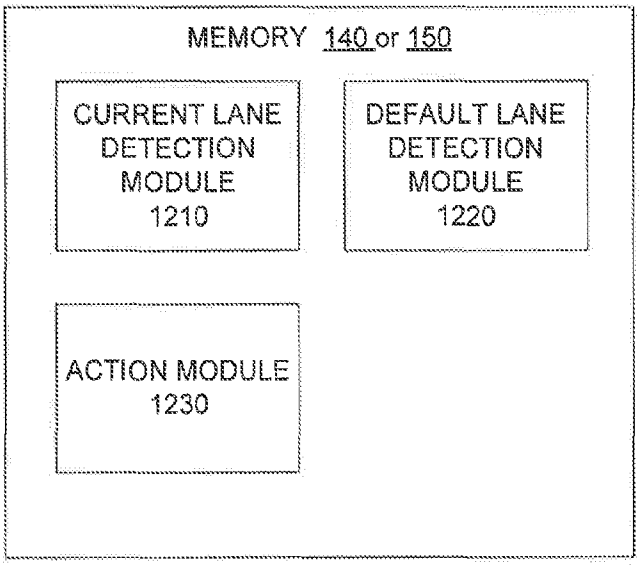


FIG. 12A

IPR2025-00944
Tesla EX1002 Page 1230

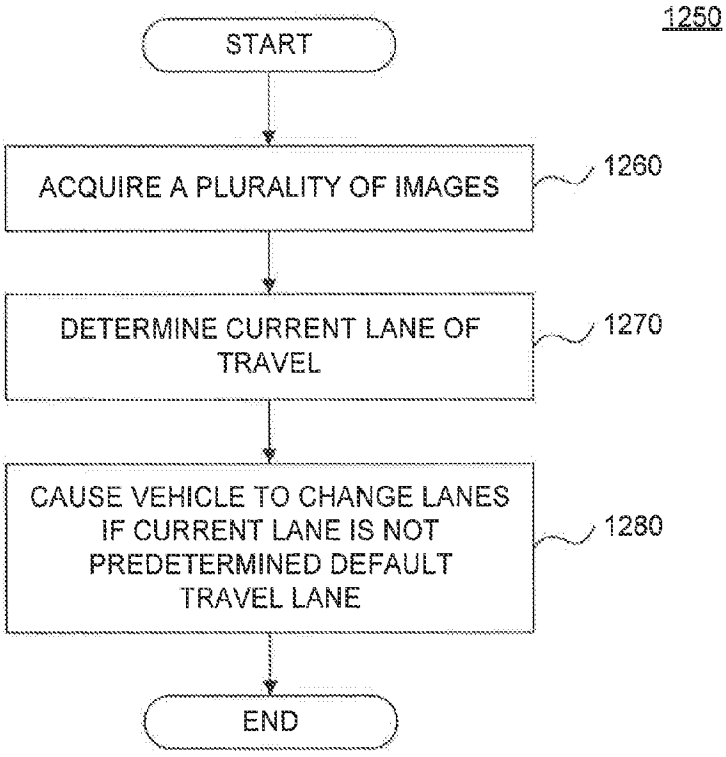


FIG. 12B

IPR2025-00944
Tesla EX1002 Page 1231

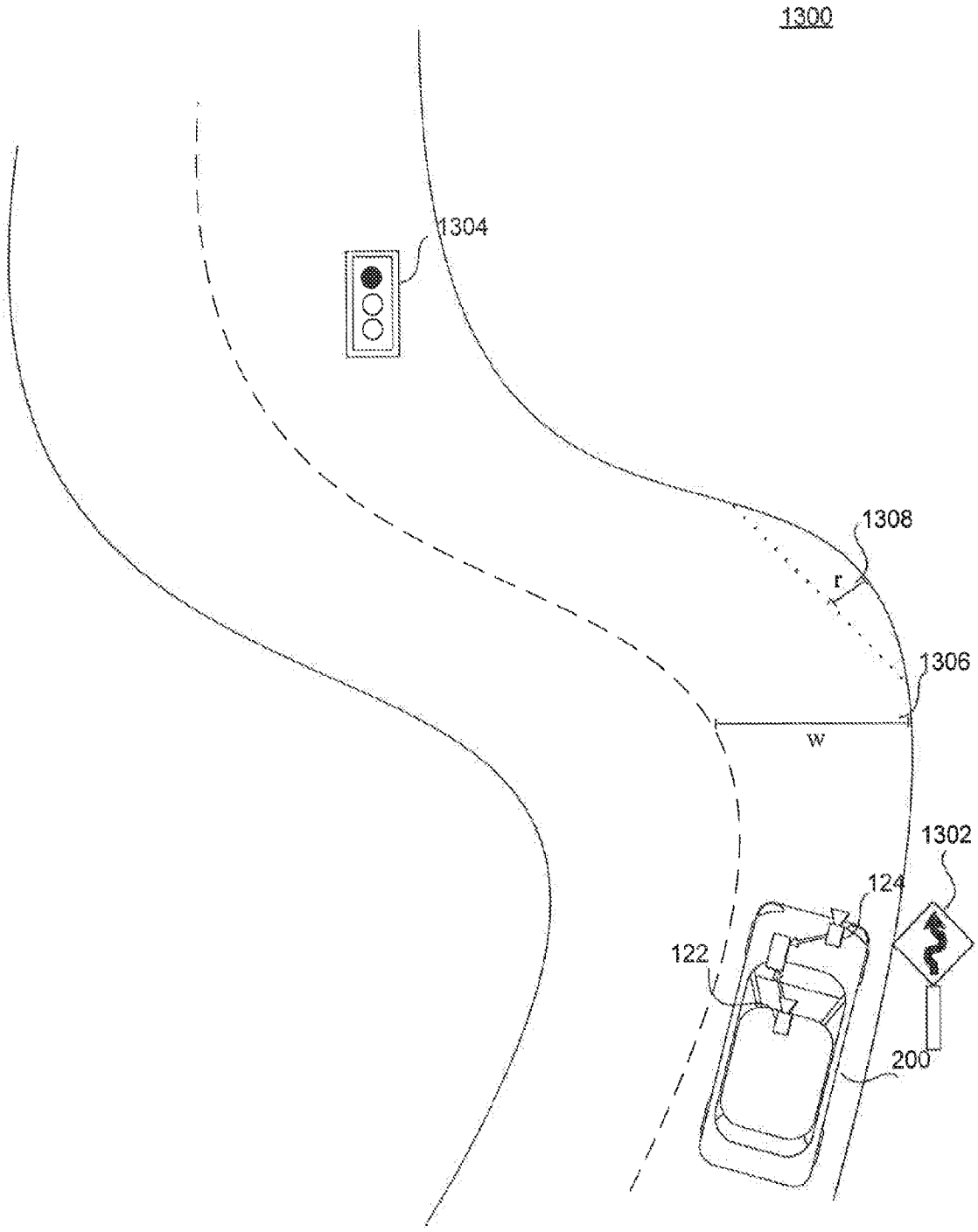


FIG. 13A

IPR2025-00944
Tesla EX1002 Page 1232

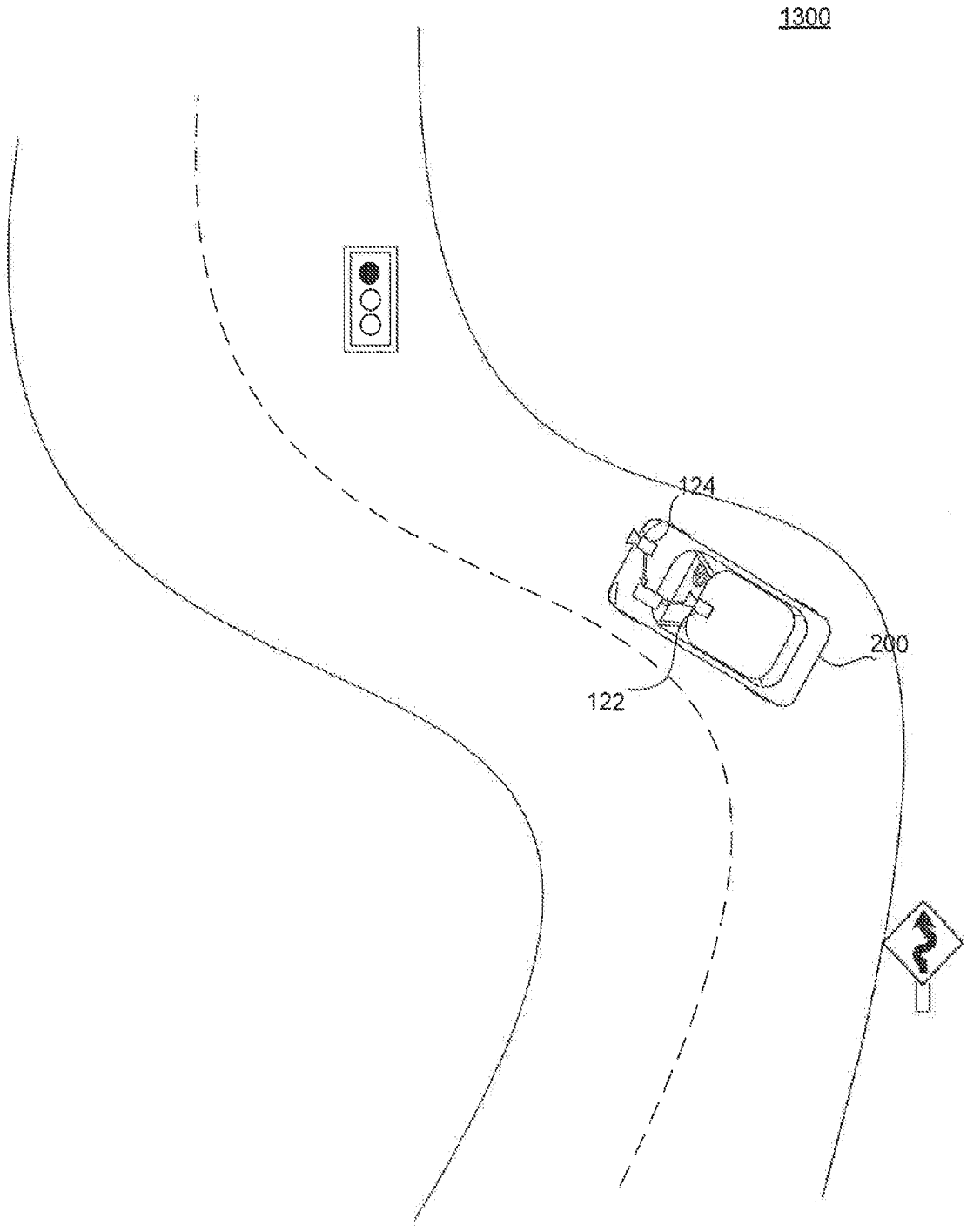


FIG. 13B

IPR2025-00944
Tesla EX1002 Page 1233

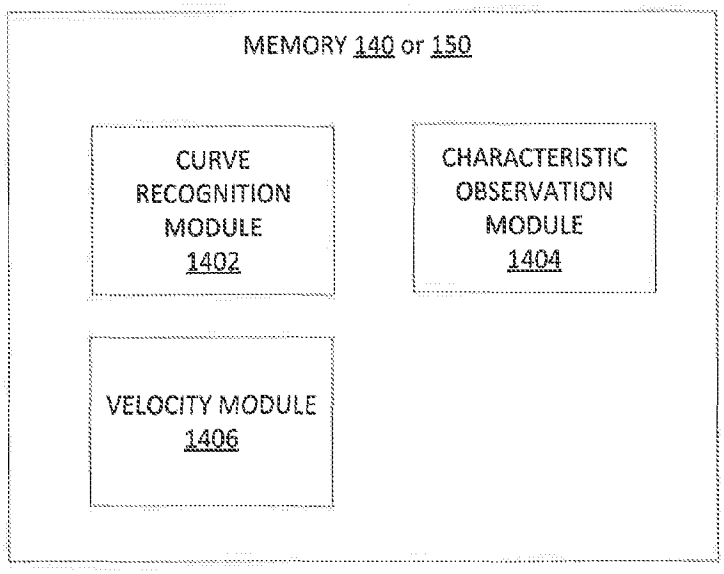


FIG. 14

IPR2025-00944
Tesla EX1002 Page 1234

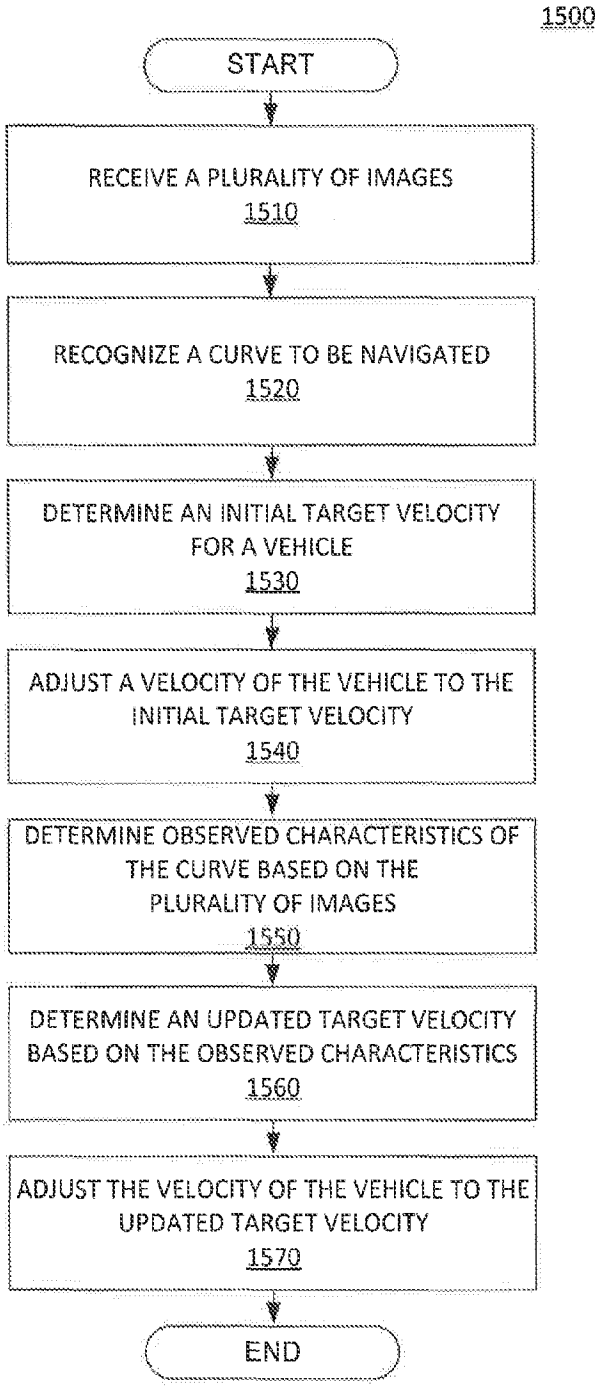


FIG. 15

IPR2025-00944
Tesla EX1002 Page 1235

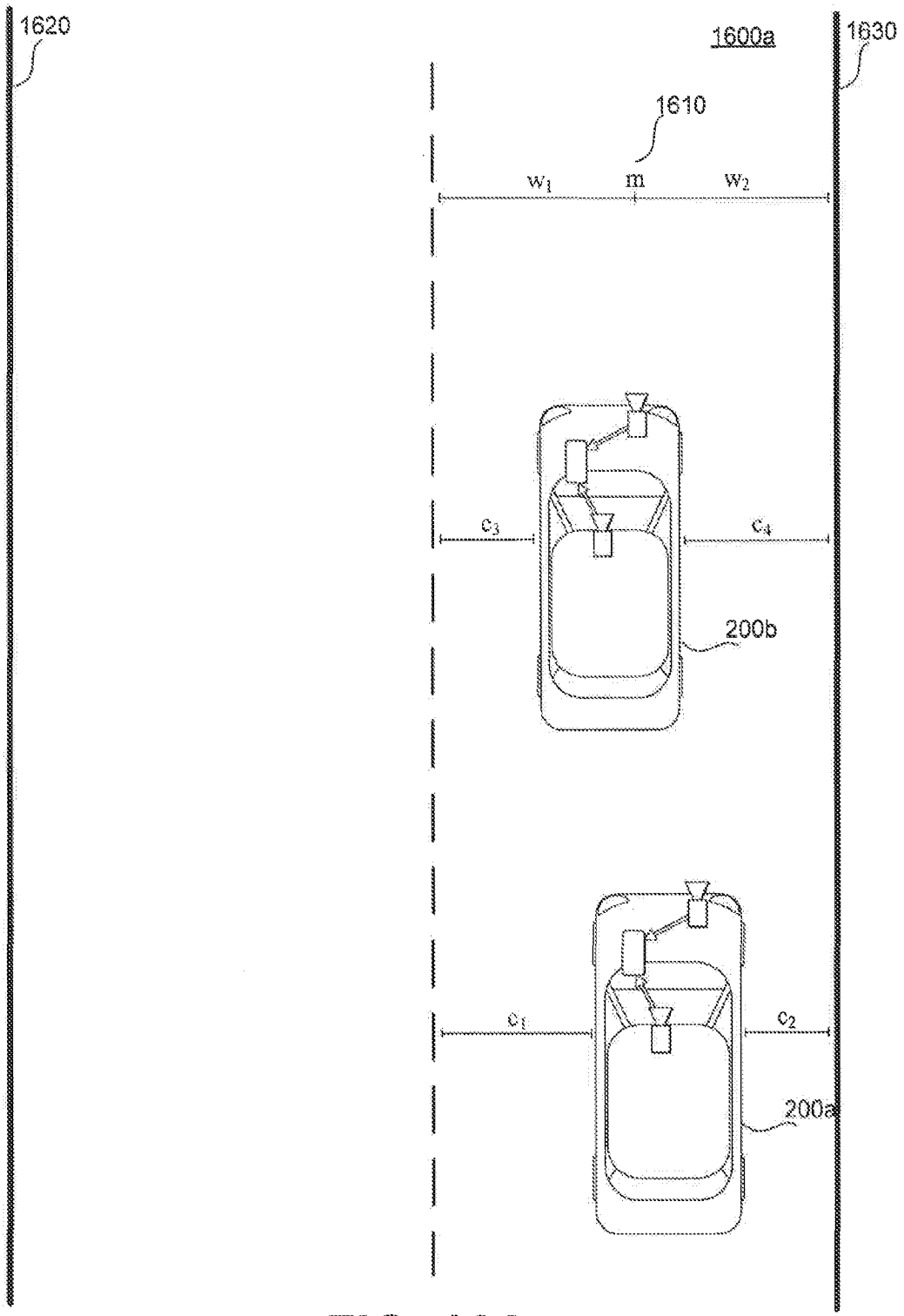


FIG. 16A

IPR2025-00944
Tesla EX1002 Page 1236

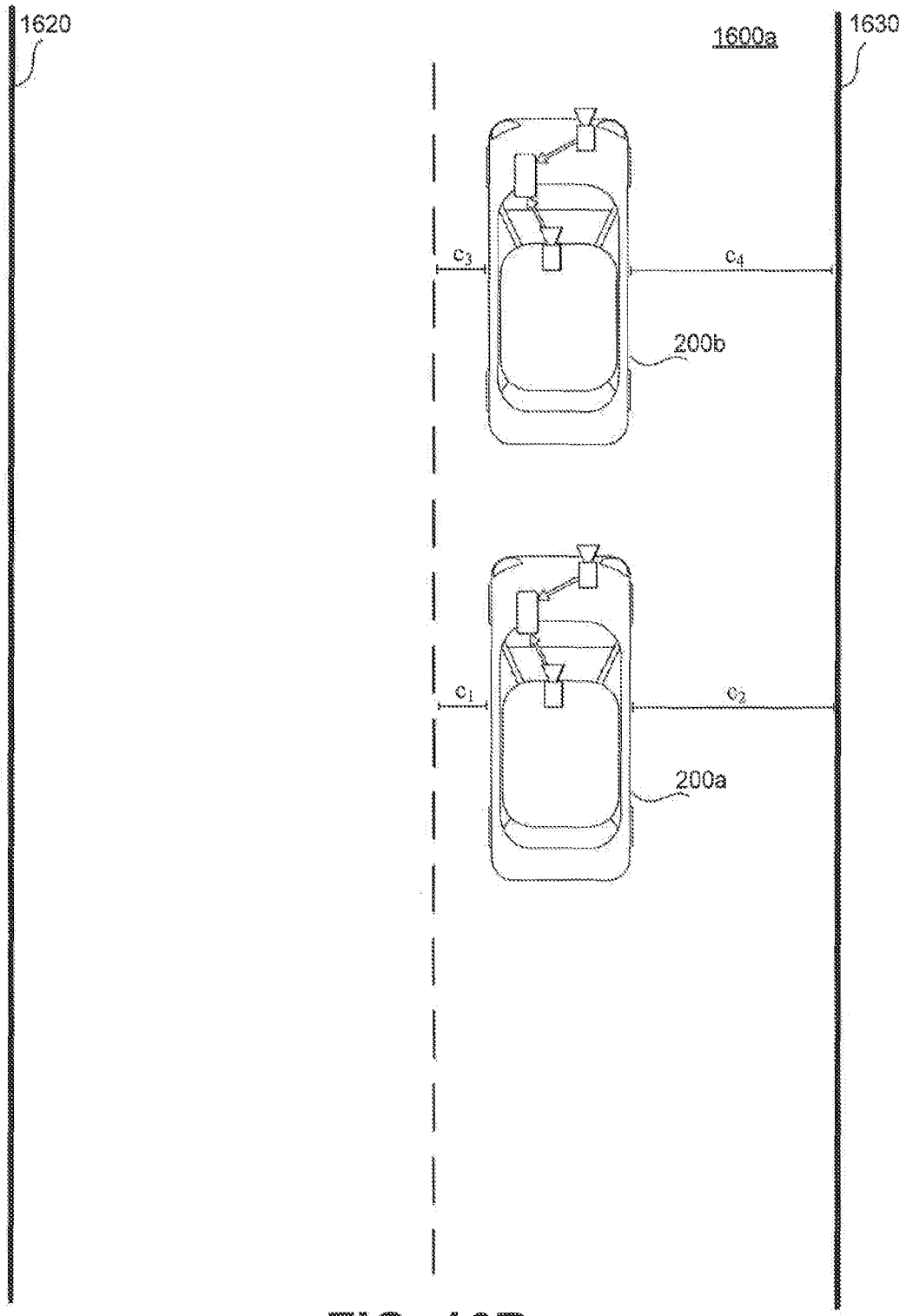


FIG. 16B

IPR2025-00944
Tesla EX1002 Page 1237

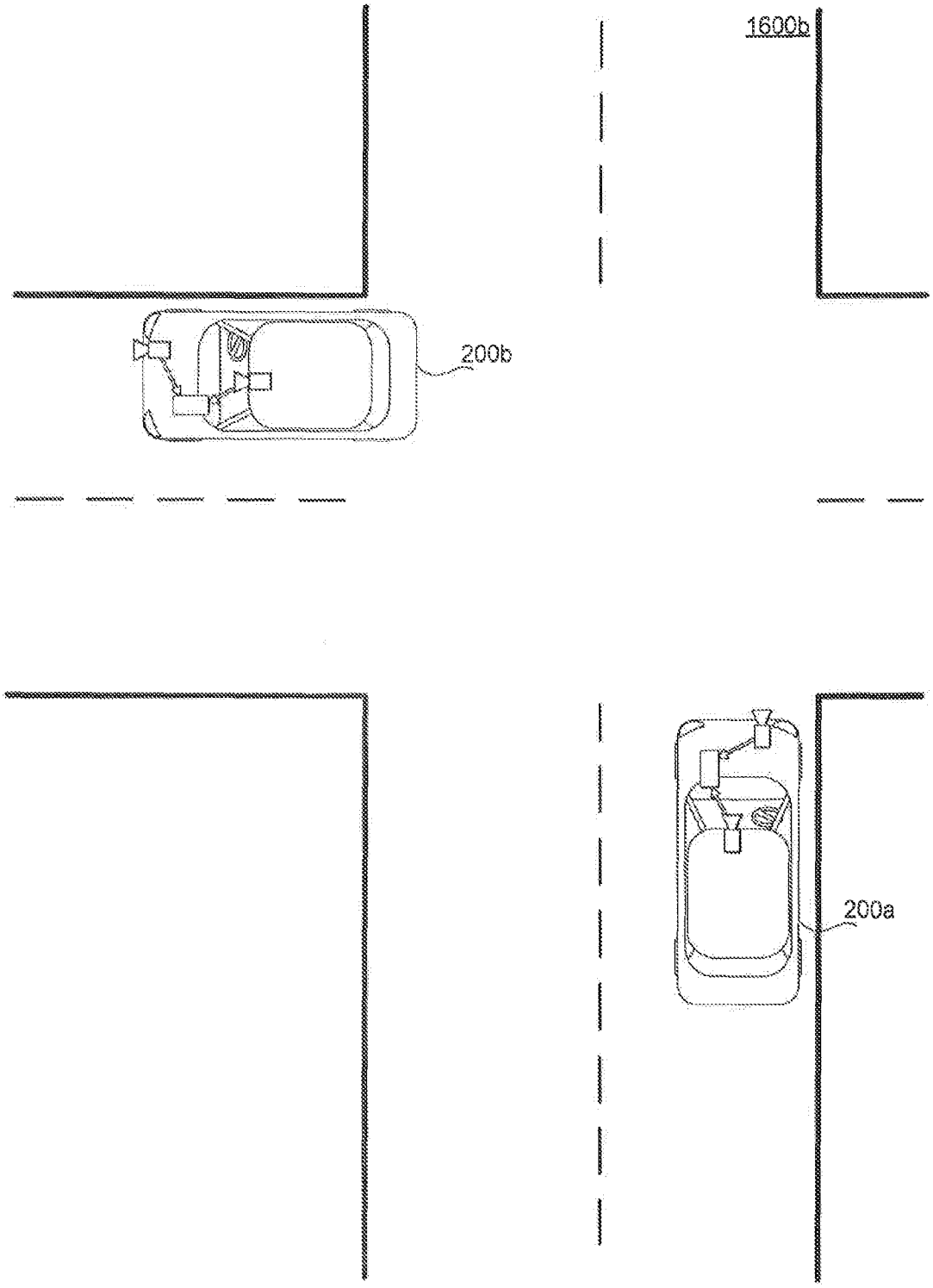


FIG. 16C

IPR2025-00944
Tesla EX1002 Page 1238

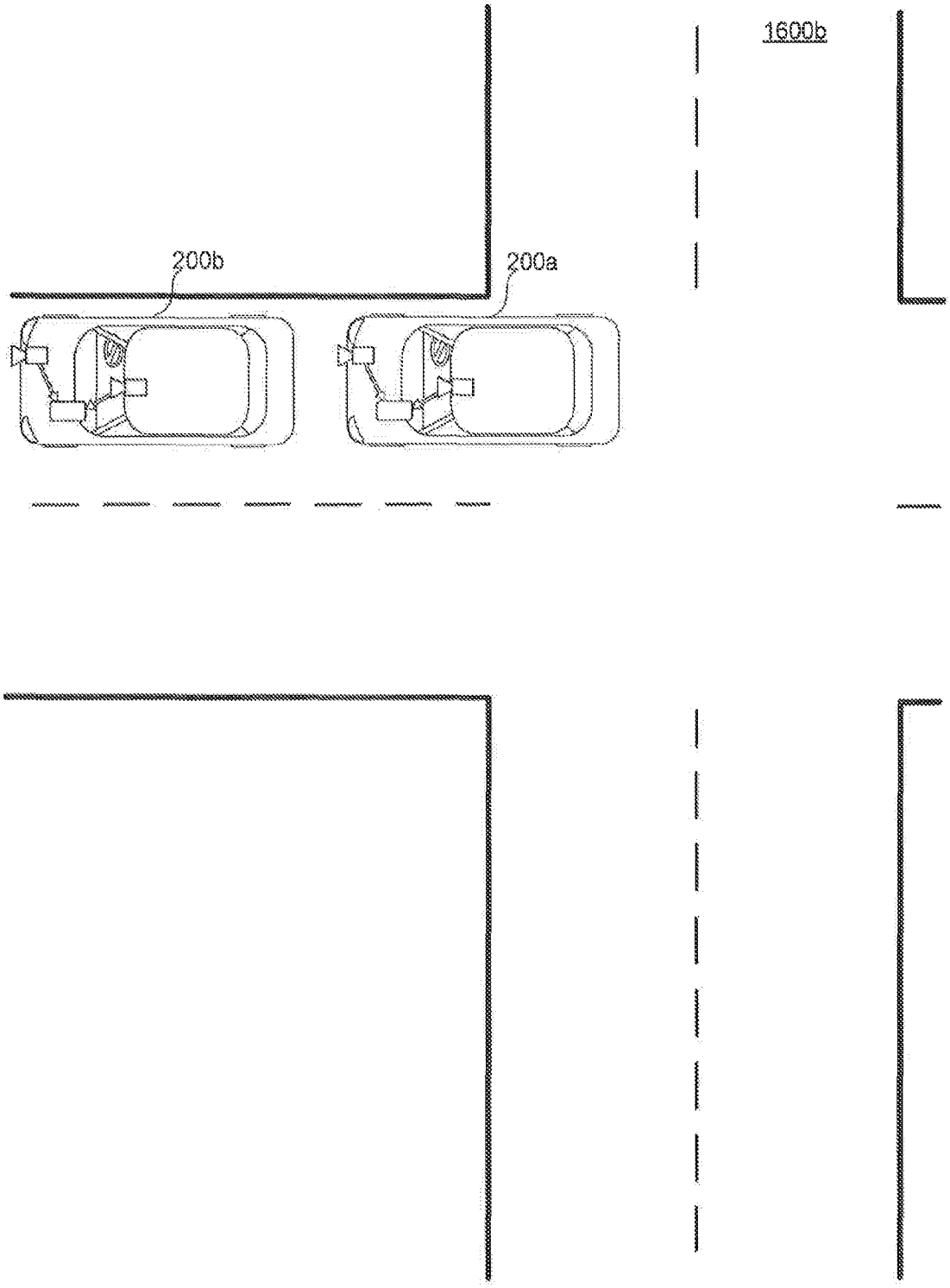


FIG. 16D

IPR2025-00944
Tesla EX1002 Page 1239

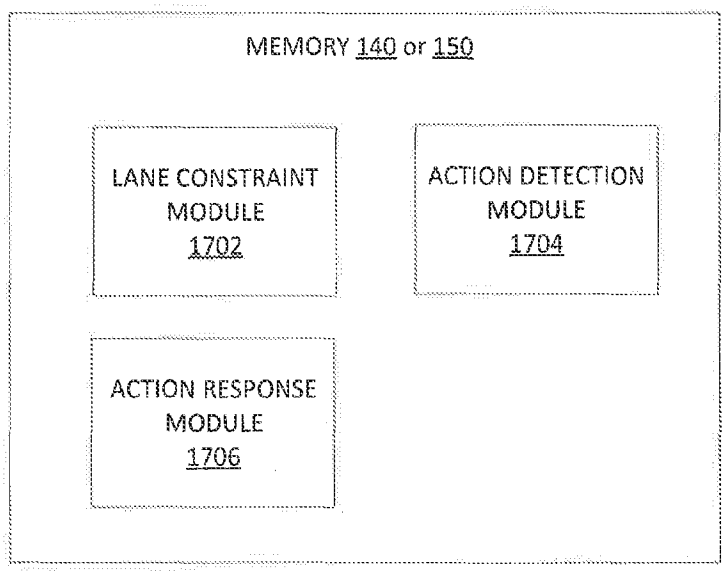


FIG. 17

IPR2025-00944
Tesla EX1002 Page 1240

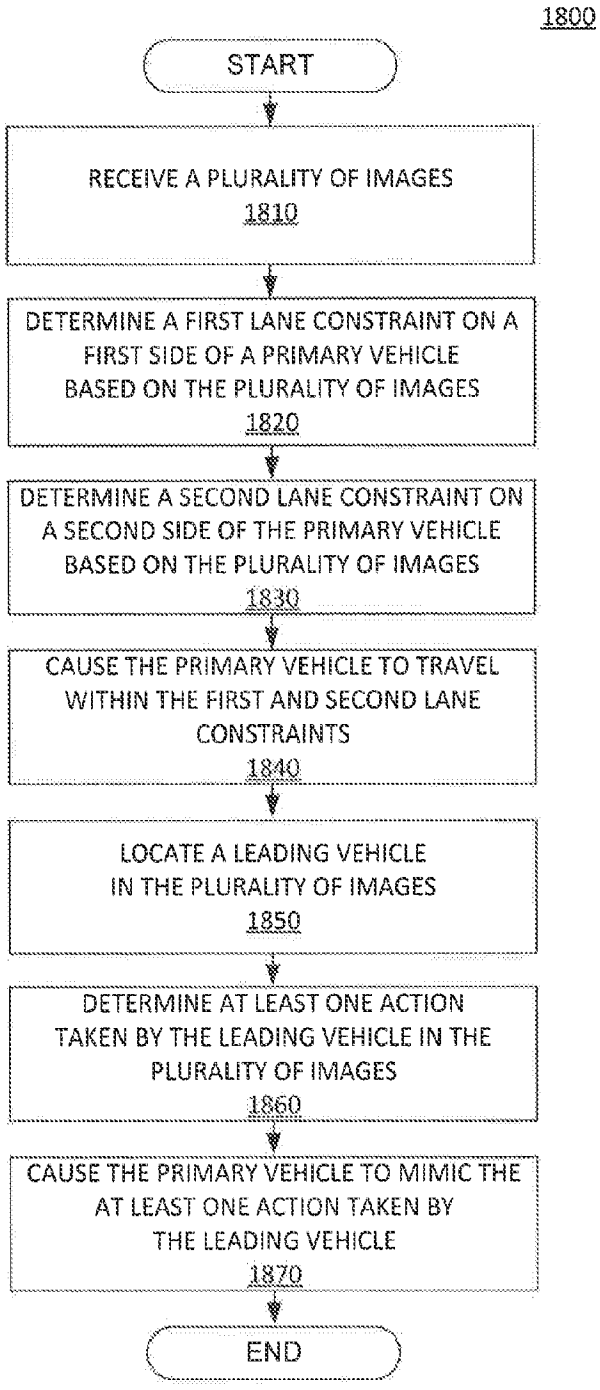


FIG. 18

IPR2025-00944
Tesla EX1002 Page 1241

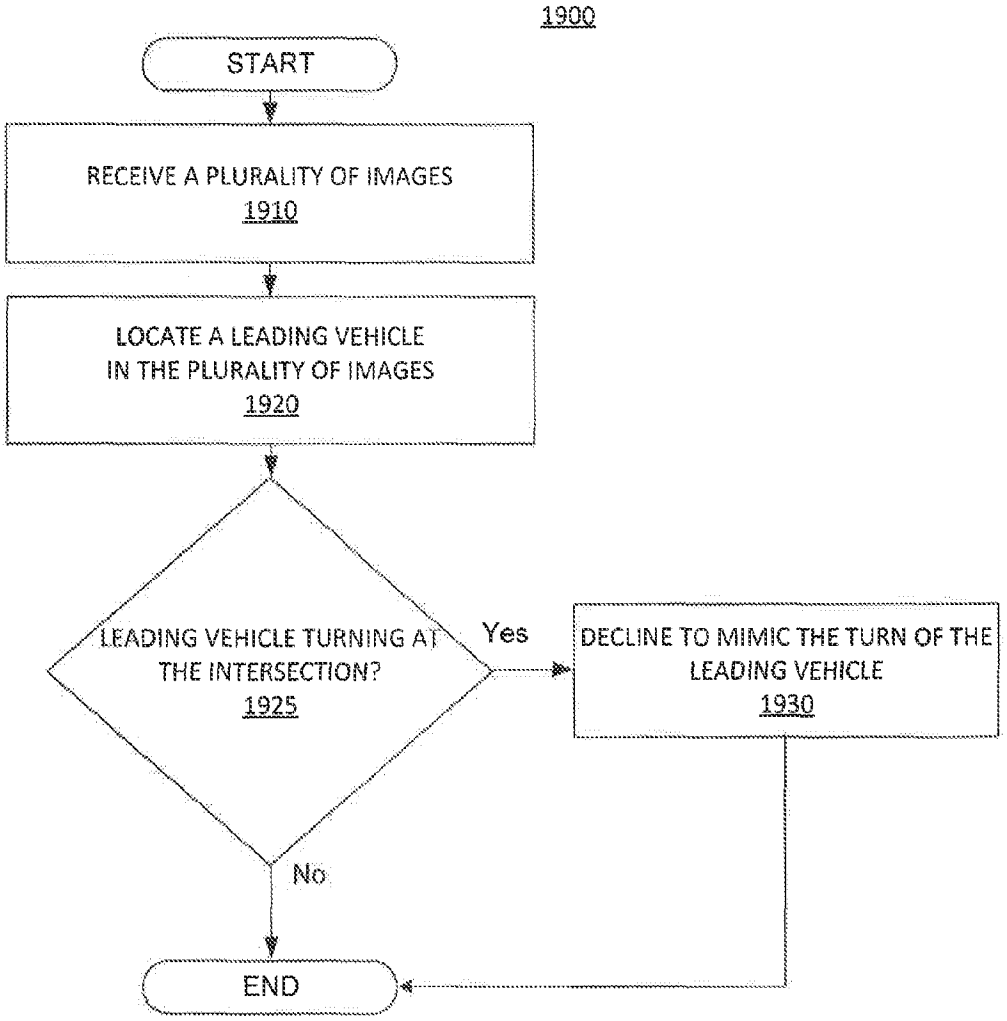


FIG. 19

IPR2025-00944
Tesla EX1002 Page 1242

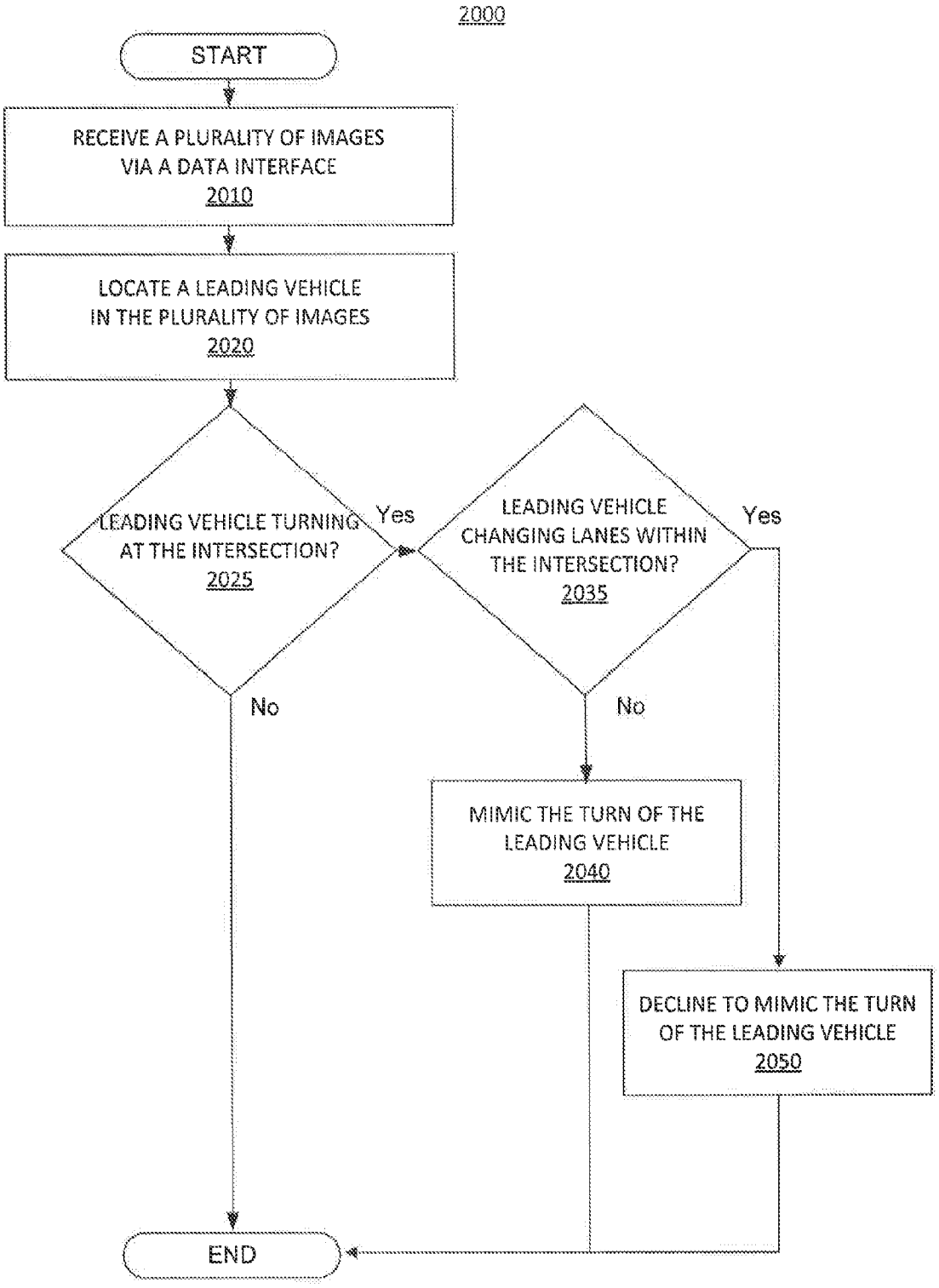


FIG. 20

IPR2025-00944
Tesla EX1002 Page 1243

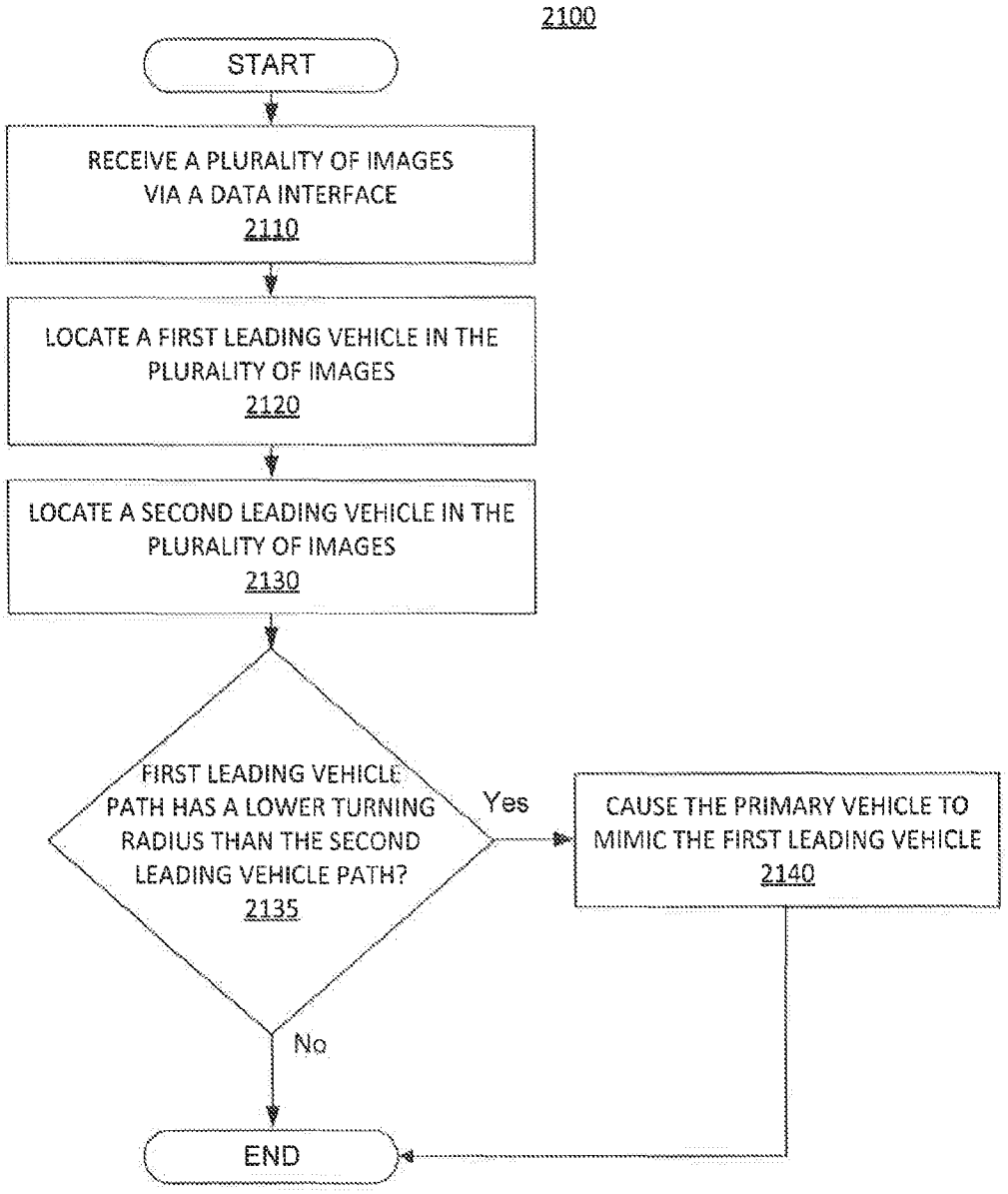


FIG. 21

IPR2025-00944
Tesla EX1002 Page 1244

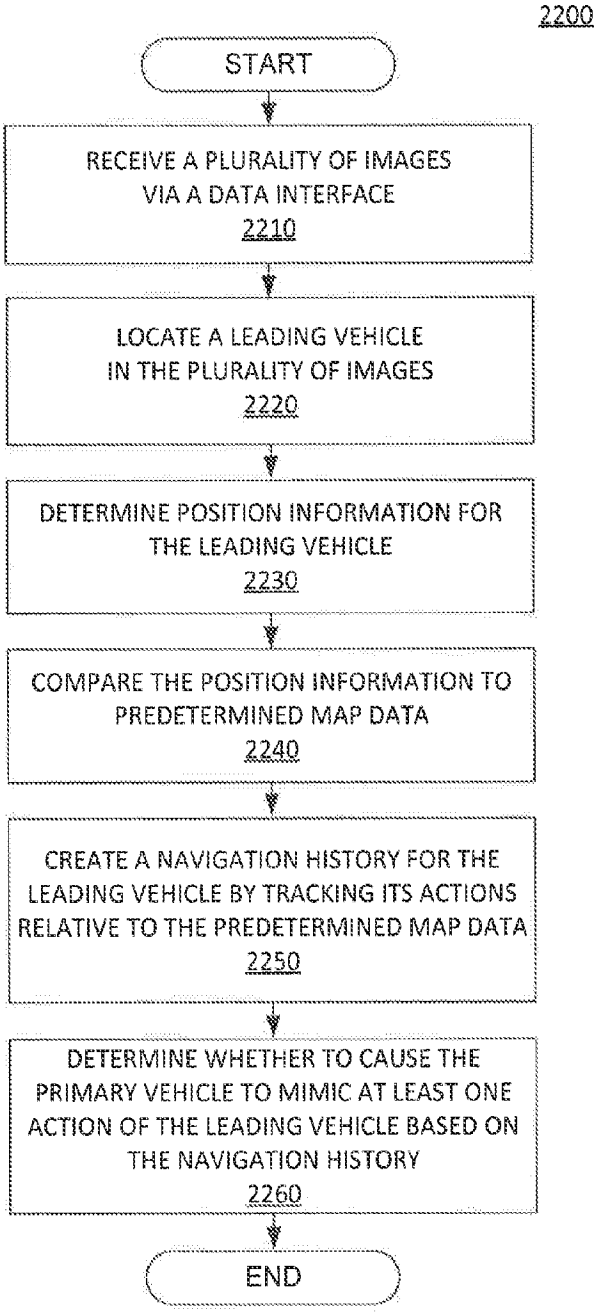


FIG. 22

IPR2025-00944
Tesla EX1002 Page 1245

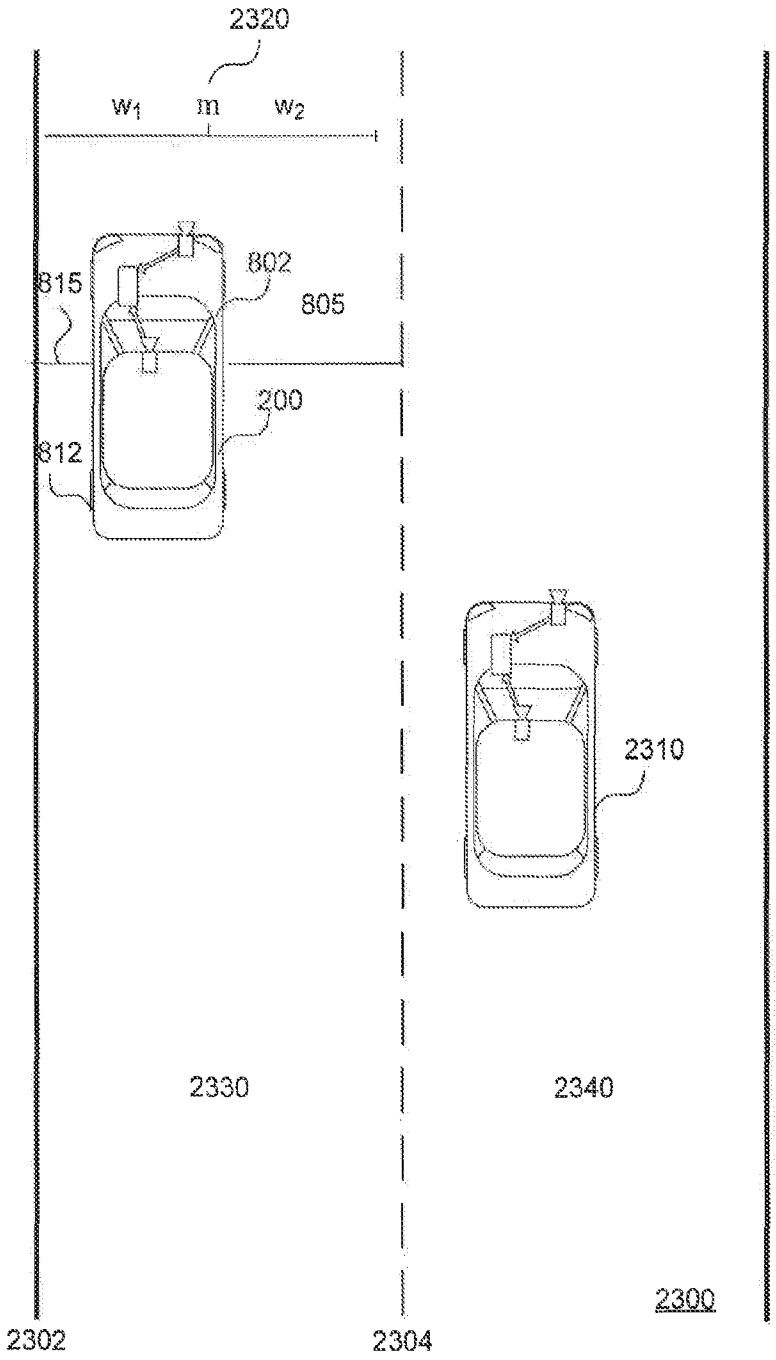


FIG. 23

IPR2025-00944
Tesla EX1002 Page 1246

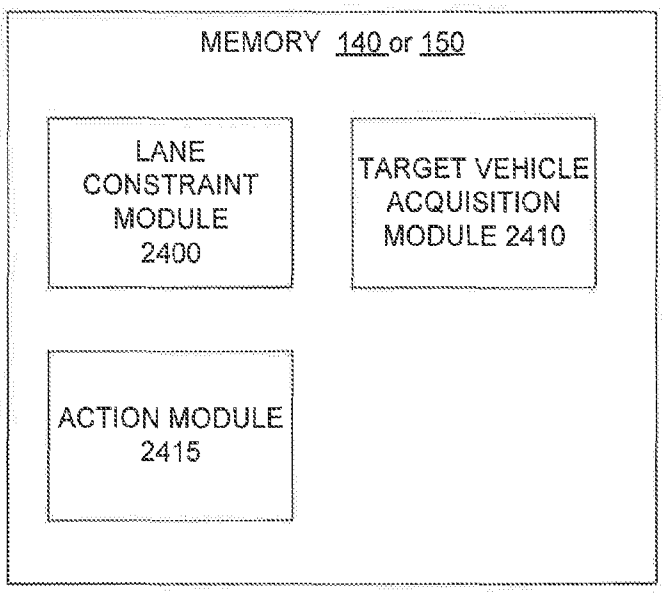


FIG. 24A

IPR2025-00944
Tesla EX1002 Page 1247

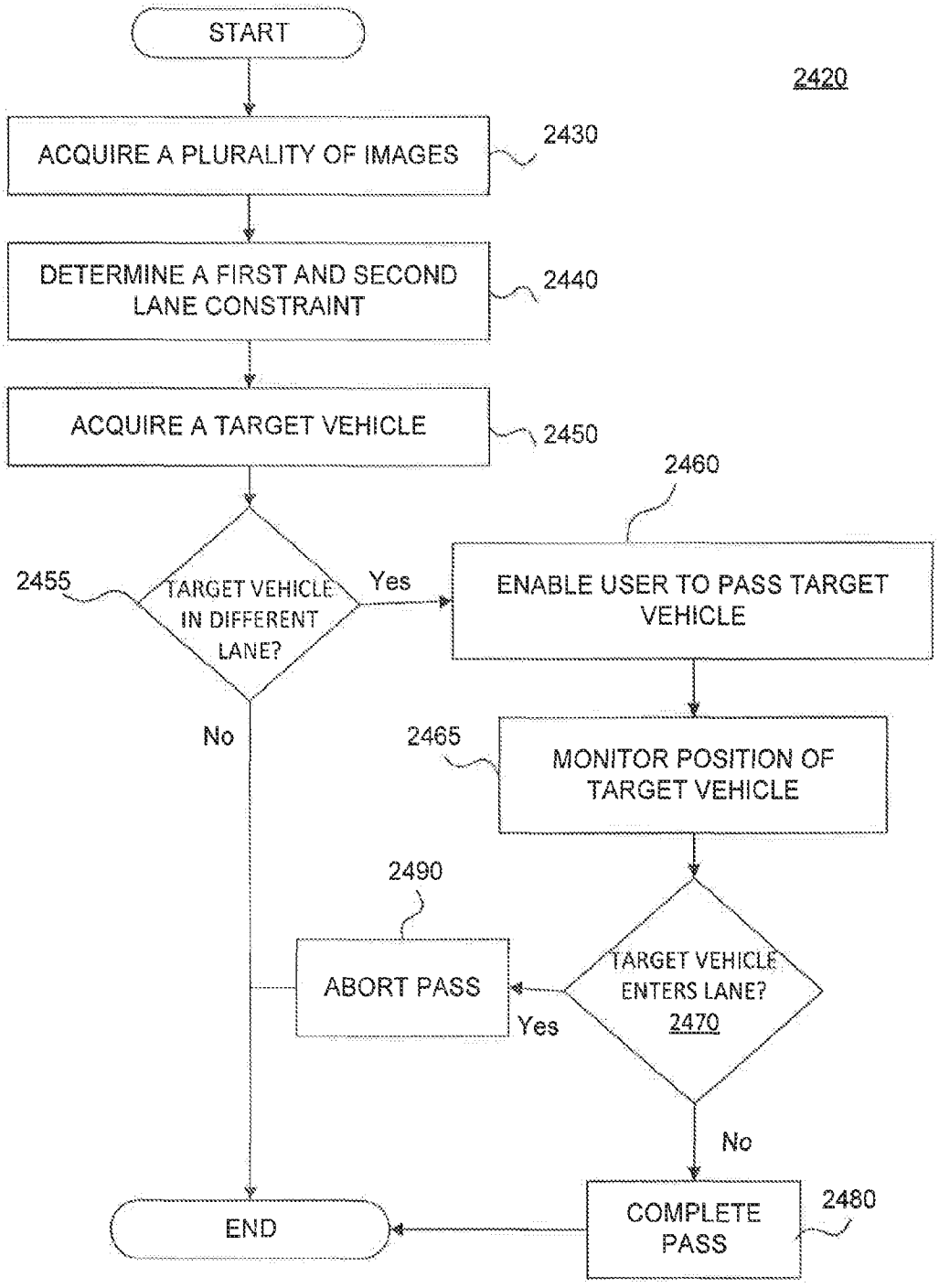


FIG. 24B

IPR2025-00944
Tesla EX1002 Page 1248

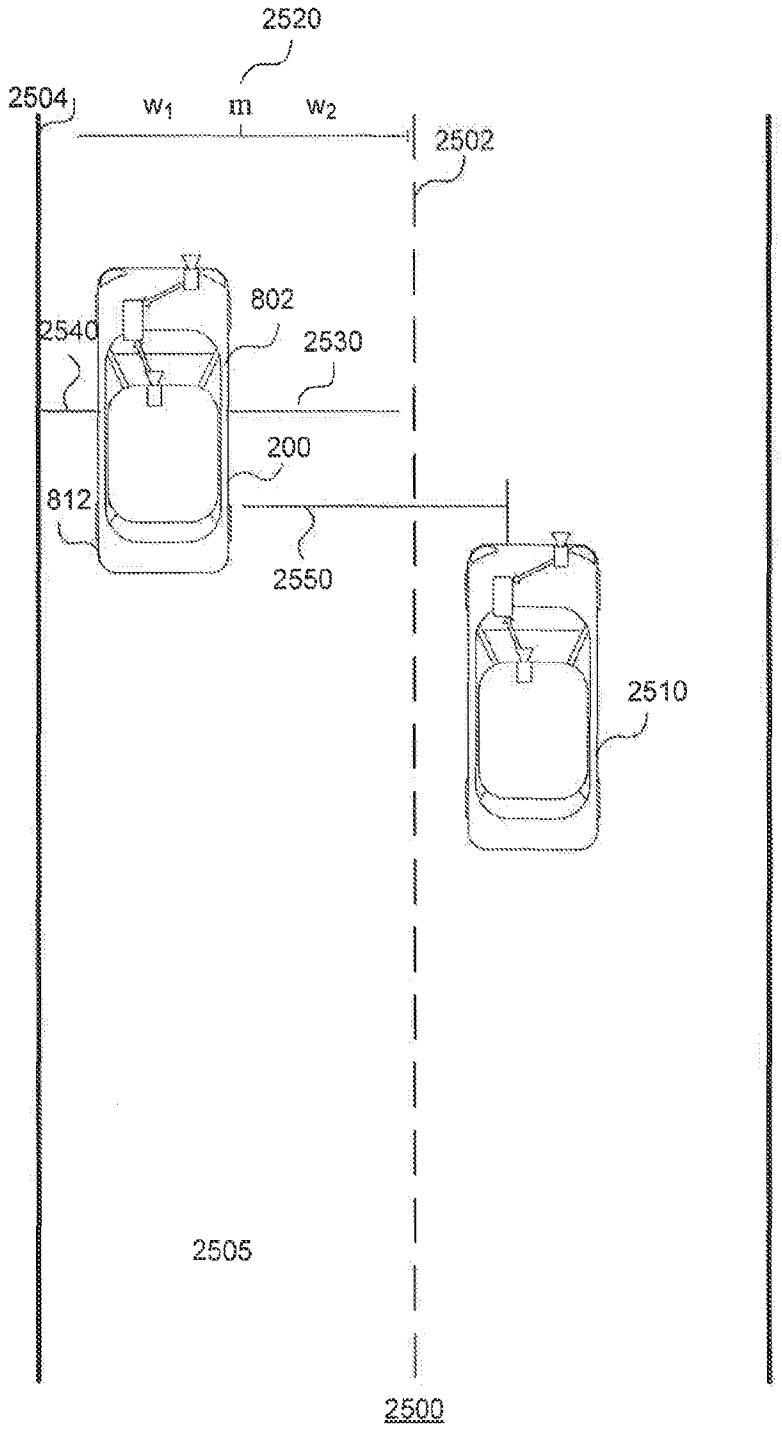


FIG. 25

IPR2025-00944
Tesla EX1002 Page 1249

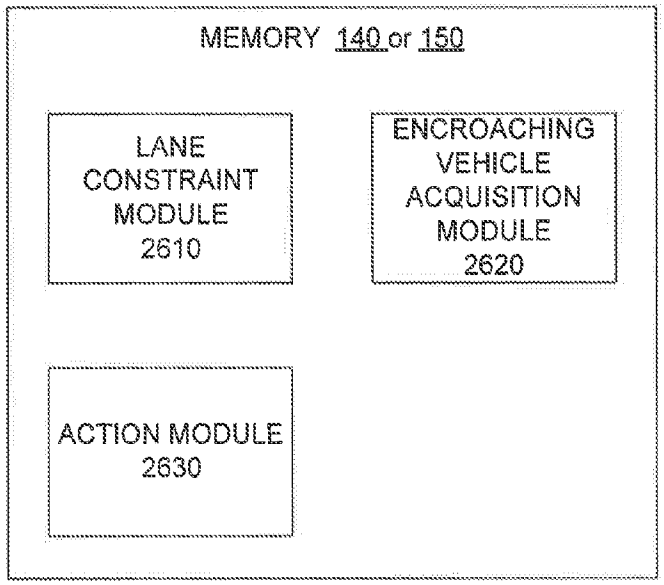


FIG. 26

IPR2025-00944
Tesla EX1002 Page 1250

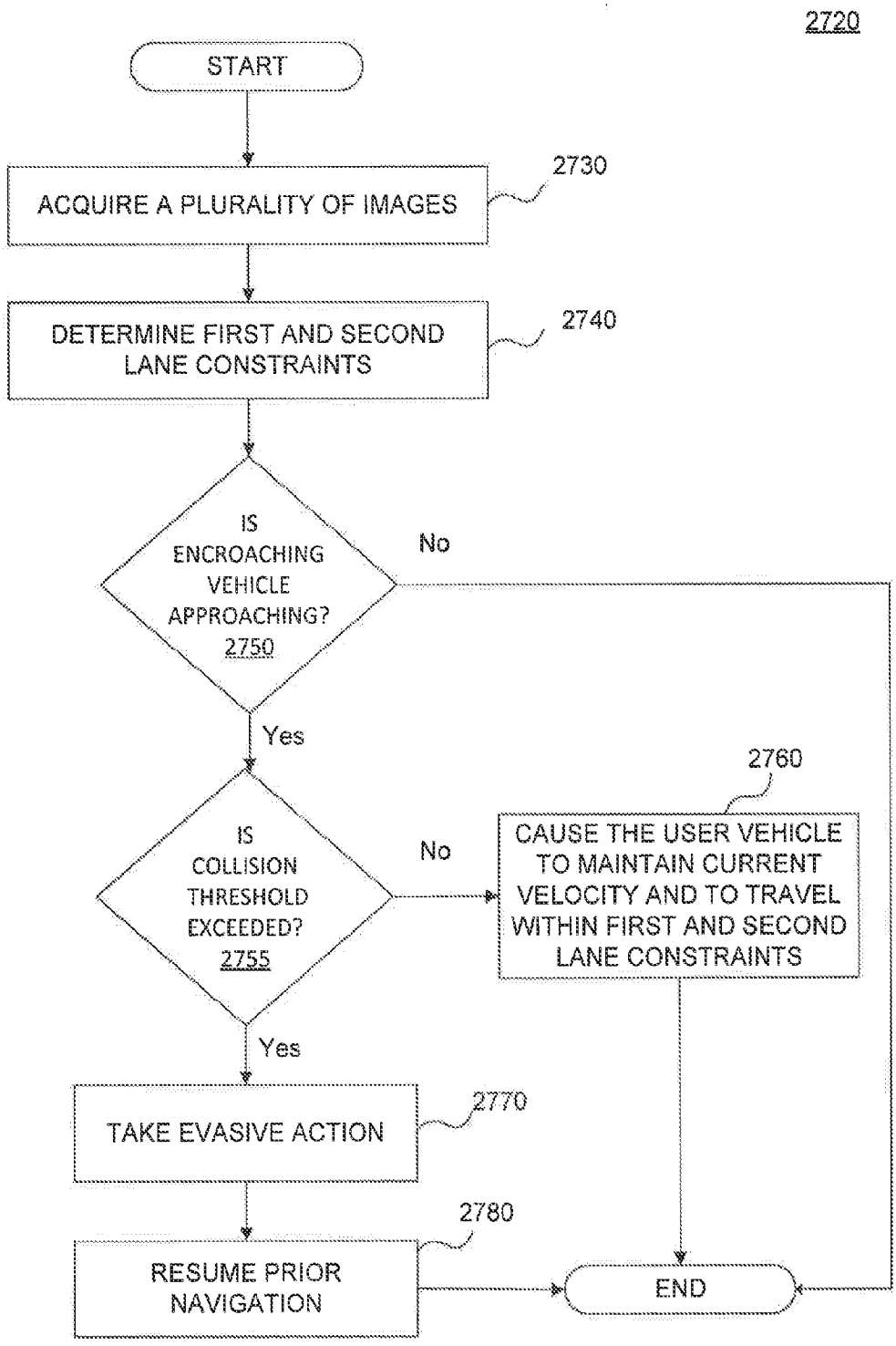


FIG. 27

IPR2025-00944
Tesla EX1002 Page 1251

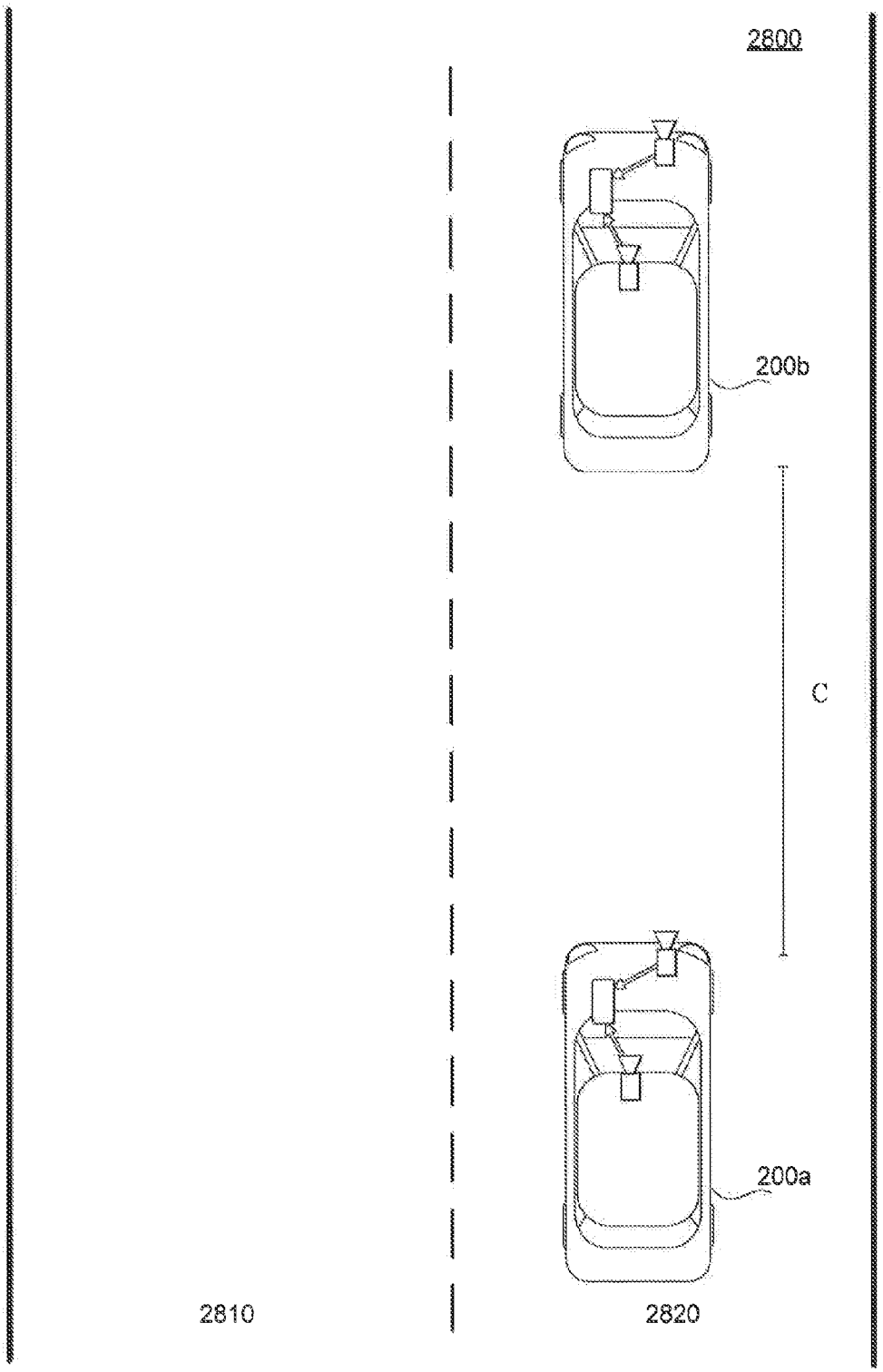


FIG. 28A

IPR2025-00944
Tesla EX1002 Page 1252

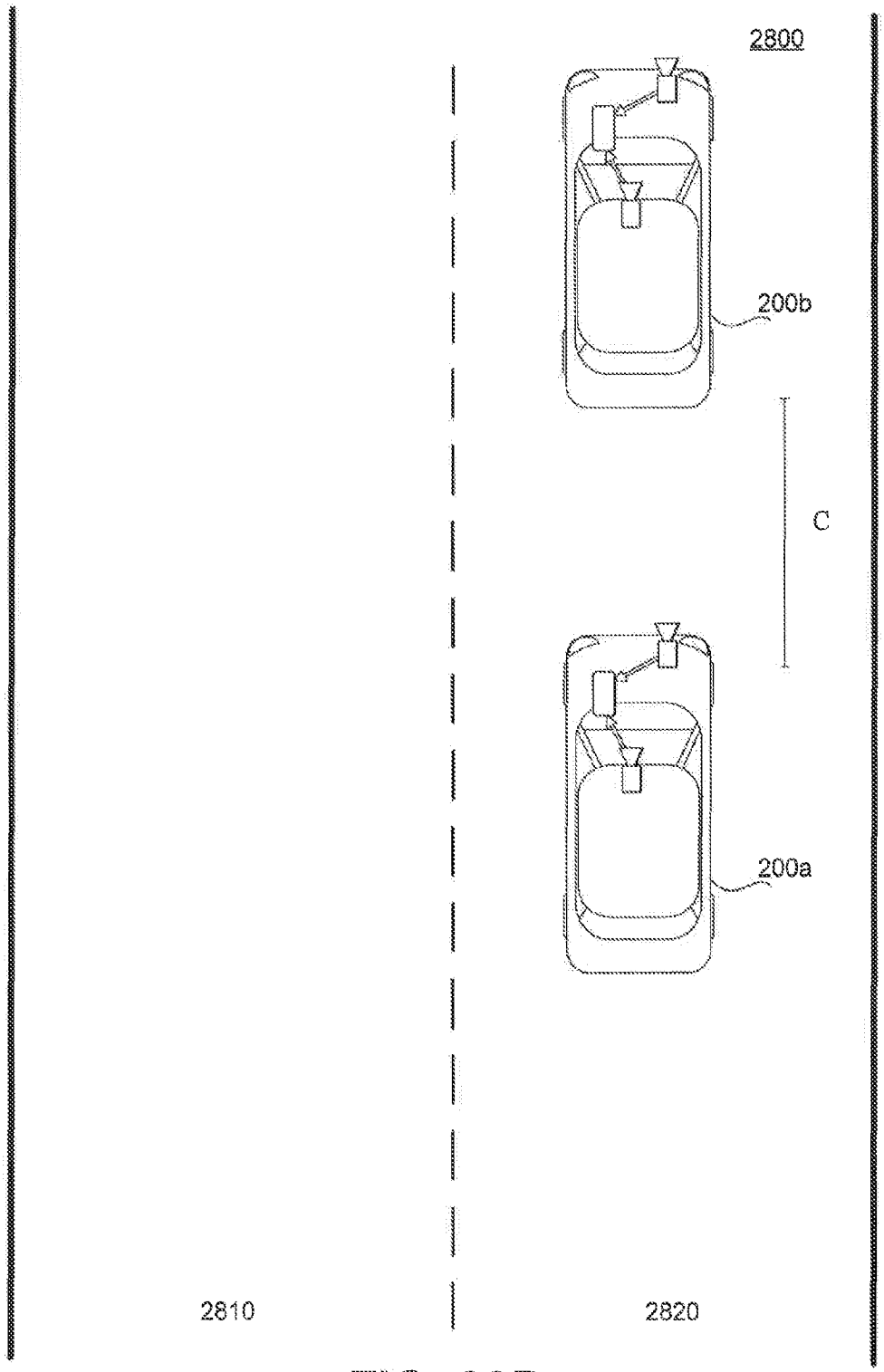


FIG. 28B

IPR2025-00944
Tesla EX1002 Page 1253

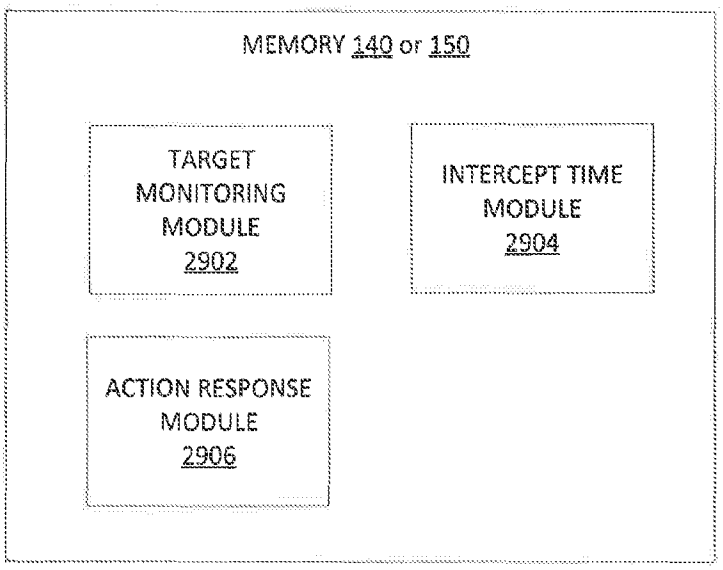


FIG. 29

IPR2025-00944
Tesla EX1002 Page 1254

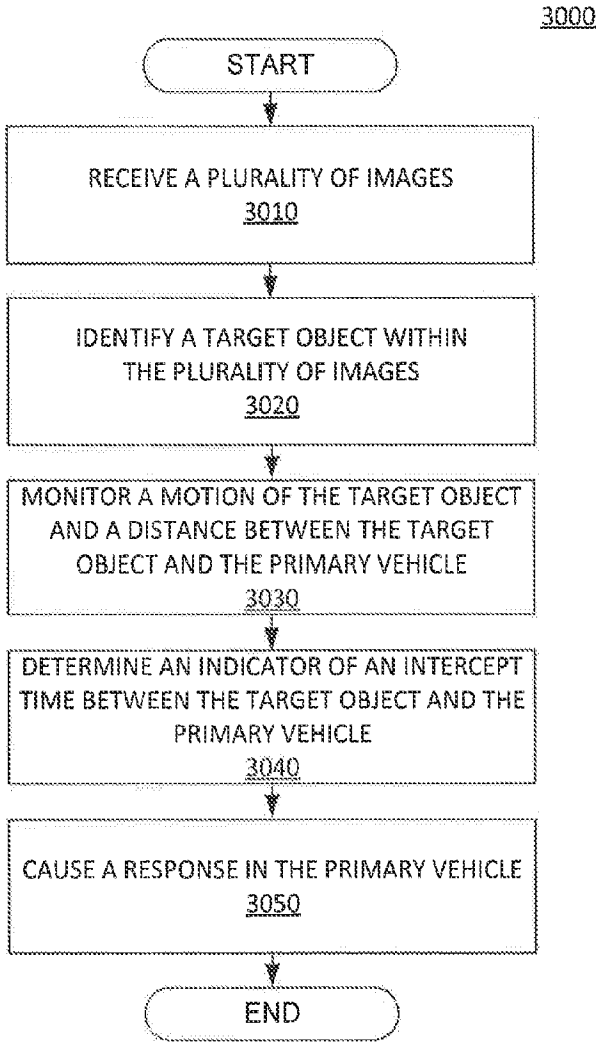


FIG. 30

IPR2025-00944
Tesla EX1002 Page 1255

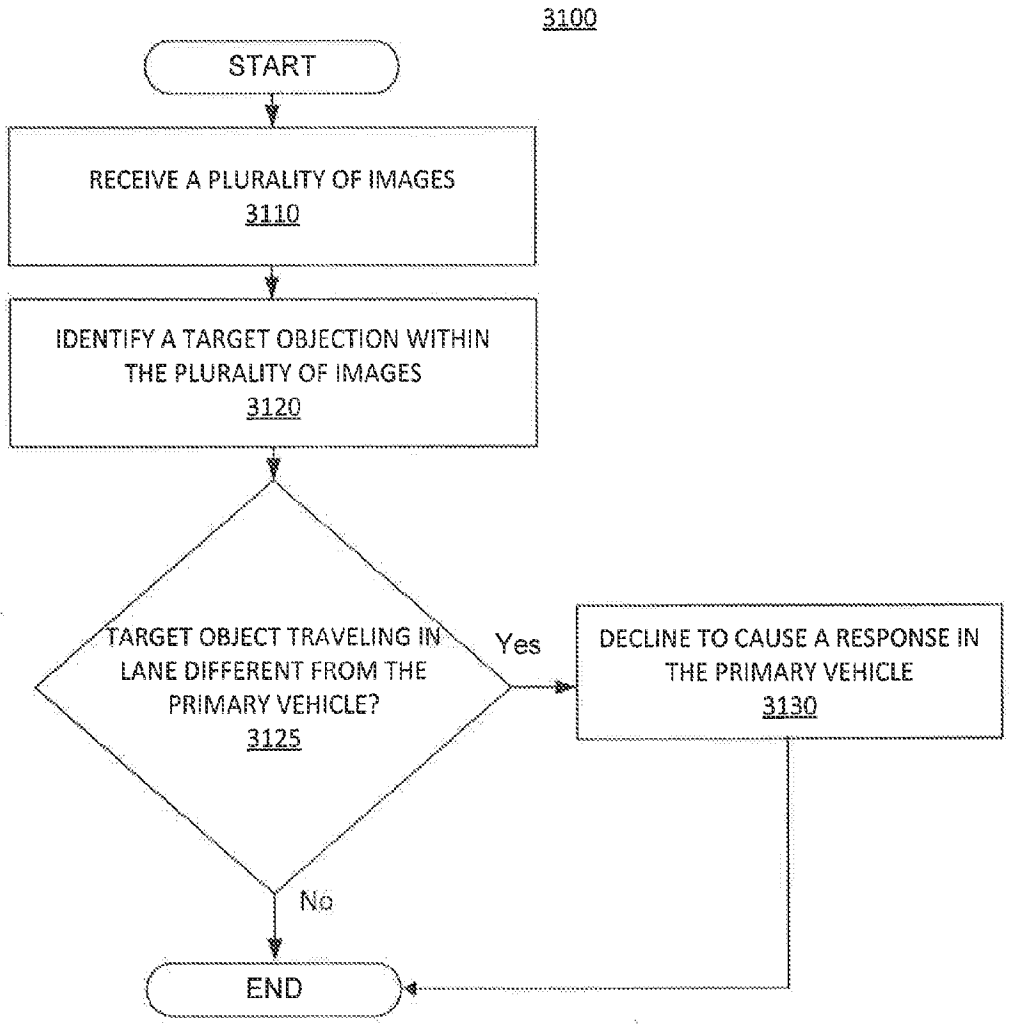


FIG. 31

IPR2025-00944
Tesla EX1002 Page 1256

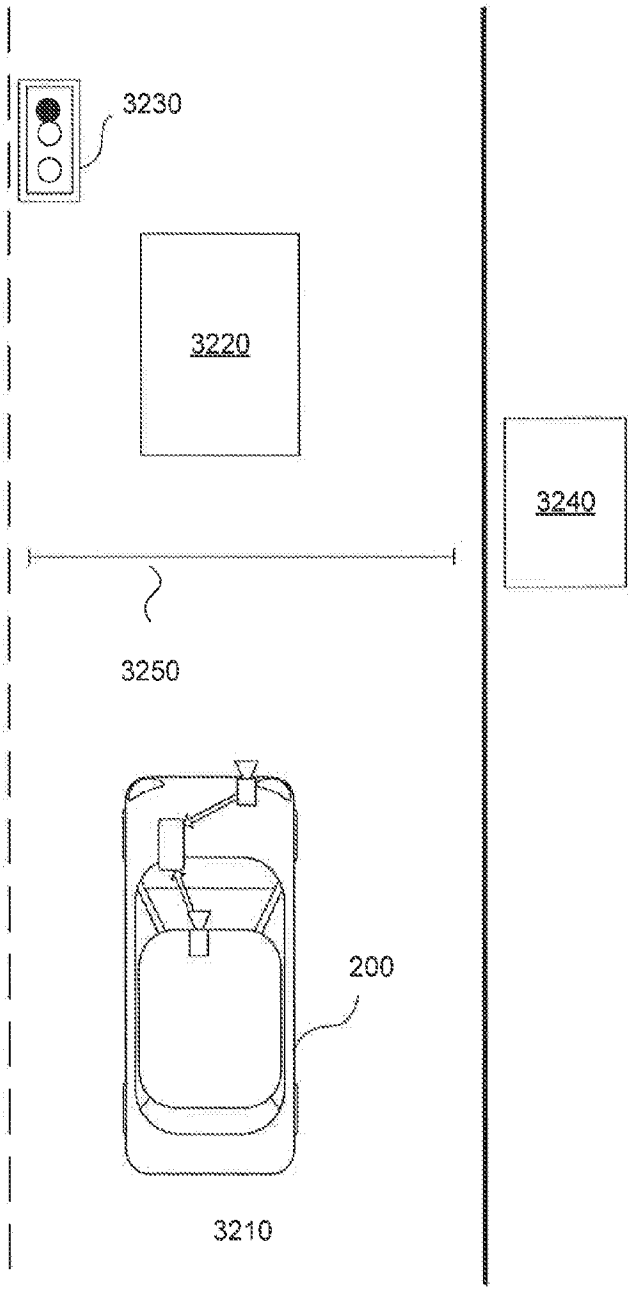


FIG. 32A

IPR2025-00944
Tesla EX1002 Page 1257

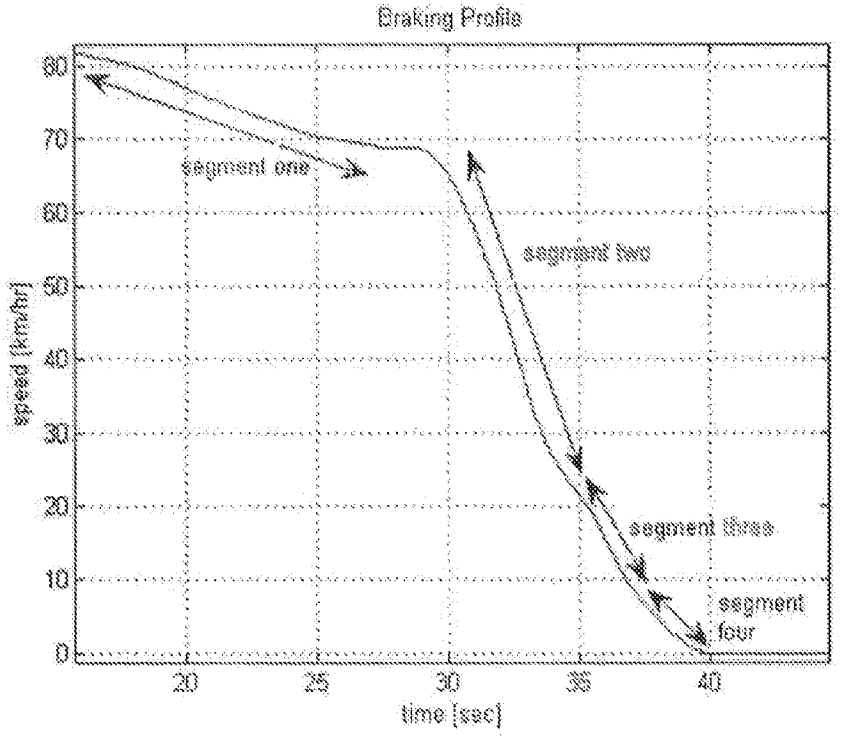


FIG. 32B

IPR2025-00944
Tesla EX1002 Page 1258

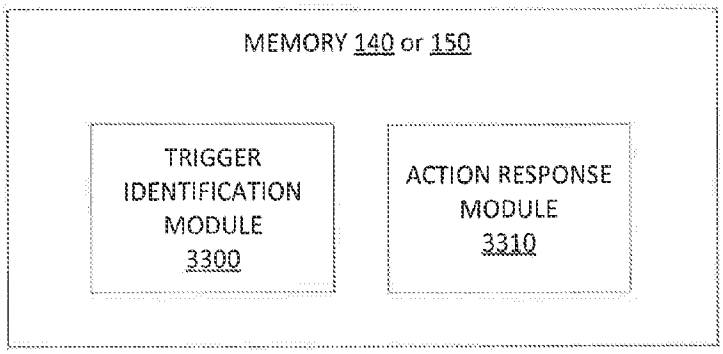


FIG. 33

IPR2025-00944
Tesla EX1002 Page 1259

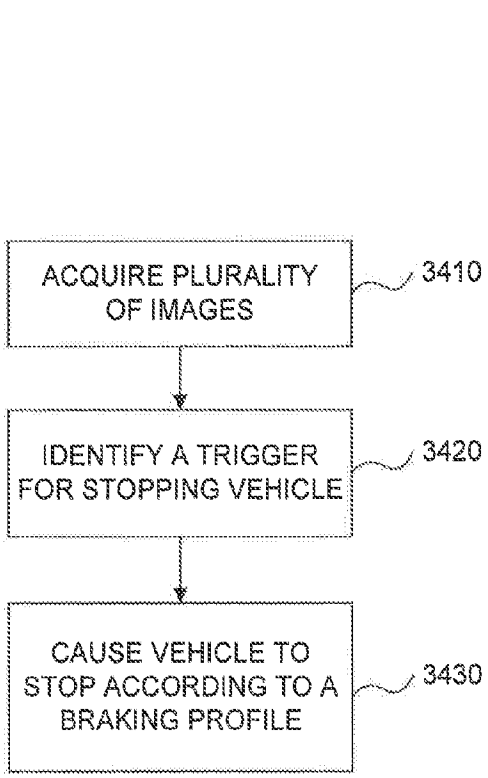


FIG. 34

IPR2025-00944
Tesla EX1002 Page 1260

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- US 20110190972 A1 **[0005]**
- US 8504233 B1 **[0006]**
- JP 2012168592 A **[0007]**

Non-patent literature cited in the description

- **XIN LIU**. *Vision-based Long-distance Lane Perception and Front Vehicle Location for Full Autonomous Vehicles on Highway Roads* **[0003]**
- **NASSER KEHTARNAVAZ**. *A Transportable Neural-Network Approach to Autonomous Vehicle Following* **[0004]**

IPR2025-00944
Tesla EX1002 Page 1261

Espacenet

Bibliographic data: EP3100206 (A1) — 2016-12-07

SYSTEMS AND METHODS FOR LANE END RECOGNITION

Inventor(s): CLARKE ANNA [IL]; BAGON EYAL [IL]; TUBIS IGOR [IL]; STEIN GIDEON [IL]; SHASHUA AMNON [IL]; BEN SHALOM ITAI [IL]; TAIEB YOAV [IL]; SHALEV SHAI [IL]; HUBERMAN DAVID [IL] ± (CLARKE, Anna, ; BAGON, Eyal, ; TUBIS, Igor, ; STEIN, Gideon, ; SHASHUA, Amnon, ; BEN SHALOM, Itai, ; TAIEB, Yoav, ; SHALEV, Shai, ; HUBERMAN, David, ; SHALEV-SCHWARZ, Shai)

Applicant(s): MOBILEYE VISION TECHNOLOGIES LTD [IL] ± (Mobileye Vision Technologies Ltd)

Classification: **- international:** ***G06K9/00***
- cooperative: **B60K31/00 (US)**; **B60T7/12 (EP, US)**; **B60W10/20 (EP, US)**; **B60W30/00 (US)**; **B60W30/09 (US)**; **B60W30/14 (US)**; **B60W30/18163 (EP, US)**; **B62D6/00 (EP, US)**; **G01S19/42 (US)**; **G05D1/0088 (US)**; **G05D1/0246 (US)**; **G05D1/0276 (US)**; **G06K9/00798 (EP, US)**; **G06K9/00805 (EP, US)**; **G06K9/00825 (EP, US)**; **G06T7/62 (US)**; **G08G1/095 (US)**; **G08G1/165 (EP)**; **G08G1/167 (EP)**; **H04N7/181 (US)**; **H04N7/183 (US)**; B60W2420/42 (EP, US); B60W2552/15 (US); B60W2552/30 (EP, US); B60W2554/00 (EP, US); B60W2554/4041 (EP, US); B60W2554/80 (EP, US); B60W2555/60 (EP, US); B60W2710/207 (EP, US); B60W30/12 (EP, US); G06T2207/30256 (EP, US); G06T2207/30261 (EP, US)

Application number: EP20150703711 20150130 Global Dossier

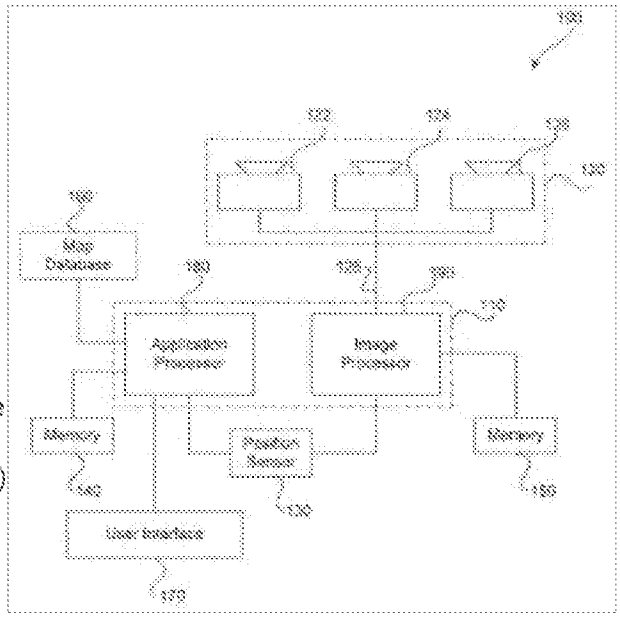
Priority number(s): US201461933325P 20140130 ; US201461993111P 20140514 ; US201462040246P 20140821 ; US201462060603P 20141007 ; US201562102669P 20150113 ; WO2015US13812 20150130

Also published as: EP3100206 (B1) CN106462727 (A) CN106462727 (B) CN111199218 (A) EP3736732 (A1) more

Abstract not available for EP3100206 (A1)
Abstract of corresponding document: US2015210274 (A1)

IPR2025-00944
Tesla EX1002 Page 1262

Systems and methods use cameras to provide autonomous navigation features. In one implementation, a lane ending detection system is provided for a vehicle. One or more processing devices associated with the system receive at least one image via a data interface. The device(s) extract lane ending information from the road sign(s) included in the image data and determine, based on at least one indicator of position of the vehicle, a distance from the vehicle to one or more lane constraints associated with the current lane. The processing device(s) determine, based on the lane ending information and the vehicle position, whether a current lane in which the vehicle is traveling is ending. Further, the system may cause the vehicle to change lanes if the lane in which the vehicle is traveling is ending.



Espacenet

Description: EP3100206 (A1) — 2016-12-07

SYSTEMS AND METHODS FOR LANE END RECOGNITION

Description of EP3100206 (B1)

A high quality text as facsimile in your desired language may be available amongst the following family members:

CN106462727 (A) US2015210274 (A1) WO2015116950 (A1) CN111199218 (A)
US2015210275 (A1) US2015210276 (A1) US2015210277 (A1) US2015210278 (A1)
US2015210312 (A1) US2016167704 (A1) US2016283807 (A1) US2016306361 (A1)
US2019294179 (A1) US2020225681 (A1)

The EPO does not accept any responsibility for the accuracy of data and information originating from other authorities than the EPO; in particular, the EPO does not guarantee that they are complete, up-to-date or fit for specific purposes.

BACKGROUND

I. Technical Field

[0001] The present disclosure relates generally to autonomous vehicle navigation and, more specifically, to systems and methods that use cameras to provide autonomous vehicle navigation features.

II. Background Information

[0002] As technology continues to advance, the goal of a fully autonomous vehicle that is capable of navigating on roadways is on the horizon. Primarily, an autonomous vehicle may be able to identify its environment and navigate without input from a human operator. Autonomous vehicles may also take into account a variety of factors and make appropriate decisions based on those factors to safely and accurately reach an intended destination. For example, various objects-such as other vehicles and pedestrians-are encountered when a vehicle typically travels a roadway. Autonomous driving systems may recognize these objects in a vehicle's environment and take appropriate and timely action to avoid collisions. Additionally, autonomous driving systems may identify other indicators-such as traffic signals, traffic signs, and lane markings-that regulate vehicle movement (e.g., when the vehicle must stop and may go, a speed at which the vehicle must not exceed, where the vehicle must be positioned on the roadway, etc.). Autonomous driving systems may need to determine

IPR2025-00944
Tesla EX1002 Page 1264

when a vehicle should change lanes, turn at intersections, change roadways, etc. As is evident from these examples, many factors may need to be addressed in order to provide an autonomous vehicle that is capable of navigating safely and accurately. Example of background art can be found in DE 10 2010 042440 A1 .

SUMMARY

[0003] The invention is set out in the appended claims.

[0004] The foregoing general description and the following detailed description are exemplary and explanatory only and are not restrictive of the claims.

BRIEF DESCRIPTION OF THE DRAWINGS

[0005] The accompanying drawings, which are incorporated in and constitute a part of this disclosure, illustrate various disclosed embodiments. In the drawings:
FIG. 1 is a diagrammatic representation of an exemplary system consistent with the disclosed embodiments.
FIG. 2A is a diagrammatic side view representation of an exemplary vehicle including a system consistent with the disclosed embodiments.
FIG. 2B is a diagrammatic top view representation of the vehicle and system shown in FIG. 2A consistent with the disclosed embodiments.
FIG. 2C is a diagrammatic top view representation of another embodiment of a vehicle including a system consistent with the disclosed embodiments.
FIG. 2D is a diagrammatic top view representation of yet another embodiment of a vehicle including a system consistent with the disclosed embodiments.
FIG. 2E is a diagrammatic top view representation of yet another embodiment of a vehicle including a system consistent with the disclosed embodiments.
FIG. 2F is a diagrammatic representation of exemplary vehicle control systems consistent with the disclosed embodiments.
FIG. 3A is a diagrammatic representation of an interior of a vehicle including a rearview mirror and a user interface for a vehicle imaging system consistent with the disclosed embodiments.
FIG. 3B is an illustration of an example of a camera mount that is configured to be positioned behind a rearview mirror and against a vehicle windshield consistent with the disclosed embodiments.
FIG. 3C is an illustration of the camera mount shown in FIG. 3B from a different perspective consistent with the disclosed embodiments.
FIG. 3D is an illustration of an example of a camera mount that is configured to be positioned behind a rearview mirror and against a vehicle windshield consistent with the disclosed embodiments.
FIG. 4 is an exemplary block diagram of a memory configured to store instructions for performing one or more operations consistent with the disclosed embodiments.
FIG. 5A is a flowchart showing an exemplary process for causing one or more navigational responses based on monocular image analysis consistent with disclosed embodiments.
FIG. 5B is a flowchart showing an exemplary process for detecting one or more vehicles and/or pedestrians in a set of images consistent with the disclosed embodiments.
FIG. 5C is a flowchart showing an exemplary process for detecting road marks and/or lane geometry information in a set of images consistent with the disclosed embodiments.

Tesla EX1002 Page 1265 IPR2025-00944

FIG. 5D is a flowchart showing an exemplary process for detecting traffic lights in a set of images consistent with the disclosed embodiments.
FIG. 5E is a flowchart showing an exemplary process for causing one or more navigational responses based on a vehicle path consistent with the disclosed embodiments.
FIG. 5F is a flowchart showing an exemplary process for determining whether a leading vehicle is changing lanes consistent with the disclosed embodiments.
FIG. 6 is a flowchart showing an exemplary process for causing one or more navigational responses based on stereo image analysis consistent with the disclosed embodiments.
FIG. 7 is a flowchart showing an exemplary process for causing one or more navigational responses based on an analysis of three sets of images consistent with the disclosed embodiments.
FIG. 8 is a diagrammatic top view representation of an exemplary vehicle including a system consistent with the disclosed embodiments in which the vehicle's current travel lane is ending.
FIG. 9 is a diagrammatic representation of the memory of an exemplary lane end recognition system consistent with the disclosed embodiments.
FIG. 10A is a diagrammatic representation of an exemplary vehicle including a lane end recognition system encountering the end of a travel lane marked by road signs consistent with the disclosed embodiments.
FIG. 10B is a diagrammatic representation of an exemplary vehicle including a lane end recognition system traveling a measurable distance behind a lead vehicle consistent with the disclosed embodiments.
FIG. 11 is a flowchart showing an exemplary process for determining whether a current lane in which the vehicle is traveling is ending, consistent with the disclosed embodiments.
FIG. 12 is a diagrammatic top view representation of an exemplary vehicle including a system consistent with the disclosed embodiments in which an object is detected within the travel lane in which the vehicle is traveling.
FIG. 13 is a diagrammatic representation of the memory of an exemplary driver-assist object detection system consistent with the disclosed embodiments.
FIG. 14A is a diagrammatic representation of two captured images indicating an object found within the travel lane among a plurality of images captured by an exemplary driver-assist object detection system consistent with the disclosed embodiments.
FIG. 14B is a diagrammatic top view representation of an exemplary vehicle including a system consistent with the disclosed embodiments in which an object is detected within the travel lane in which the vehicle is traveling.
FIG. 15 is a flowchart showing an exemplary process for determining whether a current lane in which the vehicle is traveling is ending, consistent with the disclosed embodiments.
FIG. 16 is a diagrammatic side view representation of an exemplary vehicle including an image capture device with an extended angle-of-view consistent with the disclosed embodiments.
FIG. 17 is a diagrammatic top view representation of the vehicle including the image capture device with the extended angle-of-view consistent with the disclosed embodiments.
FIG. 18 is a diagrammatic representation of a memory storing instructions to identify objects located outside of a sightline of a typical driver location consistent with the disclosed embodiments.
FIG. 19 is a flowchart showing an exemplary process for identifying objects located outside of a sightline of a typical driver location consistent with the disclosed

embodiments.
FIG. 20 is a diagrammatic representation of two exemplary vehicles approaching an intersection with a plurality of traffic lights consistent with the disclosed embodiments.
FIG. 21 is a diagrammatic representation of an exemplary vehicle approaching an intersection with a plurality of traffic lights consistent with the disclosed embodiments.
FIG. 22 is a diagrammatic representation of a memory storing instructions for detecting traffic lights consistent with the disclosed embodiments.
FIG. 23 is a flowchart showing an exemplary process for detecting traffic lights consistent with the disclosed embodiments.
FIG. 24 is a diagrammatic top view representation of an exemplary vehicle including a system consistent with the disclosed embodiments in which the vehicle is approaching an intersection with a traffic light.
FIG. 25 is a diagrammatic representation of the memory of an exemplary traffic light detection system consistent with the disclosed embodiments.
FIG. 26A is a diagrammatic representation of an exemplary vehicle including a traffic light detection system encountering an intersection with a traffic light and determining its status consistent with the disclosed embodiments.
FIG. 26B is a diagrammatic representation of an exemplary vehicle including a traffic light detection system encountering an intersection with a blinking yellow light and determining a course of action consistent with the disclosed embodiments.
FIG. 27 is a flowchart showing an exemplary process for determining the status of a traffic light at an intersection, consistent with the disclosed embodiments.
FIG. 28 is an exemplary block diagram of a memory configured to store instructions for performing one or more operations consistent with disclosed embodiments.
FIG. 29A is an illustration of an exemplary traffic light, consistent with disclosed embodiments.
FIG. 29B is another illustration of an exemplary traffic light, consistent with disclosed embodiments.
FIG. 29C is an illustration of an exemplary image of a portion of the traffic light of FIG. 29A , consistent with disclosed embodiments.
FIG. 29D is an illustration of an exemplary image of a portion of the traffic light of FIG. 29B , consistent with disclosed embodiments.
FIG. 30A is a flowchart showing an exemplary process for determining whether a traffic light includes an arrow, consistent with disclosed embodiments.
FIG. 30B is a flowchart showing an exemplary process for causing a system response based on the detection of an arrow signal in a traffic light, consistent with disclosed embodiments.

DETAILED DESCRIPTION

[0006] The following detailed description refers to the accompanying drawings. Wherever possible, the same reference numbers are used in the drawings and the following description to refer to the same or similar parts. While several illustrative embodiments are described herein, modifications, adaptations and other implementations are possible. For example, substitutions, additions or modifications may be made to the components illustrated in the drawings, and the illustrative methods described herein may be modified by substituting, reordering, removing, or adding steps to the disclosed methods. Accordingly, the following detailed description is not limited to the disclosed embodiments and examples. Instead, the proper scope is defined by the appended claims.

[0007] Disclosed embodiments provide systems and methods that use cameras to

provide autonomous navigation features. In various embodiments, the system may include one, two or more cameras that monitor the environment of a vehicle. In one embodiment, the system may recognize when a driving lane will end and, when an end is detected, make a responsive navigational change, such as a lane change. Rather than relying on map data to make this determination, the system may operate based solely on visual information. Further, while the system may recognize an end of a lane by direct observation of the lane ending, the system may recognize an approaching lane end sooner by, for example, observing and recognizing road sign alerts that provide information regarding a lane ending. In another embodiment, the system may determine the presence of obstacles in a roadway where those obstacles have a height of less than, for example, 10 cm. In another embodiment, the system may view and recognize traffic lights located above a vehicle. In another embodiment, the system may recognize a traffic light and its status from among a plurality of traffic lights and cause a navigational change based on the status of the traffic light (e.g., discontinue cruise and cause braking when a red light is recognized). For example, the system may recognize which of the imaged traffic lights are relevant to the vehicle and respond only to those traffic lights determined to be relevant. In another embodiment, the system may recognize a turn lane traffic light and its status based on analysis of road context information combined with determined characteristics of the light. In another embodiment, the system may employ a super resolution technique to recognize and analyze traffic lights. For example, the system may distinguish between round and arrow shaped traffic lights.

[0008] FIG. 1 is a block diagram representation of a system 100 consistent with the exemplary disclosed embodiments. System 100 may include various components depending on the requirements of a particular implementation. In some embodiments, system 100 may include a processing unit 110, an image acquisition unit 120, a position sensor 130, one or more memory units 140, 150, a map database 160, and a user interface 1.70. Processing unit 110 may include one or more processing devices. In some embodiments, processing unit 110 may include an applications processor 180, an image processor 190, or any other suitable processing device. Similarly, image acquisition unit 120 may include any number of image acquisition devices and components depending on the requirements of a particular application. In some embodiments, image acquisition unit 120 may include one or more image capture devices (e.g., cameras), such as image capture device 122, image capture device 124, and image capture device 126. System 100 may also include a data interface 128 communicatively connecting processing device 110 to image acquisition device 120. For example, data interface 128 may include any wired and/or wireless link or links for transmitting image data acquired by image accusation device 120 to processing unit 110.

[0009] Both applications processor 180 and image processor 190 may include various types of processing devices. For example, either or both of applications processor 180 and image processor 190 may include a microprocessor, preprocessors (such as an image preprocessor), graphics processors, a central processing unit (CPU), support circuits, digital signal processors, integrated circuits, memory, or any other types of devices suitable for running applications and for image processing and analysis. In some embodiments, applications processor 180 and/or image processor 190 may include any type of single or multi-core processor, mobile device microcontroller, central processing unit, etc. Various processing devices may be used, including, for example, processors available from manufacturers such as Intel®, AMD®, etc. and may include various architectures (e.g., x86 processor, ARM®, etc.).

IPR2025-00944
Tesla EX1002 Page 1268

[0010] In some embodiments, applications processor 180 and/or image processor 190 may include any of the EyeQ series of processor chips available from Mobileye®. These processor designs each include multiple processing units with local memory and instruction sets. Such processors may include video inputs for receiving image data from multiple image sensors and may also include video out capabilities. In one example, the EyeQ2® uses 90nm-micron technology operating at 332Mhz. The EyeQ2® architecture consists of two floating point, hyper-thread 32-bit RISC CPUs (MIPS32® 34K® cores), five Vision Computing Engines (VCE), three Vector Microcode Processors (VMP®), Denali 64-bit Mobile DDR Controller, 128-bit internal Sonics Interconnect, dual 16-bit Video input and 18-bit Video output controllers, 16 channels DMA and several peripherals. The MIPS34K CPU manages the five VCEs, three VMP™ and the DMA, the second MIPS34K CPU and the multi-channel DMA as well as the other peripherals. The five VCEs, three VMP® and the MIPS34K CPU can perform intensive vision computations required by multi-function bundle applications. In another example, the EyeQ3®, which is a third generation processor and is six times more powerful that the EyeQ2®, may be used in the disclosed embodiments.

[0011] Any of the processing devices disclosed herein may be configured to perform certain functions. Configuring a processing device, such as any of the described EyeQ processors or other controller or microprocessor, to perform certain functions may include programming of computer executable instructions and making those instructions available to the processing device for execution during operation of the processing device. In some embodiments, configuring a processing device may include programming the processing device directly with architectural instructions. In other embodiments, configuring a processing device may include storing executable instructions on a memory that is accessible to the processing device during operation. For example, the processing device may access the memory to obtain and execute the stored instructions during operation.

[0012] While FIG. 1 depicts two separate processing devices included in processing unit 110, more or fewer processing devices may be used. For example, in some embodiments, a single processing device may be used to accomplish the tasks of applications processor 180 and image processor 190. In other embodiments, these tasks may be performed by more than two processing devices.

[0013] Processing unit 110 may comprise various types of devices. For example, processing unit 110 may include various devices, such as a controller, an image preprocessor, a central processing unit (CPU), support circuits, digital signal processors, integrated circuits, memory, or any other types of devices for image processing and analysis. The image preprocessor may include a video processor for capturing, digitizing and processing the imagery from the image sensors. The CPU may comprise any number of microcontrollers or microprocessors. The support circuits may be any number of circuits generally well known in the art, including cache, power supply, clock and input-output circuits. The memory may store software that, when executed by the processor, controls the operation of the system. The memory may include databases and image processing software. The memory may comprise any number of random access memories, read only memories, flash memories, disk drives, optical storage, tape storage, removable storage and other types of storage. In one instance, the memory may be separate from the processing unit 110. In another instance, the memory may be integrated into the processing unit 110.

Tesla EX1002 Page 1269 IPR2025-00944

[0014] Each memory 140, 150 may include software instructions that when executed by a processor (e.g., applications processor 180 and/or image processor 190), may control operation of various aspects of system 100. These memory units may include various databases and image processing software. The memory units may include random access memory, read only memory, flash memory, disk drives, optical storage, tape storage, removable storage and/or any other types of storage. In some embodiments, memory units 140, 150 may be separate from the applications processor 180 and/or image processor 190. In other embodiments, these memory units may be integrated into applications processor 180 and/or image processor 190.

[0015] Position sensor 130 may include any type of device suitable for determining a location associated with at least one component of system 100. In some embodiments, position sensor 130 may include a GPS receiver. Such receivers can determine a user position and velocity by processing signals broadcasted by global positioning system satellites. Position information from position sensor 130 may be made available to applications processor 180 and/or image processor 190.

[0016] User interface 170 may include any device suitable for providing information to or for receiving inputs from one or more users of system 100. In some embodiments, user interface 170 may include user input devices, including, for example, a touchscreen, microphone, keyboard, pointer devices, track wheels, cameras, knobs, buttons, etc. With such input devices, a user may be able to provide information inputs or commands to system 100 by typing instructions or information, providing voice commands, selecting menu options on a screen using buttons, pointers, or eye-tracking capabilities, or through any other suitable techniques for communicating information to system 100.

[0017] User interface 170 may be equipped with one or more processing devices configured to provide and receive information to or from a user and process that information for use by, for example, applications processor 180. In some embodiments, such processing devices may execute instructions for recognizing and tracking eye movements, receiving and interpreting voice commands, recognizing and interpreting touches and/or gestures made on a touchscreen, responding to keyboard entries or menu selections, etc. In some embodiments, user interface 170 may include a display, speaker, tactile device, and/or any other devices for providing output information to a user.

[0018] Map database 160 may include any type of database for storing map data useful to system 100. In some embodiments, map database 160 may include data relating to the position, in a reference coordinate system, of various items, including roads, water features, geographic features, businesses, points of interest, restaurants, gas stations, etc. Map database 160 may store not only the locations of such items, but also descriptors relating to those items, including, for example, names associated with any of the stored features. In some embodiments, map database 160 may be physically located with other components of system 100. Alternatively or additionally, map database 160 or a portion thereof may be located remotely with respect to other components of system 100 (e.g., processing unit 110). In such embodiments, information from map database 160 may be downloaded over a wired or wireless data connection to a network (e.g., over a cellular network and/or the Internet, etc.).

[0019] Image capture devices 122, 124, and 126 may each include any type of device suitable for capturing at least one image from an environment. Moreover, any number

IPR2025-00944
Tesla EX1002 Page 1270

of image capture devices may be used to acquire images for input to the image processor. Some embodiments may include only a single image capture device, while other embodiments may include two, three, or even four or more image capture devices. Image capture devices 122, 124, and 126 will be further described with reference to FIGS. 2B-2E , below.

[0020] System 100, or various components thereof, may be incorporated into various different platforms. In some embodiments, system 100 may be included on a vehicle 200, as shown in FIG. 2A . For example, vehicle 200 may be equipped with a processing unit 110 and any of the other components of system 100, as described above relative to FIG. 1 . While in some embodiments vehicle 200 may be equipped with only a single image capture device (e.g., camera), in other embodiments, such as those discussed in connection with FIGS. 2B-2E , multiple image capture devices may be used. For example, either of image capture devices 122 and 124 of vehicle 200, as shown in FIG. 2A , may be part of an ADAS (Advanced Driver Assistance Systems) imaging set.

[0021] The image capture devices included on vehicle 200 as part of the image acquisition unit 120 may be positioned at any suitable location. In some embodiments, as shown in FIGS. 2A-2E and 3A-3C , image capture device 122 may be located in the vicinity of the rearview mirror. This position may provide a line of sight similar to that of the driver of vehicle 200, which may aid in determining what is and is not visible to the driver. Image capture device 122 may be positioned at any location near the rearview mirror, but placing image capture device 122 on the driver side of the mirror may further aid in obtaining images representative of the driver's field of view and/or line of sight.

[0022] Other locations for the image capture devices of image acquisition unit 120 may also be used. For example, image capture device 124 may be located on or in a bumper of vehicle 200. Such a location may be especially suitable for image capture devices having a wide field of view. The line of sight of bumper-located image capture devices can be different from that of the driver and, therefore, the bumper image capture device and driver may not always see the same objects. The image capture devices (e.g., image capture devices 122, 124, and 126) may also be located in other locations. For example, the image capture devices may be located on or in one or both of the side minors of vehicle 200, on the roof of vehicle 200, on the hood of vehicle 200, on the trunk of vehicle 200, on the sides of vehicle 200, mounted on, positioned behind, or positioned in front of any of the windows of vehicle 200, and mounted in or near light figures on the front and/or back of vehicle 200, etc.

[0023] In addition to image capture devices, vehicle 200 may include various other components of system 100. For example, processing unit 110 may be included on vehicle 200 either integrated with or separate from an engine control unit (ECU) of the vehicle. Vehicle 200 may also be equipped with a position sensor 130, such as a GPS receiver and may also include a map database 160 and memory units 140 and 150.

[0024] FIG. 2A is a diagrammatic side view representation of an exemplary vehicle imaging system consistent with the disclosed embodiments. FIG. 2B is a diagrammatic top view illustration of the embodiment shown in FIG. 2A . As illustrated in FIG. 2B , the disclosed embodiments may include a vehicle 200 including in its body a system. 100 with a first image capture device 122 positioned in the vicinity of the rearview mirror and/or near the driver of vehicle 200, a second image capture device 124 positioned on or in a bumper region (e.g., one of bumper regions 210) of vehicle 200, and a

Tesla EX1002 Page 1271 IPR2025-00944

processing unit 110.

[0025] As illustrated in FIG. 2C , image capture devices 122 and 124 may both be positioned in the vicinity of the rearview mirror and/or near the driver of vehicle 200. Additionally, while two image capture devices 122 and 124 are shown in FIGS. 2B and 2C , it should be understood that other embodiments may include more than two image capture devices. For example, in the embodiments shown in FIGS. 2D and 2E , first, second, and third image capture devices 122, 124, and 126, are included in the system 100 of vehicle 200.

[0026] As illustrated in FIG. 2D , image capture device 122 may be positioned in the vicinity of the rearview mirror and/or near the driver of vehicle 200, and image capture devices 124 and 126 may be positioned on or in a bumper region (e.g., one of bumper regions 210) of vehicle 200. And as shown in FIG. 2E , image capture devices 122, 124, and 126 may be positioned in the vicinity of the rearview mirror and/or near the driver seat of vehicle 200. The disclosed embodiments are not limited to any particular number and configuration of the image capture devices, and the image capture devices may be positioned in any appropriate location within and/or on vehicle 200.

[0027] It is to be understood that the disclosed embodiments are not limited to vehicles and could be applied in other contexts. It is also to be understood that disclosed embodiments are not limited to a particular type of vehicle 200 and may be applicable to all types of vehicles including automobiles, trucks, trailers, and other types of vehicles.

[0028] The first image capture device 122 may include any suitable type of image capture device. Image capture device 122 may include an optical axis. In one instance, the image capture device 122 may include an Aptina M9V024 WVGA sensor with a global shutter. In other embodiments, image capture device 122 may provide a resolution of 1280x960 pixels and may include a rolling shutter. Image capture device 122 may include various optical elements. In some embodiments one or more lenses may be included, for example, to provide a desired focal length and field of view for the image capture device. In some embodiments, image capture device 122 may be associated with a 6mm lens or a 12mm lens. In some embodiments, image capture device 122 may be configured to capture images having a desired field-of-view (FOV) 202, as illustrated in FIG. 2D . For example, image capture device 122 may be configured to have a regular FOV, such as within a range of 40 degrees to 56 degrees, including a 46 degree FOV, 50 degree FOV, 52 degree FOV, or greater. Alternatively, image capture device 122 may be configured to have a narrow FOV in the range of 23 to 40 degrees, such as a 28 degree FOV or 36 degree FOV. In addition, image capture device 122 may be configured to have a wide FOV in the range of 100 to 180 degrees. In some embodiments, image capture device 122 may include a wide angle bumper camera or one with up to a 180 degree FOV.

[0029] The first image capture device 122 may acquire a plurality of first images relative to a scene associated with the vehicle 200. Each of the plurality of first images may be acquired as a series of image scan lines, which may be captured using a rolling shutter. Each scan line may include a plurality of pixels.

[0030] The first image capture device 122 may have a scan rate associated with acquisition of each of the first series of image scan lines. The scan rate may refer to a rate at which an image sensor can acquire image data associated with each pixel

Tesla EX1002 Page 1272 IPR2025-00944

included in a particular scan line.

[0031] Image capture devices 122, 124, and 126 may contain any suitable type and number of image sensors, including CCD sensors or CMOS sensors, for example. In one embodiment, a CMOS image sensor may be employed along with a rolling shutter, such that each pixel in a row is read one at a time, and scanning of the rows proceeds on a row-by-row basis until an entire image frame has been captured. In some embodiments, the rows may be captured sequentially from top to bottom relative to the frame.

[0032] The use of a rolling shutter may result in pixels in different rows being exposed and captured at different times, which may cause skew and other image artifacts in the captured image frame. On the other hand, when the image capture device 122 is configured to operate with a global or synchronous shutter, all of the pixels may be exposed for the same amount of time and during a common exposure period. As a result, the image data in a frame collected from a system employing a global shutter represents a snapshot of the entire FOV (such as FOV 202) at a particular time. In contrast, in a rolling shutter application, each row in a frame is exposed and data is capture at different times. Thus, moving objects may appear distorted in an image capture device having a rolling shutter. This phenomenon will be described in greater detail below.

[0033] The second image capture device 124 and the third image capturing device 126 may be any type of image capture device. Like the first image capture device 122, each of image capture devices 124 and 126 may include an optical axis. In one embodiment, each of image capture devices 124 and 126 may include an Aptina M9V024 WVGA sensor with a global shutter. Alternatively, each of image capture devices 124 and 126 may include a rolling shutter. Like image capture device 122, image capture devices 1.24 and 126 may be configured to include various lenses and optical elements. In some embodiments, lenses associated with image capture devices 124 and 126 may provide FOVs (such as FOVs 204 and 206) that are the same as, or narrower than, a FOV (such as FOV 202) associated with image capture device 122. For example, image capture devices 124 and 126 may have FOVs of 40 degrees, 30 degrees, 26 degrees, 23 degrees, 20 degrees, or less.

[0034] Image capture devices 124 and 126 may acquire a plurality of second and third images relative to a scene associated with the vehicle 200. Each of the plurality of second and third images may be acquired as a second and third series of image scan lines, which may be captured using a rolling shutter. Each scan line or row may have a plurality of pixels. Image capture devices 124 and 126 may have second and third scan rates associated with acquisition of each of image scan lines included in the second and third series.

[0035] Each image capture device 122, 124, and 126 may be positioned at any suitable position and orientation relative to vehicle 200. The relative positioning of the image capture devices 122, 124, and 126 may be selected to aid in fusing together the information acquired from the image capture devices. For example, in some embodiments, a FOV (such as FOV 204) associated with image capture device 124 may overlap partially or fully with a FOV (such as FOV 202) associated with image capture device 122 and a FOV (such as FOV 206) associated with image capture device 126.

Tesla EX1002 Page 1273 IPR2025-00944

[0036] Image capture devices 122, 124, and 126 may be located on vehicle 200 at any suitable relative heights. In one instance, there may be a height difference between the image capture devices 122, 124, and 126, which may provide sufficient parallax information to enable stereo analysis. For example, as shown in FIG. 2A , the two image capture devices 122 and 124 are at different heights. There may also be a lateral displacement difference between image capture devices 122, 124, and 126, giving additional parallax information for stereo analysis by processing unit 110, for example. The difference in the lateral displacement may be denoted by dx, as shown in FIGS. 2C and 2D . In some embodiments, fore or aft displacement (e.g., range displacement) may exist between image capture devices 122, 124, and 126. For example, image capture device 122 may be located 0.5 to 2 meters or more behind image capture device 124 and/or image capture device 126. This type of displacement may enable one of the image capture devices to cover potential blind spots of the other image capture device(s).

[0037] Image capture devices 122 may have any suitable resolution capability (e.g., number of pixels associated with the image sensor), and the resolution of the image sensor(s) associated with the image capture device 122 may be higher, lower, or the same as the resolution of the image sensor(s) associated with image capture devices 124 and 126. In some embodiments, the image sensor(s) associated with image capture device 122 and/or image capture devices 124 and 126 may have a resolution of 640 x 480, 1024 x 768, 1280 x 960, or any other suitable resolution.

[0038] The frame rate (e.g., the rate at which an image capture device acquires a set of pixel data of one image frame before moving on to capture pixel data associated with the next image frame) may be controllable. The frame rate associated with image capture device 122 may be higher, lower, or the same as the frame rate associated with image capture devices 124 and 126. The frame rate associated with image capture devices 122, 124, and 126 may depend on a variety of factors that may affect the timing of the frame rate. For example, one or more of image capture devices 122, 124, and 126 may include a selectable pixel delay period imposed before or after acquisition of image data associated with one or more pixels of an image sensor in image capture device 122, 124, and/or 126. Generally, image data corresponding to each pixel may be acquired according to a clock rate for the device (e.g., one pixel per clock cycle). Additionally, in embodiments including a rolling shutter, one or more of image capture devices 122, 124, and 126 may include a selectable horizontal blanking period imposed before or after acquisition of image data associated with a row of pixels of an image sensor in image capture device 122, 124, and/or 126. Further, one or more of image capture devices 122, 124, and/or 126 may include a selectable vertical blanking period imposed before or after acquisition of image data associated with an image frame of image capture device 122, 124, and 126.

[0039] These timing controls may enable synchronization of frame rates associated with image capture devices 122, 124, and 126, even where the line scan rates of each are different. Additionally, as will be discussed in greater detail below, these selectable timing controls, among other factors (e.g., image sensor resolution, maximum line scan rates, etc.) may enable synchronization of image capture from an area where the FOV of image capture device 122 overlaps with one or more FOVs of image capture devices 124 and 126, even where the field of view of image capture device 122 is different from the FOVs of image capture devices 124 and 126.

[0040] Frame rate timing in image capture device 122, 124, and 126 may depend on

IPR2025-00944
Tesla EX1002 Page 1274

the resolution of the associated image sensors. For example, assuming similar line scan rates for both devices, if one device includes an image sensor having a resolution of 640 x 480 and another device includes an image sensor with a resolution of 1280 x 960, then more time will be req5ired to acquire a frame of image data from the sensor having the higher resolution.

[0041] Another factor that may affect the timing of image data acquisition in image capture devices 122, 124, and 126 is the maximum line scan rate. For example, acquisition of a row of image data from an image sensor included in image capture device 122, 124, and 126 will require some minimum amount of time. Assuming no pixel delay periods are added, this minimum amount of time for acquisition of a row of image data will be related to the maximum line scan rate for a particular device. Devices that offer higher maximum line scan rates have the potential to provide higher frame rates than devices with lower maximum line scan rates. In some embodiments, one or more of image capture devices 124 and 126 may have a maximum line scan rate that is higher than a maximum line scan rate associated with image capture device 122. In some embodiments, the maximum line scan rate of image capture device 124 and/or 126 may be 1.25, 1.5, 1.75, or 2 times or more than a maximum line scan rate of image capture device 122.

[0042] In another embodiment, image capture devices 122, 124, and 126 may have the same maximum line scan rate, but image capture device 122 may be operated at a scan rate less than or equal to its maximum scan rate. The system may be configured such that one or more of image capture devices 124 and 126 operate at a line scan rate that is equal to the line scan rate of image capture device 122. In other instances, the system may be configured such that the line scan rate of image capture device 124 and/or image capture device 126 may be 125, 1.5, 1.75, or 2 times or more than the line scan rate of image capture device 122.

[0043] In some embodiments, image capture devices 122, 124, and 126 may be asymmetric. That is, they may include cameras having different fields of view (FOV) and focal lengths. The fields of view of image capture devices 122, 124, and 126 may include any desired area relative to an environment of vehicle 200, for example. In some embodiments, one or more of image capture devices 122, 124, and 126 may be configured to acquire image data from an environment in front of vehicle 200, behind vehicle 200, to th%sides of vehicle 200, or combinations thereof.

[0044] Further, the focal length associated with each image capture device 122, 124, and/or 126 may be selectable (e.g., by inclusion of appropriate lenses etc.) such that each device acquires images of objects at a desired distance range relative to vehicle 200. For example, in some embodiments image capture devices 122,124, and 126 may acquire images of close-up objects within a few meters from the vehicle. Image capture devices 122, 124, and 126 may also be configured to acquire images of objects at ranges more distant from the vehicle (e.g., 25 m, 50 m, 100 m, 150 m, or more). Further, the focal lengths of image capture devices 122, 124, and 126 may be selected such that one image capture device (e.g., image capture device 122) can acquire images of objects relatively close to the vehicle (e.g., within 10 m or within 20 m) while the other image capture devices (e.g., image capture devices 124 and 126) can acquire images of more distant objects (e.g., greater than 20 m, 50 m, 100 m, 150 m, etc.) from vehicle 200.

[0045] According to some embodiments, the FOV of one or more image capture

Tesla EX1002 Page 1275 IPR2025-00944

devices 122, 124, and 126 may have a wide angle. For example, it may be advantageous to have a FOV of 140 degrees, especially for image capture devices 122, 124, and 126 that may be used to capture images of the area in the vicinity of vehicle 200. For example, image capture device 122 may be used to capture images of the area to the right or left of vehicle 200 and, in such embodiments, it may be desirable for image capture device 122 to have a wide FOV (e.g., at least 140 degrees).

[0046] The field of view associated with each of image capture devices 122, 124, and 126 may depend on the respective focal lengths. For example, as the focal length increases, the corresponding field of view decreases.

[0047] Image capture devices 122, 124, and 126 may be configured to have any suitable fields of view. In one particular example, image capture device 122 may have a horizontal FOV of 46 degrees, image capture device 124 may have a horizontal FOV of 23 degrees, and image capture device 126 may have a horizontal FOV in between 23 and 46 degrees. In another instance, image capture device 122 may have a horizontal FOV of 52 degrees, image capture device 124 may have a horizontal FOV of 26 degrees, and image capture device 126 may have a horizontal FOV in between 26 and 52 degrees. In some embodiments, a ratio of the FOV of image capture device 122 to the FOVs of image capture device 124 and/or image capture device 126 may vary from 1.5 to 2.0. In other embodiments, this ratio may vary between 1.25 and 2.25.

[0048] System 100 may be configured so that a field of view of image capture device 122 overlaps, at least partially or fully, with a field of view of image capture device 124 and/or image capture device 126. In some embodiments, system 100 may be configured such that the fields of view of image capture devices 124 and 126, for example, fall within (e.g., are narrower than) and share a common center with the field of view of image capture device 122. In other embodiments, the image capture devices 122, 124, and 126 may capture adjacent FOVs or may have partial overlap in their FOVs. In some embodiments, the fields of view of image capture devices 122, 124, and 126. may be aligned such that a center of the narrower FOV image capture devices 124 and/or 126 may be located in a lower half of the field of view of the wider FOV device 122.

[0049] FIG. 2F is a diagrammatic representation of exemplary vehicle control systems, consistent with the disclosed embodiments. As indicated in FIG. 2F , vehicle 200 may include throttling system 220, braking system 230, and steering system 240. System 100 may provide inputs (e.g., control signals) to one or more of throttling system 220, braking system 230, and steering system 240 over one or more data links (e.g., any wired and/or wireless link or links for transmitting data). For example, based on analysis of images acquired by image capture devices 122, 124, and/or 126, system 100 may provide control signals to one or more of throttling system 220, braking system 230, and steering system 240 to navigate vehicle 200 (e.g., by causing an acceleration, a turn, a lane shift, etc.). Further, system 100 may receive inputs from one or more of throttling system 220, braking system 230, and steering system 24 indicating operating conditions of vehicle 200 (e.g., speed, whether vehicle 200 is braking and/or turning, etc.). Further details are provided in connection with FIGS. 4-7 , below.

[0050] As shown in FIG. 3A , vehicle 200 may also include a user interface 170 for interacting with a driver or a passenger of vehicle 200. For example, user interface 170 in a vehicle application may include a touch screen 320, knobs 330, buttons 340, and a

Tesla EX1002 Page 1276 IPR2025-00944

microphone 350. A driver or passenger of vehicle 200 may also use handles (e.g., located on or near the steering column of vehicle 200 including, for example, turn signal handles), buttons (e.g., located on the steering wheel of vehicle 200), and the like, to interact with system 100. In some embodiments, microphone 350 may be positioned adjacent to a rearview mirror 310. Similarly, in some embodiments, image capture device 122 may be located near rearview mirror 310. In some embodiments, user interface 170 may also include one or more speakers 360 (e.g., speakers of a vehicle audio system). For example, system 100 may provide various notifications (e.g., alerts) via speakers 360.

[0051] FIGS. 3B-3D are illustrations of an exemplary camera mount 370 configured to be positioned behind a rearview mirror (e.g., rearview mirror 310) and against a vehicle windshield, consistent with disclosed embodiments. As shown in FIG. 3B , camera mount 370 may include image capture devices 122, 124, and 126. Image capture devices 124 and 126 may be positioned behind a glare shield 380, which may be flush against the vehicle windshield and include a composition of film and/or anti-reflective materials. For example, glare shield 380 may be positioned such that it aligns against a vehicle windshield having a matching slope. In some embodiments, each of image capture devices 122, 124, and 126 may be positioned behind glare shield 380, as depicted, for example, in FIG. 3D . The disclosed embodiments are not limited to any particular configuration of image capture devices 122, 124, and 126, camera mount 370, and glare shield 380. FIG. 3C is an illustration of camera mount 370 shown in FIG. 3B from a front perspective.

[0052] As will be appreciated by a person skilled in the art having the benefit of this disclosure, numerous variations and/or modifications may be made to the foregoing disclosed embodiments. For example, not all components are essential for the operation of system 100. Further, any component may be located in any appropriate part of system 100 and the components may be rearranged into a variety of configurations while providing the functionality of the disclosed embodiments. Therefore, the foregoing configurations are examples and, regardless of the configurations discussed above, system 100 can provide a wide range of functionality to analyze the surroundings of vehicle 200 and navigate vehicle 200 in response to the analysis.

[0053] As discussed below in further detail and consistent with various disclosed embodiments, system 100 may provide a variety of features related to autonomous driving and/or driver assist technology. For example, system 100 may analyze image data, position data (e.g., GPS location information), map data, speed data, and/or data from sensors included in vehicle 200. System 100 may collect the data for analysis from, for example, image acquisition unit 120, position sensor 130, and other sensors. Further, system 100 may analyze the collected data to determine whether or not vehicle 200 should take a certain action, and then automatically take the determined action without human intervention. For example, when vehicle 200 navigates without human intervention, system 100 may automatically control the braking, acceleration, and/or steering of vehicle 200 (e.g., by sending control signals to one or more of throttling system 220, braking system 230, and steering system 240). Further, system 100 may analyze the collected data and issue warnings and/or alerts to vehicle occupants based on the analysis of the collected data. Additional details regarding the various embodiments that are provided by system 100 are provided below.

Forward-Facing Multi-Imaging System

Tesla EX1002 Page 1277 IPR2025-00944

[0054] As discussed above, system 100 may provide drive assist functionality that uses a multi-camera system. The multi-camera system may use one or more cameras facing in the forward direction of a vehicle. In other embodiments, the multi-camera system may include one or more cameras facing to the side of a vehicle or to the rear of the vehicle. In one embodiment, for example, system 100 may use a two-camera imaging system, where a first camera and a second camera (e.g., image capture devices 122 and 124) may be positioned at the front and/or the sides of a vehicle (e.g., vehicle 200). The first camera may have a field of view that is greater than, less than, or partially overlapping with, the field of view of the second camera. In addition, the first camera may be connected to a first image processor to perform monocular image analysis of images provided by the first camera, and the second camera may be connected to a second image processor to perform monocular image analysis of images provided by the second camera. The outputs (e.g., processed information) of the first and second image processors may be combined. In some embodiments, the second image processor may receive images from both the first camera and second camera to perform stereo analysis. In another embodiment, system 100 may use a three-camera imaging system where each of the cameras has a different field of view. Such a system may, therefore, make decisions based on information derived from objects located at varying distances both forward and to the sides of the vehicle. References to monocular image analysis may refer to instances where image analysis is performed based on images captured from a single point of view (e.g., from a single camera). Stereo image analysis may refer to instances where image analysis is performed based on two or more images captured with one or more variations of an image capture parameter. For example, captured images suitable for performing stereo image analysis may include images captured: from two or more different positions, from different fields of view, using different focal lengths, along with parallax information, etc.

[0055] For example, in one embodiment, system 100 may implement a three camera configuration using image capture devices 122-126. In such a configuration, image capture device 122 may provide a narrow field of view (e.g., 34 degrees, or other values selected from a range of about 20 to 45 degrees, etc.), image capture device 124 may provide a wide field of view (e.g., 150 degrees or other values selected from a range of about 100 to about 180 degrees), and image capture device 126 may provide an intermediate field of view (e.g., 46 degrees or other values selected from a range of about 35 to about 60 degrees). In some embodiments, image capture device 126 may act as a main or primary camera. Image capture devices 122-126 may be positioned behind rearview mirror 310 and positioned substantially side-by-side (e.g., 6 cm apart). Further, in some embodiments, as discussed above, one or more of image capture devices 122-126 may be mounted behind glare shield 380 that is flush with the windshield of vehicle 200. Such shielding may act to minimize the impact of any reflections from inside the car on image capture devices 122-126.

[0056] In another embodiment, as discussed above in connection with FIGS. 3B and 3C , the wide field of view camera (e.g., image capture device 124 in the above example) may be mounted lower than the narrow and main field of view cameras (e.g., image devices 122 and 126 in the above example). This configuration may provide a free line of sight from the wide field of view camera. To reduce reflections, the cameras may be mounted close to the windshield of vehicle 200, and may include polarizers on the cameras to damp reflected light.

[0057] A three camera system may provide certain performance characteristics. For

IPR2025-00944
Tesla EX1002 Page 1278

example, some embodiments may include an ability to validate the detection of objects by one camera based on detection results from another camera. In the three camera configuration discussed above, processing unit 110 may include, for example, three processing devices (e.g., three EyeQ series of processor chips, as discussed above), with each processing device dedicated to processing images captured by one or more of image capture devices 122-126.

[0058] In a three camera system, a first processing device may receive images from both the main camera and the narrow field of view camera, and perform vision processing of the narrow FOV camera to, for example, detect other vehicles, pedestrians, lane marks, traffic signs, traffic lights, and other road objects. Further, the first processing device may calculate a disparity of pixels between the images from the main camera and the narrow camera and create a 3D reconstruction of the environment of vehicle 200. The first processing device may then combine the 3D reconstruction with 3D map data or with 3D information calculated based on information from another camera.

[0059] The second processing device may receive images from main camera and perform vision processing to detect other vehicles, pedestrians, lane marks, traffic signs, traffic lights, and other road objects. Additionally, the second processing device may calculate a camera displacement and, based on the displacement, calculate a disparity of pixels between successive images and create a 3D reconstruction of the scene (e.g., a structure from motion). The second processing device may send the structure from motion based 3D reconstruction to the first processing device to be combined with the stereo 3D images.

[0060] The third processing device may receive images from the wide FOV camera and process the images to detect vehicles, pedestrians, lane marks, traffic signs, traffic lights, and other road objects. The third processing device may further execute additional processing instructions to analyze images to identify objects moving in the image, such as vehicles changing lanes, pedestrians, etc.

[0061] In some embodiments, having streams of image-based information captured and processed independently may provide an opportunity for providing redundancy in the system. Such redundancy may include, for example, using a first image capture device and the images processed from that device to validate and/or supplement information obtained by capturing and processing image information from at least a second image capture device.

[0062] In some embodiments, system 100 may use two image capture devices (e.g., image capture devices 122 and 124) in providing navigation assistance for vehicle 200 and use a third image capture device (e.g., image capture device 126) to provide redundancy and validate the analysis of data received from the other two image capture devices. For example, in such a configuration, image capture devices 122 and 124 may provide images for stereo analysis by system 100 for navigating vehicle 200, while image capture device 126 may provide images for monocular analysis by system 100 to provide redundancy and validation of information obtained based on images captured from image capture device 122 and/or image capture device 124. That is, image capture device 126 (and a corresponding processing device) may be considered to provide a redundant sub-system for providing a check on the analysis derived from image capture devices 122 and 124 (e.g., to provide an automatic emergency braking (AEB) system).

IPR2025-00944
Tesla EX1002 Page 1279

[0063] One of skill in the art will recognize that the above camera configurations, camera placements, number of cameras, camera locations, etc., are examples only. These components and others described relative to the overall system may be assembled and used in a variety of different configurations without departing from the scope of the disclosed embodiments. Further details regarding usage of a multi-camera system to provide driver assist and/or autonomous vehicle functionality follow below.

[0064] FIG. 4 is an exemplary functional block diagram of memory 140 and/or 150, which may be stored/programmed with instructions for performing one or more operations consistent with the disclosed embodiments. Although the following refers to memory 140, one of skill in the art will recognize that instructions may be stored in memory 140 and/or 150.

[0065] As shown in FIG. 4 , memory 140 may store a monocular image analysis module 402, a stereo image analysis module 404, a velocity and acceleration module 406, and a navigational response module 408. The disclosed embodiments are not limited to any particular configuration of memory 140. Further, application processor 180 and/or image processor 190 may execute the instructions stored in any of modules 402-408 included in memory 140. One of skill in the art will understand that references in the following discussions to processing unit 110 may refer to application processor 180 and image processor 190 individually or collectively. Accordingly, steps of any of the following processes may be performed by one or more processing devices.

[0066] In one embodiment, monocular image analysis module 402 may store instructions (such as computer vision software) which, when executed by processing unit 110, performs monocular image analysis of a set of images acquired by one of image capture devices 122, 124, and 126. In some embodiments, processing unit 110 may combine information from a set of images with additional sensory information (e.g., information from radar) to perform the monocular image analysis. As described in connection with FIGS. 5A-5D below, monocular image analysis module 402 may include instructions for detecting a set of features within the set of images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, hazardous objects, and any other feature associated with an environment of a vehicle. Based on the analysis, system 100 (e.g., via processing unit 110) may cause one or more navigational responses in vehicle 200, such as a turn, a lane shift, a change in acceleration, and the like, as discussed below in connection with navigational response module 408.

[0067] In one embodiment, stereo image analysis module 404 may store instructions (such as computer vision software) which, when executed by processing unit 110, performs stereo image analysis of first and second sets of images acquired by a combination of image capture devices selected from any of image capture devices 122,124, and 126. In some embodiments, processing unit 110 may combine information from the first and second sets of images with additional sensory information (e.g., information from radar) to perform the stereo image analysis. For example, stereo image analysis module 404 may include instructions for performing stereo image analysis based on a first set of images acquired by image capture device 124 and a second set of images acquired by image capture device 126. As described in connection with FIG. 6 below, stereo image analysis module 404 may include instructions for detecting a set of features within the first and second sets of images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic

Tesla EX1002 Page 1280 IPR2025-00944

lights, hazardous objects, and the like. Based on the analysis, processing unit 110 may cause one or more navigational responses in vehicle 200, such as a turn, a lane shift, a change in acceleration, and the like, as discussed below in connection with navigational response module 408.

[0068] In one embodiment, velocity and acceleration module 406 may store software configured to analyze data received from one or more computing and electromechanical devices in vehicle 200 that are configured to cause a change in velocity and/or acceleration of vehicle 200. For example, processing unit 110 may execute instructions associated with velocity and acceleration module 406 to calculate a target speed for vehicle 200 based on data derived from execution of monocular image analysis module 402 and/or stereo image analysis module 404. Such data may include, for example, a target position, velocity, and/or acceleration, the position and/or speed of vehicle 200 relative to a nearby vehicle, pedestrian, or road object, position information for vehicle 200 relative to lane markings of the road, and the like. In addition, processing unit 110 may calculate a target speed for vehicle 200 based on sensory input (e.g., information from radar) and input from other systems of vehicle 200, such as throttling system 220, braking system 230, and/or steering system 240 of vehicle 200. Based on the calculated target speed, processing unit 110 may transmit electronic signals to throttling system 220, braking system 230, and/or steering system 240 of vehicle 200 to trigger a change in velocity and/or acceleration by, for example, physically depressing the brake or easing up off the accelerator of vehicle 200.

[0069] In one embodiment, navigational response module 408 may store software executable by processing unit 110 to determine a desired navigational response based on data derived from execution of monocular image analysis module 402 and/or stereo image analysis module 404. Such data may include position and speed information associated with nearby vehicles, pedestrians, and road objects, target position information for vehicle 200, and the like. Additionally, in some embodiments, the navigational response may be based (partially or fully) on map data, a predetermined position of vehicle 200, and/or a relative velocity or a relative acceleration between vehicle 200 and one or more objects detected from execution of monocular image analysis module 402 and/or stereo image analysis module 404. Navigational response module 408 may also determine a desired navigational response based on sensory input (e.g., information from radar) and inputs from other systems of vehicle 200, such as throttling system 220, braking system 230, and steering system 240 of vehicle 200. Based on the desired navigational response, processing unit 110 may transmit electronic signals to throttling system 220, braking system 230, and steering system 240 of vehicle 200 to trigger a desired navigational response by, for example, turning the steering wheel of vehicle 200 to achieve a rotation of a predetermined angle. In some embodiments, processing unit 110 may use the output of navigational response module 408 (e.g., the desired navigational response) as an input to execution of velocity and acceleration module 406 for calculating a change in speed of vehicle 200.

[0070] FIG. 5A is a flowchart showing an exemplary process 500A for causing one or more navigational responses based on monocular image analysis, consistent with disclosed embodiments. At step 510, processing unit 110 may receive a plurality of images via data interface 128 between processing unit 110 and image acquisition unit 120. For instance, a camera included in image acquisition unit 120 (such as image capture device 122 having field of view 202) may capture a plurality of images of an area forward of vehicle 200 (or to the sides or rear of a vehicle, for example) and transmit them over a data connection (e.g., digital, wired, USB, wireless, Bluetooth,

IPR2025-00944
Tesla EX1002 Page 1281

etc.) to processing unit 110. Processing unit 110 may execute monocular image analysis module 402 to analyze the plurality of images at step 520, as described in further detail in connection with FIGS. 5B-5D below. By performing the analysis, processing unit 110 may detect a set of features within the set of images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, and the like.

[0071] Processing unit 110 may also execute monocular image analysis module 402 to detect various road hazards at step 520, such as, for example, parts of a truck tire, fallen road signs, loose cargo, small animals, and the like. Road hazards may vary in structure, shape, size, and color, which may make detection of such hazards more challenging. In some embodiments, processing unit 110 may execute monocular image analysis module 402 to perform multi-frame analysis on the plurality of images to detect road hazards. For example, processing unit 110 may estimate camera motion between consecutive image frames and calculate the disparities in pixels between the frames to construct a 3D-map of the road. Processing unit 110 may then use the 3D-map to detect the road surface, as well as hazards existing above the road surface.

[0072] At step 530, processing unit 110 may execute navigational response module 408 to cause one or more navigational responses in vehicle 200 based on the analysis performed at step 520 and the techniques as described above in connection with FIG. 4 . Navigational responses may include, for example, a turn, a lane shift, a change in acceleration, and the like. In some embodiments, processing unit 110 may use data derived from execution of velocity and acceleration module 406 to cause the one or more navigational responses. Additionally, multiple navigational responses may occur simultaneously, in sequence, or any combination thereof. For instance, processing unit 110 may cause vehicle 200 to shift one lane over and then accelerate by, for example, sequentially transmitting control signals to steering system 240 and throttling system 220 of vehicle 200. Alternatively, processing unit 110 may cause vehicle 200 to brake while at the same time shifting lanes by, for example, simultaneously transmitting control signals to braking system 230 and steering system 240 of vehicle 200.

[0073] FIG. 5B is a flowchart showing an exemplary process 500B for detecting one or more vehicles and/or pedestrians in a set of images, consistent with disclosed embodiments. Processing unit 110 may execute monocular image analysis module 402 to implement process 500B. At step 540, processing unit 110 may determine a set of candidate objects representing possible vehicles and/or pedestrians. For example, processing unit 110 may scan one or more images, compare the images to one or more predetermined patterns, and identify within each image possible locations that may contain objects of interest (e.g., vehicles, pedestrians, or portions thereof). The predetermined patterns may be designed in such a way to achieve a high rate of "false hits" and a low rate of "misses." For example, processing unit 110 may use a low threshold of similarity to predetermined patterns for identifying candidate objects as possible vehicles or pedestrians. Doing so may allow processing unit 110 to reduce the probability of missing (e.g., not identifying) a candidate object representing a vehicle or pedestrian.

[0074] At step 542, processing unit 110 may filter the set of candidate objects to exclude certain candidates (e.g., irrelevant or less relevant objects) based on classification criteria. Such criteria may be derived from various properties associated with object types stored in a database (e.g., a database stored in memory 140). Properties may include object shape, dimensions, texture, position (e.g., relative to

IPR2025-00944
Tesla EX1002 Page 1282

vehicle 200), and the like. Thus, processing unit 110 may use one or more sets of criteria to reject false candidates from the set of candidate objects.

[0075] At step 544, processing unit 110 may analyze multiple frames of images to determine whether objects in the set of candidate objects represent vehicles and/or pedestrians. For example, processing unit 110 may track a detected candidate object across consecutive frames and accumulate frame-by-frame data associated with the detected object (e.g., size, position relative to vehicle 200, etc.). Additionally, processing unit 110 may estimate parameters for the detected object and compare the object's frame-by-frame position data to a predicted position.

[0076] At step 546, processing unit 110 may construct a set of measurements for the detected objects. Such measurements may include, for example, position, velocity, and acceleration values (relative to vehicle 200) associated with the detected objects. In some embodiments, processing unit 110 may construct the measurements based on estimation techniques using a series of time-based observations such as Kalman filters or linear quadratic estimation (LQE), and/or based on available modeling data for different object types (e.g., cars, tracks, pedestrians, bicycles, road signs, etc.). The Kalman filters may be based on a measurement of an object's scale, where the scale measurement is proportional to a time to collision (e.g., the amount of time for vehicle 200 to reach the object). Thus, by performing steps 540-546, processing unit 110 may identify vehicles and pedestrians appearing within the set of captured images and derive information (e.g., position, speed, size) associated with the vehicles and pedestrians. Based on the identification and the derived information, processing unit 110 may cause one or more navigational responses in vehicle 200, as described in connection with FIG. 5A , above.

[0077] At step 548, processing unit 110 may perform an optical flow analysis of one or more images to reduce the probabilities of detecting a "false hit" and missing a candidate object that represents a vehicle or pedestrian. The optical flow analysis may refer to, for example, analyzing motion patterns relative to vehicle 200 in the one or more images associated with other vehicles and pedestrians, and that are distinct from road surface motion. Processing unit 110 may calculate the motion of candidate objects by observing the different positions of the objects across multiple image frames, which are captured at different times. Processing unit 110 may use the position and time values as inputs into mathematical models for calculating the motion of the candidate objects. Thus, optical flow analysis may provide another method of detecting vehicles and pedestrians that are nearby vehicle 200. Processing unit 110 may perform optical flow analysis in combination with steps 540-546 to provide redundancy for detecting vehicles and pedestrians and increase the reliability of system 100.

[0078] FIG. 5C is a flowchart showing an exemplary process 500C for detecting road marks and/or lane geometry information in a set of images, consistent with disclosed embodiments. Processing unit 110 may execute monocular image analysis module 402 to implement process 500C. At step 550, processing unit 110 may detect a set of objects by scanning one or more images. To detect segments of lane markings, lane geometry information, and other pertinent road marks, processing unit 110 may filter the set of objects to exclude those determined to be irrelevant (e.g., minor potholes, small rocks, etc.). At step 552, processing unit 110 may group together the segments detected in step 550 belonging to the same road mark or lane mark. Based on the grouping, processing unit 110 may develop a model to represent the detected segments, such as a mathematical model.

IPR2025-00944
Tesla EX1002 Page 1283

[0079] At step 554, processing unit 110 may construct a set of measurements associated with the detected segments. In some embodiments, processing unit 110 may create a projection of the detected segments from the image plane onto the real-world plane. The projection may be characterized using a 3rd-degree polynomial having coefficients corresponding to physical properties such as the position, slope, curvature, and curvature derivative of the detected road. In generating the projection, processing unit 110 may take into account changes in the road surface, as well as pitch and roll rates associated with vehicle 200. In addition, processing unit 110 may model the road elevation by analyzing position and motion cues present on the road surface. Further, processing unit 110 may estimate the pitch and roll rates associated with vehicle 200 by tracking a set of feature points in the one or more images.

[0080] At step 556, processing unit 110 may perform multi-frame analysis by, for example, tracking the detected segments across consecutive image frames and accumulating frame-by-frame data associated with detected segments. As processing unit 110 performs multi-frame analysis, the set of measurements constructed at step 554 may become more reliable and associated with an increasingly higher confidence level. Thus, by performing steps 550-556, processing unit 110 may identify road marks appearing within the set of captured images and derive lane geometry information. Based on the identification and the derived information, processing unit 110 may cause one or more navigational responses in vehicle 200, as described in connection with FIG. 5A , above.

[0081] At step 558, processing unit 110 may consider additional sources of information to further develop a safety model for vehicle 200 in the context of its surroundings. Processing unit 110 may use the safety model to define a context in which system 100 may execute autonomous control of vehicle 200 in a safe manner. To develop the safety model, in some embodiments, processing unit 110 may consider the position and motion of other vehicles, the detected road edges and barriers, and/or general road shape descriptions extracted from map data (such as data from map database 160). By considering additional sources of information, processing unit 110 may provide redundancy for detecting road marks and lane geometry and increase the reliability of system 100.

[0082] FIG. 5D is a flowchart showing an exemplary process 500D for detecting traffic lights in a set of images, consistent with disclosed embodiments. Processing unit 110 may execute monocular image analysis module 402 to implement process 500D. At step 560, processing unit 110 may scan the set of images and identify objects appearing at locations in the images likely to contain traffic lights. For example, processing unit 110 may filter the identified objects to construct a set of candidate objects, excluding those objects unlikely to correspond to traffic lights. The filtering may be done based on various properties associated with traffic lights, such as shape, dimensions, texture, position (e.g., relative to vehicle 200), and the like. Such properties may be based on multiple examples of traffic lights and traffic control signals and stored in a database. In some embodiments, processing unit 110 may perform multi-frame analysis on the set of candidate objects reflecting possible traffic lights. For example, processing unit 110 may track the candidate objects across consecutive image frames, estimate the real-world position of the candidate objects, and filter out those objects that are moving (which are unlikely to be traffic lights). In some embodiments, processing unit 110 may perform color analysis on the candidate objects and identify the relative position of the detected colors appearing inside possible traffic

Tesla EX1002 Page 1284 IPR2025-00944

lights.

[0083] At step 562, processing unit 110 may analyze the geometry of a junction. The analysis may be based on any combination of: (i) the number of lanes detected on either side of vehicle 200, (ii) markings (such as arrow marks) detected on the road, and (iii) descriptions of the junction extracted from map data (such as data from map database 160). Processing unit 110 may conduct the analysis using information derived from execution of monocular analysis module 402. In addition, Processing unit 110 may determine a correspondence between the traffic lights detected at step 560 and the lanes appearing near vehicle 200.

[0084] As vehicle 200 approaches the junction, at step 564, processing unit 110 may update the confidence level associated with the analyzed junction geometry and the detected traffic lights. For instance, the number of traffic lights estimated to appear at the junction as compared with the number actually appearing at the junction may impact the confidence level. Thus, based on the confidence level, processing unit 110 may delegate control to the driver of vehicle 200 in order to improve safety conditions. By performing steps 560-564, processing unit 110 may identify traffic lights appearing within the set of captured images and analyze junction geometry information. Based on the identification and the analysis, processing unit 110 may cause one or more navigational responses in vehicle 200, as described in connection with FIG. 5A , above.

[0085] FIG. 5E is a flowchart showing an exemplary process 500E for causing one or more navigational responses in vehicle 200 based on a vehicle path, consistent with the disclosed embodiments. At step 570, processing unit 110 may construct an initial vehicle path associated with vehicle 200. The vehicle path may be represented using a set of points expressed in coordinates (x, z), and the distance di between two points in the set of points may fall in the range of 1 to 5 meters. In one embodiment, processing unit 110 may construct the initial vehicle path using two polynomials, such as left and right road polynomials. Processing unit 110 may calculate the geometric midpoint between the two polynomials and offset each point included in the resultant vehicle path by a predetermined offset (e.g., a smart lane offset), if any (an offset of zero may correspond to travel in the middle of a lane). The offset may be in a direction perpendicular to a segment between any two points in the vehicle path. In another embodiment, processing unit 110 may use one polynomial and an estimated lane width to offset each point of the vehicle path by half the estimated lane width plus a predetermined offset (e.g., a smart lane offset).

[0086] At step 572, processing unit 110 may update the vehicle path constructed at step 570. Processing unit 110 may reconstruct the vehicle path constructed at step 570 using a higher resolution, such that the distance dk between two points in the set of points representing the vehicle path is less then the distance di described above. For example, the distance dk may fall in the range of 0.1 to 0.3 meters. Processing unit 110 may reconstruct the vehicle path using a parabolic spline algorithm, which may yield a cumulative distance vector S corresponding to the total length of the vehicle path (i.e., based on the set of points representing the vehicle path).

[0087] At step 574, processing unit 110 may determine a look-ahead point (expressed in coordinates as (xl, zi)) based on the updated vehicle path constructed at step 572. Processing unit 110 may extract the look-ahead point from the cumulative distance vector S, and the look-ahead point may be associated with a look-ahead distance and look-ahead time. The look-ahead distance, which may have a lower bound ranging

IPR2025-00944
Tesla EX1002 Page 1285

from 10 to 20 meters, may be calculated as the product of the speed of vehicle 200 and the look-ahead time. For example, as the speed of vehicle 200 decreases, the look-ahead distance may also decrease (e.g., until it reaches the lower bound). The look-ahead time, which may range from 0.5 to 1.5 seconds, may be inversely proportional to the gain of one or more control loops associated with causing a navigational response in vehicle 200, such as the heading error tracking control loop. For example, the gain of the heading error tracking control loop may depend on the bandwidth of a yaw rate loop, a steering actuator loop, car lateral dynamics, and the like. Thus, the higher the gain of the heading error tracking control loop, the lower the look-ahead time.

[0088] At step 576, processing unit 110 may determine a heading error and yaw rate command based on the look-ahead point determined at step 574. Processing unit 110 may determine the heading error by calculating the arctangent of the look-ahead point, e.g., arctan (xl/ zl). Processing unit 110 may determine the yaw rate command as the product of the heading error and a high-level control gain. The high-level control gain may be equal to: (2 / look-ahead time), if the look-ahead distance is not at the lower bound. Otherwise, the high-level control gain may be equal to: (2 * speed of vehicle 200 / look-ahead distance).

[0089] FIG. 5F is a flowchart showing an exemplary process 500F for determining whether a leading vehicle is changing lanes, consistent with the disclosed embodiments. At step 580, processing unit 110 may determine navigation information associated with a leading vehicle (e.g., a vehicle traveling ahead of vehicle 200). For example, processing unit 110 may determine the position, velocity (e.g., direction and speed), and/or acceleration of the leading vehicle, using the techniques described in connection with FIGS. 5A and 5B , above. Processing unit 110 may also determine one or more road polynomials, a look-ahead point (associated with vehicle 200), and/or a snail trail (e.g., a set of points describing a path taken by the leading vehicle), using the techniques described in connection with FIG. 5E , above.

[0090] At step 582, processing unit 110 may analyze the navigation information determined at step 580. In one embodiment, processing unit 110 may calculate the distance between a snail trail and a road polynomial (e.g., along the trail). If the variance of this distance along the trail exceeds a predetermined threshold (for example, 0.1 to 0.2 meters on a straight road, 0.3 to 0.4 meters on a moderately curvy road, and 0.5 to 0.6 meters on a road with sharp curves), processing unit 110 may determine that the leading vehicle is likely changing lanes. In the case where multiple vehicles are detected traveling ahead of vehicle 200, processing unit 110 may compare the snail trails associated with each vehicle. Based on the comparison, processing unit 110 may determine that a vehicle whose snail trail does not match with the snail trails of the other vehicles is likely changing lanes. Processing unit 110 may additionally compare the curvature of the snail trail (associated with the leading vehicle) with the expected curvature of the road segment in which the leading vehicle is traveling. The expected curvature may be extracted from map data (e.g., data from map database 160), from road polynomials, from other vehicles' snail trails, from prior knowledge about the road, and the like. If the difference in curvature of the snail trail and the expected curvature of the road segment exceeds a predetermined threshold, processing unit 110 may determine that the leading vehicle is likely changing lanes.

[0091] In another embodiment, processing unit 110 may compare the leading vehicle's instantaneous position with the look-ahead point (associated with vehicle 200) over a specific period of time (e.g., 0.5 to 1.5 seconds). If the distance between the leading

vehicle's instantaneous position and the look-ahead point varies during the specific period of time, and the cumulative sum of variation exceeds a predetermined threshold (for example, 0.3 to 0.4 meters on a straight road, 0.7 to 0.8 meters on a moderately curvy road, and 1.3 to 1.7 meters on a road with sharp curves), processing unit 110 may determine that the leading vehicle is likely changing lanes. In another embodiment, processing unit 110 may analyze the geometry of the snail trail by comparing the lateral distance traveled along the trail with the expected curvature of the snail trail. The expected radius of curvature may be determined according to the calculation: (δz<2> + δx<2>) / 2 / (δx), where δx represents the lateral distance traveled and δz represents the longitudinal distance traveled. If the difference between the lateral distance traveled and the expected curvature exceeds a predetermined threshold (e.g., 500 to 700 meters), processing unit 110 may determine that the leading vehicle is likely changing lanes. In another embodiment, processing unit 110 may analyze the position of the leading vehicle. If the position of the leading vehicle obscures a road polynomial (e.g., the leading vehicle is overlaid on top of the road polynomial), then processing unit 110 may determine that the leading vehicle is likely changing lanes. In the case where the position of the leading vehicle is such that, another vehicle is detected ahead of the leading vehicle and the snail trails of the two vehicles are not parallel, processing unit 110 may determine that the (closer) leading vehicle is likely changing lanes.

[0092] At step 584, processing unit 110 may determine whether or not leading vehicle 200 is changing lanes based on the analysis performed at step 582. For example, processing unit 110 may make the determination based on a weighted average of the individual analyses performed at step 582. Under such a scheme, for example, a decision by processing unit 110 that the leading vehicle is likely changing lanes based on a particular type of analysis may be assigned a value of "1" (and "0" to represent a determination that the leading vehicle is not likely changing lanes). Different analyses performed at step 582 may be assigned different weights, and the disclosed embodiments are not limited to any particular combination of analyses and weights.

[0093] FIG. 6 is a flowchart showing an exemplary process 600 for causing one or more navigational responses based on stereo image analysis, consistent with disclosed embodiments. At step 610, processing unit 110 may receive a first and second plurality of images via data interface 128. For example, cameras included in image acquisition unit 120 (such as image capture devices 122 and 124 having fields of view 202 and 204) may capture a first and second plurality of images of an area forward of vehicle 200 and transmit them over a digital connection (e.g., USB, wireless, Bluetooth, etc.) to processing unit 110. In some embodiments, processing unit 110 may receive the first and second plurality of images via two or more data interfaces. The disclosed embodiments are not limited to any particular data interface configurations or protocols.

[0094] At step 620, processing unit 110 may execute stereo image analysis module 404 to perform stereo image analysis of the first and second plurality of images to create a 3D map of the road in front of the vehicle and detect features within the images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, road hazards, and the like. Stereo image analysis may be performed in a manner similar to the steps described in connection with FIGS. 5A-5D , above. For example, processing unit 110 may execute stereo image analysis module 404 to detect candidate objects (e.g., vehicles, pedestrians, road marks, traffic lights, road hazards, etc.) within the first and second plurality of images, filter out a subset of the candidate objects based on various criteria, and perform multi-frame analysis, construct

IPR2025-00944
Tesla EX1002 Page 1287

measurements, and determine a confidence level for the remaining candidate objects. In performing the steps above, processing unit 110 may consider information from both the first and second plurality of images, rather than information from one set of images alone. For example, processing unit 110 may analyze the differences in pixel-level data (or other data subsets from among the two streams of captured images) for a candidate object appearing in both the first and second plurality of images. As another example, processing unit 110 may estimate a position and/or velocity of a candidate object (e.g., relative to vehicle 200) by observing that the object appears in one of the plurality of images but not the other or relative to other differences that may exist relative to objects appearing if the two image streams. For example, position, velocity, and/or acceleration relative to vehicle 200 may be determined based on trajectories, positions, movement characteristics, etc. of features associated with an object appearing in one or both of the image streams.

[0095] At step 630, processing unit 110 may execute navigational response module 408 to cause one or more navigational responses in vehicle 200 based on the analysis performed at step 620 and the techniques as described above in connection with FIG. 4 . Navigational responses may include, for example, a turn, a lane shift, a change in acceleration, a change in velocity, braking, and the like. In some embodiments, processing unit 110 may use data derived from execution of velocity and acceleration module 406 to cause the one or more navigational responses. Additionally, multiple navigational responses may occur simultaneously, in sequence, or any combination thereof.

[0096] FIG. 7 is a flowchart showing an exemplary process 700 for causing one or more navigational responses based on an analysis of three sets of images, consistent with disclosed embodiments. At step 710, processing unit 110 may receive a first, second, and third plurality of images via data interface 128. For instance, cameras included in image acquisition unit 120 (such as image capture devices 122, 124, and 126 having fields of view 202, 204, and 206) may capture a first, second, and third plu2ality of images of an area forward and/or to t(e side of vehicle 200 and transmit them over a digital connection (e.g., USB, wireless, Bluetooth, etc.) to processing unit 110. In some embodiments, processing unit 110 may receive the first, second, and third plurality of images via three or more data interfaces. For example, each of image capture devices 122, 124, 126 may have an associated data interface for communicating data to processing unit 110. The disclosed embodiments are not limited to any particular data interface configurations or protocols.

[0097] At step 720, processing unit 110 may analyze the first, second, and third plurality of images to detect features within the images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, road hazards, and the like. The analysis may be performed in a manner similar to the steps described in connection with FIGS. 5A-5D and 6 , above. For instance, processing unit 110 may perform monocular image analysis (e.g., via execution of monocular image analysis module 402 and based on the steps described in connection with FIGS. 5A-5D , above) on each of the first, second, and third plurality of images. Alternatively, processing unit 110 may perform stereo image analysis (e.g., via execution of stereo image analysis module 404 and based on the steps described in connection with FIG. 6 , above) on the first and second plurality of images, the second and third plurality of images, and/or the first and third plurality of images. The processed information corresponding to the analysis of the first, second, and/or third plurality of images may be combined. In some embodiments, processing uni t 110 may perform a combination of monocular and

Tesla EX1002 Page 1288 IPR2025-00944

stereo image analyses. For example, processing unit 110 may perform monocular image analysis (e.g., via execution of monocular image analysis module 402) on the first plurality of images and stereo image analysis (e.g., via execution of stereo image analysis module 404) on the second and third plurality of images. The configuration of image capture devices 122, 124, and 126-including their respective locations and fields of view 202, 204, and 206-may influence the types of analyses conducted on the first, second, and third plurality of images. The disclosed embodiments are not limited to a particular configuration of image capture devices 122, 124, and 126, or the types of analyses conducted on the first, second, and third plurality of images.

[0098] In some embodiments, processing unit 110 may perform testing on system 100 based on the images acquired and analyzed at steps 710 and 720. Such testing may provide an indicator of the overall performance of system 100 for certain configurations of image capture devices 122, 124, and 126. For example, processing unit 110 may determine the proportion of "false hits" (e.g., cases where system 100 incorrectly determined the presence of a vehicle or pedestrian) and "misses."

[0099] At step 730, processing unit 110 may cause one or more navigational responses in vehicle 200 based on information derived from two of the first, second, and third plurality of images. Selection of two of the first, second, and third plurality of images may depend on various factors, such as, for example, the number, types, and sizes of objects detected in each of the plurality of images. Processing unit 110 may also make the selection based on image quality and resolution, the effective field of view reflected in the images, the number of captured frames, the extent to which one or more objects of interest actually appear in the frames (e.g., the percentage of frames in which an object appears, the proportion of the object that appears in each such frame, etc.), and the like.

[0100] In some embodiments, processing unit 110 may select information derived from two of the first, second, and third plurality of images by determining the extent to which information derived from one image source is consistent with information derived from other image sources. For example, processing unit 110 may combine the processed information derived from each of image capture devices 122, 124, and 126 (whether by monocular analysis, stereo analysis, or any combination of the two) and determine visual indicators (e.g., lane markings, a detected vehicle and its location and/or path, a detected traffic light, etc.) that are consistent across the images captured from each of image capture devices 122, 124, and 126. Processing unit 110 may also exclude information that is inconsistent across the captured images (e.g., a vehicle changing lanes, a lane model indicating a vehicle that is too close to vehicle 200, etc.). Thus, processing unit 110 may select information derived from two of the first, second, and third plurality of images based on the determinations of consistent and inconsistent information.

[0101] Navigational responses may include, for example, a turn, a lane shift, a change in acceleration, and the like. Processing unit 110 may cause the one or more navigational responses based on the analysis performed at step 720 and the techniques as described above in connection with FIG. 4 . Processing unit 110 may also use data derived from execution of velocity and acceleration module 406 to cause the one or more navigational responses. In some embodiments, processing unit 110 may cause the one or more navigational responses based on a relative position, relative velocity, and/or relative acceleration between vehicle 200 and an object detected within any of the first, second, and third plurality of images. Multiple

Tesla EX1002 Page 1289 IPR2025-00944

navigational responses may occur simultaneously, in sequence, or any combination thereof.

Lane End Recognition System

[0102] System 100 may recognize when a driving lane will end and, when an end is detected, make a responsive navigational change in the course of vehicle 200, such as a lane change. A "lane" may refer to a designated or intended travel path of a vehicle and may have marked (e.g., lines on a road) or unmarked boundaries (e.g., an edge of a road, a road barrier, guard rail, parked vehicles, etc.), or constraints. Rather than relying on map data to make a lane end determination, the system may operate based solely on visual information. In some cases, however, the system may rely upon a combination of map data and visual information relative to a lane of travel. Further, while the system may recognize an end of a lane by direct observation of the lane ending, the system may recognize an approaching lane end sooner by, for example, observing and recognizing road sign alerts that provide information regarding a lane ending.

[0103] FIG. 8 illustrates a vehicle 200 traveling on a roadway 800 in which the disclosed systems and methods for recognizing the end of a travel lane and operating vehicle 200 within the lanes may be used. While vehicle 200 is depicted as being equipped with image capture devices 122 and 124, more or fewer cameras may be employed on any particular vehicle 200. As shown, roadway 800 may be subdivided into lanes, such as lanes 810 and 820. Lanes 810 and 820 are shown as examples; a given roadway 800 may have additional lanes based on the size and nature of the roadway, for example, an interstate highway. In the example of FIG. 8 , vehicle 200 is traveling in lane 810, and lane 810 can be seen to be ending shortly. In response, system 100 may, as discussed in detail below, cause vehicle 200 to shift to lane 820.

[0104] Processing unit 110 may be configured to determine one or more lane constraints associated with each of lanes 810 and 820 based on a plurality of images acquired by image capture device 122-126 that processing unit 110 may receive via data interface 128. According to some embodiments, the lane constraints may be identified by visible lane boundaries, such as dashed or solid lines marked on a road surface. Additionally or alternatively, the lane constraints may include an edge of a road surface or a barrier. Additionally or alternatively, the lane constraints may include markers (e.g., Botts' dots). According to some embodiments, processing unit 110 (via lane constraint module 910 and/or lane positioning module 920, described in detail below) may determine constraints associated with lanes 810/820 by identifying a midpoint of a road surface width, such as the entirety of roadway 800 or a midpoint of one of lanes 810/820. Processing unit 110 may identify lane constraints in alternative manners, such as by estimation or extrapolation based on known roadway parameters when, for example, lines designating road lanes such as lanes 810/820 are not painted or otherwise labeled.

[0105] Detection of the constraints and pathways of roadway 800 and constituent lanes 810/820 may include processing unit 110 determining their 3D models via a camera coordinate system. For example, the 3D models of lanes 810/820 may be described by a third-degree polynomial. In addition to 3D modeling of travel lanes, processing unit 110 may perform multi-frame estimation of host motion parameters, such as the speed, yaw and pitch rates, and acceleration of vehicle 200. Processing unit 110 may receive information pertaining to these parameters from, for example, speed sensors, multi-axis

Tesla EX1002 Page 1290 IPR2025-00944

accelerometers, etc., included in vehicle 200. Optionally, processing unit 110 may detect static and moving vehicles and their position, heading, speed, and acceleration, all relative to vehicle 200, which will be described below in association with FIG. 10B . Processing unit 110 may further determine a road elevation model to transform all of the information acquired from the plurality of images into 3D space.

[0106] Generally, as a default condition, vehicle 200 may travel relatively centered within lanes 810/820, so that distances x1 and x2 from vehicle 200 to the lane constraints associated with lanes 810/820 depicted in FIG. 8 are equal or close to equal. However, in some circumstances, environmental factors may make this undesirable or unfeasible, such as when objects or structures are present on one side of the road. Thus, there may be circumstances in which it may be desirable or advantageous for vehicle 200 to be closer to one lane constraint or the other (i.e. x1 > x2 or vice versa), particularly if the lane is determined to be ending, as will be discussed below in association with FIGs. 10-11 .

[0107] FIG. 9 is an exemplary block diagram of memory 140 and/or 150, which may store instructions for performing one or more operations consistent with disclosed embodiments. As illustrated in FIG. 9 , memory 140 may store one or more modules for performing the lane end recognition and responses described herein. For example, memory 140 may store a lane constraint module 910, a lane positioning module 920, and an action module 930. Application processor 180 and/or image processor 190 may execute the instructions stored in any of modules 910-930 included in memory 140. One of skill in the art will understand that references in the following discussions to processing unit 110 may refer to application processor 180 and image processor 190 individually or collectively. Accordingly, steps of any of the following processes may be performed by one or more processing devices.

[0108] Lane constraint module 910 may store instructions which, when executed by processing unit 110, may detect and define boundaries associated with lanes such as lanes 810 and 820. In some embodiments the detected boundaries may be artificial, for example, painted lines or reflectors. In other embodiments, the detected boundaries may simply include the natural curvature of the road or a termination of a lane or road surface. Lane offset module 910 may process the plurality of images received from at least one image capture device 122-124 to detect lane constraints. As discussed above, this may include identifying painted lane lines and/or measuring a midpoint of a road surface. Additionally, lane constraint module 910 may detect the course and/or constraint of travel lanes such as lanes 810 and 820 by extracting information from one or more road signs within the field of view of image capture device(s) 122-124. Examples of these road signs may include, but are not limited to, signs containing textual messages such as "Lane Ends - Merge Left" or signs containing visual messages, such as a visual representation of a lane ending.

[0109] Lane positioning module 920 may store instructions which, when executed by processing unit 110, may assist system 100 in determining a position of vehicle 200. In some embodiments, determining the position of vehicle 200 may include determining, based on at least one indicator of vehicle position, a distance from the vehicle to one or more lane constraints associated with the current lane in which the vehicle is traveling. Such determinations of the position of vehicle 200 may be based at least in part on an analysis of image data captured by one or more of image capture devices 122-126. In some embodiments, the image data may be supplemented by or substituted with map data (e.g., GPS information) and/or data from other sensors (e.g., laser range finders,

Tesla EX1002 Page 1291 IPR2025-00944

etc.). In these embodiments, the distance may be measured relative to lane constraints identified by lane constraint module 910. For example, the distance to the one or more lane constraints may specify the distance to an end of a lane or an area in which a lane has partially ended (e.g., for a lane that includes a merging zone that diminishes over distance until the lane has ended). Lane positioning module 920 may process the plurality of images to detect features of the roadway on which vehicle 200 is traveling, such as roadway 800, in order to determine the position of vehicle 200. Additionally or alternatively, lane offset module 920 may receive information from other modules (including position sensor 130) or other systems indicative of the presence of other features in the image data, such as additional vehicles, curvature of the road, etc. Lane positioning module 920 may periodically update the positional status of the vehicle using coordinate systems or other notations of position (e.g., map data from a mapping application, GPS information, etc.). The status may be stored within the system, such as within memory 140/150. Lane positioning module 920 may update the positional status while vehicle 200 is moving at set timepoints, such as each minute, after a predetermined number of minutes, each hour, etc. Alternatively, lane positioning module 920 may update the positional status when certain vehicle operations are detected, such as changing lanes, accelerating, decelerating, or other steering operations.

[0110] Action module 930 may store instructions which, when executed by processing unit 110, may assist system 100 in taking one or more actions relative to the operation of vehicle 200 based on information received from one or more sources, such as position sensor 130, lane constraint module 910, or lane positioning module 920. In some embodiments, action module 930 may receive information indicating that the current travel lane of vehicle 200 is ending. Action module 930 may then perform one or more actions affecting the operational status of vehicle 200, such as causing system 100 to provide control signals to one or more of throttling system 220, braking system 230, and steering system 240 to navigate vehicle 200 (e.g., by causing an acceleration, a turn, a lane shift, etc.). For example, in some embodiments, action module 930 may send an instruction to steering system 240, and steering system 240 may execute the instruction to steer the vehicle into a new lane of travel. Additionally or alternatively, action module 300 may send instructions to other systems associated with vehicle 200, such as braking system 230, turn signals, throttling system 220, etc. In some embodiments, action module 930 may instead provide a human operator of the vehicle with audio, visual, haptic, or tactile feedback representative of the information gathered from the relevant systems and/or sensors. The human operator may then act on this feedback to steer the vehicle to a different lane.

[0111] FIG. 10A provides an annotated view of the situation depicted in FIG. 8 . Vehicle 200 is once again traveling in lane 810 of roadway 800, which is about to end. Vehicle 200 is again equipped with image capture devices 122 and 124; more or fewer devices may be associated with any particular vehicle 200.

[0112] To the right side of roadway 800 are road signs 1002 and 1004. These signs are exemplary, and not intended to be limiting. Road sign 1002 presents a visual depiction of a lane ending. This sign is familiar to many drivers, indicating a need to merge. Sign 1004, typically placed closer to the actual end of the lane, informs drivers that the "LANE ENDS - MERGE LEFT." Consistent with disclosed embodiments, system 100 may be configured to extract information from road signs such as signs 1002 and 1004 for purposes of determining if and when the current travel lane that a vehicle 200 is traveling in is ending. For visual signs such as sign 1002, image capture devices 122-

IPR2025-00944
Tesla EX1002 Page 1292

126 may capture images of the entire face of the sign. Image processor 190 may be configured to recognize the image pattern of sign 1002, and may transmit information to other modules or units within system 100 indicating that the sign instructs that the lane is ending. In some embodiments, constituent modules or subsystems of system 100, such as memory 140/150, may be configured to contain lists or databases of typical road signs and patterns for purposes of rapid, real-time detection and information extraction. Sign 1004 contains only text, and thus its warning may be read by system 100 through optical character recognition (OCR) capability, which may be configured within processing unit 110, memory 140/150, image processor 190, or any other constituent subsystem or module of system 100.

[0113] Extraction of information from road signs is one of several actions that system 100 may be configured to execute for purposes of lane recognition. In addition, system 100, via lane constraint module 910 and lane positioning module 920 within memory 140/150, may be configured to perform visual determinations of roadway 800 and make accompanying distance measurements based on at least one indicator of position of the vehicle. In the example of FIG. 10A , various potential measurements that may be used by system 100 are shown. System 100 may obtain the measurements by, for example, analyzing one or more images acquired by one or more of image capture devices 122-126. Measurement 1006 may represent the total width w of lane 810. Distance w may also be thought of as the sum of distances x1 and x2 from FIG. 8 plus the width of vehicle 200. Measurement 1008 may represent the radius of curvature r of roadway 800 as lane 810 comes to an end. A shorter radius r may indicate a more urgent need for system 100 to steer vehicle 200 into lane 820. Measurement 1010 may represent the distance d1 from vehicle 200 to a constraint associated with lane 810. This constraint may indicate that lane 810 is ending. As discussed above, the constraint may be a painted line, or may be the natural edge of roadway 800. Distance d1 may also be measured to any portion in three-dimensional space of a barrier, such as a guardrail, wall, barrel, cone, or any other such object. Distance d1 and other such distances, such as x1, x2, and w, may be measured from any portion of the interior or exterior vehicle 200, including but not limited to the front of vehicle 200, a portion of vehicle 200 such as a headlight or front license plate, a position as-installed of image capture devices 122 or 124, a determined centroid of vehicle 200, the rear of vehicle 200, one or more windshields or mirrors associated with vehicle 200, wheels of vehicle 200, right or left sides or windows of vehicle 200, a point associated with the roof of vehicle 200, or a point associated with the chassis of vehicle 200. As will be discussed below in association with process 1100 of FIG. 11 , system 100 may use one or both of the extracted road sign information or the distances 1006-1010 to determine if the current travel lane of vehicle 200 is ending.

[0114] FIG. 10B illustrates an additional visual determination and/or set of distance measurements that system 100 may use to determine if a current travel lane is ending on a roadway. In the example of FIG. 10B , there are two vehicles 200a and 200b driving in lane 810 of roadway 800, which is ending. In some embodiments, vehicle 200a may be equipped with a system 100 and may measure a distance d2 to the back side of vehicle 200b. Distance d2 may be measured to or from any of the portions of vehicles 200a and 200b discussed previously. Alternatively, vehicle 200b may be equipped with the system 100, and may measure distance d2 backwards to vehicle 200a. Both vehicles are depicted as being equipped with systems 100 for purposes of FIG. 10B 's exemplary illustration. In some embodiments, the system 100 (on whichever of the two vehicles illustrated in FIG. 10B) may be configured to make a visual determination of whether the current travel lane is ending based on distance d2.

IPR2025-00944
Tesla EX1002 Page 1293

In these embodiments, system 100 may be configured to recognize that the lane is ending based on an abrupt shortening of distance d2. Alternatively, system 100 may detect that distance d2 has acquired a lateral element instead of being a straight line extending normal to a front plane associated with vehicle 200a. When configured in this manner, measurements such as c1-c4 may be measured and used for positional status determination by lane positioning module 920. Distances c1-c4 may represent the distance from a side of vehicle 200a or 200b to the edge of lane 810, be it lane constraint 1014 or the dashed center line in the middle of roadway 800 dividing lanes 810/820. In other embodiments, a distance may be measured to lane constraint 1012 on the far side of lane 820 (not shown).

[0115] As in FIG. 8 and in FIG. 10A , lane 810 in FIG. 10B is ending, and both vehicles 200a and 200b will need to merge into lane 820. FIG. 10B illustrates additional distances that system 100 may be configured to calculate, extract, and/or analyze in a live driving situation. In addition to distances c1-c4 described above, system 100 may be further configured to calculate distances w1 and w2 and midpoint m of lane 810 relative to one or more lane constraints associated with that lane. When summed together, distances w1 and w2 equal measurement 1006 w as shown in FIG. 10A . System 100, via lane positioning module 920, may periodically measure distances w1 and w2 and midpoint m of lane 810 over set timepoints. Should distances w1 and w2 shrink, and/or if midpoint m shifts to the right or left, system 100 may determine that lane 810 is ending and that a merge to the other lane is needed.

[0116] FIG. 11 illustrates a process 1100 for detecting a lane ending for a vehicle, consistent with disclosed embodiments. Steps of process 1100 may be performed by one or more of processing unit 110, image acquisition unit 120, position sensor 130, lane constraint module 910, lane positioning module 920, or action module 930.

[0117] At step 1110, process 1100 may include acquiring, using at least one image capture device 122, 124, and/or 126, at least one image of an area forward of vehicle 200. The area may further include at least one road sign providing information indicative of the lane ending, such as road signs 1002 and/or 1004. Processing unit 110 may receive the plurality of images from the image capture device(s) through data interface 128. The at least one image may then be processed in real time by image processor 190 of processing unit 110.

[0118] At step 1120, process 1100 may extract lane ending information from a representation of the at least one road sign included in the at least one image. The extraction may be carried out by one or more of image processor 190, lane constraint module 910, or lane positioning module 920. A variety of lane ending information may be extracted from the image data. For example, on roadways such as roadway 800 comprising a plurality of travel lanes (such as lanes 810 and 820), the extracted lane ending information may include an identification of which of the lanes from the plurality of lanes is ending. In other embodiments, the extracted lane ending information may include a distance to the lane ending. Image processor 190, lane constraint module 910, and/or lane positioning module 920 may identify one or more objects or lane constraints to help determine the distance until the lane ends. Taking the example of FIG. 10A , one or more images received from data interface 128 may include road signs 1002 and 1004, and may factor in road curvature 1008 to determine the distance 1010 remaining in the lane. In some embodiments, a value of curvature 1008 may itself independently constitute the extracted lane ending information. As part of processing the extracted information, system 100 may determine the distance from vehicle 200 to

Tesla EX1002 Page 1294 IPR2025-00944

one or more of the identified objects (e.g., lane constraints) using position sensor 130. In some embodiments, system 100 may determine the distance from vehicle 200 to one or more of the identified objects (e.g., lane constraints) by analyzing one or more images acquired by one or more of image capture devices 122-126. In yet other embodiments, system 100 may determine the distance from vehicle 200 to one or more of the identified objects (e.g., lane constraints) by using a combination of data provided by position sensor 130 and data obtained through an analysis of one or more images acquired by one or more of image capture devices 122-126. At step 1130, process 1100 may determine a position of the vehicle, such as vehicle 200, based on at least one indicator of position of the vehicle. The at least one indicator of position of the vehicle may include information obtained by a direct measuring tool, physical device, or module, such as position sensor 130. Alternatively or additionally, the at least one indicator of position of the vehicle may include information obtained by an indirect measurement. For example, in some embodiments the at least one indicator of position may include a visual determination of which lane (such as lanes 810 and 820) among a plurality of lanes on a roadway 800 that vehicle 200 may be traveling. In other embodiments, the at least one indicator of position may include one or more visual determinations. The visual determinations may be directly determined by software instructions executed by image acquisition unit 120 in the manner of a "range finder," or may be determined from analysis of image data received from data interface 128 by image processor 190 and acquired by one or more of image capture devices 122-126. In some embodiments, the at least one indicator of position may include a visual determination of a distance to the lane ending. This determination may, for example, be made by lane constraint module 910 and lane positioning module 920 based on a determined distance from vehicle 200 to one or more lane constraints of roadway 800 associated with one or more lanes. In still other embodiments, the at least one indicator of position may include a visual determination of a distance to a second vehicle. The second vehicle may be traveling in front of vehicle 200. This is the scenario described above in association with FIG. 10B . Once determined, the positional information may be stored within system 100, such as in memory 140/150, and may be periodically updated by lane positioning module 920 as discussed above in association with FIG. 9 .

[0119] At step 1140, process 1100 may determine (based on the extracted lane information from step 1120 and the determined distance from the vehicle to the one or more roadway 800 lane constraints from step 1130), whether the current lane in which the vehicle is traveling is ending. As described generally above, system 100 may determine the ongoing status of a travel lane, such as lanes 810 and 820, within a roadway 800. For example, system 100 may extract information from road signs ahead of the vehicle, such as road signs 1002 and 1004 indicating that the lane in which vehicle 200 is traveling in may be ending shortly. Additionally, measurements of distance may be made, using various elements of system 100. For example, position sensor 130 and/or lane positioning module 920 may send information to processing unit 110 regarding the relative position of vehicle 200 within lane 810/820, including how far vehicle 200 is from the end of the lane. The end of the lane may be based on a measured distance or a visual determination (e.g., by analyzing one or more image images acquired by one or more of image capture devices 122-126).

[0120] If processing unit 110 determines from the acquired information that the current travel lane is not ending (Step 1140: NO), then process 1100 may revert back to the beginning, and system 100 may continue to monitor the status of vehicle 200 with respect to its current lane of travel. This condition may result, for example, if a road

IPR2025-00944
Tesla EX1002 Page 1295

sign indicates that a lane is ending due to road construction, but due to an overriding exception, however, such as weather, the day of the week, or other factors, no workers or obstacles are actually blocking the lane. System 100 may also determine, for example, via lane constraint module 910 that a perceived shift in one or more lane constraints may be due to a new or temporary road pattern, or may comprise a natural curvature in the road. Of course, the driver (if any) of vehicle 200 is still capable of changing lanes manually.

[0121] Alternatively, if system 100 determines that the current travel lane is ending (Step 1140: YES), then process 1100 may proceed to step 1150. At step 1150, process 1100 may cause vehicle 200 to change lanes to another lane of roadway 800 that is not ending. Via action module 930, processing unit 110 may send an instruction to one or more systems associated with vehicle 200. For example, an instruction may be sent to steering system 240, instructing vehicle 200 to merge left or right into the next lane. Additional instructions may be sent; for example, image processor 190 may determine from real-time image data that traffic prevents an immediate merge into the next lane. In these embodiments, an instruction may first be sent to braking system 230 and/or the turn signal control system, and vehicle 200 may decelerate or be brought to a stop until the desired new lane is safe to merge into. At that time, the instruction to steering system 220 may be made and the vehicle may turn. In some alternative embodiments, where vehicle 200 is partially or fully under the control of a human operator, step 1150 may instead comprise generating feedback to the operator informing that the travel lane is ending and that a lane change is required. In some embodiments, an automated system such as system 100 and its action module 930 may take over operation of the car and cause the vehicle to change lanes automatically as described above if the human operator does not respond to the feedback within a certain time period. The time period may be predetermined, or may be calculated based on visual determinations of the distance to the end of the lane such as those described above in association with step 1130.

Detection of Low-Height Objects in a Roadway

[0122] System 100 may detect when an object is situated within a current lane of travel and, if system 100 determines that the object exceeds a threshold height, the system may steer an associated vehicle around or over the object. A "lane" may refer to a designated or intended travel path of a vehicle and may have marked (e.g., lines on a road) or unmarked boundaries (e.g., an edge of a road, a road barrier, guard rail, parked vehicles, etc.), or constraints. System 100 may operate to make these detections and determinations based solely on visual information acquired by one or more image acquisition units. In some embodiments, system 100 may be configured to detect objects in a roadway that project above the roadway surface by as little as 2 cm.

[0123] FIG. 12 illustrates a vehicle 200 traveling on a roadway 800 in which the disclosed systems and methods for detecting an object in the roadway may be used. Vehicle 200 is depicted as being equipped with image capture devices 122 and 124; more or fewer image capture devices (including cameras, for example) may be employed on any particular vehicle 200. As shown, roadway 800 may be subdivided into lanes, such as lanes 810 and 820. Lanes 810 and 820 are shown as examples; a given roadway 800 may have additional lanes based on the size and nature of the roadway, for example, an interstate highway. In the example of FIG. 12 , vehicle 200 is traveling in lane 810, and two objects 1202 and 1204 are located in lane 810. System 100 may, as discussed in detail below, cause vehicle 200 to steer around or over the

IPR2025-00944
Tesla EX1002 Page 1296

object(s).

[0124] Processing unit 110 may be configured to determine one or more lane constraints associated with each of lanes 810 and 820 based on a plurality of images acquired by image capture device 122-126 that processing unit 110 may receive via data interface 128. According to some embodiments, the lane constraints may be identified by visible lane boundaries, such as dashed or solid lines marked on a road surface. Additionally or alternatively, the lane constraints may include an edge of a road surface or a barrier. Additionally or alternatively, the lane constraints may include markers (e.g., Botts' dots). According to some embodiments, processing unit 110 (via lane positioning module 1320, described in detail below) may determine constraints associated with lanes 810/820 by identifying a midpoint of a road surface width, such as the entirety of roadway 800 or a midpoint of one of lanes 810/820. Processing unit 110 may identify lane constraints in alternative manners, such as by estimation or extrapolation based on known roadway parameters when, for example, lines designating road lanes such as lanes 810/820 are not painted or otherwise labeled.

[0125] Detection of the constraints and pathways of roadway 800 and constituent lanes 810/820 may include processing unit 110 determining their 3D models via a camera coordinate system. For example, the 3D models of lanes 810/820 may be described by a third-degree polynomial. In addition to 3D modeling of travel lanes, processing unit 110 may perform multi-frame estimation of host motion parameters, such as the speed, yaw and pitch rates, and/or acceleration of vehicle 200. Processing unit 110 may detect static and moving objects and their position, all relative to vehicle 200, which will be described below in association with FIGs. 14A , 14B , and 15 . For example, processing unit 110 may use image data, GPS data, and/or data from sensors (e.g. position sensor 130, speed sensors, etc.) to determine the position of vehicle 200. Processing unit 110 may further determine a road elevation model to transform a portion or all of the information acquired from the plurality of images into 3D space.

[0126] In the example of FIG. 12 , several measurements are represented that system 100 may take into account when determining potential actions to take using action module 1330, which will be described in detail below. Objects 1202 and 1204 are represented in the illustration of FIG. 12 as a cardboard box and a ladder, respectively. These are common items that may be found as obstacles in a typical roadway, for example, after falling from an unseen second vehicle. These objects are meant as examples only. System 100 may be configured to detect the presence of any object that may be found in a roadway, including but not limited to branches, tires or portions thereof, scraps of metal or wood, speed bumps within a travel lane, a human or animal, a pipe, a piece of another vehicle, etc.

[0127] As part of a driver-assist object detection system consistent with disclosed embodiments, system 100 may determine, for example, distances l1 and l2, which may represent the distance of a surface of vehicle 200 to various features of roadway 800 and/or lanes 810/820, such as lane constraints, road signs, or any other such feature. Further, system 100 may determine distances d3 and d4, which are distances measured from a surface of vehicle 200 to each of objects 1202 and 1204, respectively. Distances such as l1, l2, d3, and d4, may be measured from any portion of the interior or exterior vehicle 200, including but not limited to the front of vehicle 200, a portion of vehicle 200 such as a headlight or front license plate, a position as-installed of image capture devices 122 or 124, a determined centroid of vehicle 200, the rear of vehicle 200, one or more windshields or mirrors associated with vehicle 200,

IPR2025-00944
Tesla EX1002 Page 1297

wheels of vehicle 200, right or left sides or windows of vehicle 200, a point associated with the roof of vehicle 200, or a point associated with the chassis of vehicle 200. Further detail on how these exemplary distance measurements may be used to determine potential actions taken by system 100 will now be discussed in detail.

[0128] FIG. 13 is an exemplary block diagram of memory 140 and/or 150, which may store instructions for performing one or more operations consistent with disclosed embodiments. As illustrated in FIG. 13 , memory 140 may store one or more modules for performing the object detection and responses described herein. For example, memory 140 may store an object detection module 1310, a lane positioning module 1320, and an action module 1330. Application processor 180 and/or image processor 190 may execute the instructions stored in any of modules 1310-1330 included in memory 140. One of skill in the art will understand that references in the following discussions to processing unit 110 may refer to application processor 180 and image processor 190 individually or collectively. Accordingly, steps of any of the following processes may be performed by one or more processing devices.

[0129] Object detection module 1310 may store instructions which, when executed by processing unit 110, may detect one or more objects found within a roadway 800, such as within lanes 810 and 820. As will be discussed below in association with FIGs. 14A , 14B , and 15 , object detection module 1310 may identify one or more objects within image data captured by one or more of image capture devices 122-126 and determine that those objects are resting on the roadway surface. Object detection module 1310 may further determine a reference plane corresponding to the road surface, and using the reference plane, may determine dimensions and attributes of the identified object (s), such as their height, length, and width. The object detection process may further include identifying painted lane lines and/or measuring a midpoint of a road surface such as roadway 800, for purposes of determining the reference plane and for determining the location of the objects within the roadway/lanes. In some embodiments, the field of view of the image capture device(s) used to generate the image data may range from about 35 to about 55 degrees.

[0130] Lane positioning module 1320 may store instructions which, when executed by processing unit 110, may assist system 100 in determining a position of vehicle 200. In some embodiments, determining the position of vehicle 200 may include determining, based on at least one indicator of vehicle position, a distance from the vehicle to one or more lane constraints associated with the current lane in which the vehicle is traveling. In these embodiments, the distance may also be measured relative to identified objects detected by object detection module 1310. Consistent with disclosed embodiments, the distance(s) measured to the one or more lane constraints may be used to determine if a change in a directional course of vehicle 200 is required or recommended. Lane positioning module 1320 may process the plurality of images to detect features of the roadway on which vehicle 200 is traveling, such as roadway 800, in order to determine the position of vehicle 200. Additionally or alternatively, lane positioning module 1320 may receive information from other modules (including position sensor 130) or other systems indicative of the presence of other features in the image data, such as additional vehicles, curvature of the road, etc.

[0131] Action module 1330 may store instructions which, when executed by processing unit 110, may assist system 100 in taking one or more actions relative to the operation of vehicle 200 based on information received from one or more sources, such as position sensor 130, object detection module 1310, or lane positioning module 1320. In

Tesla EX1002 Page 1298 IPR2025-00944

IPR2025-00944 Tesla EX1002 Page 1299

some embodiments, action module 1330 may receive information (from, e.g., object detection module 1310) indicating that one or more objects have been detected in roadway 800. The information may further include one or more dimensions or attributes of the detected object, such as its height relative to the roadway surface. Action module 1330 may then perform one or more actions affecting the operational status of vehicle 200, such as causing system 100 to provide control signals to one or more of throttling system 220, braking system 230, and steering system 240 to navigate vehicle 200 (e.g., by causing an acceleration, a turn, a swerve, a lane shift, etc.). For example, in some embodiments, action module 1330 may send an instruction to steering system 240, and steering system 240 may execute the instruction to steer the vehicle into a new lane of travel. In other embodiments, action module 1330 may determine that a safer and/or better course of action is to simply steer vehicle 200 such that it passes over top of the object, assuming the height of the object is less than the road clearance height of vehicle 200. Additionally or alternatively, action module 1330 may send instructions to other systems associated with vehicle 200, such as braking system 230, turn signals, throttling system 220, etc. In some embodiments, action module 1330 may instead provide a human operator of the vehicle with audio, visual, or tactile feedback representative of the information gathered from the relevant systems and/or sensors. The human operator may then act on this feedback to steer the vehicle to a different lane.

[0132] FIG. 14A illustrates examples of two images among the plurality of captured images that may be acquired by system 100 via image capture devices 122-126. In the example of FIG. 14A , image 1402 (top) may be the first image acquired, and image 1404 (bottom) may be a second image taken at a later time, such as a few seconds later. As can be seen in images 1402 and 1404, a vehicle equipped with a system 100, such as vehicle 200 above, is driving on a roadway, such as roadway 800. The roadway is aligned straight ahead of the vehicle, with the view of the road ending at horizon 1406. Based on the view illustrated in image 1402, the roadway may be cresting a hill. Any roadway configuration or alignment may be capable of being analyzed by system 100. As can be seen in both images 1402 and 1404, object 1202, representing a cardboard box of initially unknown dimensions, is lying in the middle of the right travel lane of the illustrated roadway.

[0133] In some embodiments, system 100 may be configured to analyze captured images, such as images 1402 and 1404, to determine a reference plane corresponding to a road plane, such as the roadway surface. This process will be described in more detail below in association with FIG. 15 and process 1500. Briefly, system 100 (via processing unit 110 and image processor 190) may determine the reference plane corresponding to the roadway surface as an average surface level associated with the road surface. This may account for variations in the roadway surface due to the composition of the roadway (i.e., asphalt, concrete, or even gravel) or changes in the pitch and/or curvature of the roadway. Multiple images from among the plurality of images may be used to determine the reference plane. In the illustration of FIG. 14A , the determined reference plane is represented by reference plane 1408.

[0134] System 100, via processing unit 110 and object detection module 1310, may locate one or more objects lying in the roadway, such as object 1202. The detection may occur, for example, when the object appears in the same location in multiple images among the captured plurality of images. Objects, such as object 1202, may be moving or stationary. Object 1202 is stationary in the roadway in the illustration of FIG. 14A . The locations) of the object(s) may be determined by object detection module

1310 by various methods. In some embodiments, the location of the target object(s) may be determined by GPS and/or other coordinate-based systems. Further, object detection module 1310 may make a visual determination of the location of object 1202. In some embodiments, one or more of object detection module 1310, position sensor 130, and image processor 190 may further determine how far away object 1202 is from vehicle 200, as discussed above in association with FIG. 12 . When configured in this manner, measurements such as d3-d4 as shown in FIG. 12 may be measured and used for positional status determination by lane positioning module 1320.

[0135] Distances from vehicle 200 to objects detected within the roadway, such as object 1202, may be calculated in the manner described above in association with FIGs. 5-7 . For example, system 100 may analyze image data, position data (e.g., GPS location information), map data, speed data, and/or data from sensors included in vehicle 200 to calculate the distances. System 100 may collect the data for analysis from, for example, image acquisition unit 120, position sensor 130, and other sensors.

[0136] In some embodiments, processing unit 110 may construct the measurements based on estimation techniques using a series of time-based observations such as Kalman filters or linear quadratic estimation (LQE), and/or based on available modeling data for different object types (e.g., cars, trucks, pedestrians, bicycles, road signs, etc.). The Kalman filters may be based on a measurement of an object's scale, where the scale measurement is proportional to a time to collision (e.g., the amount of time for vehicle 200 to reach the object).

[0137] Processing unit 110 may be configured in some embodiments to create a 3D reconstruction of the visible portion of the roadway, using one or more mathematical algorithms or models as discussed previously. For example, as discussed above in association with FIG. 5 , processing unit 110 may determine a look-ahead distance for vehicle 200, which may have a lower bound ranging from 10 to 20 meters, may be calculated as the product of the speed of vehicle 200 and a determined look-ahead time. For example, as the speed of vehicle 200 decreases, the look-ahead distance may also decrease (e.g., until it reaches the lower bound). Additionally or alternatively, processing unit 110 may execute stereo image analysis module 404 to perform stereo image analysis of the first and second plurality of images to create a 3D map of the road in front of the vehicle and detect features within the images, such as object 1202. Processing unit 110 may estimate a position of object 1202 relative to vehicle 200 by observing differences that may exist relative to objects appearing in at least two of the plurality of captured images within the image data. For example, position, velocity, and/or acceleration relative to vehicle 200 may be determined based on trajectories, positions, movement characteristics, etc. of features associated with object 1202.

[0138] Object 1202 may be a relatively flat object, or it may be three dimensional. As discussed in further detail below, if the object is flat or relatively flat (e.g., 2 cm or less, 1 cm or less, etc.), system 100 may determine that no further action or change in course of vehicle 200 is necessary. In the illustration of FIG. 14A , object 1202 is again a cardboard box. In first image 1402, object 1202 has initial dimensions ox1, oy1, and oz1. System 100, via object detection module 1310, may be configured to determine these dimensions based on the reference plane 1408, the detected lane constraints of the roadway, or other reference points of known size and/or location within the image data. In second image 1404, taken later, the object 1202 is now closer to the vehicle, and appears larger. Object 1202 now has second dimensions ox2, oy2, and oz2. System 100 may determine this second set of dimensions in the same manner as the

Tesla EX1002 Page 1300 IPR2025-00944

first.

[0139] In some embodiments, system 100 may use the determined dimensions of object 1202 in first image 1402 and at least a second image, such as image 1404, to gather more information about the object and determine one or more attributes of the object. In some embodiments, system 100 (via processing unit 110 and/or image processor 190) may be configured to automatically recognize the object based on known objects previously programmed into an integrated database or a database accessible by the Internet. In these embodiments, system 100 may be configured to generate audio or visual feedback to a human operator of vehicle 200.

[0140] System 100 may determine, for example, that one or more dimensions ox2, oy2, and oz2 from image 1404 are different from dimensions ox1, oy1, and oz1 from image 1402. If the second set of dimensions from image 1404 are smaller than those taken from the first image 1402, the object is likely moving away from vehicle 200 and is generally not of concern. If the second set of dimensions is determined to be larger than the first set of dimensions, then that determination may indicate that object 1202 is either stationary in the roadway or moving towards vehicle 200, and additional actions may be taken. For example, system 100 may determine the height of object 1202 (essentially, dimension ox1/ox2) relative to reference plane 1408 to determine the approximate height of object 1202 relative to the roadway surface. In some embodiments, system 100 may be configured to detect objects as small as 2 cm tall. Based on the determined height of the object, system 100 may take one of several alternative actions, as will now be discussed in further detail.

[0141] FIG. 14B provides an annotated view of the situation depicted in FIG. 12 . Vehicle 200 (here represented as vehicle 200c) is once again traveling in lane 810 of roadway 800. Vehicle 200c is again equipped with image capture devices 122 and 124; more or fewer devices may be associated with any particular vehicle 200.

[0142] Object 1202 is located in lane 810 of roadway 800. Using the process described above in association with FIG. 14A (and in further detail below in association with FIG. 15 and process 1500), system 100 may be configured to determine multiple attributes of object 1202 based on a plurality of images acquired from one or more of image capture devices 122-126, and received via data interface 128. These attributes may include the position and dimensions of object 1202, including its height relative to the surface of roadway 800.

[0143] As an initial step, system 100, via position sensor 130 and/or lane positioning module 1320, may determine relative positions of vehicle 200c and object 1202. This determination may include determining lane constraints or other such positional indicators associated with roadway 800 and/or lanes 810/820. System 100 via processing unit 110 may then determine the position of vehicle 200c within its current lane of travel (here, lane 810) by calculating distances from the determined constraints. In the illustration of FIG. 14B , system 100 may for example determine distances q1 and q2 from one or more surfaces of vehicle 200 to the lane constraints. Further, system 100 may determine distance d2 from a surface of vehicle 200c to a surface of object 1202. System 100 may also determine relative positioning of object 1202 laterally within the lane, such as by measuring distance q3 from a surface of the object 1202 to the lane constraint separating lanes 810 and 820. Each of these determined distances, such as q1, q2, etc., may be measured from any portion of the interior or exterior of vehicle 200c, including but not limited to the front of vehicle 200c, a portion

Tesla EX1002 Page 1301 IPR2025-00944

of vehicle 200c such as a headlight or front license plate, a position as-installed of image capture devices 122 or 124, a determined centroid of vehicle 200c, the rear of vehicle 200c, one or more windshields or mirrors associated with vehicle 200c, wheels of vehicle 200c, right or left sides or windows of vehicle 200c, a point associated with the roof of vehicle 200c, or a point associated with the chassis of vehicle 200c. Distances d2 and q3 may also be measured to any portion in three-dimensional space of object 1202.

[0144] Based at least on the dimensions of object 1202 determined as discussed above in association with FIG. 14A , as well as the determined distances illustrated in FIG. 14B , system 100 may determine one or more actions to take. System 100 may act via action module 1330 to send information and instructions to various associated subsystems of vehicle 200c to carry out the actions, including but not limited to throttling system 220, braking system 230, and steering system 240. The action taken by action module 1330 may be further informed by a determination of the roadway clearance height of vehicle 200c. This clearance height may be a known value stored in memory 140/150 of system 100, or it may be determined anew in each driving situation based on various factors such as the inflation status of the tires, the road surface, weather conditions, etc. In some embodiments, the clearance height may be measured using optics and a reference point (e.g., from vehicle 200c to a point on the road surface) or by a sound echo, optical echo, radio frequency from the road surface, etc. Action module 1330 may then determine one of several alternative actions to undertake based on the determined height of object 1202 and the roadway clearance height of vehicle 200c. For example, if the height of object 1202 is less than the roadway clearance height of vehicle 200c, action module 1330 may determine that the safest course of action is to simply do nothing, and allow vehicle 200c to drive over object 1202 with the object oriented between the wheels of vehicle 200c. Alternatively, action module 1330 may determine that the height of object 1202 is less than the roadway clearance height of vehicle 200c, but that a minor change in the directional course of vehicle 200c may permit vehicle 200c to safely drive over object 1202. Further, system 100 and action module 1330 may determine that the height of object 1202 exceeds the roadway clearance height of vehicle 200c, and that corrective adjustments to the directional course of vehicle 200c may be required in order to avoid object 1202. These adjustments may further require instructions to be sent to throttling system 220 and/or braking system 230 in order to safely undertake the change in course.

[0145] This latter scenario is illustrated in FIG. 14B . This illustration is an example only, and is not intended to be limiting. In the example of FIG. 14B , system 100 has determined that the height of object 1202 is larger than the roadway clearance height of vehicle 200c. Accordingly, action module 1330 may determine, based on detected lane constraints and measured distances as described above, that the best course of action for vehicle 200c is to move slightly to the right to avoid object 1202. Action module 1330 may then send instructions to throttling system 220, braking system 230, and steering system 240 to effect this change in directional course.

[0146] The projected location of vehicle 200c as it passes object 1202 is represented by vehicle 200d. System 100 may determine new and additional distances in order to successfully complete the pass of object 1202. For example, system 100 may determine a distance q4 as shown in FIG. 14B , corresponding to a safe distance needed from object 1202 to a surface of vehicle 200d. System 100 may further calculate distances q5 and q6 via lane positioning module 1320, representing the position of vehicle 200d relative to detected lane constraints. These distances may be

Tesla EX1002 Page 1302 IPR2025-00944

sent to action module 1330 to assist in determining the proper directional course for vehicle 200d. Ideally, vehicle 200d would successfully swerve around object 1202 while staying within the constraints of lane 810. In some embodiments, however, it may be necessary for vehicle 200d to move outside of lane 810's constraints in order to avoid object 1202, or it may be necessary for vehicle 200d to simply change lanes, such as to lane 820. Once object 1202 is passed, action module 1330 may send an indication to throttling system 220, braking system 230, and steering system 240 that normal operation of vehicle 200 may be resumed.

[0147] FIG. 15 illustrates a process 1500 for detecting an object in a roadway, consistent with disclosed embodiments. Steps of process 1500 may be performed by one or more of processing unit 110, image acquisition unit 120, position sensor 130, image processor 190, object detection module 1310, lane positioning module 1320, or action module 1330.

[0148] At step 1510, process 1500 may include acquiring, using at least one image capture device 122, 124, and/or 126, a plurality of images of the area surrounding vehicle 200. In some embodiments, the plurality of images may be of an area forward of vehicle 200. Processing unit 110 may receive the plurality of images from the image capture device(s) through data interface 128. The plurality of images may then be processed in real time by image processor 190 of processing unit 110. Image processor 190 may process the first image captured by the at least one image capture device, as well as at least one additional image captured at a later point in time. The at least one image capture device may be configured to have a field of view of between about 35 degrees and about 55 degrees. Examples of acquired images and potential fields of view for system 100 are illustrated above in association with FIG. 14A .

[0149] At step 1520, process 1500 may determine a reference plane corresponding to the surface of roadway 800. The determination may be carried out by one or more of image processor 190, object detection module 1310, or lane positioning module 1320. The reference plane may be similar to the plane 1408 described above in association with FIG. 14A . The reference plane may be determined based on the plurality of images captured by the at least one image capture device. In these embodiments, the reference plane may be determined by analyzing the first image and at least a second image from among the plurality of images. In some embodiments, the location of a roadway surface may be determined by analysis of the plurality of images by image processor 190.

[0150] The reference plane may be detected by way of parallax information, or by way of detecting one or more lane constraints on the roadway surface. In some embodiments, the reference plane may be determined by color or darkness differences in the image data. Processing unit 110 may generate and propose multiple potential reference planes, with image processor 190, and then select the best candidate. Once a reference plane is selected, information about its location may be stored within system 100, such as in memory 140/150. In some embodiments, processing unit 110 and/or image processor 190 may periodically redetermine the reference plane based on changes in the roadway surface, such as elevation or curvature changes, or when vehicle 200 turns onto a different roadway. In some embodiments, processing unit 110 may be configured to determine the reference plane 14-08, corresponding to the surface of roadway 800 as an average surface level associated with a road surface appearing in the plurality of images. This method may be useful, for example, when roadway 800 is not perfectly flat or is otherwise unusual, such as if the road surface

includes one or more of gravel, potholes, snow or ice, mud, vegetation, etc.

[0151] At step 1530, process 1500 may locate a target object in the first image and at least one second image taken at a later time within the plurality of images. Object detection module 1310 may locate the images based on their presence in multiple images among the plurality of images acquired by the system, as illustrated above in association with FIG. 14A . Object detection module 1310 may determine one or more positional measurements associated with the detected object. These measurements, as discussed above in association with FIGs. 14A-14B , may include relative measurements to other objects, such as vehicles or lane constraints. The determination of object position may be assisted by a direct measuring tool, physical device, or module, such as position sensor 130. In other embodiments, the determinations may include one or more visual determinations. The visual determinations may be directly determined by software instructions executed by image acquisition unit 120 in the manner of a "range finder," or may be determined from analysis of image data received from data interface 128 by image processor 190 and acquired by one or more of image capture devices 122-126. Once determined, the object's position may be stored within system 100, such as in memory 140/150, and may be periodically updated by object detection module 1310.

[0152] At step 154.0, process 1500 may determine one or more attributes of the detected object(s). In some embodiments, the object may be recognized as a particular object, based on information stored in a database associated with system 100. As discussed above, these determined attributes may include dimensions of the detected object, such as its height relative to the detected reference plane. System 100 may be configured to determine the height of any object within a roadway that extends at least about 2 cm above the reference plane. In alternative embodiments, system 100 may be configured to detect and determine attributes for any object extending between 2 and 10 cm above the reference plane. In other embodiments, the determination of target object height may also be based on a detected velocity of the vehicle. The velocity may be determined based on information received from one or more of throttling system 220 or braking system 230. Additional details on the calculation of a reference plane for the road surface and calculating the height of a target object may be found in U.S. Patent Application No. 14/554,500, filed November 26, 2014 , which is hereby incorporated by reference in its entirety.

[0153] System 100 may be configured to determine actions associated with the detected object based on the determined attributes. For example, system 100 and action module 1330 may be configured to take no substantive action unless the determined height of the detected object exceeds a threshold level. In some embodiments, the threshold level may be about 2 cm, as discussed above, or may be higher (e.g., 3 cm, 4 cm, 5 cm, 6 cm, etc.) depending on user or system preferences. For example, in some embodiments, the threshold may be associated with a detected roadway clearance height of vehicle 200. If processing unit 110 determines from the acquired information that height of the object does not exceed the configured threshold (Step 1550: NO), then process 1500 may revert back to the beginning, and system 100 may continue to monitor the status of potential objects within a roadway.

[0154] Alternatively, if system 100 determines that an action may be needed in response to the presence of a detected object in the roadway (Step 1550: YES), then process 1100 may proceed to step 1560. At step 1560, process 1500 may cause vehicle 200 to change directional course to evade or otherwise pass a detected object.

IPR2025-00944
Tesla EX1002 Page 1304

Via action module 1330, processing unit 110 may send an instruction to one or more systems associated with vehicle 200. For example, an instruction may be sent to steering system 240, instructing vehicle 200 to merge left or right into the next lane. Alternatively, instructions may be sent to swerve around the object, or to change directional course such that the vehicle simply drives over the object with the object situated between the wheels. Additional instructions may be sent; for example, image processor 190 may determine from real-time image data that traffic or other obstacles prevent an immediate change in directional course. In these embodiments, an instruction may first be sent to braking system 230 and/or the turn signals, and vehicle 200 may decelerate or be brought to a stop until the directional course change is possible. At that time, the instruction to steering system 220 may be safely made and the vehicle may turn to change course.

[0155] In some alternative embodiments, where vehicle 200 is partially or fully under the control of a human operator, step 1560 may instead comprise generating feedback to the operator informing that an object is detected in the roadway and that a change in directional course is suggested or required. In some embodiments, an automated system such as system 100 and its action module 1330 may take over operation of the car and cause the vehicle to change course automatically as described above if the human operator does not respond to the feedback within a certain time period. The time period may be predetermined, or may be calculated based on visual determinations such as those described above in association with FIGs. 12 and 14A-14B .

Ultra-Wide Traffic Light Detection System

[0156] In some embodiments, system 100 may identify objects located outside of a sightline of a typical driver location in vehicle 200. The typical driver location in vehicle 200 is on the left side of the front seat in countries that drive on the right-hand side of the road, and on the right side of the front seat in countries that drive on the left-hand side of the road. The U.S. National Highway Traffic Safety Administration recommends that drivers place themselves 10 inches (25.4 centimeters) from the driver air bag for safety reasons. This distance is measured from the center of the steering wheel to the driver's breastbone. Because vehicle 200 is not transparent, portions of vehicle 200 (e.g., the roof, doors, windshield frame, etc.) may obscure the angle-of-view from the typical driver location. In addition, while humans have an almost 180-degree forward-facing horizontal field of view, the vertical range of the field of view in humans is significantly less. However, a driver assistance system may use an image capture device (e.g., image capture device 122) that has an extended angle-of-view that exceeds the angle-of-view of a driver who is seated in the typical driver location. The extended angle-of-view may enable system 100 to identify objects in the environment of vehicle 200 that are otherwise undetectable by the driver. Moreover, system 100 may determine a change associated with the identified objects, and act accordingly. For example, system 100 may determine a change in status of a traffic light located in a region above vehicle 200. As another example, system 100 may determine the trajectory of different objects (e.g., pedestrians) located in a region beside vehicle 200.

[0157] FIGS. 16 and 17 are diagrammatic representations of vehicle 200 in two situations, consistent with disclosed embodiments. As illustrated in FIGS. 16 and 17 , vehicle 200 may include image capture device 122 associated with a first angle-of-view. The first angle-of-view is vertically represented by FOV 1610, as shown in FIG. 16 , and horizontally represented by FOV 1710, as shown in FIG. 17 . Also illustrated in

Tesla EX1002 Page 1305 IPR2025-00944

FIGS. 16 and 17 is a second angle-of-view from the typical driver location in vehicle 200. The second angle-of-view is vertically represented by FOV 1620, as shown in FIG. 16 , and horizontally represented by FOV 1720, as shown in FIG. 17 . Consistent with embodiments of this disclosure, the first angle-of-view associated with image capture device 122 may be broader than the second angle-of-view associated with the typical driver location in vehicle 200, The delta between the broader first angle-of-view and the second angle-of-view is referenced in this disclosure as an extended zone. In some embodiments, the extended zone may include objects located outside the sightline of the typical driver location in vehicle 200, which system 100 may identify. In addition, system 100 may trigger one or more actions (e.g., cause navigational responses) based on a determined change associated with the objects identified in the extended zone.

[0158] FIG. 16 illustrates the first example for determining a change associated with one or more objects located outside of the sightline of the typical driver location in vehicle 200. The angular range of FOV 1610 and FOV 1620 is used in FIG. 16 for illustrative purposes. In some embodiments, FOV 1620 may be, for example, about 65 degrees, and FOV 1610 may be, for example, higher than 65 degrees. In other embodiments, FOV 1620 may be, for example, about 135 degrees, and FOV 1610 may be, for example, higher than 150 degrees, higher than 160 degrees, or higher than 170 degrees. In some embodiments, FOV 1610 and FOV 1620 may be measured in an upward direction from the horizon. As shown in FIG. 16 , the extended zone may include a traffic lamp fixture 1630 located in a region above vehicle 200. Traffic lamp fixture 1630, which is located outside of the sightline of the typical driver location, displays a red light. In some embodiments, system 100 may be able to detect traffic lights and changes in the displayed color of traffic lights in situations where the driver of vehicle 200 cannot even see the traffic light fixture.

[0159] One example of these situations is when a driver of vehicle 200 applies the brakes in a manner that causes vehicle 200 to stop over the stopping line of the intersection and/or positions vehicle 200 too close to traffic light fixture 1360. Another example of these situations is when vehicle 200 reaches the stopping line at an intersection, and traffic lamp fixture 1630 is placed on a high pole so it can be seen from a distance. Because traffic lamp fixture 1630 is located outside FOV 1620, when the light in traffic lamp fixture 1630 changes to green, a driver who sits in the typical driver location would not be able to detect the status change. However, system 100 would be able to detect the status change of the traffic light because traffic lamp fixture 1630 is located within FOV 1610, and thus in view of image capture device 122. Identifying traffic lights located in a region above vehicle 200 may improve safety and driver convenience. For example, when considering the safety of the driver and the passengers of vehicle 200, if traffic lamp fixture 1630 is not inside FOV 1620, it may be unsafe to back up vehicle 200 until traffic lamp fixture 1630 is within FOV 1620. In addition, even if backing up vehicle 200 is not dangerous, it might be inconvenient for the driver to back up vehicle 200. In one embodiment, after determining that the light in traffic lamp fixture 1630 has turned to green, system 100 may cause vehicle 200 to accelerate to cross the intersection.

[0160] FIG. 17 illustrates the second example for determining a change associated with the objects located outside of the sightline of the typical driver location in vehicle 200. The angular range of FOV 1710 and FOV 1720 is used in FIG. 17 for illustrative purposes. In some embodiments, FOV 1720 may be, for example, about 180 degrees, and FOV 1710 may be, for example, higher than 200 degrees, higher than 240

IPR2025-00944
Tesla EX1002 Page 1306

degrees, or higher than 300 degrees. As shown in FIG. 17 , the extended zone may include a pedestrian 1730 located in a region beside vehicle 200. Pedestrian 1730 may be moving in a certain direction. In some cases, pedestrian 1730 may be located at the edge of FOV 1720, but because the driver is focused on the road ahead, he or she may fail to detect pedestrian 1730. In contrast, system 100 not only has a broader FOV, it does not have to focus on the road ahead and can identify objects located in any region of its FOV. Because pedestrian 213 is located within FOV 1710, system 100 may detect pedestrian 1730 and determine the pedestrian's walking direction trajectory. In other embodiments, system 100 may identify an object located beside vehicle 200, determine the object's moving direction trajectory, determine whether the object is on a possible collision course with vehicle 200, and act accordingly. For example, after determining that pedestrian 1730 is on a possible collision course with vehicle 200, system 100 may control the vehicle's horn to generate a type of audio alarm.

[0161] As discussed above, FIGS. 3B-3D illustrate exemplary camera mount 370, which is configured to be positioned behind rearview mirror 310 and against a vehicle windshield. As shown in FIGS. 3B-3D , camera mount 370 may include a plurality of image capture devices (122,124, and 126). In one embodiment, the combined angle-of-view of image capture devices 122, 124, and 126 forms the first angle-of-view. For example, in order for FOV 1710 to be higher than 300 degrees, a combination of six image capture devices may be used. In a different embodiment, at least one of the image capture devices depicted in FIGS. 3B-3D (e.g., image capture device 122) includes an ultra-wide lens. The ultra-wide lens may enable image capture device 122 to have the first angle-of-view. In some embodiments, the ultra-wide lens may have a focal length shorter than the short side of the image sensor of image capture device 122. Alternatively, the ultra-wide lens may have a focal length shorter than the long side of the image sensor, but longer then the short side of the image sensor. Consistent with an embodiment of the present disclosure, image capture device 122 may be configured to capture images of traffic lights located in a conic section zone of, for example, between 150 to 180 degrees, 150 to 190 degrees, 160 to 180 degrees, or 170 to 180 degrees relative to a location of image capture device 122.

[0162] FIG. 18 is an exemplary block diagram of a memory configured to store instructions for performing one or more operations, consistent with the disclosed embodiments. As shown in FIG. 18 , memory 140 or 150 may store an object identifying module 1802, a status determining module 1804, and a system response module 1806. The disclosed embodiments are not limited to any particular configuration of memory 140 or 150. Further, processing unit 110 may execute the instructions stored in any of modules 1802-1806 included in memory 140 or 150.

[0163] In one embodiment, object identifying module 1802 may store software instructions (such as computer vision software) which, when executed by processing unit 110, identifies an object located outside of the sightline of the typical driver location in vehicle 200. In one embodiment, processing unit 110 may scan a plurality of images, compare portions of the images to one or more predetermined patterns, and identify a traffic light located in a region above vehicle 200. In another embodiment, processing unit 110 may scan a plurality of images, compare portions of the images to one or more predetermined patterns, and identify an object located in a region beside vehicle 200. In some embodiments, module 1802 may store software instructions, which, when executed by processing unit 110, identifies intersections and determines that a traffic light may be nearby. One way for processing unit 110 to identify an intersection is by using historical image data. For example, processing unit 110 may determine that a

Tesla EX1002 Page 1307 IPR2025-00944

traffic light may be nearby even when no traffic light is visible in a current image, but the traffic light was visible in a prior image (e.g., a prior frame, five frames ago, etc.). Additionally or alternatively, processing unit 110 may identify in image data objects indicative of a close intersection. For example, processing unit 110 may determine that a traffic light may be nearby by identifying cross roads in image data of the road ahead. In another example, processing unit 110 may identify in image data the back of other traffic lights regulating traffic going in the opposite direction to the driving direction of vehicle 200. Additionally or alternatively, processing unit 110 may use map data (such as data from map database 160) to identify an intersection. For example, processing unit 110 may consider general road shape descriptions extracted from map database 160 to determine that a traffic light may be nearby.

[0164] In another embodiment, module 1802 may store software instructions, which, when executed by processing unit 110, corrects optical distortion in the plurality of images to enable identifying the objects. The optical distortion may be a result of image capture device 122 using an ultra-wide lens. For example, the optical distortion may be a barrel distortion in which the image magnification decreases with distance from the optical axis, and thus the distortion may be most significant close to the edge of the image frame. Consistent with the present disclosure, a plurality of algorithms may be used to correct the optical distortion that may result from using an ultra-wide lens. For example, an algorithm may be used to correct the optical distortion to identify an object close to the edge of the image frame. As another example, a different algorithm may be used to identify the status of an object located at the edge of the image frame. In another embodiment, the image capture device may include a plurality of image capture devices (e.g., image capture devices 122, 124, and 126), and each image capture device may be associated with a different field-of-view. Module 1802 may store software instructions, which, when executed by processing unit 110, correctly combines images retrieved from the plurality of image capture devices. For example, even when there may be some overlap in the field-of-view of two image capture devices, different parts of an object may be captured by the different image capture devices. Accordingly, object identifying module 1802 may include instructions for identifying objects in regions that span adjoining images captured by different image capture devices.

[0165] In one embodiment, status determining module 1804 may store software instructions which, when executed by processing unit 110, determines a status of an object identified in the image data, and located at least in part within an extended zone. Consistent with the present disclosure, the extended zone includes the region between the field-of-view of an image capture device (e.g., FOV 1610 and 1710) and the field-of-view of a typical driver location in vehicle 200 (e.g., FOV 1620 and 1720). In one embodiment, status determining module 1804 may store software instructions, which, when executed by processing unit 110, determines the status of a traffic light located in a region above vehicle 200. For example, processing unit 110 may determine the status of a traffic light by comparing one or more images to stored patterns. In addition, a change in the status of the traffic light may be determined based on the differences from between the stored patterns and two or more images taken at different times. In another embodiment, status determining module 1804 may store software instructions, which, when executed by processing unit 110, determines the trajectory of moving objects located in a region beside vehicle 200. For example, processing unit 110 may determine the trajectory of the moving objects by comparing a plurality of images to each other. In addition, processing unit 110 may determine whether the moving object is on a possible collision course with vehicle 200.

IPR2025-00944
Tesla EX1002 Page 1308

[0166] In one embodiment, system response module 1806 may store software instructions which, when executed by processing unit 110, causes one or more responses in vehicle 200. One type of system response may include navigational responses, as described above in detail with reference to navigational response module 408. Other types of system responses may involve control throttling system 220, braking system 230, and/or steering system 240. For example, processing unit 110 may transmit electronic signals that cause system 100 to physically depress the brake by a predetermined amount or ease partially off the accelerator of vehicle 200. Further, processing unit 110 may transmit electronic signals that cause system 100 to steer vehicle 200 in a particular direction. Such responses may be based on the determination of the relevant traffic light. Additional system responses may include providing various notifications (e.g., warnings and/or alerts). The notifications may include reporting the status of the traffic light located above vehicle 200, reporting a change in status of the traffic light located above vehicle 200, reporting an existence of an object beside a vehicle, and reporting that the object is on a collision course with vehicle 200.

[0167] FIG. 19 is a flowchart showing an exemplary process 1900 for navigating a vehicle based on a change in status of a traffic light located in a region above vehicle 200, as depicted in FIG. 16 . A similar process is contemplated for controlling a vehicle based on the trajectory of moving objects located in a region beside vehicle 200, as depicted in FIG. 17 . Accordingly, exemplary process 1900 may be applied to a variety of situations, including the situation depicted in FIG. 17 .

[0168] As shown in FIG. 19 , at step 1902, image capture device 122 may acquire a plurality of images of a traffic light located in a region above vehicle 200, outside of the sightline of a typical driver location. At step 1904, processing unit 110 may determine the relevancy of the traffic light captured in the plurality of images based on at least one indicator of the vehicle position, for example, lane markers. In some embodiments, this step may be optional and omitted. At step 1906, processing unit 110 may determine the status of the traffic light based on an analysis of at least a first image from among the plurality of images. At step 1908, processing unit 110 may determine a change in the status of the traffic light based on at least a second image from among the plurality of images. For example, the change in the status of the traffic light may be from a red light to a green light. After determining that the status of a traffic light has changed, at step 1910, processing unit 110 may cause a system response based on the determined change in status. These steps of process 1900 are discussed in further detail below.

[0169] At step 1902, image capture device 122 may acquire a plurality of images of a traffic light located in a region above vehicle 200. Consistent with disclosed embodiments, image capture device 122 may include one or more image capture devices. Consistent with some embodiments, the traffic light may be located outside of a sightline of a typical driver location in vehicle 200. For example, the traffic light may be located in the extended zone as described with reference to FIG. 16 . In some embodiments, image capture device 122 may be located within vehicle 200. In addition, although in some embodiments image capture device 122 may be located behind a rearview mirror and against a vehicle windshield, processing unit 110 may determine the status of the traffic light through the windshield.

[0170] At step 1904, processing unit 110 may determine the relevancy of the traffic light located in the region above vehicle 200 based on at least one indicator of vehicle position. A detailed disclosure of how system 100 may determine the relevancy of

Tesla EX1002 Page 1309 IPR2025-00944

traffic lights is provided below in connection with FIGS. 20-23 . In some embodiments, the indicator of vehicle position may be derived from analyzing the plurality of images. For example, processing unit 110 may determine the relative position of the traffic light to vehicle 200, based on identification of a lane marker in the plurality of images. Additionally or alternatively, the indicator of vehicle position may be derived from geographic position data retrieved from position sensor 130. For example, processing unit 110 may determine the relative position of the traffic light to vehicle 200, based on GPS signals or local positioning signals. In some embodiments, processing unit 110 may compare the GPS acquired vehicle location to map data to determine the relevance of the traffic light. For example, processing unit 110 may conduct an analysis using information derived from image data and the map data. Using the derived information, processing unit 110 may determine a correspondence between the detected traffic light located in the region above vehicle 200 and the lane vehicle 200 is currently driving.

[0171] At step 1906, processing unit 110 may determine the status of the traffic light based on an analysis of at least a first image from among the plurality of images. In some embodiments, the status of the traffic light located in the region above vehicle 200 is associated with the color the traffic light indicates. For example, a steady green light may indicate that vehicle 200 can continue after yielding to other vehicles, bicycles, or pedestrians in the road; a flashing yellow light is a warning and may indicate that vehicle 200 should proceed with caution; and a steady red light may indicate that vehicle 200 needs to stop. Additionally or alternatively, the status of the traffic light may also include additional information associated with the traffic light. The additional information may be the type of the traffic light. For example, the traffic light located in the region above vehicle 200 may be a pedestrian crossing signal. Accordingly, processing unit 110 may determine that pedestrians may cross the road in front of vehicle 200. In some embodiments, in order to determine the status of the traffic light located in the region above vehicle 200, processing device 110 may correct optical distortion in the at least a first image. The optical distortion may occur when image capture device 122 includes an ultra-wide lens.

[0172] At step 1908, processing unit 110 may determine a change in status of the traffic light based on at least a second image from among the plurality of images. In some embodiments, the status of the traffic light located in the region above vehicle 200 may change, for example, from a red light to a green light. In one embodiment, the at least second image may include two or more images subsequent to the at least a first image. In addition, processing unit 110 may process only a portion of the at least a second image. For example, when vehicle 200 is not moving the position of the traffic light in the frame remains the same. Accordingly, in some embodiments, upon identifying the portion of the frame that includes the traffic lamp fixture in the at least a first image, processing unit 110 may determine the change in status by processing only the same portion of the frame in the at least a second image. For example, in some embodiments, by processing only the portion of the frame that includes the traffic lamp fixture, processing unit 110 may save processing power.

[0173] At step 1910, processing unit 110 may cause a system response based on the determined change in status. In some embodiments, processing unit 110 may cause one or more system responses (e.g., two or more responses), including responses of different types. One type of system response may include providing various notifications (e.g., warnings and/or alerts) to the driver of vehicle 200. The notifications may be provided via speakers 360 or via a touch screen 320. For example, the system

Tesla EX1002 Page 1310 IPR2025-00944

response may include reporting the change in status to the driver of vehicle 200. Another type of system response may include navigational responses, the navigational responses may include, for example, accelerating to a driver-specified cruise speed.

Identifying Relevant Traffic Lights

[0174] In some embodiments, system 100 may distinguish between relevant and irrelevant (or less relevant) traffic lights. During a typical driving session, vehicle 200 may cross one or more junctions, such as intersections, having multiple traffic lights. For example, one or more of the traffic lights at an intersection may regulate traffic traveling toward the intersection in a particular direction. Accordingly, one or more traffic lights may regulate whether vehicle 200 may continue to travel through the intersection or whether vehicle 200 must stop at the intersection. However, in addition to the traffic lights that regulate the lane in which vehicle 200 is traveling, other traffic lights that regulate traffic in other lanes may be visible to vehicle 200. Navigating vehicle 200 according to any of the traffic lights that regulate lanes other than the lane in which vehicle 200 is traveling may result in navigational responses undesirable or inapplicable to the intended route of vehicle 20. Accordingly, to enable autonomous control of vehicle 200 in a manner appropriate to the intended navigational path of vehicle 200, system 100 may identify which of a plurality of traffic lights is regulating traffic in the lane in which vehicle 200 is traveling while disregarding (or placing less emphasis on) other traffic lights that regulate other lanes of traffic. Further, after system 100 identifies a relevant traffic light, system 100 may identify a status of the traffic light (e.g., red, yellow, green) and implement an appropriate navigational response. For example, system 100 may discontinue cruise control and apply the brakes when a red light is recognized that regulates the lane in which vehicle 200 is traveling or when a yellow light is recognized that regulates the lane in which vehicle 200 is traveling and vehicle 200 is beyond a predetermined distance of a junction.

[0175] Distinguishing between relevant and irrelevant (or less relevant) traffic lights on a road may be complex. FIG. 20 and FIG. 21 illustrate examples of situations where driving based on information derived from an inapplicable traffic light may be undesirable. In FIG. 20 , vehicle 200a is traveling on a multilane road. Each lane of the road is associated with a different traffic lamp fixture. Vehicle 200a is approaching an intersection and is traveling in a lane designated for proceeding through the intersection and to the other side of the intersection. Also shown in FIG. 20 , vehicle 200b is traveling in a lane that allows traffic to continue to travel straight and through the intersection or that allows traffic to make a right turn. The traffic lamp fixture associated with the lane in which vehicle 200b is traveling includes a right-turn traffic light. As another example, in FIG. 21 , vehicle 200a reaches a junction of non-perpendicular roads. Multiple traffic lamp fixtures (e.g., traffic lamp fixtures 2122 and 2124), including some that do not regulate the lane in which vehicle 200a is traveling, may be visible to vehicle 200a due to the orientation of the junction.

[0176] Returning to FIG. 20 , an intersection 2000 is shown that has the following driving options: road A has lanes leading to roads B, D, and F; road C has lanes leading to roads D and F; road E has lanes leading to roads F, H, and B; and road G has lanes leading to roads H, B, and D. Road A is associated with four traffic lamp fixtures 2012, 2014, 2016, and 2018. In the situation shown, each traffic lamp fixture regulates a different lane. Road C is associated with two traffic lamp fixtures 2022 and 2024. Traffic lamp fixture 2024 includes a traffic light 2024a for continuing straight and a traffic light 2024b for right turns (e.g., displaying a green arrow when a right turn is

Tesla EX1002 Page 1311 IPR2025-00944

authorized). Road E is associated with three traffic lamp fixtures 2032, 2034, end 2036. And road G is associated with two traffic lamp fixtures 2042 and 2044. Traffic lamp fixture 2042 includes a traffic light 2042a for continuing straight and a traffic light 2042b for left turns. In the situation illustrated in FIG. 20 , traffic light 2024b and traffic lamp fixtures 2012, 2018, and 2032 display green lights while all the other traffic lamp fixtures display red lights. Of course, many other road variations and relative traffic light configurations may exist in addition to those shown in FIG. 20 .

[0177] In the situation illustrated in FIG. 20 , vehicle 200a is located on road A in a lane that continues straight through intersection 2000 to road D. However, the field-of-view of an image capture device included in vehicle 200a may include both traffic lamp fixtures 2012 and traffic lamp fixture 2014 (or even additional fixtures). Vehicle 200a arrived at intersection 2000 when both traffic lamp fixtures 2012 and 2014 displayed red lights, and only recently the light in traffic lamp fixture 2012 has turned green. In this situation, it is important that the traffic light detection system 100 of vehicle 200a recognizes that the green light of fixture 2012 is not applicable to vehicle 200a. Rather, system 100 should base any determined navigational response on the status of the lamp associated with the more relevant fixture 2014.

[0178] In another aspect of the situation illustrated in FIG. 20 , vehicle 200b is located on road C in a lane that continues straight through intersection 2000 to road F, and that lane also allows traffic to make a right turn to road D. Vehicle 200b faces traffic lamp fixture 2024 that includes a traffic light 2024a for continuing straight and a traffic light 2024b for right turns. Vehicle 200b arrived at intersection 2000 when traffic lamp fixture 2024 displayed red lights for continuing straight (traffic light 2024a) and for right turns (traffic light 2024b), and only recently traffic light 2024b has turned green. This means that the current status of traffic lamp fixture 2024 prohibits traffic from driving straight and turning right. An undesirable situation might occur if vehicle 200b acts en the status of light 2024b without recognizing and accounting for the status of light 2024a. For example, if vehicle 200b drives to road F (i.e., drives straight) based on the information of traffic light 2024b (showing a green light), vehicles driving from road E to road B may create a hazard for vehicle 200b.

[0179] Turning to FIG. 21, FIG. 21 illustrates an intersection 2100 with non-perpendicular roads that has the following driving options: road A has two lanes, one lane leading to road C and the other lane leading to road E; road B has two lanes leading to road C, and one lane that can also lead to road E; and road D has a single lane leading to road E. Road A is associated with two traffic lamp fixtures 2112 and 211, and each traffic lamp fixture regulates a different lane. Road B is also associated with two traffic lamp fixtures 2122 and 2124, and each regulates a different lane. Road D is associated with a single traffic lamp fixture 2132. In the situation illustrated in FIG. 21 , traffic lamp fixture 2122 displays a green light, while all the other traffic lamp fixtures display red lights. As shown in FIG. 21 , vehicle 200a is located on road A in a lane that continues straight through intersection 2000 to road C. However, the field-of-view of image capture device 122 included in vehicle 200a may include both traffic lamp fixture 2114 and traffic lamp fixture 2122. In this example, vehicle 200a has arrived at intersection 2100 when both traffic lamp fixtures 2114 and 2122 were red, and only recently the light in traffic lamp fixture 2122 has turned green. In such a situation, it may be important that vehicle 200a determines its navigational response based on the status of traffic fixture 2114, rather than on the status of traffic fixture 2122.

Tesla EX1002 Page 1312 IPR2025-00944

[0180] The three situations depicted in FIGS. 20 and 21 provide just a few examples of road situations in which it may be helpful to have a driver assistance system that can distinguish between relevant and irrelevant traffic lamp fixtures, to determine the status of the traffic lights included in the relevant traffic lamp fixture(s), and to generate and take an appropriate navigational response based on the status. Further, these examples demonstrate that the system may need to evaluate multiple traffic lamp fixtures that may face a vehicle in order to identify the traffic lamp fixture that is most applicable to the lane in which the vehicle is traveling or to the intended travel direction of the vehicle (especially where, for example, multiple traffic lights or traffic lamp fixtures may be associated with a single lane in which the vehicle is traveling).

[0181] As discussed above, system 100 may distinguish between relevant and irrelevant traffic lamp fixtures, determine the status of the traffic lights included in the relevant traffic lamp fixture(s), and generate and take an appropriate navigational response based on the status in various driving scenarios. For example, as vehicle 200 approaches an intersection, system 100 may determine which traffic light is relevant, determine a status of that traffic light, and find any other relevant information in images captured by one or more of image capture devices 122-126. If the traffic light is red, system 100 may cause vehicle 200 to apply its brakes. If the traffic light is green, system 100 may cause vehicle 200 to continue. If the traffic light is yellow, system 100 may determine a distance to the intersection and/or an estimate time to the intersection based on analysis of images, the speed of vehicle 200, and/or positional data (e.g., GPS data). If vehicle 200 is within a predetermined time (e.g., five seconds, ten seconds, etc.) and/or distance (e.g., one meter, five meters, ten meters, etc.) threshold, system 100 may cause vehicle 200 to continue. If vehicle 200 is not within the predetermined time threshold and/or distance threshold, system 100 may cause vehicle 200 to stop before reaching the intersection. As another example, when vehicle 200 is stopped at a traffic light, after the traffic light changes its status from red to green, system 100 may cause a navigational response that includes applying the accelerator, releasing the brakes, and steering through an intersection, for example.

[0182] FIG. 22 is an exemplary block diagram of a memory 140/150 configured to store instructions for performing one or more operations, consistent with the disclosed embodiments. Memory 140/150 may be accessed by processing unit 110. As discussed above, processing unit 110 may include various devices, such as a controller, an image preprocessor, a central processing unit (CPU), support circuits, digital signal processors, integrated circuits, memory, or any other types of devices for image processing and analysis. Accordingly, memory 140/150 may be accessed by processing unit 110, may be integrated with processing unit 110, or many included in an embedded system together with processing unit 110. As shown in FIG. 22 , memory 140 or 150 may store a relevancy determining module 2202, a status determining module 2204, and a system response module 2206. The disclosed embodiments are not limited to any particular configuration of memory 140. Further, processing unit 110 may execute the instructions stored in any of modules 2202-2206 included in memory 140 or 150.

[0183] In one embodiment, relevancy determining module 2202 may store software instructions (such as computer vision software) which, when executed by processing unit 110, determines a relevance of each of the plurality of traffic lamp fixtures to vehicle 200. For purposes of this disclosure, determining the relevancy of traffic lamp fixtures may include executing one or more assessments. In one embodiment, processing unit 110 may assess the orientation of each of a plurality of traffic lamp

IPR2025-00944
Tesla EX1002 Page 1313

fixtures with respect to the vehicle. For example, in the situation illustrated in FIG. 21 the field-of-view of image capture device 122 includes both traffic lamp fixture 2114 and traffic lamp fixture 2122. However, the orientation of traffic lamp fixture 2114 suggests that traffic lamp fixture 2114 is more relevant to vehicle 200a than fixture 2122.

[0184] Other assessments may be applicable to the decision making process. For example, processing unit 110 may assess the distance of each of the plurality of traffic lamp fixtures with respect to the vehicle. Additionally or alternatively, processing unit 110 may assess which traffic lamp is facing a front portion of vehicle 200. Additionally or alternatively, processing unit 110 may use the identified lane markers to divide the area forward of vehicle 200 to a plurality of zones, associate each identified traffic lamp fixture with a zone, and assess which zone is the most relevant for vehicle 200. Additionally or alternatively, processing unit 110 may compare a vehicle location that was acquired via GPS to map data to determine the relevance of the traffic light. For example, vehicle 200 may access map data that includes information about the possible driving options at a number of locations. By using the GPS acquired vehicle location, processing unit 110 may determine which driving options are available to vehicle 200 approaching a junction, and use this information to determine the relevance of the traffic lights at the junction to vehicle 200.

[0185] In one embodiment, status determining module 2204 may store software instructions which, when executed by processing unit 110, determine a status of a traffic light included in at least one traffic lamp determined to be relevant to the vehicle. As described above, the status of the traffic light may be also associated with the information the traffic light provides. In a typical traffic light, the color being displayed and the relative location of the illuminated traffic light may provide basic information relevant to vehicle 200.

[0186] Status determining module 2204 may derive additional information from the environment of the relevant traffic lamp fixture and from non-typical traffic light included in the relevant traffic lamp fixture. For example, the relevant traffic lamp fixture may have in its proximity a sign that includes relevant text, e.g., a sign stating specific days and hours. Accordingly, in some embodiments, status determining module 2204 may derive information from the text included in signs associated with the relevant traffic fixture. For example, status determining module 2204 may implement optical character recognition techniques to recognize text in the signs. Status determining module 2204 may then compare the recognized text to a database to determine the information provided by the sign. As another example, the relevant traffic lamp fixture may include a pedestrian crossing signal. Status determining module 2204 may determine that the status of the pedestrian crossing signal means that pedestrians may cross the road in front of vehicle 200, while vehicle 200 has a green light to turn right.

[0187] In one embodiment, system response module 2206 may store software instructions (such as computer vision software) which, when executed by processing unit 110, causes one or more responses in vehicle 200. One type of system response may include one or more navigational responses, as described above in detail with reference to navigational response module 408. Other types of system responses may provide control signals to throttling system 220, braking system 230, and/or steering system 240. For example, processing unit 110 may transmit electronic signals that cause system 100 to physically apply the brakes by a predetermined amount or ease partially off the accelerator of vehicle 200. Further, processing unit 110 may transmit electronic signals that cause system 100 to steer vehicle 200 in a particular direction.

IPR2025-00944
Tesla EX1002 Page 1314

Such responses may be based on the determination of the relevant traffic light. Further, if pedestrians may cross the road in front of vehicle 200, then the system responses may include executing additional processing of the image data. For example, processing unit 110 may confirm that a pedestrian in the proximity of vehicle 200 is not walking in a trajectory that would place the pedestrian in danger while vehicle 200 is moving.

[0188] The location of the traffic lamp fixtures in a junction can be before, after, or in the middle of the junction. Identifying the position of each traffic light fixture in the junction may be used, for example, to determine the relevancy of the traffic lamp fixtures. In some embodiments, processing unit 110 may estimate the distance of one or more traffic lamp fixtures with respect to vehicle 200 to create a 3D model of a junction. In one embodiment, the 3D model of the junction may be stored for future usage. The 3D model of the junction may include one or more of the following: a 3D position for one or more traffic light fixtures, the relative distance between each traffic light fixture and other traffic light fixtures in the junction, the direction(s) each traffic light fixture refers to, and the relative distance between each traffic light fixture and the stop line of the junction. In addition, the 3D model may be periodically updated using details recognizable when vehicle 200 approaches the junction. Examples of the recognizable details may include arrows in traffic lights, the lane marking in the junction, etc. In one example, when time vehicle 200 passes the junction, recognized details are compared to the information stored in the 3D model, and if appropriate, the 3D model is updated.

[0189] Consistent with embodiments of the present disclosure, the 3D model may be used to determine the relevancy of one or more traffic lamp fixtures. In order to activate braking system 230 with sufficient time to stop vehicle 200 before the junction, system 100 may determine which traffic light is relevant to vehicle 200 at a distance of about eighty meters from the junction. As vehicle 200 approaches a junction, information derived from the captured images may be compared to the 3D model to find a match. For example, processing unit 110 may compare the relative distance between recognized traffic light fixtures in the junction to the 3D model to determine the relevancy of one or more traffic light fixtures. Using the 3D model, processing unit 110 may identify a relevant traffic light when vehicle 200 is more than 50 meters, 75 meters, 100 meters, or 125 meters from the junction. As another example, processing unit 110 may compare the relative distance between recognized traffic light fixtures in the junction to the 3D model to determine the distance to the junction's stop line, even when the stop line is not visible from the current location of vehicle 200. Using the 3D model, processing unit 110 may determine the distance to the junction's stop line when vehicle 200 is more than 50 meters, 75 meters, 100 meters, or 125 meters from the junction. Further details regarding relevancy determining module 2202, a status determining module 2204, and a system response module 2206 are provided below in connection with the discussion of FIG. 23 .

[0190] FIG. 23 is a flowchart showing an exemplary process 2300 for identifying relevant traffic lights, consistent with disclosed embodiments. System 100 may implement process 2300 to address the situations depicted in, for example, Figs. 20 and 21 .

[0191] As illustrated in FIG. 23 , at step 2310, image capture device 122 may acquire images of an area forward of vehicle 200. The area may include a plurality of traffic lamp fixtures, each including at least one traffic light. At step 2320, processing unit 110 may determine at least one indicator of vehicle position, for example, based on

IPR2025-00944
Tesla EX1002 Page 1315

identified features from the acquired images (e.g., lane markers). At step 2330, processing unit 110 may use the at least one indicator to determine a relevance of each of the plurality of traffic lamp fixtures. At step 2340, processing unit 110 may determine a status of a traffic light included in a relevant traffic lamp fixture. For example, the status of the traffic light may indicate that vehicle 200 should stop (e.g., due to a red light) or can turn left (if there are no crossing pedestrians). After determining the status of a traffic light included in a relevant traffic lamp fixture, at step 2350, processing unit 110 may cause a system response based on the determined status. These steps of process 2300 are discussed in further detail below.

[0192] For example, at step 2310, image capture device 122 may acquire at least one image of an area forward of vehicle 200, An "area forward of the vehicle" includes any geographical region located in front of a vehicle, relative to its moving direction. The region may include a junction (for example intersections 2000 in FIG. 20 and 2100 in FIG. 21), an intersection, a crossroad, a traffic circle, a street, a road, etc. In some cases, the area may include a plurality of traffic lamp fixtures each including at least one traffic light. A "traffic lamp fixture" includes any form of structure housing one or more light-producing devices used to regulate traffic and/or to provide road-related information. In some cases, two or more traffic lamp fixtures may be joined in a single traffic lamp assembly, but each traffic lamp fixture may be associated with a different lane. A typical traffic lamp fixture may include three circular traffic lights: a green traffic light, a yellow traffic light, and a red traffic light. A "traffic light" includes a device having at least one light source capable of displaying a distinctive color. In some cases, vehicle 200 may encounter a non-typical traffic lamp fixture. A non-typical traffic lamp fixture may include one or more non-circular traffic lights having different colors. For example, a right-turn arrow traffic light, a left-turn arrow traffic light, a public transportation traffic light, a pedestrian crossing traffic light, a bicycle crossing traffic light, etc.

[0193] In some embodiments, processing unit 110 may receive the at least one acquired image via data interface 128. As discussed in detail with reference to FIG. 5D , processing unit 110 may identify traffic lights in the at least one acquired image. For example, processing unit 110 may filter objects identified in the at least one acquired image to construct a set of candidate objects, excluding those objects unlikely to correspond to traffic lights. The filtering may be done based on various properties associated with traffic lights, such as shape, dimensions, texture, position (e.g., relative to vehicle 200), and the like. In addition, processing unit 110 may analyze the geometry of the area forward of vehicle 200. The analysis may be based on any combination of: (i) the number of lanes detected on either side of vehicle 200, (ii) markings (such as arrow marks) detected on the road, and (iii) descriptions of the area extracted from map data. For example, if lane markings defining a lane of travel are recognized on a road, and a traffic light fixture is within the boundaries of the lane markings, system 100 may conclude that the traffic light fixture is associated with the lane associated with the lane markings.

[0194] At step 2320, processing unit 110 may determine at least one indicator of vehicle position from the images. An "indicator of vehicle position" includes any form of information related to the physical location of a vehicle. The indicator of vehicle position may be derived from analyzing the image data (e.g., the at least one acquired image). Additionally, the indicator of vehicle position may be derived from geographic position data (e.g., GPS signals, local positioning signals, and/or map data) or from data indicative of the position of vehicle 200 relative to other vehicles on the road. In some

Tesla EX1002 Page 1316 IPR2025-00944

embodiments, the indicator of vehicle position may include the distance to at least one traffic lamp fixture derived from the at least one image. In other embodiments, the at least one indicator of vehicle position may include a lane marker recognized based on analysis of the at least one image. For example, processing unit 110 may conduct an analysis using information derived from the at least one image to identify one or more lane markers. Using the identified lane markers, processing unit 110 may determine a correspondence between the detected traffic lights and the lane vehicle 200 is currently driving.

[0195] At step 2330, processing unit 110 (e.g., via relevancy determining module 2202) may use the at least one indicator of the vehicle position, as determined at step 2320, to determine the relevancy of each of the plurality of traffic lamp fixtures to vehicle 200. In some embodiments, processing unit 110 may rank the relevancy of the traffic lamp fixtures identified in the at least one acquired image. The traffic lamp fixture having the highest value of relevancy ranking may be determined to be the relevant traffic lamp fixture. For example, in the situation illustrated in FIG. 20 the field-of-view of image capture device 122 includes both traffic lamp fixture 2012 and traffic lamp fixture 2014, but by using one or more assessments, processing unit 110 may determine that traffic lamp fixture 2014 has a higher relevancy ranking value. One way for processing unit 110 to determine that traffic lamp fixture 2014 has a higher relevancy ranking than traffic lamp fixture 2012 is by assessing the distance of each of the plurality of traffic lamp fixtures with respect to vehicle 200a. For example, in the situation depicted in FIG. 20 , traffic lamp fixture 2014 is closer than traffic lamp fixture 2012. Thus, traffic lamp fixture 2014 is more likely to be relevant than traffic lamp fixture 2012. Other ways to determine relevancy are described above in reference to relevancy determining module 2202.

[0196] In some embodiments, the relevancy determination of each of the plurality of traffic lamp fixtures to vehicle 200 may include a preliminary examination to eliminate improbable traffic lamp fixtures. For example, when the at least one acquired image includes three close traffic lamp fixtures and two distant traffic lamp fixtures, the two distant traffic lamp fixtures may be classified as improbable to be relevant to vehicle 200. By eliminating improbable traffic lamp fixtures and ranking the relevancy of a subject of the traffic lamp fixtures identified in the at least one acquired image, processing unit 110 may save processing power. In some embodiments, the relevancy ranking may change when vehicle 200 approaches the junction. For instance, the orientation of the traffic lamp fixture may change from different points on the road, thus, the distance to the junction may impact the probability that a given traffic lamp fixture is relevant to vehicle 200. Accordingly, the relevancy ranking may be associated with a confidence level, which may take into account factors, such as distance to a junction, when assessing traffic lamp fixtures. Further, processing unit 110 may periodically or constantly update the relevancy ranking of the traffic lamp fixtures when the confidence level is below a certain predetermined threshold.

[0197] Relevancy determining module 2202 may further use navigation information previously stored within system 100, such as within memory 140/150, in order to determine a traffic light relevant to vehicle 200. Based on a determined 3D position of vehicle 200 and/or image capture devices 122-126 as discussed above, relevancy determining module 2202 may perform a registration between navigational map data and vehicle 200. Based on the determination of relative distance measurements as described above, relevancy determining module 2202 may use the navigational map data to determine the point in 3D space at which vehicle 200 (via braking system 230)

IPR2025-00944
Tesla EX1002 Page 1317

should brake in order to stop at each detected traffic light fixture at the junction. According to the 3D registration result, relevancy determining module 2202 may determine a lane assignment for each traffic light detected at the intersection using the navigational map data. Relevancy determining module 2202 may then determine the lane assignments within system 100 and image processor 190, then perform the registration.

[0198] At step 2340, processing unit 110 (e.g., via status determining module 2204) may determine, based on the at least one acquired image, a status of a traffic light included in at least one traffic lamp fixture determined to be relevant to vehicle 200. In some cases, the relevant traffic lamp fixture may include a plurality of illuminated traffic lights, and the status of each traffic light may depend on the type of the traffic light. In one embodiment, the status of a traffic light means simply the color it indicates, e.g., green, yellow, and red. System 100 may identify the status of a traffic light using a variety of techniques. For example, system 100 may identify an area of one or more images that include a traffic light and perform an analysis of the pixels in the area to determine the colors of the pixels. After analyzing at least a threshold number of pixels (e.g., two pixels, ten pixels, twenty pixels, etc.) in the area, system 100 may, for example, determine the color of the traffic light by finding an average value of the pixels in the area.

[0199] Additionally or alternatively, the status of the traffic light may include the direction the traffic light refers to. In a typical traffic light, a green color indicates that vehicle 200 is allowed to proceed. However, sometimes this information by itself is insufficient to decide whether it is safe to drive in a certain direction, such as when a green light only authorizes a turn. One way to determine which direction the traffic light refers to includes accessing a database that correlates each traffic light with one or more directions. Another way includes identifying, from the image data, the type of the traffic light and determining from the contextual situation the direction the traffic light refers to. For example, in the second situation depicted in FIG. 20 relative to vehicle 200b, traffic lamp fixture 2024 may be determined as the relevant traffic fixture, but traffic lamp fixture 2024 includes two illuminated traffic lights 2024a and 2024b. Accordingly, in step 2340, processing unit 110 may determine that the status of traffic light 2024a is a red light for continuing straight, and the status of traffic light 2024b is a green light for turning right. In some embodiments, the determination of the status of the traffic light includes one or more of determining a location of the traffic light within a relevant traffic lamp, determining whether the traffic light is illuminated, determining a color of the traffic light, and determining whether the traffic light includes an arrow.

[0200] At step 2350, processing unit 110 (e.g., via system response module 2206) may cause a system response based on the determined status. In some embodiments, processing unit 110 may cause one or more system responses (e.g., two or more responses), including responses of different types. One type of system response may include navigational responses, the navigational response may include, for example, starting to drive, a change in acceleration, a change in velocity, applying vehicle brakes, discontinuing cruise control, and the like. For example, these system responses may include providing control signals to one or more of throttling system 220, braking system 230, and steering system 240 to navigate vehicle 200. Another type of system response may include initiating a timer and/or incrementing a counter in order to provide statistical information about one or more driving sessions to the driver of vehicle 200. For example, the statistical information may indicate how many times vehicle 200 has encountered red lights in a driving session, and/or the duration of a

Tesla EX1002 Page 1318 IPR2025-00944

driving session that vehicle 200 spent waiting at red lights. Another type of system response may include providing various notifications (e.g., warnings and/or alerts) to the driver of vehicle 200. The warnings and/or alerts may include, for example, announcing the color of a relevant traffic light and/or a distance to a junction. The notifications may be provided via speakers 360 or via an associated display (e.g., touch screen 320).

Turn Lane Traffic Light Detection

[0201] System 100 may detect that a vehicle, such as vehicle 200, is traveling in a turn lane. A "lane" may refer to a designated or intended travel path of a vehicle and may have marked (e.g., lines on a road) or unmarked boundaries (e.g., an edge of a road, a road barrier, guard rail, parked vehicles, etc.), or constraints. System 100 may operate to make these detections and determinations based on visual information acquired via one or more image capture devices, for example. In some embodiments, these detections and determinations may also be made at least in part on map data and/or sensed vehicle position. In addition to determining the status of vehicle 200 as being in a turn lane, system 100 may recognize a traffic light associated with the lane, and may be configured to determine the status of the traffic light based on analysis of road context information and determined characteristics of the traffic light.

[0202] FIG. 24 illustrates a vehicle 200 traveling on a roadway 800 in which the disclosed systems and methods for detecting an object in the roadway may be used. Vehicle 200 is depicted as being equipped with image capture devices 122 and 124; more or fewer cameras may be employed on any particular vehicle 200. As shown, roadway 800 may be subdivided into lanes, such as lanes 810 and 820. Lanes 810 and 820 are shown as examples; a given roadway 800 may have additional lanes based on the size and nature of the roadway. In the example of FIG. 24 , vehicle 200 is traveling in lane 820, and can be seen to be approaching intersection 2402 with cross street 2404. Traffic in lane 820 at intersection 2402 is regulated by traffic light fixture 2406. System 100 may, as discussed in detail below, determine the status of traffic light fixture 2406 and cause a system response affecting operations of vehicle 200. Further illustrated on roadway 800 are warning lines 2410 leading to stop line 2412.

[0203] Processing unit 110 may be configured to determine one or more lane constraints associated with each of lanes 810 and 820, intersection 2402, or cross street 2404 based on a plurality of images acquired by one or more of image capture device 122-126 that processing unit 110 may receive via data interface 128. According to some embodiments, the lane constraints may be identified by visible lane boundaries, such as dashed or solid lines marked on a road surface. Additionally or alternatively, the lane constraints may include an edge of a road surface or a barrier. Additionally or alternatively, the lane constraints may include markers (e.g., Botts' dots). According to some embodiments, processing unit 110 (via lane positioning module 2520, described in detail below) may determine constraints associated with lanes 810/820, intersection 2402, or cross street 2404 by identifying a midpoint of a road surface width, such as the entirety of roadway 800 or cross street 2404 or a midpoint of one of lanes 810/820. Processing unit 110 may identity lane constraints in alternative manners, such as by estimation or extrapolation based on known roadway parameters when, for example, lines designating road lanes such as lanes 810/820 are not painted or otherwise labeled. Processing unit 110 may also determine the physical relative distance between the various constraints and detected objects, such as traffic light lamps.

Tesla EX1002 Page 1319 IPR2025-00944

[0204] Distance estimation for a junction with traffic lights may be a challenge when developing autonomous vehicle or red traffic light warning systems, because the location of the traffic light may be after or in the middle of the junction. For example, determining the 3D position for each traffic light at an intersection may imply a braking point for the vehicle but the stop line may provide an accurate position.

[0205] In the present system, detection of the constraints and pathways of roadway 800, cross street 2404, or intersection 2402, as well as constituent lanes 810/820 may include processing unit 110 determining their 3D models via a camera coordinate system. For example, the 3D models of lanes 810/820 may be described by a third-degree polynomial. In addition to 3D modeling of travel lanes, processing unit 110 may perform multi-frame estimation of host motion parameters, such as the speed, yaw and pitch rates, and acceleration of vehicle 200. Processing unit 110 may further determine a road elevation model to transform the information acquired from the plurality of images into 3D space. One feature of the present system is that a global model may be created for static objects at a traffic intersection simultaneously rather than creating one standalone model for every detected object. System 100 may thus be configured to determine absolute distances for objects or lane constraints at a distance, for example, within about one hundred meters. In some cases, system 100 may be configured to determine absolute distances for objects or lane constraints at other distances (e.g., within 125 meters, within 150 meters, etc.). Further, system 100 may be configured to determine which traffic light fixture (e.g., traffic light fixture 2406) is relevant to vehicle 200 within a distance of about eighty meters from stop line 2412.

[0206] Accurate distance estimation to the traffic light(s) at an intersection may be achieved through processes referred to as expansion and scaling. The expansion process may use relative distances to determine a distance from vehicle 200 to an object, such as stop line 2412. For example, if two traffic lamp fixtures are detected in the image data, and the distance between them increases by 5% when the vehicle moves a distance of ten meters, then the system calculates that the distance is 200 meters.

[0207] FIG. 25 is an exemplary block diagram of memory 140 and/or 150, which may store instructions for performing one or more operations consistent with disclosed embodiments. As illustrated in FIG. 25 , memory 140/150 may store one or more modules for performing the object detection and responses described herein. For example, memory 140/150 may store a traffic light detection module 2510, a lane positioning module 2520, and an action module 2530. Application processor 180 and/or image processor 190 may execute the instructions stored in any of modules 2510-2530 included in memory 140/150. One of skill in the art will understand that references in the following discussions to processing unit 110 may refer to application processor 180 and image processor 190 individually or collectively. Accordingly, steps of any of the following processes may be performed by one or more processing devices.

[0208] Traffic light detection module 2510 may store instructions which, when executed by processing unit 110, may detect the presence and status of a traffic light fixture, such as traffic light fixture 2406. As will be discussed below in association with FIGs. 26A , 26B , and 27 , traffic light detection module 2510, along with image processor 190, may perform image processing on one or more images acquired by one or more of image capture devices 122-126. Traffic light detection module 2510 may further determine the status of traffic light fixture 2406, including a determination of whether

IPR2025-00944
Tesla EX1002 Page 1320

one of the lights within traffic light fixture 2406 includes an arrow. Traffic light detection module 2510 may determine other information relevant to the status of traffic light fixture 2406, including but not limited to whether any of the traffic lights associated within traffic light fixture 2406 are illuminated (in either a solid or blinking manner), determining positions of the traffic lights within the traffic lamp fixture (i.e. a horizontal orientation of traffic lights versus a vertical orientation), or determining a color associated with the traffic light. In some embodiments, traffic light detection module 2510 may store information determined for a particular heading, a particular intersection 2402, and/or a particular traffic light fixture 2406 within processing unit 110, such as in memory 140/150. In these embodiments, previously determined and saved information may be used in the future when vehicle 200 returns to the same intersection.

[0209] In some embodiments, a blinking traffic light may be determined through image analysis of multiple images acquired at a predetermined or known capture rate (e.g., analyzing images acquired 1 second, 1.5 seconds, 2 seconds, etc., apart). For example, system 100 may analyze image data to identify a pattern in illumination among a plurality of images. System 100 may further determine a region of a captured image determined to be within the boundaries of a particular traffic light of interest in a traffic light fixture. System 100 may then determine the color of a traffic light through pixel analysis in the region determined to be within the boundaries of the traffic light of interest.

[0210] Lane positioning module 2520 may store instructions which, when executed by processing unit 110, may assist system 100 in determining a position of vehicle 200. In some embodiments, determining the position of vehicle 200 may include determining at least one indicator of vehicle position, either via a visual determination or through analysis of the at least one image received from image capture devices 122-126 via data interface 128. In these embodiments, the at least one indicator of vehicle position may include a distance from the vehicle to one or more lane constraints or lane markers associated with the current lane in which the vehicle is traveling, or to markers associated with an intersection, such as warning lines 2410 and stop line 2412. The distance from the vehicle to a location or object may be determined based on, for example, one or more of an analysis of image data, GPS information, or data from a position sensor. Further, the at least one indicator of vehicle position may include an arrow associated with the traffic lamp fixture having at least one associated traffic light. Additionally or alternatively, the at least one indicator of vehicle position may include a vehicle location acquired by GPS or a like coordinate system. In some embodiments, lane positioning module 2520 may store lane constraint information determined for a particular roadway 800 and its lanes 810/820. For example, processing unit 110 may store the information in memory 140/150. In these embodiments, previously determined and saved information may be used in the future when vehicle 200 returns to the same intersection. For example, GPS information may be used to determine that vehicle 200 has returned to the same intersection.

[0211] Consistent with disclosed embodiments, lane positioning module 2520 may use the information from the at least one indicator of vehicle position to determine if a system response changing the operation of vehicle 200 is required or recommended. Additionally or alternatively, lane positioning module 2520 may receive information from other modules (including position sensor 130) or other systems indicative of the presence of other features in the image data, such as additional vehicles, curvature of the road, etc.

IPR2025-00944
Tesla EX1002 Page 1321

[0212] Action module 2530 may store instructions which, when executed by processing unit 110, may assist system 100 in taking one or more actions relative to the operation of vehicle 200 based on information received from one or more sources, such as position sensor 130, image processor 190, traffic light detection module 2510, or lane positioning module 2520. In some embodiments, action module 2530 may receive information (from, e.g., traffic light detection module 2510) regarding the status of a traffic light at an intersection, such as traffic light fixture 2406 at intersection 2402 discussed above. Other information received by action module 2530 may include a determination of whether the vehicle is in a turn lane, and whether the traffic light fixture 2406 includes an arrow.

[0213] Based on this received information, action module 2530 may then cause a system response affecting the operational status of vehicle 200, such as causing system 100 to provide control signals to one or more of throttling system 220, braking system 230, and steering system 240 to navigate vehicle 200 (e.g., by accelerating, turning, etc.). For example, in some embodiments, system 100 may determine that the turn lane traffic light authorizes vehicle 200 to make a turn. In these embodiments, action module 2530 may send an instruction to steering system 240, and steering system 240 may execute the instruction to turn vehicle 200 through intersection 2402 into a new lane of travel associated with cross street 2404. In other embodiments, the system response initiated and executed by action module 2530 may include any or all of providing a visual or audible notice to the operator of the vehicle, applying vehicle brakes, discontinuing a cruise control function, or initiating one or more automated turning maneuvers.

[0214] Additionally or alternatively, action module 2530 may send instructions to other systems associated with vehicle 200, such as braking system 230, turn signals, throttling system 220, etc. In some embodiments, action module 2530 may instead provide a human operator of the vehicle with audio, visual, or tactile feedback representative of the information gathered from the relevant systems and/or sensors. The human operator may then act on this feedback to turn the vehicle.

[0215] FIG. 26A provides an annotated view of the situation depicted in FIG. 24 . Vehicle 200 is once again traveling in lane 820 of roadway 800, and is approaching intersection 2402 with cross street 2404. Vehicle 200 is again equipped with image capture devices 122 and 124; more or fewer devices may be associated with any particular vehicle 200. For simplicity of illustration, roadway 800 in FIG. 26A is depicted as a one-way street oriented from the top to the bottom of the page with two travel lanes 810 and 820. Cross street 2404 is a two-way street oriented left-right across the page with one travel lane going in each direction.

[0216] To the left side of roadway 800 at intersection 2402 is traffic light fixture 2406. Painted on the surface of lane 820 in front of vehicle 200 is lane arrow 2408, indicating that lane 820 is a left turn lane. Also painted or otherwise affixed on the surface of roadway 800 are warning lines 2410 leading to stop line 2412.

[0217] Consistent with disclosed embodiments, system 100 may be configured to determine whether vehicle 200's travel lane approaching an intersection (here, lane 820) is a turn lane; determine whether a traffic light fixture (here, traffic light fixture 2406) regulates the intersection; determine the status of a traffic light in the traffic light fixture; and determine whether that traffic light includes an arrow. This process will be

Tesla EX1002 Page 1322 IPR2025-00944

described in further detail below in association with FIG. 27 and process 2700.

[0218] Briefly, system 100 associated with vehicle 200 may determine the position of vehicle 200 within roadway 800 via one or more of processing unit 110, position sensor 130, traffic light detection module 2510, or lane position module 2520. Additionally or alternatively, system 100 may gather information from at least one indicator of vehicle position. As discussed above in association with lane positioning module 2520, the at least one indicator of vehicle position may include a distance from vehicle 200 to one or more lane constraints or lane markers associated with the current lane in which the vehicle is traveling (such as lane 820), or to markers associated with an intersection 2401, such as warning lines 2410 and stop line 2412. Further, the at least one indicator of vehicle position may include an arrow associated with traffic light fixture 2406. Additionally or alternatively, the at least one indicator of vehicle position may include a vehicle location acquired by GPS or a like coordinate system.

[0219] In the illustration of FIG. 26A , multiple indicators of the position of vehicle 200 are present. One or more of image capture devices 122-126 associated with vehicle 200 may be configured to capture a plurality of images of the area in front of vehicle 200 that may assist in determining a position of vehicle 200. For example, the images may include lane arrow 2408, which indicates that lane 820 is a left turn lane for intersection 2402. The images may also include traffic light fixture 2406, and may indicate that one or more of the individual lamps within traffic light fixture 2406 includes an arrow suggestive of a turn lane situation. One or more of image capture devices 122-126 may further capture images of warning lines 2410 or stop lines 2412 associated with either roadway 800 or cross street 2404. Further, system 100 via lane positioning module 2520 may determine a position of vehicle 200 within the turn lane by measuring distance s1 from a surface of vehicle 200 to stop line 2412. Still further, system 100 may implement an optical character recognition (OCR) process to obtain text included in one or more captured images (e.g., text from signs and/or road markings). System 100 may then use the text information as part of or as the basis of determining whether vehicle 200 is within a turn lane. For example, system 100 may identify certain words indicative of a turn lane (e.g., "turn, "right," left," etc.).

[0220] The measurement(s) may be direct measurements, such as via position sensor 130, or may be determined via analysis of captured image data by image processor 190 and may be used in connection with map data. The measurement(s) may be measured from any portion of the interior or exterior vehicle 200, including but not limited to the front of vehicle 200, a portion of vehicle 200 such as a headlight or front license plate, a position as-installed of image capture devices 122-126, a determined centroid of vehicle 200, the rear of vehicle 200, one or more windshields or mirrors associated with vehicle 200, wheels of vehicle 200, right or left sides or windows of vehicle 200, a point associated with the roof of vehicle 200, or a point associated with the chassis of vehicle 200.

[0221] In some embodiments, determining the distance from vehicle 200 to traffic light fixture 2406 may be sufficient to assist processing unit 110 in calculating a braking distance or other such measurement. In other embodiments, however, traffic light fixture 2406 may be located past the point at which vehicle 200 would be required to stop. In these embodiments, the distance to the intersection may be determined by using one or more of warning lines 2410 or stop line 2412.

[0222] As discussed above in association with FIG. 24 , these distances may be

Tesla EX1002 Page 1323 IPR2025-00944

calculated using an estimation process employing expansion and scaling. At an intersection 2402 with two traffic light fixtures, for example, system 100 may measure the relative distance between the two fixtures over a period of time as captured in the image data, and then use that relative distance to estimate the distance to stop line 2412. In some embodiments, these measurements may be repeated over time for increased precision. Precision may also be increased by other methods; for example, if three or four traffic light fixtures are situated at intersection 2402 relative distances may be calculated between each of the fixtures and averaged. Additionally or alternatively, the distance may be estimated using a Kalman filter, as discussed above.

[0223] System 100 may still need additional information or inputs to determine the location of a stop line 2412 relative to an intersection 2402 and/or traffic light fixtures 2406. In some embodiments, system 100 may use previously-stored map data gathered from prior trips to the intersection. This information may be received from traffic light detection module 2510, or lane positioning module 2520. A distance measure Z may also be derived from the image data using the equation Z = fW / w Image available on "Original document"

where W is the known distance between two traffic light fixtures 2406, w is the distance in the image data in pixels, and f is the focal length of the particular image capture device 122-126 in pixels.

[0224] In FIG. 26A , inset 2602 shows the two statuses of traffic light fixture 2406 relevant to the example. Based on image data from at least one image received via data interface 128 from image capture devices 122-126, system 100 may be configured to determine the status of traffic light fixture 2406. This determination will be described in further detail below in association with FIG. 27 and process 2700. Upon approaching intersection 2402, system 100 may first determine that traffic light fixture 2406 is displaying a solid red light, as seen in image 1 of inset 2602. In the example of FIG. 26A , lane 820 is a left turn lane, so unless roadway 800 is located in a jurisdiction where left turns on red are legal (or if cross street 2404 were configured as a one way street with both lanes flowing from right to left on the page), vehicle 200 must stop at stop line 2412 and remain stationary until traffic light fixture 2406 is determined to have changed. In alternative embodiments (not shown), vehicle 200 may be traveling in a lane that is determined to be a right turn lane, and assuming that right terms on red are legally permissible in that jurisdiction, system 100 may cause a system response enabling vehicle 200 to turn right while traffic light fixture 2406 remains red.

[0225] Subsequently, as seen in image 2 of inset 2602, traffic light fixture 2406 shifts from a solid red light to a solid green arrow indicating a protected left turn. System 100 may be configured to detect not only that the color of the light has changed, but also that an arrow has appeared within the lamp based on pixel-level image analysis by image processor 190 of acquired images of traffic light fixture 2406. In some embodiments, in order for system 100 to detect that the color of a traffic light has changed, system 100 may analyze captured images over a period of time (e.g., 1 second, 5 seconds, 10 seconds, 15, seconds, 20 seconds, 30 seconds, 1 minute, etc.) in order to recognize the status change of the traffic light. System 100 may further determine once traffic light fixture 2406 changes to display the green arrow whether it is safe or authorized to proceed in the direction indicated by the status of the light. Even though the light may show a green arrow, various situations (not shown) may require vehicle 200 to further delay before completing the turn. For example, a pedestrian may be crossing cross street 2404. Although roadway 800 in FIG. 26A is a one-way street,

Tesla EX1002 Page 1324 IPR2025-00944

in alternative embodiments (such as that described below and shown in FIG. 26B), the road may be a two-way street and vehicle 200 may further assess whether a second vehicle coming from the opposite direction is making a right turn on red onto cross street 2404 from the other direction. Further, other obstacles may prevent a timely left turn; for example, a set of railroad tracks may be oriented parallel to roadway 800 to its left side, and although a left turn green arrow may be illuminated, barriers associated with the train track may be down to allow a train to pass through. These examples are not intended to be limiting. Once system 100 determines that both traffic light fixture 2406 and the situation at the intersection enables a safe turn, system 100 via action module 2530 may cause a system response involving one or more of throttling system 220, braking system 230, steering system 240, turn signals, and/or other subsystems and modules that automatically causes vehicle 200 to turn left.

[0226] FIG. 26B illustrates an additional intersection scenario consistent with disclosed embodiments. In the example of FIG. 26B , vehicle 200 (here represented as vehicle 200e) is now traveling in lane 810 of roadway 800, and is approaching intersection 2402 with cross street 2404. Vehicle 200 is again equipped with image capture devices 122 and 124; more or fewer devices may be associated with any particular vehicle 200. Unlike the illustration of FIG. 26A , roadway 800 is now a two-way street in FIG. 26B , and traffic may proceed in both directions. Cross street 2404 remains a two-way street oriented left-right across the page with one travel lane going in each direction.

[0227] To the right side of roadway 800 at intersection 2402 is traffic light fixture 2604. Painted on the surface of lane 810 in front of vehicle 200 is lane arrow 2408, indicating that lane 810 is a left turn lane. As before, consistent with disclosed embodiments, system 100 may be configured to determine whether vehicle 200's travel lane approaching an intersection (this time, lane 810) is a turn lane; determine whether a traffic light fixture (here, traffic light fixture 2604) regulates the intersection; determine the status of the traffic light; and determine whether that traffic light includes an arrow.

[0228] System 100 may operate to make these determinations in much the same manner as described above. However, in the example of FIG. 26B , inset 2606 demonstrates that as vehicle 200e approaches intersection 2402, traffic light fixture 2604 is displaying a flashing or blinking yellow light. The non-solid nature is indicated in inset 2606 by the dashed lines of the arrow. As may be familiar to one of ordinary skill in the relevant art, some jurisdictions use blinking yellow lights to essentially serve the same functions as "Yield" or "Stop" signs during certain times of the day, while still permitting the lanes of travel to be fully stopped with a red light or fully authorized to proceed with a green light.

[0229] As outlined above, system 100 may determine that traffic light fixture 2604 is displaying a blinking yellow light, and may further determine that lane 810 is a turn lane. With the situation depicted in FIG. 26B , however, system 100 may be required to make additional determinations in order for vehicle 200e to successfully traverse intersection 2402. As one of ordinary skill would understand, a blinking yellow light authorizes an approaching vehicle to turn at will, but only if it is safe to do so. Drivers are expected to use caution at the intersection and determine that no potential collisions or other hazards would result from making the turn too early.

[0230] One such determination that system 100 may make is whether one or more vehicles are approaching the intersection from the other direction. For example, while travel lanes traveling one direction of a roadway may have a traffic light indicating a

IPR2025-00944
Tesla EX1002 Page 1325

blinking yellow light in a turn lane, traffic traveling the opposite direction may have a green light to proceed straight through the intersection. Accordingly, a collision could occur if the operator of the vehicle with the blinking yellow signal fails to yield. In the example of FIG. 26B , a second vehicle 200f is indeed passing straight through intersection 2402. System 100 may be configured to detect the presence of vehicle 200f via image data captured from one or more of image capture devices 122-126, as well as determine that vehicle's location within the roadway/intersection and its velocity using one or more of position sensor 130, image processor 190, traffic light detection module 2510, or lane positioning module 2520. Upon detecting and analyzing vehicle 200f, system 100 (via action module 2530) may cause braking system 230 to remain engaged until vehicle 200f has passed.

[0231] Once no additional vehicles are approaching the intersection from the opposite direction, system 100, via action module 2530, may cause a system response based on that determination. The system response may include one or more of providing feedback to an operator of vehicle 200, applying vehicle brakes via braking system 230, discontinuing cruise control, or initiating an automated turning maneuver onto cross street 2404 using one or more of throttling system 220, braking system 230, or steering system 240. First, however, system 100 may analyze image data for the presence of any additional obstacles or objects associated with intersection 2402 or cross street 2404. For example, in the illustration of FIG. 26B , pedestrian 2608 can be seen walking across cross street 2404, since the other lane of roadway 800 has a green light in that direction. System 100 may detect the presence of pedestrian 2608 via image processor 190 by analyzing one or more of the plurality of images captured by one or more of image capture devices 122-126. As described above, other subsystems and modules of system 100 may coordinate to detect the velocity of pedestrian 2608, and may notify system 100 when pedestrian 2608 is no longer in the roadway and it is safe to turn.

[0232] FIG. 27 illustrates a process 2700 for detecting the status of a traffic light, consistent with disclosed embodiments. Steps of process 2700 may be performed by one or more of processing unit 110, image acquisition unit 120, position sensor 130, image processor 190, traffic light detection module 2510, lane positioning module 2520, or action module 2530.

[0233] At step 2710, process 2700 may include acquiring, using at least one image capture device 122, 124, and/or 126, a plurality of images of an area forward of vehicle 200. The area may contain one or more traffic lamp fixtures having at least one traffic light, such as traffic light fixtures 2406 and 2604 described above. Processing unit 110 may receive the plurality of images from the image capture device(s) through data interface 128. The plurality of images may then be processed in real time by image processor 190 of processing unit 110. Image processor 190 may process at least one image captured by the at least one image capture device(s), as well as additional images captured at later points in time.

[0234] At step 2720, process 2700 may determine, based on at least one indicator of vehicle position, whether vehicle 200 is traveling in a turn lane. As described generally above, various indicators of position may be used to determine whether a travel lane is a turn lane. In some embodiments, the position indication may include a vehicle location acquired by GPS. In other embodiments, system 100 may recognize lane markers or constraints, such as lane arrow 2408 indicating that the lane is a turn lane. Alternatively, image processor 190 may determine that an arrow is associated with the

traffic lamp fixture regulating the lane travel. Additionally, measurements of distance may be made, using various elements of system 100. For example, position sensor 130 and/or lane positioning module 2520 may send information to processing unit 110 regarding the relative position of vehicle 200 within lane 810/820, including how far vehicle 200 is from intersection 2402 (distance s1 described above). The end of the lane may be based on a measured distance or a visual determination (e.g., by analyzing one or more image images acquired by one or more of image capture devices 122-126 and detecting the presence of warning lines 2410 or stop line 2412).

[0235] At step 2730, process 2700 may perform image processing on the at least one image received via data interface 128 to determine whether an arrow exists in the image data associated with the traffic light. Traffic light detection module 2510 and/or image processor 190 may be configured to analyze the pixels of the image corresponding to the locations representative of the individual traffic lights within the traffic lamp fixture. One or more image processing algorithms may be applied to the pixels to determine if either a solid or blinking/flashing arrow is present in the relevant regions. Further examples of how system 100 may analyze one or more images to determine whether an arrow is displayed by a traffic light are discussed below in connection with FIGs. 28-30B .

[0236] At step 2740, process 2700 may determine a status of the traffic light. As discussed above, determining the status of the traffic light may include, but not be limited to, determining whether the traffic light is illuminated (and/or functional at all), determining a position of the illuminated traffic light within the traffic lamp fixture, or determining a color associated with the traffic light (i.e. red, yellow, or green). This list is intended as exemplary only, and any information relevant to the traffic light may be determined and processed by system 100 and processing unit 110. For example, the system may be configured to determine whether the traffic light is solid or flashing based on image processing over a series of acquired images. Image processor 190 may determine in these embodiments that the same traffic light within the fixture is illuminated in one image, darkened in the next in the sequence, and then immediately illuminated again; this would indicate that the particular light is flashing. In some embodiments, system 100 may be configured to recognize traffic light situations particular to the specific jurisdiction in which the vehicle is traveling. This information may be pre-programmed into system 100, such as in a database stored within memory 140/150 (not shown).

[0237] At step 2750, process 2700 may cause a system response via system 100 impacting the operation of vehicle 200. The response initiated by system 100 may be based on any or all of the determination of the status of the traffic light, whether the traffic light includes an arrow, and whether the vehicle is in a turn lane. As discussed above, the particular action(s) initiated by system 100 via action module 2530 may include, but not be limited to, providing a visual or audible feedback notice to an operator of vehicle 200 regarding the traffic light, applying vehicle brakes via braking system 230, discontinuing a cruise control function, or proceeding to initiate an automated turning maneuver and proceeding through the intersection using one or more of throttling system 220, braking system 230, or steering system 240.

[0238] Additionally, in some embodiments as part of the function of system 100, action module 2530 and/or processing unit 110 may be further configured to determine whether the turn lane in which the vehicle is located is authorized to proceed in the direction indicated by the determined status of the traffic light. In other words, even if

IPR2025-00944
Tesla EX1002 Page 1327

the traffic light indicates that a particular system response is legally permissible, system 100 may provide the occupants of vehicle 200 an extra level of safety protection by further determining whether such an action is actually possible and safe. For example, as discussed above in association with FIG. 26B , system 100 and image processor 190 may be configured to detect the presence of other obstacles associated with the roadway(s) or the intersection, such as pedestrians present in the street or the presence of other vehicles. In some embodiments, the other vehicles may be proceeding normally through the intersection, such as in the scenario depicted in FIG. 26B where vehicle 200e has a blinking yellow light. In these embodiments, vehicle 200e could theoretically turn onto cross street 2404, but system 100 may further account for the fact that vehicle 200f is proceeding on a green light through the intersection. Alternatively, a second vehicle may be present in an unusual or illegal manner, such as if the operator of the other vehicle ran through a stop sign or a red light, or if a vehicle has become disabled in the intersection or one of the roadways.

[0239] In still other embodiments, determining whether the turn lane in which the vehicle is located is authorized to proceed in the direction indicated by the determined status of the traffic light may further comprise determining whether to actually follow the indicated status of the light based on a predetermined destination for the vehicle. In embodiments involving autonomous vehicles, it may be presumed that the operators and/or passengers have indicated their desired destination to the processing unit 110 of system 100. Accordingly, using GPS location technology or other such information, system 100 may be configured to determine if making the turn indicated by a traffic light is actually relevant to the predetermined destination. For example, in some embodiments the vehicle 200 may have ended up in the turn lane unintentionally, and proceeding to the predetermined destination may involve merging to an adjacent lane (with the help of lane positioning module 2520) and continuing to proceed straight through the intersection. Accordingly, system 100 may be configured to determine whether to turn left, turn right, or continue straight through the intersection based on this predetermined destination information.

Traffic Light Detail Detection

[0240] In some embodiments, system 100 may provide additional functionality related to the recognition of traffic signals. For example, system 100 may provide a traffic light detail detection function to provide feedback based on analysis of the physical features of a traffic light. To implement such functionality, processing unit 110 may process images captured by at least one of image capture devices 122, 124, 126 and employ a super resolution technique to recognize a specific feature of a traffic light in the images, where the specific feature may indicate useful information to system 100. For example, system 100 may recognize and distinguish between round and arrow shaped signals of a traffic light that have been illuminated.

[0241] In some embodiments, memory 140 and/or 150 may store instructions programmed such that upon execution by a processing device will provide the traffic light detail detection function. As shown in FIG. 28 , memory 140 and/or 150 may store an image alignment module 2810, a pixel expansion module 2820, an image analysis module 2830, and a database 2840. Image alignment module 2810 may store instructions for aligning a plurality of images captured by image capture devices 122, 124, 126. Pixel expansion module 2820 may store instructions for increasing the resolution of at least a portion of an image to allow images to be scaled to a standard size. Image analysis module 2830 may store instructions for analyzing the images to

Tesla EX1002 Page 1328 IPR2025-00944

detect and identify features of the image. Database 2830 may be configured to store data associated with providing traffic light detail detection functionality and provide data when requested. Further, image alignment module 2810, pixel expansion module 2820, and image analysis module 2830 may store instructions executable by one or more processors (e.g., processing unit 110), alone or in various combinations with each other. For example, image alignment module 2810, pixel expansion module 2820, and image analysis module 2830 may be configured to interact with each other and/or other modules of system 100 to perform functions consistent with disclosed embodiments.

[0242] Database 2840 may include one or more memory devices that store information and are accessed and/or managed through a computing device, such as processing unit 110. In some embodiments, database 2840 may be located in memory 150, as shown in FIG. 28 . In other embodiments, database 2840 may be located remotely from memory 150, and be accessible to other components of system 100 (e.g., processing unit 120) via one or more wireless connections (e.g., a wireless network). While one database 2840 is shown, it should be understood that several separate and/or interconnected databases may make up database 2840. Database 630 may include computing components (e.g., database management system, database server, etc.) configured to receive and process requests for data stored in memory devices associated with database 2840 and to provide data from database 2840 (e.g., to processing unit 110).

[0243] In some embodiments, database 2840 may be configured to store data associated with providing traffic light detail detection functionality. For example, database 2840 may store data, such as images, maps, algorithms, sets of values, or the like, which may allow processing unit 110 to identify information detected in an image. For example, database 2840 may store a set of average pixel values corresponding to each of a round signal and an arrow signal of a traffic light.

[0244] In some embodiment, system 100 may provide feedback based on a detail of a traffic light detected in one or more capture images. For example, system 100 may be configured to differentiate between colors, shapes, and positions of a signal associated with a traffic light in order to provide specific feedback based on a meaning of the detected detail (e.g., an arrow signal controls turning movement, while a round signal control straight forward movement).

[0245] FIGS. 29A and 29B depict exemplary objects that may be detected by system 100. FIG. 29A depicts a traffic lamp fixture 2900A that includes a round signal 2910. FIG. 29B depicts a traffic lamp fixture that includes an arrow signal 2920. While both objects include traffic lights, they include different details that indicate different messages. In an exemplary embodiment, system 100 may be configured to analyze the pixels of one or more images to identify differentiating features, such as a round signal 2910 and arrow signal 2920.

[0246] FIG. 29C depicts an exemplary portion of an image 2915 of round signal 2910, including a plurality of pixels 2930. System 100 may be configured to analyze pixels 2930 to identify an object and/or detail found in the image. For example, image analysis module 2830 may be configured to identify features such as brightness, color, contrast, tint, opacity, and the like, associated with each pixel 2930, and determine an object and/or detail associated with the image based on the pixel analysis. Each feature that image analysis module 2830 identifies may be given a value. For example, image analysis module 2830 may assign a value of one to pixels that are identified as the

color red and a value of zero to any other pixel.

[0247] In some embodiments, system 100 may be configured to use a super resolution technique to allow for a more accurate analysis of an object in plurality of images. For example, pixel expansion module 2820 may modify image 2915 to divide one or more of pixels 2930 into a subset of pixels 2940. While only some pixels 2930 are shown as expanded into pixels 2940 in FIG. 29C , it should be understood that any portion or all of pixels 2930 may be divided in this way. Thus, pixel expansion module 2820 may increase the number of pixels associated with a particular portion of an image.

[0248] In an exemplary embodiment, pixel expansion module 2820 may increase the number of pixels associated with a particular object in an image such that the section of the image including the object includes a selected number of pixels (e.g., a standard number of pixels). In this way, even if a plurality of images include the same object at different sizes (e.g., because the images were taken at different distances from the object), the images can be modified such that the portions of the images that include the object include the same number of pixels, which can be directly compared and averaged to more reliably determine a feature (e.g., color) of the part of the object located within the pixel.

[0249] Scaling selected portions of images to associate the same number of pixels with an object in each image may allow the images to be directly compared (e.g., compare the value of one pixel in an image with the same pixel in another image). In this way, multiple images may be considered and analyzed by image analysis module 2830, thus increasing accuracy as compared to analysis of one image. For example, the super resolution technique may increase the likelihood that image analysis module 2830 may correctly distinguish between round signal 2920 and arrow signal 2930 in images.

[0250] FIG. 29D depicts an exemplary portion of an image 2925 of arrow signal 2920. As described above with respect to image 2915, image 2925 may be divided into pixels 2930, which each may be expanded into a subset of pixels 2940. As can be seen in FIGS. 29C and 29D , pixels 2930 differ around the edges of the respective signals 2920, 2930. Image analysis module 2830 may be configured to identify one or more characteristics of pixels 2930 to determine whether the pixels correspond to round signal 2910 or arrow signal 2920. Expansion of pixels 2930 into pixels 2940 may allow multiple images to be aligned and considered, thus increasing the reliability of detection.

[0251] FIG. 30A is a flowchart showing an exemplary process 3000A for determining whether a traffic light includes an arrow, consistent with disclosed embodiments. In an exemplary embodiment, system 100 may perform process 3000A to determine whether a portion of a traffic light detected in a plurality of images includes an arrow signal. An arrow signal (or "arrow") may be a type of signal of a traffic light in which a visual component of the signal takes the shape of an arrow. The arrow may be a signal indicating that a vehicle turn (e.g., a turn in a direction of the arrow) is allowable at that time.

[0252] At step 3002, processing unit 110 may receive a plurality of images of a traffic light. For example, one or more of image capture devices 122, 124, 126 may acquire a plurality of images of an area forward of vehicle 200. In some instances, such as when vehicle 200 is approaching a junction, the area forward of vehicle 200 may include one or more traffic lamp fixtures that include at least one traffic light. In these instances, one

or more of image capture devices 122, 124, 126 may acquire a plurality of images of the traffic light. Processing unit 110 may receive the plurality of images from image capture devices 122, 124,126 through data interface 128.

[0253] Processing unit 110 may analyze the received plurality of images. In some embodiments, processing unit 110 may analyze the plurality of images to determine that the plurality of images include a traffic light. For example, processing unit 110 may recognize one or more objects as a candidate for being a traffic light and perform one or more processes (e.g., image matching) to determine whether the candidate corresponds to a traffic light. It should be understood, however, that processing unit 110 may perform process 3000A without the distinct step of identifying an object as a traffic light. In addition, processing unit 110 may analyze portions of the plurality of images to determine characteristics of the traffic light. These characteristics may help system 100 determine a response to the traffic light (e.g., the light is green - move through the junction, the light is red - stop at the junction).

[0254] At step 3004, processing unit 110 (e.g., via image alignment module 2810) may align areas of the plurality of images that correspond to the traffic light. For example, processing unit 110 may use an area of a traffic light as a common point to align the plurality of images. The alignment may allow processing unit 110 to analyze a common area of multiple images.

[0255] In order to align the areas of the plurality of images that correspond to the traffic light, processing unit 110 may select a trait or characteristic of a traffic light that may allow for an area of interest to be identified. In an exemplary embodiment, processing unit 110 may align areas of the plurality of images corresponding to the traffic light based on a determined center of brightness (e.g., of a lit signal of the traffic light). In some embodiments, the center of brightness may be determined by finding, for example, a center of mass of brightness of a traffic light. In other embodiments, processing unit 110 may use color, shape, position, or any other trait or characteristic of a traffic light to align the plurality of images. Further, in some instances, processing unit 110 may rotate the areas of the plurality of images to align the images in the plane of the image (e.g., to match the rotational position of the areas).

[0256] As described above, the aligned areas of the plurality of images may correspond to an area of interest of the traffic light. For example, by aligning the plurality of images at areas corresponding to the traffic light based on brightness, a lit signal of the traffic light may be identified and placed within the aligned areas. A lit signal may indicate to a viewer an action that is acceptable at that time (e.g., stop, go, turn, etc.). Therefore, aligning areas that correspond to a lit signal may help processing unit 110 isolate an area of interest of the traffic light.

[0257] In step 3006, processing unit 110 (e.g., via pixel expansion module 2820) may expand pixels of the plurality of images within the aligned areas to obtain a standard size (e.g., the same number of pixels). For example, processing unit 110 may perform an image modification process in which processing unit 110 replaces each pixel corresponding to the aligned areas of the plurality images with a plurality of pixels. In this way, the aligned areas of one or more images may be expanded to a standard size. The expansion of the pixels may allow for direct comparison of pixels between the plurality of images.

[0258] In some embodiments, processing unit 110 may expand the pixels of the

IPR2025-00944
Tesla EX1002 Page 1331

plurality of images within the aligned areas by replacing each pixel with a matrix of pixels. Each pixel in the matrix of pixels may be attributed with the same value as the parent pixel. In this way, the expanded pixels may also include a value associated with a feature of the image at the pixel (e.g., color, brightness, etc.).

[0259] In one embodiment, processing unit 110 may expand each pixel within the aligned areas into a 3 x 3 matrix of pixels. In another embodiment, processing unit 110 may expand each pixel within the aligned areas into a 4 x 4 matrix of pixels. In other embodiments, other matrix sizes may be used (e.g., 5 x 5, 6 x 6, 8 x 8, etc.).

[0260] In some embodiments, pixel expansion module 2820 may expand pixels based on a scaling factor (e.g., 2, 2.5, 3, 4, 5, etc.) that transforms each pixel into a number of pixels corresponding to the scaling factor. Pixel expansion module 2820 may determine a scaling factor based on a size of an object in an image and a standard size that is used. In this way, the objects in the images may be standardized to include the same number of pixels.

[0261] Processing unit 110 may analyze the pixels across the plurality of images. For example, in step 3008, processing unit 110 (e.g., via image analysis module 2830) may determine a set of average pixel values including an average pixel value for each pixel, including expanded pixels. As described above, expanded pixels include values associated with a parent pixel. The pixel values may be any value that defines a selected pixel and allows processing unit 110 to determine a characteristic of the traffic light in the aligned areas. For example, the pixel values may correspond to a color or brightness value associated with each pixel. In another embodiment, the pixel values may be binary values indicating whether each pixel (or pixel matrix average) includes a color or brightness value above a threshold value. Processing unit 110 may determine an average pixel value for each pixel across all of the plurality of images.

[0262] Processing unit 110 may use the pixel values to determine a characteristic of the traffic light in the aligned areas. For example, in step 3010, processing unit 110 may determine whether the traffic light includes an arrow based on the set of average pixel values. In one example, processing unit 110 may compare the set of average pixel values to stored criteria (e.g., information stored in database 2840) to determine whether the pixel values correspond to an arrow. In another example, processing unit may map the average pixel values and compare it to a stored map to determine whether the pixel values correspond to the shape of an arrow.

[0263] In some embodiments, processing unit 110 may use an accuracy factor to determine whether the pixel values correspond to an arrow. For example, processing unit 110 may determine whether the set of average pixel values match stored criteria (e.g., a set of pixel values, a stored image, a pixel map, etc.) to within a degree of accuracy greater than a threshold value. In some embodiments, the accuracy factor may change depending on a quality of the images. For example, a relatively low accuracy factor may be used if vehicle 200 is far away from a traffic light and/or if only one of image capture devices 122, 124, 126 is able to capture images of the traffic light. Similarly, a relatively high accuracy factor may be used if vehicle 200 is close to the traffic light and/or more than one (or all) of image capture devices 122, 124, 126 are able to capture images of the traffic light. In some embodiments, image analysis module 2830 may classify an object for one or more features. For example, image analysis module 2830 may classify an object as an arrow or round signal.

IPR2025-00944
Tesla EX1002 Page 1332

[0264] Through exemplary process 3000A, system 100 may be configured to identify an arrow signal associated with a traffic light. The ability to identify the arrow signal provides additional functionality to system 100 that enhances the capabilities of system 100. In particular, since an arrow signal may indicate different information than a non-arrow signal (e.g., a plain round light) system 100 may be configured to provide a particularized system response (e.g., acceptable to turn left at this time, not acceptable to turn left at this time, etc. that may otherwise be incorrect (e.g., green indicates go in all directions). In another example, system 100 may differentiate between two identified traffic signal lights (e.g., a green round light and a red arrow light) to determine which traffic signal light should be followed. An exemplary process for using the detection of an arrow signal to achieve this functionality is described in more detail below.

[0265] FIG. 30B is a flowchart showing an exemplary process 3000B for causing a system response based on the detection of an arrow signal in a traffic light, consistent with disclosed embodiments. In an exemplary embodiment, processing unit 110 may perform process 3000B to use a result of process 3000A to control vehicle 200. For example, based on a detection of an arrow signal in a traffic light, processing unit 110 may command vehicle 200 to accelerate, turn, stop, change lanes, etc.

[0266] In step 3012, processing unit 110 may identify a traffic signal light. For example, processing unit 110 may identify a traffic signal light in one or more images captured by one or more image capture devices 122, 124, 126. In some instances, more than one traffic signal light may be identified in an image (e.g., a green round light and a red arrow light).

[0267] At step 3014, processing unit 110 may determine a color of the identified traffic signal light. Processing unit 110 may use a color identification process to analyze one or more of the plurality of images of the traffic light to determine and categorize a color of the traffic signal light. For example, processing unit 110 may match image data to stored criteria to determine a color from a plurality of color possibilities (e.g., green, red, yellow, orange, etc.).

[0268] In one embodiment, processing unit 110 may identify a color of the pixels that include a pixel value indicating that the pixel is part of the shape of the traffic signal light. In another embodiment, one or more image capture devices 122, 124, 126 may use clear pixels and red (or other color) pixels to assist with color differentiation. For example, in some embodiments, when a red light is present in an image, both the red and clear pixels may be used, whereas when a green (or other color) light is present, only the clear pixels may be used while the red pixels are ignored (e.g., defined as black, no color, not red, etc.). For example, when the light is red, the red pixels may match the color and be "used" to determine the color. As another example, when the light is green, the clear pixels may be used to identify the color and/or the red pixels are ignored to define the color as not red. In this way, processing unit 110 may differentiate between colors in images depending on the type of pixels that are used to define a value of the light within the pixel. Further, since the images may be aligned and averaged with other images to achieve a higher resolution image, values for ignored red pixels (or other pixels missing values) may be supplied based on information from other images.

[0269] It should be understood that the foregoing are examples and other color identification processes and/or techniques may be used. In some embodiments, processing unit 110 may be configured to determine one or more other characteristics

IPR2025-00944
Tesla EX1002 Page 1333

of the arrow and/or traffic signal. For example, processing unit 110 may determine whether the arrow is blinking and/or steady. In another example, processing unit 110 may determine whether the arrow has recently changed from another signal. For example, the plurality of images may include two identified signals, a previous signal and a current signal (e.g., the identified arrow). These other characteristics may provide processing unit 110 with additional information that may be helpful in determining how to interpret the arrow signal.

[0270] In step 3018, processing unit 110 may cause one or more system responses based on the color of the arrow (and/or other characteristic of the arrow and/or traffic signal). In one example, the system response may correspond to an indication of the arrow (e.g., green arrow - acceptable to turn, red arrow - not acceptable to turn, etc.). In another example, the system response may correspond to an indication that a particular traffic signal does not apply because it is an arrow (e.g., the vehicle is traveling straight). The system response may be a navigational response configured to change or continue an operating parameter of vehicle 200. For example, processing unit 110 may provide a notice to the driver of the vehicle (e.g., an audible sound or visual indicator via user interface 170), apply vehicle brakes, discontinue cruise control, and/or initiate an automated turning maneuver. For example, processing unit 110 may provide control signals to one or more of throttling system 220, braking system 230, and steering system 240 to navigate vehicle 200. In some embodiments, processing unit 110 may continue to analyze images to determine whether the system response should be interrupted (e.g., the arrow is no longer green - discontinue cruise control and apply brakes).

[0271] Through the exemplary disclosed processes 3000A and 3000B, system 100 may have enhanced functionality in providing autonomous control of vehicle 200. In particular, system 100 may be configured to react to a traffic signal based on detailed information provided by the traffic signal and provide an appropriate system response that allows for safe operation of vehicle 200. Further, while the above processes have been described in the context of an arrow signal, it should be understood that one or more of the steps thereof may be used in similar processes to determine a system response to other objects and/or signals, such as pedestrian signals, no-turn-on-red signals, fixed road signs, street signs, traffic advisory boards, billboards, etc.

[0272] The foregoing description has been presented for purposes of illustration. It is not exhaustive and is not limited to the precise forms or embodiments disclosed. Modifications and adaptations will be apparent to those skilled in the art from consideration of the specification and practice of the disclosed embodiments. Additionally, although aspects of the disclosed embodiments are described as being stored in memory, one skilled in the art will appreciate that these aspects can also be stored on other types of computer readable media, such as secondary storage devices, for example, hard disks or CD ROM, or other forms of RAM or ROM, USB media, DVD, Blu-ray, or other optical drive media.

[0273] Computer programs based on the written description and disclosed methods are within the skill of an experienced developer. The various programs or program modules can be created using any of the techniques known to one skilled in the art or can be designed in connection with existing software. For example, program sections or program modules can be designed in or by means of .Net Framework, .Net Compact Framework (and related languages, such as Visual Basic, C, etc.), Java, C++, Objective-C, HTML, HTML/AJAX combinations, XML, or HTML with included Java

Tesla EX1002 Page 1334 IPR2025-00944

applets.

[0274] Moreover, while illustrative embodiments have been described herein, the scope of any and all embodiments having equivalent elements, modifications, omissions, combinations (e.g., of aspects across various embodiments), adaptations and/or alterations as would be appreciated by those skilled in the art based on the present disclosure. The limitations in the claims are to be interpreted broadly based on the language employed in the claims and not limited to examples described in the present specification or during the prosecution of the application. The examples are to be construed as non-exclusive, Furthermore, the steps of the disclosed methods may be modified in any manner, including by reordering steps and/or inserting or deleting steps. It is intended, therefore, that the specification and examples be considered as illustrative only, with a true scope being indicated by the following claims.

Espacenet

Claims: EP3100206 (A1) — 2016-12-07

SYSTEMS AND METHODS FOR LANE END RECOGNITION

Claims of EP3100206 (B1)

A high quality text as facsimile in your desired language may be available amongst the following family members:

CN106462727 (A) US2015210274 (A1) WO2015116950 (A1) CN111199218 (A) US2015210275 (A1) US2015210276 (A1) US2015210277 (A1) US2015210278 (A1) US2015210312 (A1) US2016167704 (A1) US2016283807 (A1) US2016306361 (A1) US2019294179 (A1) US2020225681 (A1)

- Original claims
- Claims tree

The EPO does not accept any responsibility for the accuracy of data and information originating from other authorities than the EPO; in particular, the EPO does not guarantee that they are complete, up-to-date or fit for specific purposes.

1. A lane ending detection system (100) for a vehicle (200), the system comprising:
at least one image capture device (122, 124, 126) configured to acquire at least one image of an area forward of the vehicle, the area including at least one road sign providing information indicative of a lane ending;
a data interface; and
at least one processing device configured to:
receive (1110) the at least one image via the data interface;
extract (1120) lane ending information from a representation of the at least one road sign included in the at least one image;
determine (1130), based on at least one indicator of position of the vehicle, a distance from the vehicle to one or more marked or unmarked lane boundaries associated with the current lane in which the vehicle is traveling, wherein the indicator of position includes information obtained by direct or indirect measurement of the position of the vehicle;
determine (1140), based on the extracted lane ending information and based on the determined distance from the vehicle to the one or more marked or unmarked lane boundaries, that the current lane in which the vehicle is traveling is ending;
cause (1150) the vehicle to change lanes to another lane that is not ending based on the determination of that the current lane in which the vehicle is traveling is ending; and
update, based on the at least one indicator of position of the vehicle, the distance from the vehicle to one or more marked or unmarked lane boundaries associated with the lane that is not ending.

2. The lane ending detection system of claim 1, wherein the at least one indicator of position of the vehicle includes a visual determination of which lane among a plurality of lanes the vehicle is traveling.

3. The lane ending detection system of claim 1, wherein the extracted lane ending information includes:
an identification of which lane from among a plurality of lanes is ending;
a distance to the lane ending; and/or
a curvature of the roadway prior to the lane ending.

4. The lane ending detection system of claim 3, wherein the extracted lane ending information includes the curvature of the roadway prior to the lane ending, indicating the urgency of a need to move the vehicle from the current lane.

5. The lane ending detection system of claim 1, wherein the at least one indicator of position of the vehicle includes a visual determination of a distance to the lane ending.

6. The lane ending detection system of claim 1, wherein the at least one indicator of position of the vehicle includes; a visual determination of a distance from the vehicle to a second vehicle; optionally wherein the second vehicle is traveling in front of the vehicle.

7. The lane ending detection system of claim 1, wherein causing the vehicle to change lanes includes sending an instruction to a steering system of the vehicle.

8. The lane ending detection system of claim 1, wherein the at least one processing device is further configured to:
determine, from at least one image, that traffic prevents causing the vehicle to change lanes; and
in response to the determination, apply brakes of the vehicle and/or cause the vehicle to decelerate until causing the vehicle to change lanes to another lane that is not ending is safe.

9. A vehicle (200), comprising:
a vehicle body; and
a lane ending detection system (100) according to any preceding claim.

10. A method for detecting a lane ending for a vehicle, the method comprising:
acquiring (1110), via at least one image capture device (122, 124, 126), at least one image of an area forward of the vehicle (200), the area including at least one road sign providing information indicative of the lane ending;
extracting (1120) lane ending information from a representation of the at least one road sign included in the at least one image;
determining (1130), based on at least one indicator of position of the vehicle, a distance from the vehicle to one or more marked or unmarked lane boundaries associated with the current lane in which the vehicle is traveling, wherein the indicator of position includes information obtained by direct or indirect measurement of the position of the vehicle;
determining (1140), based on the extracted lane ending information and based on the determined distance from the vehicle to the one or more marked or unmarked lane boundaries, that a current lane in which the vehicle is traveling is ending;
causing (1150) the vehicle to change lanes to another lane that is not ending based on

IPR2025-00944
Tesla EX1002 Page 1337

the determination of that the current lane in which the vehicle is traveling is ending; and updating, based on the at least one indicator of position of the vehicle, the distance from the vehicle to one or more marked or unmarked lane boundaries associated with the lane that is not ending.

11. The method of claim 10, wherein the at least one indicator of position of the vehicle includes a visual determination of which lane among a plurality of lanes the vehicle is traveling.

12. The method of claim 10, wherein the extracted lane ending information includes:
an identification of which lane from among a plurality of lanes is ending;
a distance to the lane ending; and/or
a curvature of the roadway prior to the lane ending.

13. The method of claim 12, wherein the extracted lane ending information includes the curvature of the roadway prior to the lane ending, indicating the urgency of a need to move the vehicle from the current lane.

14. The method of claim 10, wherein the at least one indicator of position of the vehicle includes a visual determination of a distance to the lane ending.

15. The method of claim 10, wherein the at least one indicator of position of the vehicle includes a visual determination of a distance from the vehicle to a second vehicle; optionally wherein the second vehicle is traveling in front of the vehicle.

16. The method of claim 10, wherein causing the vehicle to change lanes includes sending an instruction to a steering system of the vehicle.

17. The method of claim 10, further comprising:
determining, from at least one image, that traffic prevents causing the vehicle to change lanes; and
in response to the determination, applying brakes of the vehicle and/or causing the vehicle to decelerate until causing the vehicle to change lanes to another lane that is not ending is safe.

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Veröffentlichungsnummer:
(11) Publication number: **EP 3 100 206 A0**
(11) Numéro de publication:

Internationale Anmeldung veröffentlicht durch die Weltorganisation für geistiges Eigentum unter der Nummer:

WO 2015/116950 (Art. 153(3) EPÜ).

International application published by the World Intellectual Property Organization under number:

WO 2015/116950 (Art. 153(3) EPC).

Demande internationale publiée par l'Organisation Mondiale de la Propriété Intellectuelle sous le numéro:

WO 2015/116950 (art. 153(3) CBE).

Espacenet

Bibliographic data: EP3108264 (A2) — 2016-12-28

ADVANCED DRIVER ASSISTANCE SYSTEM BASED ON RADAR-CUED VISUAL IMAGING

Inventor(s): ROSENBLUM KEVIN [IL]; DAGAN EREZ [IL] ± (ROSENBLUM, Kevin, ; DAGAN, Erez)

Applicant(s): MOBILEYE VISION TECHNOLOGIES LTD [IL] ± (Mobileye Vision Technologies Ltd)

Classification: **- international:** ***G01S13/86***
- cooperative: **G01S13/86 (EP, CN, US)**; **G01S13/867 (EP, CN, US)**; **G01S13/931 (EP, CN, US)**; **G06K9/00791 (EP, US)**; **G06T7/215 (EP, US)**; G01S13/426 (EP, CN, US); G01S13/87 (EP, CN, US); G01S2013/9318 (EP, CN, US); G01S2013/93185 (EP, CN, US); G01S2013/9319 (EP, CN, US); G01S2013/9322 (EP, CN, US); G01S2013/93273 (EP, CN, US); G06T2207/30261 (EP, US); G06T7/00 (US)

Application number: EP20150726290 20150220 Global Dossier

Priority number(s): US201461942088P 20140220 ; WO2015IB00673 20150220

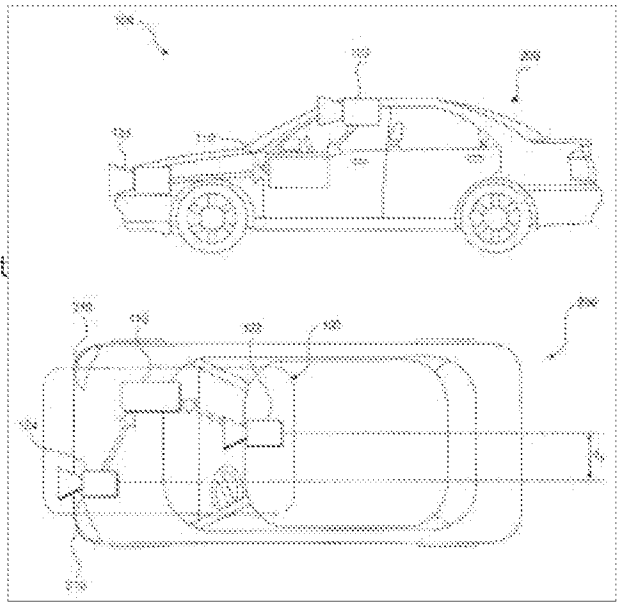
Also published as: CN106461774 (A) CN106461774 (B) CN109884618 (A) US10274598 (B2) US10690770 (B2) more

Abstract not available for EP3108264 (A2)
Abstract of corresponding document: US2015234045 (A1)

A navigation system for a vehicle may include at least one image capture device configured to acquire a plurality of images of an environment of a vehicle and a radar sensor to detect an object in the environment of the vehicle and to provide and output including range information indicative of at least one of a range or range rate between the vehicle and the object. The system may also include at least one processing device programmed to: receive the plurality of images from the at least one image capture device; receive the output from the radar sensor; determine, for each of a plurality of image segments in a first image, from among the plurality of images, and corresponding image segments in a second image, from among the plurality of images, an indicator of optical flow; use range information determined based on the output of

IPR2025-00944
Tesla EX1002 Page 1340

the radar sensor together with the indicators of optical flow determined for each of the plurality of image segments in the first image and the corresponding image segments in the second image to calculate for each of a plurality of imaged regions at least one value indicative of a focus of expansion; identify a target object region, including at least a subset of the plurality of imaged regions that share a substantially similar focus of expansion; and cause a system response based on the identified target object region.



Espacenet

Description: EP3108264 (A2) — 2016-12-28

ADVANCED DRIVER ASSISTANCE SYSTEM BASED ON RADAR-CUED VISUAL IMAGING

Description not available for EP3108264 (A2)
Description of corresponding document: US2015234045 (A1)

A high quality text as facsimile in your desired language may be available amongst the following family members:

CN106461774 (A) US2015234045 (A1) WO2015125022 (A2) CN109884618 (A) US2017246694 (A1) US2019235073 (A1)

The EPO does not accept any responsibility for the accuracy of data and information originating from other authorities than the EPO; in particular, the EPO does not guarantee that they are complete, up-to-date or fit for specific purposes.

CROSS REFERENCE TO RELATED APPLICATION

[0001] This application claims the benefit of priority of U.S. Provisional Patent Application No. 61/942,088, filed on Feb. 20, 2014. The foregoing application is incorporated herein by reference in its entirety.

BACKGROUND

[0002] I. Technical Field

[0003] The present disclosure relates generally to vehicle navigation and, more specifically, to systems and methods that fuse radar and visual information to navigate relative to detected objects.

[0004] II. Background Information

[0005] Various sensors may be employed, whether in a driver assist system or an autonomous vehicle, in order to aid in navigation of the vehicle. Such sensors may include radar units and cameras, among others. These sensors may collect information from the environment of a vehicle and use the collected information to make navigational decisions relative to various objects, hazards, etc. present in the environment of the vehicle. For example, the sensors may collect information associated with objects in a roadway, other vehicles, light poles, guard rails,

pedestrians, among others. Information collected from the onboard sensor systems may enable driver assist or autonomous driving systems to recognize objects in a vehicle's environment and take appropriate action to reduce or minimize the risk of collisions.

SUMMARY

[0006] Embodiments consistent with the present disclosure provide systems and methods for autonomous vehicle navigation. The disclosed embodiments may use cameras and/or radar to provide autonomous vehicle navigation features.

[0007] Consistent with a disclosed embodiment, a navigation system for a vehicle may include at least one image capture device configured to acquire a plurality of images of an environment of a vehicle and a radar sensor to detect an object in the environment of the vehicle and to provide and output indicative of a range and/or range rate between the vehicle and the object. The system may also include at least one processing device programmed to: receive the plurality of images from the at least one image capture device; receive the output from the radar sensor; determine, for each of a plurality of image segments in a first image, from among the plurality of images, and corresponding image segments in a second image, from among the plurality of images, an indicator of optical flow; determine a value indicative of an expected optical inflation associated with the object in the plurality of images based at least upon range and/or range rate information determined from the output of the radar sensor; determine one or more areas of the second image where the value indicative of an expected optical inflation associated with the object substantially matches the indicator of optical flow; identify a target object region, including at least a subset of the determined one or more areas of the second image; and cause a system response based on the identified target object region.

[0008] The disclosed embodiments may also include a vehicle comprising a navigation system. The navigation system may include at least one image capture device configured to acquire a plurality of images of an environment of a vehicle and a radar sensor to detect an object in the environment of the vehicle and to provide and output indicative of a range and/or range rate between the vehicle and the object. The system may also include at least one processing device programmed to: receive the plurality of images from the at least one image capture device; receive the output from the radar sensor; determine, for each of a plurality of image segments in a first image, from among the plurality of images, and corresponding image segments in a second image, from among the plurality of images, an indicator of optical flow; determine a value indicative of an expected optical inflation associated with the object in the plurality of images based at least upon range and/or range rate information determined from the output of the radar sensor; determine one or more areas of the second image where the value indicative of an expected optical inflation associated with the object substantially matches the indicator of optical flow; identify a target object region, including at least a subset of the determined one or more areas of the second image; and cause a system response based on the identified target object region.

[0009] The disclosed embodiments may also include a method for navigating a vehicle based on radar detection and visual image data. The method may include acquiring a plurality of images of an environment of the vehicle; detecting an object in the environment of the vehicle using a radar sensor and providing an output of the radar sensor indicative of a range and/or range rate between the object and the vehicle;

Tesla EX1002 Page 1343 IPR2025-00944

receiving at one or more processing devices the plurality of images from the at least one image capture device and the output from the radar sensor; determining an indicator of optical flow, for each of a plurality of image segments in a first image, from among the plurality of images, and corresponding image segments in a second image, from among the plurality of images; identifying a target object region having a substantially common focus of expansion, based on range information provided by the radar sensor and the determined indicators of optical flow; and causing a system response based on the identified target object region.

[0010] The disclosed embodiments may include a navigation system for a vehicle. The system may include at least one image capture device configured to acquire a plurality of images of an environment of a vehicle; a radar sensor to detect an object in the environment of the vehicle and to provide and output indicative of a range and/or range rate between the vehicle and the object; and at least one processing device. The at least one processing device may be programmed to receive the plurality of images from the at least one image capture device; receive the output from the radar sensor; determine, for each of a plurality of image segments in a first image, from among the plurality of images, and corresponding image segments in a second image, from among the plurality of images, an indicator of optical flow; identify a target object region having a substantially common focus of expansion, based on range and/or range rate information provided by the radar sensor and the determined indicators of optical flow; and cause a system response based on the identified target object region.

[0011] Consistent with other disclosed embodiments, non-transitory computer-readable storage media may store program instructions, which are executed by at least one processing device and perform any of the methods described herein.

[0012] The foregoing general description and the following detailed description are exemplary and explanatory only and are not restrictive of the claims.

BRIEF DESCRIPTION OF THE DRAWINGS

[0013] The accompanying drawings, which are incorporated in and constitute a part of this disclosure, illustrate various disclosed embodiments. In the drawings:

[0014] FIG. 1 is a diagrammatic representation of an exemplary system consistent with the disclosed embodiments.

[0015] FIG. 2A is a diagrammatic side view representation of an exemplary vehicle including a system consistent with the disclosed embodiments.

[0016] FIG. 2B is a diagrammatic top view representation of the vehicle and system shown in FIG. 2A consistent with the disclosed embodiments.

[0017] FIG. 2C is a diagrammatic top view representation of another embodiment of a vehicle including a system consistent with the disclosed embodiments.

[0018] FIG. 2D is a diagrammatic top view representation of yet another embodiment of a vehicle including a system consistent with the disclosed embodiments.

[0019] FIG. 2E is a diagrammatic top view representation of yet another embodiment of a vehicle including a system consistent with the disclosed embodiments.

Tesla EX1002 Page 1344 IPR2025-00944

[0020] FIG. 2F is a diagrammatic representation of exemplary vehicle control systems consistent with the disclosed embodiments.

[0021] FIG. 3A is a diagrammatic representation of an interior of a vehicle including a rearview mirror and a user interface for a vehicle imaging system consistent with the disclosed embodiments.

[0022] FIG. 3B is an illustration of an example of a camera mount that is configured to be positioned behind a rearview mirror and against a vehicle windshield consistent with the disclosed embodiments.

[0023] FIG. 3C is an illustration of the camera mount shown in FIG. 3B from a different perspective consistent with the disclosed embodiments.

[0024] FIG. 3D is an illustration of an example of a camera mount that is configured to be positioned behind a rearview mirror and against a vehicle windshield consistent with the disclosed embodiments.

[0025] FIG. 4 is an exemplary block diagram of a memory configured to store instructions for performing one or more operations consistent with the disclosed embodiments.

[0026] FIG. 5A is a flowchart showing an exemplary process for causing one or more navigational responses based on monocular image analysis consistent with disclosed embodiments.

[0027] FIG. 5B is a flowchart showing an exemplary process for detecting one or more vehicles and/or pedestrians in a set of images consistent with the disclosed embodiments.

[0028] FIG. 5C is a flowchart showing an exemplary process for detecting road marks and/or lane geometry information in a set of images consistent with the disclosed embodiments.

[0029] FIG. 5D is a flowchart showing an exemplary process for detecting traffic lights in a set of images consistent with the disclosed embodiments.

[0030] FIG. 5E is a flowchart showing an exemplary process for causing one or more navigational responses based on a vehicle path consistent with the disclosed embodiments.

[0031] FIG. 5F is a flowchart showing an exemplary process for determining whether a leading vehicle is changing lanes consistent with the disclosed embodiments.

[0032] FIG. 6 is a flowchart showing an exemplary process for causing one or more navigational responses based on stereo image analysis consistent with the disclosed embodiments.

[0033] FIG. 7 is a flowchart showing an exemplary process for causing one or more navigational responses based on an analysis of three sets of images consistent with the disclosed embodiments.

Tesla EX1002 Page 1345 IPR2025-00944

[0034] FIG. 8 is a diagrammatic representation of an exemplary system consistent with the disclosed embodiments.

[0035] FIG. 9A is a diagrammatic side view representation of an exemplary vehicle including a system consistent with the disclosed embodiments.

[0036] FIG. 9B is a diagrammatic top view representation of the vehicle and system shown in FIG. 9A consistent with the disclosed embodiments.

[0037] FIG. 9C is a diagrammatic top view representation of another embodiment of a vehicle including a system consistent with the disclosed embodiments.

[0038] FIG. 9D is a diagrammatic top view representation of yet another embodiment of a vehicle including a system consistent with the disclosed embodiments.

[0039] FIG. 9E is a diagrammatic top view representation of yet another embodiment of a vehicle including a system consistent with the disclosed embodiments.

[0040] FIG. 10 is a diagrammatic representation of an environment in which a vehicle may travel.

[0041] FIG. 11A represents an image of an object within an environment of the vehicle.

[0042] FIG. 11B represents another image of the object within an environment of the vehicle, from which optical flow may be determined relative to the image of FIG. 11A.

[0043] FIG. 12A represents a horizontal optical flow map relative to the object of FIG. 11A.

[0044] FIG. 12B represents a vertical optical flow map relative to the object of FIG. 11A.

[0045] FIG. 13 represents a target object region derived from optical flow information and radar range information according to exemplary disclosed embodiments.

DETAILED DESCRIPTION

[0046] The following detailed description refers to the accompanying drawings. Wherever possible, the same reference numbers are used in the drawings and the following description to refer to the same or similar parts. While several illustrative embodiments are described herein, modifications, adaptations and other implementations are possible. For example, substitutions, additions or modifications may be made to the components illustrated in the drawings, and the illustrative methods described herein may be modified by substituting, reordering, removing, or adding steps to the disclosed methods. Accordingly, the following detailed description is not limited to the disclosed embodiments and examples. Instead, the proper scope is defined by the appended claims.

[0047] FIG. 1 is a block diagram representation of a system 100 consistent with the exemplary disclosed embodiments. System 100 may include various components depending on the requirements of a particular implementation. In some embodiments, system 100 may include a processing unit 110, an image acquisition unit 120, a

Tesla EX1002 Page 1346 IPR2025-00944

position sensor 130, one or more memory units 140, 150, a map database 160, and a user interface 170. Processing unit 110 may include one or more processing devices. In some embodiments, processing unit 110 may include an applications processor 180, an image processor 190, or any other suitable processing device. Similarly, image acquisition unit 120 may include any number of image acquisition devices and components depending on the requirements of a particular application. In some embodiments, image acquisition unit 120 may include one or more image capture devices (e.g., cameras), such as image capture device 122, image capture device 124, and image capture device 126. System 100 may also include a data interface 128 communicatively connecting processing device 110 to image acquisition device 120. For example, data interface 128 may include any wired and/or wireless link or links for transmitting image data acquired by image accusation device 120 to processing unit 110.

[0048] Both applications processor 180 and image processor 190 may include various types of processing devices. For example, either or both of applications processor 180 and image processor 190 may include a microprocessor, preprocessors (such as an image preprocessor), graphics processors, a central processing unit (CPU), support circuits, digital signal processors, integrated circuits, memory, or any other types of devices suitable for running applications and for image processing and analysis. In some embodiments, applications processor 180 and/or image processor 190 may include any type of single or multi-core processor, mobile device microcontroller, central processing unit, etc. Various processing devices may be used, including, for example, processors available from manufacturers such as Intel®, AMD®, etc. and may include various architectures (e.g., x86 processor, ARM®, etc.).

[0049] In some embodiments, applications processor 180 and/or image processor 190 may include any of the EyeQ series of processor chips available from Mobileye®. These processor designs each include multiple processing units with local memory and instruction sets. Such processors may include video inputs for receiving image data from multiple image sensors and may also include video out capabilities. In one example, the EyeQ2® uses 90 nm-micron technology operating at 332 Mhz. The EyeQ2® architecture consists of two floating point, hyper-thread 32-bit RISC CPUs (MIPS32® 34K® cores), five Vision Computing Engines (VCE), three Vector Microcode Processors (VMP®), Denali 64-bit Mobile DDR Controller, 128-bit internal Sonics Interconnect, dual 16-bit Video input and 18-bit Video output controllers, 16 channels DMA and several peripherals. The MIPS34K CPU manages the five VCEs, three VMP™ and the DMA, the second MIPS34K CPU and the multi-channel DMA as well as the other peripherals. The five VCEs, three VMP® and the MIPS34K CPU can perform intensive vision computations required by multi-function bundle applications. In another example, the EyeQ3®, which is a third generation processor and is six times more powerful that the EyeQ2®, may be used in the disclosed embodiments.

[0050] Any of the processing devices disclosed herein may be configured to perform certain functions. Configuring a processing device, such as any of the described EyeQ processors or other controller or microprocessor, to perform certain functions may include programming of computer executable instructions and making those instructions available to the processing device for execution during operation of the processing device. In some embodiments, configuring a processing device may include programming the processing device directly with architectural instructions. In other embodiments, configuring a processing device may include storing executable instructions on a memory that is accessible to the processing device during operation.

Tesla EX1002 Page 1347 IPR2025-00944

For example, the processing device may access the memory to obtain and execute the stored instructions during operation.

[0051] While FIG. 1 depicts two separate processing devices included in processing unit 110, more or fewer processing devices may be used. For example, in some embodiments, a single processing device may be used to accomplish the tasks of applications processor 180 and image processor 190. In other embodiments, these tasks may be performed by more than two processing devices.

[0052] Processing unit 110 may comprise various types of devices. For example, processing unit 110 may include various devices, such as a controller, an image preprocessor, a central processing unit (CPU), support circuits, digital signal processors, integrated circuits, memory, or any other types of devices for image processing and analysis. The image preprocessor may include a video processor for capturing, digitizing and processing the imagery from the image sensors. The CPU may comprise any number of microcontrollers or microprocessors. The support circuits may be any number of circuits generally well known in the art, including cache, power supply, clock and input-output circuits. The memory may store software that, when executed by the processor, controls the operation of the system. The memory may include databases and image processing software. The memory may comprise any number of random access memories, read only memories, flash memories, disk drives, optical storage, tape storage, removable storage and other types of storage. In one instance, the memory may be separate from the processing unit 110. In another instance, the memory may be integrated into the processing unit 110.

[0053] Each memory 140, 150 may include software instructions that when executed by a processor (e.g., applications processor 180 and/or image processor 190), may control operation of various aspects of system 100. These memory units may include various databases and image processing software. The memory units may include random access memory, read only memory, flash memory, disk drives, optical storage, tape storage, removable storage and/or any other types of storage. In some embodiments, memory units 140, 150 may be separate from the applications processor 180 and/or image processor 190. In other embodiments, these memory units may be integrated into applications processor 180 and/or image processor 190.

[0054] Position sensor 130 may include any type of device suitable for determining a location associated with at least one component of system 100. In some embodiments, position sensor 130 may include a GPS receiver. Such receivers can determine a user position and velocity by processing signals broadcasted by global positioning system satellites. Position information from position sensor 130 may be made available to applications processor 180 and/or image processor 190.

[0055] User interface 170 may include any device suitable for providing information to or for receiving inputs from one or more users of system 100. In some embodiments, user interface 170 may include user input devices, including, for example, a touchscreen, microphone, keyboard, pointer devices, track wheels, cameras, knobs, buttons, etc. With such input devices, a user may be able to provide information inputs or commands to system 100 by typing instructions or information, providing voice commands, selecting menu options on a screen using buttons, pointers, or eye-tracking capabilities, or through any other suitable techniques for communicating information to system 100.

Tesla EX1002 Page 1348 IPR2025-00944

[0056] User interface 170 may be equipped with one or more processing devices configured to provide and receive information to or from a user and process that information for use by, for example, applications processor 180. In some embodiments, such processing devices may execute instructions for recognizing and tracking eye movements, receiving and interpreting voice commands, recognizing and interpreting touches and/or gestures made on a touchscreen, responding to keyboard entries or menu selections, etc. In some embodiments, user interface 170 may include a display, speaker, tactile device, and/or any other devices for providing output information to a user.

[0057] Map database 160 may include any type of database for storing map data useful to system 100. In some embodiments, map database 160 may include data relating to the position, in a reference coordinate system, of various items, including roads, water features, geographic features, businesses, points of interest, restaurants, gas stations, etc. Map database 160 may store not only the locations of such items, but also descriptors relating to those items, including, for example, names associated with any of the stored features. In some embodiments, map database 160 may be physically located with other components of system 100. Alternatively or additionally, map database 160 or a portion thereof may be located remotely with respect to other components of system 100 (e.g., processing unit 110). In such embodiments, information from map database 160 may be downloaded over a wired or wireless data connection to a network (e.g., over a cellular network and/or the Internet, etc.).

[0058] Image capture devices 122, 124, and 126 may each include any type of device suitable for capturing at least one image from an environment. Moreover, any number of image capture devices may be used to acquire images for input to the image processor. Some embodiments may include only a single image capture device, while other embodiments may include two, three, or even four or more image capture devices. Image capture devices 122, 124, and 126 will be further described with reference to FIGS. 2B-2E, below.

[0059] System 100, or various components thereof, may be incorporated into various different platforms. In some embodiments, system 100 may be included on a vehicle 200, as shown in FIG. 2A. For example, vehicle 200 may be equipped with a processing unit 110 and any of the other components of system 100, as described above relative to FIG. 1. While in some embodiments vehicle 200 may be equipped with only a single image capture device (e.g., camera), in other embodiments, such as those discussed in connection with FIGS. 2B-2E, multiple image capture devices may be used. For example, either of image capture devices 122 and 124 of vehicle 200, as shown in FIG. 2A, may be part of an ADAS (Advanced Driver Assistance Systems) imaging set.

[0060] The image capture devices included on vehicle 200 as part of the image acquisition unit 120 may be positioned at any suitable location. In some embodiments, as shown in FIGS. 2A-2E and 3A-3C, image capture device 122 may be located in the vicinity of the rearview mirror. This position may provide a line of sight similar to that of the driver of vehicle 200, which may aid in determining what is and is not visible to the driver. Image capture device 122 may be positioned at any location near the rearview mirror, but placing image capture device 122 on the driver side of the mirror may further aid in obtaining images representative of the driver's field of view and/or line of sight.

[0061] Other locations for the image capture devices of image acquisition unit 120 may

IPR2025-00944
Tesla EX1002 Page 1349

also be used. For example, image capture device 124 may be located on or in a bumper of vehicle 200. Such a location may be especially suitable for image capture devices having a wide field of view. The line of sight of bumper-located image capture devices can be different from that of the driver and, therefore, the bumper image capture device and driver may not always see the same objects. The image capture devices (e.g., image capture devices 122, 124, and 126) may also be located in other locations. For example, the image capture devices may be located on or in one or both of the side mirrors of vehicle 200, on the roof of vehicle 200, on the hood of vehicle 200, on the trunk of vehicle 200, on the sides of vehicle 200, mounted on, positioned behind, or positioned in front of any of the windows of vehicle 200, and mounted in or near light figures on the front and/or back of vehicle 200, etc.

[0062] In addition to image capture devices, vehicle 200 may include various other components of system 100. For example, processing unit 110 may be included on vehicle 200 either integrated with or separate from an engine control unit (ECU) of the vehicle. Vehicle 200 may also be equipped with a position sensor 130, such as a GPS receiver and may also include a map database 160 and memory units 140 and 150.

[0063] FIG. 2A is a diagrammatic side view representation of an exemplary vehicle imaging system consistent with the disclosed embodiments. FIG. 2B is a diagrammatic top view illustration of the embodiment shown in FIG. 2A. As illustrated in FIG. 2B, the disclosed embodiments may include a vehicle 200 including in its body a system 100 with a first image capture device 122 positioned in the vicinity of the rearview mirror and/or near the driver of vehicle 200, a second image capture device 124 positioned on or in a bumper region (e.g., one of bumper regions 210) of vehicle 200, and a processing unit 110.

[0064] As illustrated in FIG. 2C, image capture devices 122 and 124 may both be positioned in the vicinity of the rearview minor and/or near the driver of vehicle 200. Additionally, while two image capture devices 122 and 124 are shown in FIGS. 2B and 2C, it should be understood that other embodiments may include more than two image capture devices. For example, in the embodiments shown in FIGS. 2D and 2E, first, second, and third image capture devices 122, 124, and 126, are included in the system 100 of vehicle 200.

[0065] As illustrated in FIG. 2D, image capture device 122 may be positioned in the vicinity of the rearview mirror and/or near the driver of vehicle 200, and image capture devices 124 and 126 may be positioned on or in a bumper region (e.g., one of bumper regions 210) of vehicle 200. And as shown in FIG. 2E, image capture devices 122, 124, and 126 may be positioned in the vicinity of the rearview mirror and/or near the driver seat of vehicle 200. The disclosed embodiments are not limited to any particular number and configuration of the image capture devices, and the image capture devices may be positioned in any appropriate location within and/or on vehicle 200.

[0066] It is to be understood that the disclosed embodiments are not limited to vehicles and could be applied in other contexts. It is also to be understood that disclosed embodiments are not limited to a particular type of vehicle 200 and may be applicable to all types of vehicles including automobiles, trucks, trailers, and other types of vehicles.

[0067] The first image capture device 122 may include any suitable type of image capture device. Image capture device 122 may include an optical axis. In one instance,

Tesla EX1002 Page 1350 IPR2025-00944

the image capture device 122 may include an Aptina M9V024 WVGA sensor with a global shutter. In other embodiments, image capture device 122 may provide a resolution of 1280×960 pixels and may include a rolling shutter. Image capture device 122 may include various optical elements. In some embodiments one or more lenses may be included, for example, to provide a desired focal length and field of view for the image capture device. In some embodiments, image capture device 122 may be associated with a 6 mm lens or a 12 mm lens. In some embodiments, image capture device 122 may be configured to capture images having a desired field-of-view (FOV) 202, as illustrated in FIG. 2D. For example, image capture device 122 may be configured to have a regular FOV, such as within a range of 40 degrees to 56 degrees, including a 46 degree FOV, 50 degree FOV, 52 degree FOV, or greater. Alternatively, image capture device 122 may be configured to have a narrow FOV in the range of 23 to 40 degrees, such as a 28 degree FOV or 36 degree FOV. In addition, image capture device 122 may be configured to have a wide FOV in the range of 100 to 180 degrees. In some embodiments, image capture device 122 may include a wide angle bumper camera or one with up to a 180 degree FOV.

[0068] The first image capture device 122 may acquire a plurality of first images relative to a scene associated with the vehicle 200. Each of the plurality of first images may be acquired as a series of image scan lines, which may be captured using a rolling shutter. Each scan line may include a plurality of pixels.

[0069] The first image capture device 122 may have a scan rate associated with acquisition of each of the first series of image scan lines. The scan rate may refer to a rate at which an image sensor can acquire image data associated with each pixel included in a particular scan line.

[0070] Image capture devices 122, 124, and 126 may contain any suitable type and number of image sensors, including CCD sensors or CMOS sensors, for example. In one embodiment, a CMOS image sensor may be employed along with a rolling shutter, such that each pixel in a row is read one at a time, and scanning of the rows proceeds on a row-by-row basis until an entire image frame has been captured. In some embodiments, the rows may be captured sequentially from top to bottom relative to the frame.

[0071] The use of a rolling shutter may result in pixels in different rows being exposed and captured at different times, which may cause skew and other image artifacts in the captured image frame. On the other hand, when the image capture device 122 is configured to operate with a global or synchronous shutter, all of the pixels may be exposed for the same amount of time and during a common exposure period. As a result, the image data in a frame collected from a system employing a global shutter represents a snapshot of the entire FOV (such as FOV 202) at a particular time. In contrast, in a rolling shutter application, each row in a frame is exposed and data is capture at different times. Thus, moving objects may appear distorted in an image capture device having a rolling shutter. This phenomenon will be described in greater detail below.

[0072] The second image capture device 124 and the third image capturing device 126 may be any type of image capture device. Like the first image capture device 122, each of image capture devices 124 and 126 may include an optical axis. In one embodiment, each of image capture devices 124 and 126 may include an Aptina M9V024 WVGA sensor with a global shutter. Alternatively, each of image capture devices 124 and 126

Tesla EX1002 Page 1351 IPR2025-00944

may include a rolling shutter. Like image capture device 122, image capture devices 124 and 126 may be configured to include various lenses and optical elements. In some embodiments, lenses associated with image capture devices 124 and 126 may provide FOVs (such as FOVs 204 and 206) that are the same as, or narrower than, a FOV (such as FOV 202) associated with image capture device 122. For example, image capture devices 124 and 126 may have FOVs of 40 degrees, 30 degrees, 26 degrees, 23 degrees, 20 degrees, or less.

[0073] Image capture devices 124 and 126 may acquire a plurality of second and third images relative to a scene associated with the vehicle 200. Each of the plurality of second and third images may be acquired as a second and third series of image scan lines, which may be captured using a rolling shutter. Each scan line or row may have a plurality of pixels. Image capture devices 124 and 126 may have second and third scan rates associated with acquisition of each of image scan lines included in the second and third series.

[0074] Each image capture device 122, 124, and 126 may be positioned at any suitable position and orientation relative to vehicle 200. The relative positioning of the image capture devices 122, 124, and 126 may be selected to aid in fusing together the information acquired from the image capture devices. For example, in some embodiments, a FOV (such as FOV 204) associated with image capture device 124 may overlap partially or fully with a FOV (such as FOV 202) associated with image capture device 122 and a FOV (such as FOV 206) associated with image capture device 126.

[0075] Image capture devices 122, 124, and 126 may be located on vehicle 200 at any suitable relative heights. In one instance, there may be a height difference between the image capture devices 122, 124, and 126, which may provide sufficient parallax information to enable stereo analysis. For example, as shown in FIG. 2A, the two image capture devices 122 and 124 are at different heights. There may also be a lateral displacement difference between image capture devices 122, 124, and 126, giving additional parallax information for stereo analysis by processing unit 110, for example. The difference in the lateral displacement may be denoted by dx, as shown in FIGS. 2C and 2D. In some embodiments, fore or aft displacement (e.g., range displacement) may exist between image capture devices 122, 124, and 126. For example, image capture device 122 may be located 0.5 to 2 meters or more behind image capture device 124 and/or image capture device 126. This type of displacement may enable one of the image capture devices to cover potential blind spots of the other image capture device(s).

[0076] Image capture devices 122 may have any suitable resolution capability (e.g., number of pixels associated with the image sensor), and the resolution of the image sensor(s) associated with the image capture device 122 may be higher, lower, or the same as the resolution of the image sensor(s) associated with image capture devices 124 and 126. In some embodiments, the image sensor(s) associated with image capture device 122 and/or image capture devices 124 and 126 may have a resolution of 640×480, 1024×768, 1280×960, or any other suitable resolution.

[0077] The frame rate (e.g., the rate at which an image capture device acquires a set of pixel data of one image frame before moving on to capture pixel data associated with the next image frame) may be controllable. The frame rate associated with image capture device 122 may be higher, lower, or the same as the frame rate associated

IPR2025-00944
Tesla EX1002 Page 1352

with image capture devices 124 and 126. The frame rate associated with image capture devices 122, 124, and 126 may depend on a variety of factors that may affect the timing of the frame rate. For example, one or more of image capture devices 122, 124, and 126 may include a selectable pixel delay period imposed before or after acquisition of image data associated with one or more pixels of an image sensor in image capture device 122, 124, and/or 126. Generally, image data corresponding to each pixel may be acquired according to a clock rate for the device (e.g., one pixel per clock cycle). Additionally, in embodiments including a rolling shutter, one or more of image capture devices 122, 124, and 126 may include a selectable horizontal blanking period imposed before or after acquisition of image data associated with a row of pixels of an image sensor in image capture device 122, 124, and/or 126. Further, one or more of image capture devices 122, 124, and/or 126 may include a selectable vertical blanking period imposed before or after acquisition of image data associated with an image frame of image capture device 122, 124, and 126.

[0078] These timing controls may enable synchronization of frame rates associated with image capture devices 122, 124, and 126, even where the line scan rates of each are different. Additionally, as will be discussed in greater detail below, these selectable timing controls, among other factors (e.g., image sensor resolution, maximum line scan rates, etc.) may enable synchronization of image capture from an area where the FOV of image capture device 122 overlaps with one or more FOVs of image capture devices 124 and 126, even where the field of view of image capture device 122 is different from the FOVs of image capture devices 124 and 126.

[0079] Frame rate timing in image capture device 122, 124, and 126 may depend on the resolution of the associated image sensors. For example, assuming similar line scan rates for both devices, if one device includes an image sensor having a resolution of 640×480 and another device includes an image sensor with a resolution of 1280×960, then more time will be required to acquire a frame of image data from the sensor having the higher resolution.

[0080] Another factor that may affect the timing of image data acquisition in image capture devices 122, 124, and 126 is the maximum line scan rate. For example, acquisition of a row of image data from an image sensor included in image capture device 122, 124, and 126 will require some minimum amount of time. Assuming no pixel delay periods are added, this minimum amount of time for acquisition of a row of image data will be related to the maximum line scan rate for a particular device. Devices that offer higher maximum line scan rates have the potential to provide higher frame rates than devices with lower maximum line scan rates. In some embodiments, one or more of image capture devices 124 and 126 may have a maximum line scan rate that is higher than a maximum line scan rate associated with image capture device 122. In some embodiments, the maximum line scan rate of image capture device 124 and/or 126 may be 1.25, 1.5, 1.75, or 2 times or more than a maximum line scan rate of image capture device 122.

[0081] In another embodiment, image capture devices 122, 124, and 126 may have the same maximum line scan rate, but image capture device 122 may be operated at a scan rate less than or equal to its maximum scan rate. The system may be configured such that one or more of image capture devices 124 and 126 operate at a line scan rate that is equal to the line scan rate of image capture device 122. In other instances, the system may be configured such that the line scan rate of image capture device 124 and/or image capture device 126 may be 1.25, 1.5, 1.75, or 2 times or more than the

Tesla EX1002 Page 1353 IPR2025-00944

line scan rate of image capture device 122.

[0082] In some embodiments, image capture devices 122, 124, and 126 may be asymmetric. That is, they may include cameras having different fields of view (FOV) and focal lengths. The fields of view of image capture devices 122, 124, and 126 may include any desired area relative to an environment of vehicle 200, for example. In some embodiments, one or more of image capture devices 122, 124, and 126 may be configured to acquire image data from an environment in front of vehicle 200, behind vehicle 200, to the sides of vehicle 200, or combinations thereof.

[0083] Further, the focal length associated with each image capture device 122, 124, and/or 126 may be selectable (e.g., by inclusion of appropriate lenses etc.) such that each device acquires images of objects at a desired distance range relative to vehicle 200. For example, in some embodiments image capture devices 122, 124, and 126 may acquire images of close-up objects within a few meters from the vehicle. Image capture devices 122, 124, and 126 may also be configured to acquire images of objects at ranges more distant from the vehicle (e.g., 25 m, 50 m, 100 m, 150 m, or more). Further, the focal lengths of image capture devices 122, 124, and 126 may be selected such that one image capture device (e.g., image capture device 122) can acquire images of objects relatively close to the vehicle (e.g., within 10 m or within 20 m) while the other image capture devices (e.g., image capture devices 124 and 126) can acquire images of more distant objects (e.g., greater than 20 m, 50 m, 100 m, 150 m, etc.) from vehicle 200.

[0084] According to some embodiments, the FOV of one or more image capture devices 122, 124, and 126 may have a wide angle. For example, it may be advantageous to have a FOV of 140 degrees, especially for image capture devices 122, 124, and 126 that may be used to capture images of the area in the vicinity of vehicle 200. For example, image capture device 122 may be used to capture images of the area to the right or left of vehicle 200 and, in such embodiments, it may be desirable for image capture device 122 to have a wide FOV (e.g., at least 140 degrees).

[0085] The field of view associated with each of image capture devices 122, 124, and 126 may depend on the respective focal lengths. For example, as the focal length increases, the corresponding field of view decreases.

[0086] Image capture devices 122, 124, and 126 may be configured to have any suitable fields of view. In one particular example, image capture device 122 may have a horizontal FOV of 46 degrees, image capture device 124 may have a horizontal FOV of 23 degrees, and image capture device 126 may have a horizontal FOV in between 23 and 46 degrees. In another instance, image capture device 122 may have a horizontal FOV of 52 degrees, image capture device 124 may have a horizontal FOV of 26 degrees, and image capture device 126 may have a horizontal FOV in between 26 and 52 degrees. In some embodiments, a ratio of the FOV of image capture device 122 to the FOVs of image capture device 124 and/or image capture device 126 may vary from 1.5 to 2.0. In other embodiments, this ratio may vary between 1.25 and 2.25.

[0087] System 100 may be configured so that a field of view of image capture device 122 overlaps, at least partially or fully, with a field of view of image capture device 124 and/or image capture device 126. In some embodiments, system 100 may be configured such that the fields of view of image capture devices 124 and 126, for

Tesla EX1002 Page 1354 IPR2025-00944

example, fall within (e.g., are narrower than) and share a common center with the field of view of image capture device 122. In other embodiments, the image capture devices 122, 124, and 126 may capture adjacent FOVs or may have partial overlap in their FOVs. In some embodiments, the fields of view of image capture devices 122, 124, and 126 may be aligned such that a center of the narrower FOV image capture devices 124 and/or 126 may be located in a lower half of the field of view of the wider FOV device 122.

[0088] FIG. 2F is a diagrammatic representation of exemplary vehicle control systems, consistent with the disclosed embodiments. As indicated in FIG. 2F, vehicle 200 may include throttling system 220, braking system 230, and steering system 240. System 100 may provide inputs (e.g., control signals) to one or more of throttling system 220, braking system 230, and steering system 240 over one or more data links (e.g., any wired and/or wireless link or links for transmitting data). For example, based on analysis of images acquired by image capture devices 122, 124, and/or 126, system 100 may provide control signals to one or more of throttling system 220, braking system 230, and steering system 240 to navigate vehicle 200 (e.g., by causing an acceleration, a turn, a lane shift, etc.). Further, system 100 may receive inputs from one or more of throttling system 220, braking system 230, and steering system 24 indicating operating conditions of vehicle 200 (e.g., speed, whether vehicle 200 is braking and/or turning, etc.). Further details are provided in connection with FIGS. 4-7, below.

[0089] As shown in FIG. 3A, vehicle 200 may also include a user interface 170 for interacting with a driver or a passenger of vehicle 200. For example, user interface 170 in a vehicle application may include a touch screen 320, knobs 330, buttons 340, and a microphone 350. A driver or passenger of vehicle 200 may also use handles (e.g., located on or near the steering column of vehicle 200 including, for example, turn signal handles), buttons (e.g., located on the steering wheel of vehicle 200), and the like, to interact with system 100. In some embodiments, microphone 350 may be positioned adjacent to a rearview mirror 310. Similarly, in some embodiments, image capture device 122 may be located near rearview mirror 310. In some embodiments, user interface 170 may also include one or more speakers 360 (e.g., speakers of a vehicle audio system). For example, system 100 may provide various notifications (e.g., alerts) via speakers 360.

[0090] FIGS. 3B-3D are illustrations of an exemplary camera mount 370 configured to be positioned behind a rearview mirror (e.g., rearview mirror 310) and against a vehicle windshield, consistent with disclosed embodiments. As shown in FIG. 3B, camera mount 370 may include image capture devices 122, 124, and 126. Image capture devices 124 and 126 may be positioned behind a glare shield 380, which may be flush against the vehicle windshield and include a composition of film and/or anti-reflective materials. For example, glare shield 380 may be positioned such that it aligns against a vehicle windshield having a matching slope. In some embodiments, each of image capture devices 122, 124, and 126 may be positioned behind glare shield 380, as depicted, for example, in FIG. 3D. The disclosed embodiments are not limited to any particular configuration of image capture devices 122, 124, and 126, camera mount 370, and glare shield 380. FIG. 3C is an illustration of camera mount 370 shown in FIG. 3B from a front perspective.

[0091] As will be appreciated by a person skilled in the art having the benefit of this disclosure, numerous variations and/or modifications may be made to the foregoing disclosed embodiments. For example, not all components are essential for the

Tesla EX1002 Page 1355 IPR2025-00944

operation of system 100. Further, any component may be located in any appropriate part of system 100 and the components may be rearranged into a variety of configurations while providing the functionality of the disclosed embodiments. Therefore, the foregoing configurations are examples and, regardless of the configurations discussed above, system 100 can provide a wide range of functionality to analyze the surroundings of vehicle 200 and navigate vehicle 200 in response to the analysis.

[0092] As discussed below in further detail and consistent with various disclosed embodiments, system 100 may provide a variety of features related to autonomous driving and/or driver assist technology. For example, system 100 may analyze image data, position data (e.g., GPS location information), map data, speed data, and/or data from sensors included in vehicle 200. System 100 may collect the data for analysis from, for example, image acquisition unit 120, position sensor 130, and other sensors. Further, system 100 may analyze the collected data to determine whether or not vehicle 200 should take a certain action, and then automatically take the determined action without human intervention. For example, when vehicle 200 navigates without human intervention, system 100 may automatically control the braking, acceleration, and/or steering of vehicle 200 (e.g., by sending control signals to one or more of throttling system 220, braking system 230, and steering system 240). Further, system 100 may analyze the collected data and issue warnings and/or alerts to vehicle occupants based on the analysis of the collected data. Additional details regarding the various embodiments that are provided by system 100 are provided below.

[0093] Forward-Facing Multi-Imaging System

[0094] As discussed above, system 100 may provide drive assist functionality that uses a multi-camera system. The multi-camera system may use one or more cameras facing in the forward direction of a vehicle. In other embodiments, the multi-camera system may include one or more cameras facing to the side of a vehicle or to the rear of the vehicle. In one embodiment, for example, system 100 may use a two-camera imaging system, where a first camera and a second camera (e.g., image capture devices 122 and 124) may be positioned at the front and/or the sides of a vehicle (e.g., vehicle 200). The first camera may have a field of view that is greater than, less than, or partially overlapping with, the field of view of the second camera. In addition, the first camera may be connected to a first image processor to perforin monocular image analysis of images provided by the first camera, and the second camera may be connected to a second image processor to perform monocular image analysis of images provided by the second camera. The outputs (e.g., processed information) of the first and second image processors may be combined. In some embodiments, the second image processor may receive images from both the first camera and second camera to perform stereo analysis. In another embodiment, system 100 may use a three-camera imaging system where each of the cameras has a different field of view. Such a system may, therefore, make decisions based on information derived from objects located at varying distances both forward and to the sides of the vehicle. References to monocular image analysis may refer to instances where image analysis is performed based on images captured from a single point of view (e.g., from a single camera). Stereo image analysis may refer to instances where image analysis is performed based on two or more images captured with one or more variations of an image capture parameter. For example, captured images suitable for performing stereo image analysis may include images captured: from two or more different positions, from different fields of view, using different focal lengths, along with parallax information, etc.

IPR2025-00944
Tesla EX1002 Page 1356

[0095] For example, in one embodiment, system 100 may implement a three camera configuration using image capture devices 122-126. In such a configuration, image capture device 122 may provide a narrow field of view (e.g., 34 degrees, or other values selected from a range of about 20 to 45 degrees, etc.), image capture device 124 may provide a wide field of view (e.g., 150 degrees or other values selected from a range of about 100 to about 180 degrees), and image capture device 126 may provide an intermediate field of view (e.g., 46 degrees or other values selected from a range of about 35 to about 60 degrees). In some embodiments, image capture device 126 may act as a main or primary camera. Image capture devices 122-126 may be positioned behind rearview mirror 310 and positioned substantially side-by-side (e.g., 6 cm apart). Further, in some embodiments, as discussed above, one or more of image capture devices 122-126 may be mounted behind glare shield 380 that is flush with the windshield of vehicle 200. Such shielding may act to minimize the impact of any reflections from inside the car on image capture devices 122-126.

[0096] In another embodiment, as discussed above in connection with FIGS. 3B and 3C, the wide field of view camera (e.g., image capture device 124 in the above example) may be mounted lower than the narrow and main field of view cameras (e.g., image devices 122 and 126 in the above example). This configuration may provide a free line of sight from the wide field of view camera. To reduce reflections, the cameras may be mounted close to the windshield of vehicle 200, and may include polarizers on the cameras to damp reflected light.

[0097] A three camera system may provide certain performance characteristics. For example, some embodiments may include an ability to validate the detection of objects by one camera based on detection results from another camera. In the three camera configuration discussed above, processing unit 110 may include, for example, three processing devices (e.g., three EyeQ series of processor chips, as discussed above), with each processing device dedicated to processing images captured by one or more of image capture devices 122-126.

[0098] In a three camera system, a first processing device may receive images from both the main camera and the narrow field of view camera, and perform vision processing of the narrow FOV camera to, for example, detect other vehicles, pedestrians, lane marks, traffic signs, traffic lights, and other road objects. Further, the first processing device may calculate a disparity of pixels between the images from the main camera and the narrow camera and create a 3D reconstruction of the environment of vehicle 200. The first processing device may then combine the 3D reconstruction with 3D map data or with 3D information calculated based on information from another camera.

[0099] The second processing device may receive images from main camera and perform vision processing to detect other vehicles, pedestrians, lane marks, traffic signs, traffic lights, and other road objects. Additionally, the second processing device may calculate a camera displacement and, based on the displacement, calculate a disparity of pixels between successive images and create a 3D reconstruction of the scene (e.g., a structure from motion). The second processing device may send the structure from motion based 3D reconstruction to the first processing device to be combined with the stereo 3D images.

[0100] The third processing device may receive images from the wide FOV camera and

process the images to detect vehicles, pedestrians, lane marks, traffic signs, traffic lights, and other road objects. The third processing device may further execute additional processing instructions to analyze images to identify objects moving in the image, such as vehicles changing lanes, pedestrians, etc.

[0101] In some embodiments, having streams of image-based information captured and processed independently may provide an opportunity for providing redundancy in the system. Such redundancy may include, for example, using a first image capture device and the images processed from that device to validate and/or supplement information obtained by capturing and processing image information from at least a second image capture device.

[0102] In some embodiments, system 100 may use two image capture devi£es (e.g., image capture devices 122 and 124) in providing navigation assistance for vehicle 200 and use a third image capture device (e.g., image capture device 126) to provide redundancy and validate the analysis of data received from the other two image capture devices. For example, in such a configuration, image capture devices 122 and 124 may provide images for stereo analysis by system 100 for navigating vehicle 200, while image capture device 126 may provide images for monocular analysis by system 100 to provide redundancy and validation of information obtained based on images captured from image capture device 122 and/or image capture device 124. That is, image capture device 126 (and a corresponding processing device) may be considered to provide a redundant sub-system for providing a check on the analysis derived from image capture devices 122 and 124 (e.g., to provide an automatic emergency braking (AEB) system).

[0103] One of skill in the art will recognize that the above camera configurations, camera placements, number of cameras, camera locations, etc., are examples only. These components and others described relative to the overall system may be assembled and used in a variety of different configurations without departing from the scope of the disclosed embodiments. Further details regarding usage of a multi-camera system to provide driver assist and/or autonomous vehicle functionality follow below.

[0104] FIG. 4 is an exemplary functional block diagram of memory 140 and/or 150, which may be stored/programmed with instructions for performing one or more operations consistent with the disclosed embodiments. Although the following refers to memory 140, one of skill in the art will recognize that instructions may be stored in memory 140 and/or 150.

[0105] As shown in FIG. 4, memory 140 may store a monocular image analysis module 402, a stereo image analysis module 404, a velocity and acceleration module 406, and a navigational response module 408. The disclosed embodiments are not limited to any particular configuration of memory 140. Further, application processor 180 and/or image processor 190 may execute the instructions stored in any of modules 402-408 included in memory 140. One of skill in the art will understand that references in the following discussions to processing unit 110 may refer to application processor 180 and image processor 190 individually or collectively. Accordingly, steps of any of the following processes may be performed by one or more processing devices.

[0106] In one embodiment, monocular image analysis module 402 may store instructions (such as computer vision software) which, when executed by processing unit 110, performs monocular image analysis of a set of images acquired by one of

image capture devices 122, 124, and 126. In some embodiments, processing unit 110 may combine information from a set of images with additional sensory information (e.g., information from radar) to perform the monocular image analysis. As described in connection with FIGS. 5A-5D below, monocular image analysis module 402 may include instructions for detecting a set of features within the set of images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, hazardous objects, and any other feature associated with an environment of a vehicle. Based on the analysis, system 100 (e.g., via processing unit 110) may cause one or more navigational responses in vehicle 200, such as a turn, a lane shift, a change in acceleration, and the like, as discussed below in connection with navigational response module 408.

[0107] In one embodiment, stereo image analysis module 404 may store instructions (such as computer vision software) which, when executed by processing unit 110, performs stereo image analysis of first and second sets of images acquired by a combination of image capture devices selected from any of image capture devices 122, 124, and 126. In some embodiments, processing unit 110 may combine information from the first and second sets of images with additional sensory information (e.g., information from radar) to perform the stereo image analysis. For example, stereo image analysis module 404 may include instructions for performing stereo image analysis based on a first set of images acquired by image capture device 124 and a second set of images acquired by image capture device 126. As described in connection with FIG. 6 below, stereo image analysis module 404 may include instructions for detecting a set of features within the first and second sets of images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, hazardous objects, and the like. Based on the analysis, processing unit 110 may cause one or more navigational responses in vehicle 200, such as a turn, a lane shift, a change in acceleration, and the like, as discussed below in connection with navigational response module 408.

[0108] In one embodiment, velocity and acceleration module 406 may store software configured to analyze data received from one or more computing and electromechanical devices in vehicle 200 that are configured to cause a change in velocity and/or acceleration of vehicle 200. For example, processing unit 110 may execute instructions associated with velocity and acceleration module 406 to calculate a target speed for vehicle 200 based on data derived from execution of monocular image analysis module 402 and/or stereo image analysis module 404. Such data may include, for example, a target position, velocity, and/or acceleration, the position and/or speed of vehicle 200 relative to a nearby vehicle, pedestrian, or road object, position information for vehicle 200 relative to lane markings of the road, and the like. In addition, processing unit 110 may calculate a target speed for vehicle 200 based on sensory input (e.g., information from radar) and input from other systems of vehicle 200, such as throttling system 220, braking system 230, and/or steering system 240 of vehicle 200. Based on the calculated target speed, processing unit 110 may transmit electronic signals to throttling system 220, braking system 230, and/or steering system 240 of vehicle 200 to trigger a change in velocity and/or acceleration by, for example, physically depressing the brake or easing up off the accelerator of vehicle 200.

[0109] In one embodiment, navigational response module 408 may store software executable by processing unit 110 to determine a desired navigational response based on data derived from execution of monocular image analysis module 402 and/or stereo image analysis module 404. Such data may include position and speed information

IPR2025-00944
Tesla EX1002 Page 1359

associated with nearby vehicles, pedestrians, and road objects, target position information for vehicle 200, and the like. Additionally, in some embodiments, the navigational response may be based (partially or fully) on map data, a predetermined position of vehicle 200, and/or a relative velocity or a relative acceleration between vehicle 200 and one or more objects detected from execution of monocular image analysis module 402 and/or stereo image anal9sis mo$ule 404. Navigational response module 408 may also determine a desired navigational response based on sensory input (e.g., information from radar) and inputs from other systems of vehicle 200, such as throttling system 220, braking system 230, and steering system 240 of vehicle 200. Based on the desired navigational response, processing unit 110 may transmit electronic signals to throttling system 220, braking system 230, and steering system 240 of vehicle 200 to trigger a desired navigational response by, for example, turning the steering wheel of vehicle 200 to achieve a rotation of a predetermined angle. In some embodiments, processing unit 110 may use the output of navigational response module 408 (e.g., the desired navigational response) as an input to execution of velocity and acceleration module 406 for calculating a change in speed of vehicle 200.

[0110] FIG. 5A is a flowchart showing an exemplary process 500A for causing one or more navigational responses based on monocular image analysis, consistent with disclosed embodiments. At step 510, processing unit 110 may receive a plurality of images via data interface 128 between processing unit 110 and image acquisition unit 120. For instance, a camera included in image acquisition unit 120 (such as image capture device 122 having field of view 202) may capture a plurality of images of an area forward of vehicle 200 (or to the sides or rear of a vehicle, for example) and transmit them over a data connection (e.g., digital, wired, USB, wireless, Bluetooth, etc.) to processing unit 110. Processing unit 110 may execute monocular image analysis module 402 to analyze the plurality of images at step 520, as described in further detail in connection with FIGS. 5B-5D below. By performing the analysis, processing unit 110 may detect a set of features within the set of images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, and the like.

[0111] Processing unit 110 may also execute monocular image analysis module 402 to detect various road hazards at step 520, such as, for example, parts of a truck tire, fallen road signs, loose cargo, small animals, and the like. Road hazards may vary in structure, shape, size, and color, which may make detection of such hazards more challenging. In some embodiments, processing unit 110 may execute monocular image analysis module 402 to perforin multi-frame analysis on the plurality of images to detect road hazards. For example, processing unit 110 may estimate camera motion between consecutive image frames and calculate the disparities in pixels between the frames to construct a 3D-map of the road. Processing unit 110 may then use the 3D-map`t/ detect the road surface, as well as hazards existing above the road surface.

[0112] At step 530, processing unit 110 may execute navigational response module 408 to cause one or more navigational responses in vehicle 200 based on the analysis performed at step 520 and the techniques as described above in connection with FIG. 4. Navigational responses may include, for example, a turn, a lane shift, a change in acceleration, and the like. In some embodiments, processing unit 110 may use data derived from execution of velocity and acceleration module 406 to cause the one or more navigational responses. Additionally, multiple navigational responses may occur simultaneously, in sequence, or any combination thereof. For instance, processing unit 110 may cause vehicle 200 to shift one lane over and then accelerate by, for example,

IPR2025-00944
Tesla EX1002 Page 1360

sequentially transmitting control signals to steering system 240 and throttling system 220 of vehicle 200. Alternatively, processing unit 110 may cause vehicle 200 to brake while at the same time shifting lanes by, for example, simultaneously transmitting control signals to braking system 230 and steering system 240 of vehicle 200.

[0113] FIG. 5B is a flowchart showing an exemplary process 500B for detecting one or more vehicles and/or pedestrians in a set of images, consistent with disclosed embodiments. Processing unit 110 may execute monocular image analysis module 402 to implement process 500B. At step 540, processing unit 110 may determine a set of candidate objects representing possible vehicles and/or pedestrians. For example, processing unit 110 may scan one or more images, compare the images to one or more predetermined patterns, and identify within each image possible locations that may contain objects of interest (e.g., vehicles, pedestrians, or portions thereof). The predetermined patterns may be designed in such a way to achieve a high rate of "false hits" and a low rate of "misses." For example, processing unit 110 may use a low threshold of similarity to predetermined patterns for identifying candidate objects as possible vehicles or pedestrians. Doing so may allow processing unit 110 to reduce the probability of missing (e.g., not identifying) a candidate object representing a vehicle or pedestrian.

[0114] At step 542, processing unit 110 may filter the set of candidate objects to exclude certain candidates (e.g., irrelevant or less relevant objects) based on classification criteria. Such criteria may be derived from various properties associated with object types stored in a database (e.g., a database stored in memory 140). Properties may include object shape, dimensions, texture, position (e.g., relative to vehicle 200), and the like. Thus, processing unit 110 may use one or more sets of criteria to reject false candidates from the set of candidate objects.

[0115] At step 544, processing unit 110 may analyze multiple frames of images to determine whether objects in the set of candidate objects represent vehicles and/or pedestrians. For example, processing unit 110 may track a detected candidate object across consecutive frames and accumulate frame-by-frame data associated with the detected object (e.g., size, position relative to vehicle 200, etc.). Additionally, processing unit 110 may estimate parameters for the detected object and compare the object's frame-by-frame position data to a predicted position.

[0116] At step 546, processing unit 110 may construct a set of measurements for the detected objects. Such measurements may include, for example, position, velocity, and acceleration values (relative to vehicle 200) associated with the detected objects. In some embodiments, processing unit 110 may construct the measurements based on estimation techniques using a series of time-based observations such as Kalman filters or linear quadratic estimation (LQE), and/or based on available modeling data for different object types (e.g., cars, trucks, pedestrians, bicycles, road signs, etc.). The Kalman filters may be based on a measurement of an object's scale, where the scale measurement is proportional to a time to collision (e.g., the amount of time for vehicle 200 to reach the object). Thus, by performing steps 540-546, processing unit 110 may identify vehicles and pedestrians appearing within the set of captured images and derive information (e.g., position, speed, size) associated with the vehicles and pedestrians. Based on the identification and the derived information, processing unit 110 may cause one or more navigational responses in vehicle 200, as described in connection with FIG. 5A, above.

IPR2025-00944
Tesla EX1002 Page 1361

[0117] At step 548, processing unit 110 may perform an optical flow analysis of one or more images to reduce the probabilities of detecting a "false hit" and missing a candidate object that represents a vehicle or pedestrian. The optical flow analysis may refer to, for example, analyzing motion patterns relative to vehicle 200 in the one or more images associated with other vehicles and pedestrians, and that are distinct from road surface motion. Processing unit 110 may calculate the motion of candidate objects by observing the different positions of the objects across multiple image frames, which are captured at different times. Processing unit 110 may use the position and time values as inputs into mathematical models for calculating the motion of the candidate objects. Thus, optical flow analysis may provide another method of detecting vehicles and pedestrians that are nearby vehicle 200. Processing unit 110 may perform optical flow analysis in combination with steps 540-546 to provide redundancy for detecting vehicles and pedestrians and increase the reliability of system 100.

[0118] FIG. 5C is a flowchart showing an exemplary process 500C for detecting road marks and/or lane geometry information in a set of images, consistent with disclosed embodiments. Processing unit 110 may execute monocular image analysis module 402 to implement process 500C. At step 550, processing unit 110 may detect a set of objects by scanning one or more images. To detect segments of lane markings, lane geometry information, and other pertinent road marks, processing unit 110 may filter the set of objects to exclude those determined to be irrelevant (e.g., minor potholes, small rocks, etc.). At step 552, processing unit 110 may group together the segments detected in step 550 belonging to the same road mark or lane mark. Based on the grouping, processing unit 110 may develop a model to represent the detected segments, such as a mathematical model.

[0119] At step 554, processing unit 110 may construct a set of measurements associated with the detected segments. In some embodiments, processing unit 110 may create a projection of the detected segments from the image plane onto the real-world plane. The projection may be characterized using a 3rd-degree polynomial having coefficients corresponding to physical properties such as the position, slope, curvature, and curvature derivative of the detected road. In generating the projection, processing unit 110 may take into account changes in the road surface, as well as pitch and roll rates associated with vehicle 200. In addition, processing unit 110 may model the road elevation by analyzing position and motion cues present on the road surface. Further, processing unit 110 may estimate the pitch and roll rates associated with vehicle 200 by tracking a set of feature points in the one or more images.

[0120] At step 556, processing unit 110 may perform multi-frame analysis by, for example, tracking the detected segments across consecutive image frames and accumulating frame-by-frame data associated with detected segments. As processing unit 110 performs multi-frame analysis, the set of measurements constructed at step 554 may become more reliable and associated with an increasingly higher confidence level. Thus, by performing steps 550-556, processing unit 110 may identify road marks appearing within the set of captured images and derive lane geometry information. Based on the identification and the derived information, processing unit 110 may cause one or more navigational responses in vehicle 200, as described in connection with FIG. 5A, above.

[0121] At step 558, processing unit 110 may consider additional sources of information to further develop a safety model for vehicle 200 in the context of its surroundings. Processing unit 110 may use the safety model to define a context in which system 100

Tesla EX1002 Page 1362 IPR2025-00944

may execute autonomous control of vehicle 200 in a safe manner. To develop the safety model, in some embodiments, processing unit 110 may consider the position and motion of other vehicles, the detected road edges and barriers, and/or general road shape descriptions extracted from map data (such as data from map database 160). By considering additional sources of information, processing unit 110 may provide redundancy for detecting road marks and lane geometry and increase the reliability of system 100.

[0122] FIG. 5D is a flowchart showing an exemplary process 500D for detecting traffic lights in a set of images, consistent with disclosed embodiments. Processing unit 110 may execute monocular image analysis module 402 to implement process 500D. At step 560, processing unit 110 may scan the set of images and identify objects appearing at locations in the images likely to contain traffic lights. For example, processing unit 110 may filter the identified objects to construct a set of candidate objects, excluding those objects unlikely to correspond to traffic lights. The filtering may be done based on various properties associated with traffic lights, such as shape, dimensions, texture, position (e.g., relative to vehicle 200), and the like. Such properties may be based on multiple examples of traffic lights and traffic control signals and stored in a database. In some embodiments, processing unit 110 may perform multi-frame analysis on the set of candidate objects reflecting possible traffic lights. For example, processing unit 110 may track the candidate objects across consecutive image frames, estimate the real-world position of the candidate objects, and filter out those objects that are moving (which are unlikely to be traffic lights). In some embodiments, processing unit 110 may perform color analysis on the candidate objects and identify the relative position of the detected colors appearing inside possible traffic lights.

[0123] At step 562, processing unit 110 may analyze the geometry of a junction. The analysis may be based on any combination of: (i) the number of lanes detected on either side of vehicle 200, (ii) markings (such as arrow marks) detected on the road, and (iii) descriptions of the junction extracted from map data (such as data from map database 160). Processing unit 110 may conduct the analysis using information derived from execution of monocular analysis module 402. In addition, Processing unit 110 may determine a correspondence between the traffic lights detected at step 560 and the lanes appearing near vehicle 200.

[0124] As vehicle 200 approaches the junction, at step 564, processing unit 110 may update the confidence level associated with the analyzed junction geometry and the detected traffic lights. For instance, the number of traffic lights estimated to appear at the junction as compared with the number actually appearing at the junction may impact the confidence level. Thus, based on the confidence level, processing unit 110 may delegate control to the driver of vehicle 200 in order to improve safety conditions. By performing steps 560-564, processing unit 110 may identify traffic lights appearing within the set of captured images and analyze junction geometry information. Based on the identification and the analysis, processing unit 110 may cause one or more navigational responses in vehicle 200, as described in connection with FIG. 5A, above.

[0125] FIG. 5E is a flowchart showing an exemplary process 500E for causing one or more navigational responses in vehicle 200 based on a vehicle path, consistent with the disclosed embodiments. At step 570, processing unit 110 may construct an initial vehicle path associated with vehicle 200. The vehicle path may be represented using a set of points expressed in coordinates (x, z), and the distance di between two points in

IPR2025-00944
Tesla EX1002 Page 1363

the set of points may fall in the range of 1 to 5 meters. In one embodiment, processing unit 110 may construct the initial vehicle path using two polynomials, such as left and right road polynomials. Processing unit 110 may calculate the geometric midpoint between the two polynomials and offset each point included in the resultant vehicle path by a predetermined offset (e.g., a smart lane offset), if any (an offset of zero may correspond to travel in the middle of a lane). The offset may be in a direction perpendicular to a segment between any two points in the vehicle path. In another embodiment, processing unit 110 may use one polynomial and an estimated lane width to offset each point of the vehicle path by half the estimated lane width plus a predetermined offset (e.g., a smart lane offset).

[0126] At step 572, processing unit 110 may update the vehicle path constructed at step 570. Processing unit 110 may reconstruct the vehicle path constructed at step 570 using a higher resolution, such that the distance dk between two points in the set of points representing the vehicle path is less than the distance di described above. For example, the distance dk may fall in the range of 0.1 to 0.3 meters. Processing unit 110 may reconstruct the vehicle path using a parabolic spline algorithm, which may yield a cumulative distance vector S corresponding to the total length of the vehicle path (i.e., based on the set of points representing the vehicle path).

[0127] At step 574, processing unit 110 may determine a look-ahead point (expressed in coordinates as (xl, zl) based on the updated vehicle path constructed at step 572. Processing unit 110 may extract the look-ahead point from the cumulative distance vector 5, and the look-ahead point may be associated with a look-ahead distance and look-ahead time. The look-ahead distance, which may have a lower bound ranging from 10 to 20 meters, may be calculated as the product of the speed of vehicle 200 and the look-ahead time. For example, as the speed of vehicle 200 decreases, the look-ahead distance may also decrease (e.g., until it reaches the lower bound). The look-ahead time, which may range from 0.5 to 1.5 seconds, may be inversely proportional to the gain of one or more control loops associated with causing a navigational response in vehicle 200, such as the heading error tracking control loop. For example, the gain of the heading error tracking control loop may depend on the bandwidth of a yaw rate loop, a steering actuator loop, car lateral dynamics, and the like. Thus, the higher the gain of the heading error tracking control loop, the lower the look-ahead time.

[0128] At step 576, processing unit 110 may determine a heading error and yaw rate command based on the look-ahead point determined at step 574. Processing unit 110 may determine the heading error by calculating the arctangent of the look-ahead point, e.g., arctan (xl/zl). Processing unit 110 may determine the yaw rate command as the product of the heading error and a high-level control gain. The high-level control gain may be equal to: (2/look-ahead time), if the look-ahead distance is not at the lower bound. Otherwise, the high-level control gain may be equal to: (2*speed of vehicle 200/look-ahead distance).

[0129] FIG. 5F is a flowchart showing an exemplary process 500F for determining whether a leading vehicle is changing lanes, consistent with the disclosed embodiments. At step 580, processing unit 110 may determine navigation information associated with a leading vehicle (e.g., a vehicle traveling ahead of vehicle 200). For example, processing unit 110 may determine the position, velocity (e.g., direction and speed), and/or acceleration of the leading vehicle, using the techniques described in connection with FIGS. 5A and 5B, above. Processing unit 110 may also determine one or more road polynomials, a look-ahead point (associated with vehicle 200), and/or a

IPR2025-00944
Tesla EX1002 Page 1364

IPR2025-00944 Tesla EX1002 Page 1365

snail trail (e.g., a set of points describing a path taken by the leading vehicle), using the techniques described in connection with FIG. 5E, above.

[0130] At step 582, processing unit 110 may analyze the navigation information determined at step 580. In one embodiment, processing unit 110 may calculate the distance between a snail trail and a road polynomial (e.g., along the trail). If the variance of this distance along the trail exceeds a predetermined threshold (for example, 0.1 to 0.2 meters on a straight road, 0.3 to 0.4 meters on a moderately curvy road, and 0.5 to 0.6 meters on a road with sharp curves), processing unit 110 may determine that the leading vehicle is likely changing lanes. In the case where multiple vehicles are detected traveling ahead of vehicle 200, processing unit 110 may compare the snail trails associated with each vehicle. Based on the comparison, processing unit 110 may determine that a vehicle whose snail trail does not match with the snail trails of the other vehicles is likely changing lanes. Processing unit 110 may additionally compare the curvature of the snail trail (associated with the leading vehicle) with the expected curvature of the road segment in which the leading vehicle is traveling. The expected curvature may be extracted from map data (e.g., data from map database 160), from road polynomials, from other vehicles' snail trails, from prior knowledge about the road, and the like. If the difference in curvature of the snail trail and the expected curvature of the road segment exceeds a predetermined threshold, processing unit 110 may determine that the leading vehicle is likely changing lanes.

[0131] In another embodiment, processing unit 110 may compare the leading vehicle's instantaneous position with the look-ahead point (associated with vehicle 200) over a specific period of time (e.g., 0.5 to 1.5 seconds). If the distance between the leading vehicle's instantaneous position and the look-ahead point varies during the specific period of time, and the cumulative sum of variation exceeds a predetermined threshold (for example, 0.3 to 0.4 meters on a straight road, 0.7 to 0.8 meters on a moderately curvy road, and 1.3 to 1.7 meters on a road with sharp curves), processing unit 110 may determine that the leading vehicle is likely changing lanes. In another embodiment, processing unit 110 may analyze the geometry of the snail trail by comparing the lateral distance traveled along the trail with the expected curvature of the snail trail. The expected radius of curvature may be determined according to the calculation: (δz<2>+δx<2>)/2/(δx), where δx represents the lateral distance traveled and δz represents the longitudinal distance traveled. If the difference between the lateral distance traveled and the expected curvature exceeds a predetermined threshold (e.g., 500 to 700 meters), processing unit 110 may determine that the leading vehicle is likely changing lanes. In another embodiment, processing unit 110 may analyze the position of the leading vehicle. If the position of the leading vehicle obscures a road polynomial (e.g., the leading vehicle is overlaid on top of the road polynomial), then processing unit 110 may determine that the leading vehicle is likely changing lanes. In the case where the position of the leading vehicle is such that, another vehicle is detected ahead of the leading vehicle and the snail trails of the two vehicles are not parallel, processing unit 110 may determine that the (closer) leading vehicle is likely changing lanes.

[0132] At step 584, processing unit 110 may determine whether or not leading vehicle 200 is changing lanes based on the analysis performed at step 582. For example, processing unit 110 may make the determination based on a weighted average of the individual analyses performed at step 582. Under such a scheme, for example, a decision by processing unit 110 that the leading vehicle is likely changing lanes based on a particular type of analysis may be assigned a value of "1" (and "0" to represent a

determination that the leading vehicle is not likely changing lanes). Different analyses performed at step 582 may be assigned different weights, and the disclosed embodiments are not limited to any particular combination of analyses and weights.

[0133] FIG. 6 is a flowchart showing an exemplary process 600 for causing one or more navigational responses based on stereo image analysis, consistent with disclosed embodiments. At step 610, processing unit 110 may receive a first and second plurality of images via data interface 128. For example, cameras included in image acquisition unit 120 (such as image capture devices 122 and 124 having fields of view 202 and 204) may capture a first and second plurality of images of an area forward of vehicle 200 and transmit them over a digital connection (e.g., USB, wireless, Bluetooth, etc.) to processing unit 110. In some embodiments, processing unit 110 may receive the first and second plurality of images via two or more data interfaces. The disclosed embodiments are not limited to any particular data interface configurations or protocols.

[0134] At step 620, processing unit 110 may execute stereo image analysis module 404 to perform stereo image analysis of the first and second plurality of images to create a 3D map of the road in front of the vehicle and detect features within the images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, road hazards, and the like. Stereo image analysis may be performed in a manner similar to the steps described in connection with FIGS. 5A-5D, above. For example, processing unit 110 may execute stereo image analysis module 404 to detect candidate objects (e.g., vehicles, pedestrians, road marks, traffic lights, road hazards, etc.) within the first and second plurality of images, filter out a subset of the candidate objects based on various criteria, and perforin multi-frame analysis, construct measurements, and determine a confidence level for the remaining candidate objects. In performing the steps above, processing unit 110 may consider information from both the first and second plurality of images, rather than information from one set of images alone. For example, processing unit 110 may analyze the differences in pixel-level data (or other data subsets from among the two streams of captured images) for a candidate object appearing in both the first and second plurality of images. As another example, processing unit 110 may estimate a position and/or velocity of a candidate object (e.g., relative to vehicle 200) by observing that the object appears in one of the plurality of images but not the other or relative to other differences that may exist relative to objects appearing if the two image streams. For example, position, velocity, and/or acceleration relative to vehicle 200 may be determined based on trajectories, positions, movement characteristics, etc. of features associated with an object appearing in one or both of the image streams.

[0135] At step 630, processing unit 110 may execute navigational response module 408 to cause one or more navigational responses in vehicle 200 based on the analysis performed at step 620 and the techniques as described above in connection with FIG. 4. Navigational responses may include, for example, a turn, a lane shift, a change in acceleration, a change in velocity, braking, and the like. In some embodiments, processing unit 110 may use data derived from execution of velocity and acceleration module 406 to cause the one or more navigational responses. Additionally, multiple navigational responses may occur simultaneously, in sequence, or any combination thereof.

[0136] FIG. 7 is a flowchart showing an exemplary process 700 for causing one or more navigational responses based on an analysis of three sets of images, consistent with disclosed embodiments. At step 710, processing unit 110 may receive a first,

Tesla EX1002 Page 1366 IPR2025-00944

second, and third plurality of images via data interface 128. For instance, cameras included in image acquisition unit 120 (such as image capture devices 122, 124, and 126 having fields of view 202, 204, and 206) may capture a first, second, and third plurality of images of an area forward and/or to the side of vehicle 200 and transmit them over a digital connection (e.g., USB, wireless, Bluetooth, etc.) to processing unit 110. In some embodiments, processing unit 110 may receive the first, second, and third plurality of images via three or more data interfaces. For example, each of image capture devices 122, 124, 126 may have an associated data interface for communicating data to processing unit 110. The disclosed embodiments are not limited to any particular data interface configurations or protocols.

[0137] At step 720, processing unit 110 may analyze the first, second, and third plurality of images to detect features within the images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, road hazards, and the like. The analysis may be performed in a manner similar to the steps described in connection with FIGS. 5A-5D and 6, above. For instance, processing unit 110 may perform monocular image analysis (e.g., via execution of monocular image analysis module 402 and based on the steps described in connection with FIGS. 5A-5D, above) on each of the first, second, and third plurality of images. Alternatively, processing unit 110 may perform stereo image analysis (e.g., via execution of stereo image analysis module 404 and based on the steps described in connection with FIG. 6, above) on the first and second plurality of images, the second and third plurality of images, and/or the first and third plurality of images. The processed information corresponding to the analysis of the first, second, and/or third plurality of images may be combined. In some embodiments, processing unit 110 may perform a combination of monocular and stereo image analyses. For example, processing unit 110 may perform monocular image analysis (e.g., via execution of monocular image analysis module 402) on the first plurality of images and stereo image analysis (e.g., via execution of stereo image analysis module 404) on the second and third plurality of images. The configuration of image capture devices 122, 124, and 126—including their respective locations and fields of view 202, 204, and 206—may influence the types of analyses conducted on the first, second, and third plurality of images. The disclosed embodiments are not limited to a particular configuration of image capture devices 122, 124, and 126, or the types of analyses conducted on the first, second, and third plurality of images.

[0138] In some embodiments, processing unit 110 may perform testing on system 100 based on the images acquired and analyzed at steps 710 and 720. Such testing may provide an indicator of the overall performance of system 100 for certain configurations of image capture devices 122, 124, and 126. For example, processing unit 110 may determine the proportion of "false hits" (e.g., cases where system 100 incorrectly determined the presence of a vehicle or pedestrian) and "misses."

[0139] At step 730, processing unit 110 may cause one or more navigational responses in vehicle 200 based on information derived from two of the first, second, and third plurality of images. Selection of two of the first, second, and third plurality of images may depend on various factors, such as, for example, the number, types, and sizes of objects detected in each of the plurality of images. Processing unit 110 may also make the selection based on image quality and resolution, the effective field of view reflected in the images, the number of captured frames, the extent to which one or more objects of interest actually appear in the frames (e.g., the percentage of frames in which an object appears, the proportion of the object that appears in each such frame, etc.), and the like.

IPR2025-00944
Tesla EX1002 Page 1367

[0140] In some embodiments, processing unit 110 may select information derived from two of the first, second, and third plurality of images by determining the extent to which information derived from one image source is consistent with information derived from other image sources. For example, processing unit 110 may combine the processed information derived from each of image capture devices 122, 124, and 126 (whether by monocular analysis, stereo analysis, or any combination of the two) and determine visual indicators (e.g., lane markings, a detected vehicle and its location and/or path, a detected traffic light, etc.) that are consistent across the images captured from each of image capture devices 122, 124, and 126. Processing unit 110 may also exclude information that is inconsistent across the captured images (e.g., a vehicle changing lanes, a lane model indicating a vehicle that is too close to vehicle 200, etc.). Thus, processing unit 110 may select information derived from two of the first, second, and third plurality of images based on the determinations of consistent and inconsistent information.

[0141] Navigational responses may include, for example, a turn, a lane shift, braking, a change in acceleration, and the like. Processing unit 110 may cause the one or more navigational responses based on the analysis performed at step 720 and the techniques as described above in connection with FIG. 4. Processing unit 110 may also use data derived from execution of velocity and acceleration module 406 to cause the one or more navigational responses. In some embodiments, processing unit 110 may cause the one or more navigational responses based on a relative position, relative velocity, and/or relative acceleration between vehicle 200 and an object detected within any of the first, second, and third plurality of images. Multiple navigational responses may occur simultaneously, in sequence, or any combination thereof.

[0142] Radar-Cued Visual System for Vehicle Navigation

[0143] Radar systems offer the potential for detecting objects in an environment surrounding a vehicle, even in conditions where visibility may be poor. Therefore, there is an interest in including radar sensor systems on vehicles for purposes of aiding in navigation. For example, driver-assist systems, collision warning/avoidance systems, autonomous vehicles, and semi-autonomous vehicles may all benefit from the potential target object detection capabilities that radar systems offer.

[0144] While radar sensors may provide accurate ranging information, they also can exhibit certain deficiencies. First, radar sensors may generate frequent "ghost" detections, which may cause false positives in a vehicle navigation or collision avoidance system. Such unwanted ghost detections may result, for example, when objects are present in the vicinity of the vehicle, and those objects exhibit good radar reflectivity, but pose little or no threat to the vehicle. Such objects may include those with non-upright profiles or insignificant size. Manhole covers, soda cans, small debris, etc. may result in radar ghost detections. In some instances, ghost detections may occur even when no object is present.

[0145] Second, radar sensors may offer angular resolution and angular accuracy characteristics that may be unsuitable for certain vehicle applications. For example, forward collision warning systems and vehicle navigation systems (either autonomous or driver assist) may require accurate information regarding the angular location of a target object relative to the vehicle or a path of travel of the vehicle.

IPR2025-00944
Tesla EX1002 Page 1368

[0146] To enhance the reliability and usability of radar-based sensor systems, such radar systems may be deployed on a vehicle in combination with an image sensor system including one or more image acquisition devices. The radar and image systems may be used in tandem to reduce or eliminate the occurrence of ghost detections by the radar system and to accurately determine one or more of the angular position of a target object relative to the vehicle and the solid viewing angle that a target object occupies within the field of view of an image acquisition device. From a combination of range information from the radar, as well as spatial location information derived from the image system, performance of collision warning systems, driver assist systems, and other types of vehicle navigation systems may be enhanced.

[0147] FIG. 8 provides a diagrammatic representation of a system 100 along with all of the same parts and features associated with system 100 as described above relative to FIGS. 1, 2A, 2B, 2C, 2D, 2E, and others. In the embodiment shown in FIG. 8, a radar unit 801, including one or more radar sensors 802, is included in system 100. Radar unit 801 may provide an output either directly or indirectly to application processor 180 (or any other suitable logic based device) via, for example, data interface 128.

[0148] In some embodiments, radar unit 801 may be combined with an image acquisition unit 120 that includes a single image capture device or multiple image capture devices. In some cases, radar unit 801 may be included on a vehicle equipped with two image capture devices (e.g., image capture device 122 and image capture device 124) as shown in FIGS. 9A, 9B, and 9C. In other embodiments, radar unit 801 may be included on a vehicle equipped with three image capture devices (e.g., image capture device 122, image capture device 124, and image capture device 126) as shown in FIGS. 8, 9D, and 9E.

[0149] Radar unit 801 may be located at any suitable location relative to vehicle 200. In some cases, as shown in FIGS. 9A-9E, radar unit 801 may be located on the roof of the vehicle. In other cases, radar unit 801 may be located in front of or behind the windshield, in the grill, in or on the forward bumper (in or on the rear bumper, in some rear facing cases), in or on a side panel (e.g., for blind spot detection), or at any other viable location on vehicle 200.

[0150] Radar unit 801 may include any suitable type of radar useful in identifying the presence of target objects in a vehicle navigation application. For example, in some embodiments, radar unit 801 may include mechanically scanning radar. In other embodiments, electronically scanning radar may be used that reduce or eliminate moving parts from radar unit 801. Any suitable frequencies or frequency ranges may be employed. For example, radar units operating at frequencies of 79 GHz, 77 GHz, 25 GHz, or 24 GHz may be used. Such units may offer various detection ranges along with various scan angles. In some embodiments, radar unit 801 may have a detection range of 100 m, 200 m, 250 m, or more and may offer scan angles of up to 360 degrees.

[0151] In the disclosed embodiments, radar unit 801 may provide information relating to one or multiple target objects in an environment of vehicle 200. Such information may include, for example, an indication of the presence of one or more target objects in the vicinity of vehicle 200 and an indication of a distance (range) to the one or more target objects. In some cases, additional information may be available from radar unit 801, such as a radar range rate (e.g., a rate that a target object is moving toward or away relative to vehicle 200). The radar range and/or radar range rate information provided

from radar unit 801 may serve as an indicator of a time to contact between vehicle 200 and a target object.

[0152] During operation, radar unit 801 may provide a radar range, radar angle, and/or a radar range rate relative to a target object either directly or indirectly to application processor 180. Using this information, application processor 180 may determine a time to contact between vehicle 200 and a particular target object.

[0153] In addition to radar information, application processor 180 and/or image processor 190 may receive image data captured by one or more of image capture device 122, 124, or 126. This image data may be received in the form of a series of image frames captured at predetermined time intervals.

[0154] Using the received radar information together with the plurality of images, processing unit 110 may determine whether a particular radar detection correlates to a real world target object. If a target object is validated, processing unit 110 may identify physical attributes of the target object such as edges of the object, height of the object relative to a road surface or other surface, and angular location of the target object relative to vehicle 200.

[0155] More specifically, a process for validating and locating a target object may include the identification of a rough region of interest (ROI) in one or more of the received plurality of images from the image acquisition unit 120. For example, based on the angle and range associated with a particular radar detection, and knowing values of various parameters associated with the image acquisition device that supplied a certain image (e.g., the camera/radar calibration values, camera focal length, and height of the camera above the road), processing unit 110 may determine what region of an acquired image is likely to include a target object associated with a particular radar detection. Such a rough region of interest in the acquired image may be designated by left, right, bottom, and top boundaries of a region of the acquired image where the target object may be found.

[0156] Next, the optical flow between the rough region of interest in the acquired image and a corresponding region of a second acquired image may be determined. Conceptually, the optical flow between at least a pair of acquired images may be associated with the changing angular position of corresponding points in the images. As an example, as a vehicle travels directly toward a target object on a roadway, the target object will appear to the driver as expanding radially in all directions as the vehicle moves closer to the target object. Similarly, in images captured over time from an image captured device located in or on a vehicle, a target object captured in the images will appear to expand radially in all directions relative to images captured earlier in a sequence of images. The non-moving point directly in front of the vehicle at the center of the horizon, from which the objects in a sequence of images captured from a moving vehicle appear to emanate, is referred to as the focus of expansion.

[0157] Regarding objects in motion, if the target object is moving relative to the stationary world, the FOE will reflect the relative motion between the host vehicle and the target object. For example, if a vehicle is approaching from the right (at an intersection for example) the FOE will be to the right of the image. If the vehicle and the target object are on a collision course, the FOE will be on the target object. Because the FOE depends on the motion of the target, the FOE may be useful in isolating the target from its surroundings within one or more images of a scene.

IPR2025-00944
Tesla EX1002 Page 1370

[0158] In practice, the optical flow between two images and, more specifically, between a particular feature point within a region of interest in a first image and a corresponding feature point in a region of interest in a second image may be determined in a number of ways. As a first step, pairs of corresponding feature points from the two images may be determined. In one technique, a pair of corresponding feature points may be determined by defining a patch around a particular feature point in the region of interest of the first image (or about every pixel in the region of interest of the first image). Then, every defined patch from the first image may be compared with patches from the second image within a predetermined search area (e.g., within a predetermined or selectable range of pixels) relative to a selected patch from the first image. From every patch comparison, a correlation score may be determined based on the similarity of the compared patches. The patch from the second image that provides the highest correlation score with the patch from the first image containing the feature point of interest determines the location of that feature point in the second image.

[0159] Another technique for determining pairs of corresponding feature points from the two images involves matching or registration based on detected feature points. For example, certain feature points (e.g., corners of objects or other well-defined points in the image data) may be detected in the first image through image analysis. Similar feature points may also be detected in a second image. Patches may be defined around the detected feature points from both the first and second images. Any number of feature points may be detected within a region of interest (e.g., a region that includes 10 feature points, 20 feature points, 50 feature points, or more). A correlation score may be calculated for each patch defined for the first image by comparing each individual patch of the first image with each patch of the second image. The pair of patches (one from the first image and one from the second image) with the highest correlation score indicates the location of a particular, detected feature point in both the first and second images.

[0160] Once the locations are known for a set of feature points (or patches or other suitable image segment) in the first image and the corresponding set of feature points (or patches or any other suitable image segment) in the second image, the optical flow between the images may be determined. For example, for each corresponding pair of feature points (patches or image segments) in the pair of images, the motion that resulted in movement of the feature points from their location in the first image to their location in the second image may be determined. This motion represents the optical flow between the corresponding points of the pair of images. The optical flow inside the region of interest defined for the first image may be represented as a set of {(x,y,u,v)i} quadruples representing image positions (x,y)i and the 2D flow (u,v)i measured at those positions.

[0161] If a rolling shutter is employed on one or more of the image capture devices, an observed vertical motion in the image may indicate or result from the fact that the two points were imaged at a slightly different time. If the rolling shutter scans from top to bottom an upward flow of N lines, for example, means that the time between imaging is less than a frame (by N line scan times). To get more accurate estimates of the optical flow, the flow vectors can be corrected by extending in the vector direction by frameTime/(frameTime-N*linetime).

[0162] Next, an expected optical inflation value (ds) may be determined based on information obtained from radar unit 801. For example, using the radar range to the

Tesla EX1002 Page 1371 IPR2025-00944

target object, radar range rate between the target object and vehicle, and the time between captured images (dt), the expected optical inflation value in the region of interest may be represented as:

Optical inflation(ds)=dt/TTC=dt*(range rate)/range;

[0163] where TTC is the time to contact between the target object and the vehicle.

[0164] Using the expected optical inflation value and the determined optical flow information, the target object may be isolated within the region of interest. For example, using as input the set of {(x,y,u,v)i} quadruples along with the optical inflation (ds) expectation, a consensus set of pixels {(x,y)} may be determined where the optical flow substantially matches the expected optical inflation. This consensus set of pixels (image segments, points, or other portions of the region of interest in the plurality of images) may represent a target object region in the plurality of images occupied by the target object. The points, pixels, image segments, etc. associated with this target object region may share a common focus of expansion, as opposed to other image areas within the region of interest that do not correspond to the target object. For those other image areas (e.g., image portions associated with the road or background objects), the determined optical flow for those areas will likely not match the expected optical inflation determined based on the target object range and range rate, because the background objects and road (aside from the portion of the road located at the same range as the target object) will have ranges different from the target object.

[0165] Regarding the focus of expansion (FOE) and its relationship to the TTC or time to contact, all stationary points in a captured image will share the same true FOE, which corresponds to the direction that vehicle 200 is traveling. For an upright object, if a TTC to that object is available (e.g., based on the output of radar unit 801), the estimated FOE for every point in the object's image may be derived based on the optical flow (i.e., the points' motion in the image). If the TTC for every point in an image was available, and every point in the image could be tracked from frame to frame, then the same FOE may be derived for all points in the image. In practice, however, a TTC may be available only for a single radar-detected object or a group of radar-detected objects and not for every point associated with a particular image. Deriving the FOE based on the TTC value for a single detected object may yield the same FOE for all imaged points associated with that object (assuming the TTC is valid for all points of the object). On the other hand, using the TTC associated with a target object to derive and FOE associated with a background object or a foreground object at a range different from the detected target object may result in an FOE different from the one derived for the target object. These differences in derived FOE values may be used to isolate a target object within an image to both confirm that a radar detection corresponds to an actual object and to determine the size, shape, and spatial orientation of the target object relative to vehicle 200.

[0166] Examining the process for isolating the target object in the image frames in further detail, when approaching an upright target object in a roadway, the optical flow of a pixel/patch in the image can be described using a model as a function of the position in the image, TTC, and the FOE. The radar may detect an object that moves from distance Z1 in frame 1 to distance Z2 in frame 2. A feature on the target object having coordinates (x1, y1) may be detected in frame 1 and tracked to coordinates (x2, y2) in frame 2. The point may be assumed as lying at a constant lateral distance X from the optical axis and at a constant height Y above the road. Using a pinhole camera

Tesla EX1002 Page 1372 IPR2025-00944

model, the motion (Ax, Ay) of the feature point in the image may be given by:

x1=x0+X·f/Z1,y1=y0−(cH−Y)·f/Z1

x2=X0+X·f/Z2,y2=y0−(cH−Y)·f/Z2

Image available on "Original document"
Δx≡x2−x1=X·f·(1/Z2−1/Z1)=(x1−x0)·Z1·(1/Z2−1/Z1)=(x1−x0)·(Z1−Z2)/Z2

Image available on "Original document"
Δy≡x2−y1=−(cH−Y)·f·(1/Z2−1/Z1)=(y1−y0)·Z1·(1/Z2−1/Z1)=(y1−y0)·(Z1−Z2)/Z2

[0167] where the camera's focal length is given by f and its height by cH. The FOE is unknown and given by (x0, y0).

[0168] The motion equations can therefore be parametrized as follows:

Δx=S·(x−x0)

Δy=S·(y−y0)

[0169] where the scale S=(Z1−Z2)/Z2 describes the expansion in the image of the object relative to the FOE. The time to contact (TTC) with the object, assuming that the relative speed between the target object and the vehicle remains constant, can be derived from the scale S as follows:

S≡−ΔZ/Z

Image available on "Original document"
TTC=−Z/Vrel=−Z·Δt/ΔZ=Δt/S

[0170] In practice, the consensus set of feature points may be described by the largest union of points whose observed motion can be described by a certain fixed FOE, given the TTC obtained from the radar target. The size of this consensus set will define the probability that there is an actual upright object with a TTC as obtained from the radar target within the region of interest as defined by the radar target.

[0171] Any suitable approach may be used for detecting the set of pixels in the image that correspond to the expected TTC, as obtained from the radar target. In one embodiment, an optimization technique may be used, such that given a set of N feature points {(xi, yi)}i=1<N>, observed motions {(Δxi, Δyi)}i=1<N >and radar scale (minus the range divided by the range-rate), we can derive the estimated FOE ({circumflex over (x)}0,ŷ0) that best describes the observed motion (i.e. the model that predicts motions (
Image available on "Original document"
i,
Image available on "Original document"
i) that are optimal in a sense that they are 'closest' to the observed motions). The number of points whose observed motion vectors are sufficiently close (euclidian distance measure) to the expected motion vectors, will define the consensus set.

[0172] In another embodiment, a Hough transform approach may be used. In this approach, for each tracked point i, the FOE {x0(S,Δxi, xi), y0(S, Δyi, yi)} may be derived

as a function of the point's position, measured translation in the image, and the expected TTC or scale. A two-dimensional histogram (Hough transformation) may be constructed by discretizing the obtained FOE. Then, the value in the histogram that obtained the highest number of votes may be determined in order to define the consensus set of points that contributed to that value.

[0173] As an output of the target identification process based on the radar information together with the visual image information, a target object region may be identified in the images. This target object region may be delineated by an image bounding box (or other shape) encompassing the consensus set {(x,y)} of pixels that comply with the expected optical inflation and share a common FOE in agreement with a TTC determined based on the detected radar information. Using this target object region, the edges of a physical target object may be identified and located relative to vehicle 200. Further, the height, shape, width or any other dimensional characteristics of the target object may be determined. Combining the target object region in an acquired image with known parameters of the image acquisition device (e.g., the focal length, height above the road, orientation relative to vehicle 200, etc.) may enable determination of the angular orientation of the target object relative to vehicle 200.

[0174] Such information may be helpful in determining whether the target object resides in a path of vehicle 200. For example, if based on the location in a certain image of the target object region, processing unit 110 determines that the target object falls within a travel path of vehicle 200, processing unit 200 may cause any of a number of system responses. In some embodiments the system response may include a navigation response, such as braking, changing steering direction, etc. If processing unit 110 determines that allowing vehicle 200 to continue along the current travel direction may result in a collision with the target object, then processing unit may warn a driver of vehicle 200, automatically change a steering direction, automatically apply the vehicle's brakes, etc. For example, processing unit 110 may provide a notice to the driver of the vehicle (e.g., an audible sound or visual indicator via user interface 170) and/or provide control signals to one or more of throttling system 220, braking system 230, and steering system 240 to navigate vehicle 200. In some embodiments, processing unit 110 may continue to analyze images and/or radar output to determine whether the system response should be interrupted (e.g., if the object has moved out of the path of vehicle 200, discontinue applying the brakes).

[0175] FIGS. 10 through 13 represent various phases of the process for identifying a target object region within one or more images. For example, FIG. 10 depicts a scene 1001 including a road 1002 on which vehicle 200 is traveling. FIG. 10 may represent scene 1001 from, for example, a forward facing perspective of vehicle 200. Consistent with the discussion above, the true FOE for scene 1001 as vehicle 200 travels along road 1002 will appear to be located at point 1004. Positioned in one of the lanes of road 1002 is an upright hazard in the form of a trash can 1006. As vehicle 200 approaches trash can 1006, radar unit 801 may acquire range and range rate information relative to vehicle 200 and trash can 1006. From this information, a TTC between vehicle 200 and trash can 1006 may be determined.

[0176] As vehicle 200 approaches trash can 1006, images of the scene, including trash can 1006, may be acquired by one or more of image acquisition devices 122, 124, and/or 126. As noted above, the radar information may also be used to identify regions of interest within captured images where the detected target object may be located. FIGS. 11A and 11B show representative regions of interest from each of a pair of

IPR2025-00944
Tesla EX1002 Page 1374

acquired images, as determined based on the information associated with detection by the radar unit 801 of trash can 1006. As vehicle 200 travels along the center of road 1002, a series of images may be captured showing trash can 1006. Due to the motion of vehicle along road 1002 in the direct of FOE 1004, images of trash can 1006 captured later in the sequence will show trash can 1006 expanding within the image frame as well as moving down and to the right within the image frame (as represented by FIGS. 11A and 11B).

[0177] The optical flow from the image region shown in FIG. 11A to the image region in FIG. 11B may be determined. For example, based on any of the methods discussed above, the motion of any of selected target or reference points 1008, 1010, and 1012 (or any other number of target or reference points) on trash may 1006 may be determined from the frame shown in FIG. 11A to the frame shown in FIG. 11B. In some cases, the target or reference points may coincide with features in the image distinguishable from the immediate surroundings (e.g., a corner of an object, small dark or light areas of an object, a light source, a handle, bump, etc.). Such target or reference points may be readily identified in subsequent images and may provide a high level of confidence that the identified reference point in one or more subsequent images corresponds with the reference point chosen from the original image. The determined motion constitutes the measured optical flow, which is a 2D vector, (u, v).

[0178] FIG. 12A provides a diagrammatic representation of plotted values of the horizontal optical flow between FIGS. 11A and 11B. Similarly, FIG. 12B provides a diagrammatic representation of plotted values of the vertical optical flow between FIGS. 11A and 11B. Using the radar range and radar range rate information, an expected optical inflation magnitude may be determined as a scale factor. And, given the scale from the radar, all pixels, points, image segments, patches, etc. that share a common FOE may be determined. The optical flow determination is separable, meaning that it can be determined or solved separately for the measured vertical flow and the measured horizontal flow. Thus, for a given radar-determined scale and vertical flow, for example, associated with vertical position (y) in the image of FIG. 12B, the following expression may be used:

[0179] FOE.y=y−V/(scale−1), where V is vertical flow and the scale is the optical inflation magnitude derived from the radar range and range rate values.

[0180] Thus, applying this expression to the vertical flow data represented by FIG. 12B and scaling by the radar-determined scaling factor, a plot similar to the one shown in FIG. 13 may be obtained. FIG. 13 includes a target object region 1301 coinciding with points in the image that, for the measured radar range and radar rate values, share a common FOE. Target object region 1301 coincides with trash can 1006 and may be used to confirm the presence of an actual target object as well as locate the trash can spatially relative to vehicle 200 (using not only the radar range and radar rate measurements, but also known characteristics associated with the image acquisition device(s) that captured the images of trash can 1006(. As a result, edge locations, size and shape characteristics, etc. may all be determined for trash can 1006, and if trash can 1006 resides in the path of vehicle 200, system 100 may provide a warning to a driver of vehicle 200 or cause one or more navigational changes to occur.

[0181] The presently described embodiments may operate in a stand-alone capacity in a navigational system for a vehicle. In some embodiments, though, the radar-cued navigational system may be integrated with or operated in conjunction with one or more

Tesla EX1002 Page 1375 IPR2025-00944

other navigational systems for a vehicle. For example, in some cases the radar-cued visual navigation system may operate with or be integrated with an advanced driver assist system (ADAS) and may be used to provide at least some functionality to the ADAS system. Navigational assist modules for traffic sign recognition (TSR), lane departure warning (LDW), intelligent headlight control (IHC), and/or traffic light detection may all benefit from radar-cued output. For example, in some embodiments, as vehicle 200 approaches an intersection, the optical flow of identified light sources present in two or more images may be observed. The optical flow of such light sources may be compared to time to contact (TTC) values available from radar unit 801 for stationary objects in the surrounding environment of vehicle 200 (such TTC values may be determined, for example, based on radar range rate and vehicle speed). Where a TTC value for a stationary object agrees with the optical flow observed for a light source in the acquired images, application processor 180, e.g., may determine that the light source is a traffic light or is a potential traffic light. If the light source is determined to be a traffic light, the system may further determine the color associated with the light source. If the light source is red or yellow, then a warning may be issued for the driver to slow to a stop. Alternatively, the system could automatically apply the brakes and slow the vehicle to a stop. If the light source is green, then no navigational change may be deemed necessary.

[0182] The presently disclosed system may also operate according to different modes. For example, in some situations there may be benefits from operating the navigational system in a wide detect mode. For example, in an urban environment where vehicle speeds may be reduced, but the number of potential target objects surrounding vehicle 200, such as pedestrians, lights, intersections, trash cans, etc., may be higher, the system may select a wide mode of operation. In other environments, such as a rural road, for example, where speeds may be higher than in a city, and the number of potential target objects may be fewer, the navigational system may select a narrow mode of operation. In some embodiments, the wide mode of operation may be associated with analysis of images from an image capture device having a wider field of view than another available image capture device having a narrower field of view. The navigational system may select a certain camera (or lens where lenses are selectable or adjustable) for image capture or image analysis based on any suitable factor. For example, applications processor may select image acquisition device 122 for image capture or for image stream analysis in certain situations and may select image acquisition device 124 or 126 in other situations. In some embodiments, a camera, lens, or image stream may be selected based on radar range or range rate information obtained from radar unit 801, based on vehicle speed, based on a recognized environment or target object density derived from the captured images and/or radar information, or using any other suitable parameter.

[0183] The foregoing description has been presented for purposes of illustration. It is not exhaustive and is not limited to the precise forms or embodiments disclosed. Modifications and adaptations will be apparent to those skilled in the art from consideration of the specification and practice of the disclosed embodiments. Additionally, although aspects of the disclosed embodiments are described as being stored in memory, one skilled in the art will appreciate that these aspects can also be stored on other types of computer readable media, such as secondary storage devices, for example, hard disks or CD ROM, or other forms of RAM or ROM, USB media, DVD, Blu-ray, or other optical drive media.

[0184] Computer programs based on the written description and disclosed methods are

Tesla EX1002 Page 1376 IPR2025-00944

within the skill of an experienced developer. The various programs or program modules can be created using any of the techniques known to one skilled in the art or can be designed in connection with existing software. For example, program sections or program modules can be designed in or by means of .Net Framework, .Net Compact Framework (and related languages, such as Visual Basic, C, etc.), Java, C++, Objective-C, HTML, HTML/AJAX combinations, XML, or HTML with included Java applets.

[0185] Moreover, while illustrative embodiments have been described herein, the scope of any and all embodiments having equivalent elements, modifications, omissions, combinations (e.g., of aspects across various embodiments), adaptations and/or alterations as would be appreciated by those skilled in the art based on the present disclosure. The limitations in the claims are to be interpreted broadly based on the language employed in the claims and not limited to examples described in the present specification or during the prosecution of the application. The examples are to be construed as non-exclusive. Furthermore, the steps of the disclosed methods may be modified in any manner, including by reordering steps and/or inserting or deleting steps. It is intended, therefore, that the specification and examples be considered as illustrative only, with a true scope and spirit being indicated by the following claims and their full scope of equivalents.

Espacenet

Claims: EP3108264 (A2) — 2016-12-28

ADVANCED DRIVER ASSISTANCE SYSTEM BASED ON RADAR-CUED VISUAL IMAGING

Claims not available for EP3108264 (A2)
Claims of corresponding document: US2015234045 (A1)

A high quality text as facsimile in your desired language may be available amongst the following family members:

CN106461774 (A) US2015234045 (A1) WO2015125022 (A2) CN109884618 (A) US2017248694 (A1) US2019235073 (A1)

- Original claims
- Claims tree

The EPO does not accept any responsibility for the accuracy of data and information originating from other authorities than the EPO; in particular, the EPO does not guarantee that they are complete, up-to-date or fit for specific purposes.

1. What is claimed is: 1. A navigation system for a vehicle, the system comprising:
at least one image capture device configured to acquire a plurality of images of an environment of a vehicle;
a radar sensor to detect an object in the environment of the vehicle and to provide and output including range information indicative of at least one of a range or range rate between the vehicle and the object; and
at least one processing device programmed to:
receive the plurality of images from the at least one image capture device;
receive the output from the radar sensor;
determine, for each of a plurality of image segments in a first image, from among the plurality of images, and corresponding image segments in a second image, from among the plurality of images, an indicator of optical flow;
determine a value indicative of an expected optical inflation associated with the object in the plurality of images based at least upon the range information from the output of the radar sensor;
determine one or more areas of the second image where the value indicative of an expected optical inflation associated with the object substantially matches the indicator of optical flow;
identify a target object region, including at least a subset of the determined one or more areas of the second image; and
cause a system response based on the identified target object region.

2. The navigation system of claim 1, wherein the target object region is associated with

a common focus of expansion.

3. The navigation system of claim 1, wherein the at least one processing device is further programmed to determine an angular position of the object relative to the vehicle based on a determined location of the target object region.

4. The navigation system of claim 1, wherein the at least one processing device is further programmed to determine a location of the object relative to the vehicle based on the identified target object region.

5. The navigation system of claim 1, wherein the system response includes at least one of braking or a change in current travel direction of the vehicle if the at least one processing device determines that the current travel direction may result in a collision with the object.

6. The navigation system of claim 1, wherein the system response is a navigational response including braking.

7. The navigation system of claim 1, wherein the system response is a navigational response including a change in a direction of travel of the vehicle.

8. The navigation system of claim 1, wherein the at least one processing device is further programmed to identify edges of the object based on the identified target object region.

9. The navigation system of claim 1, wherein the at least one processing device is further programmed to determine a height of the object based on the identified target object region.

10. The navigation system of claim 1, wherein the output of the radar sensor includes an indicator of a time to contact between the vehicle and the object.

11. The navigation system of claim 1, wherein the at least one processing device is configured to account for at least one effect caused by use of a rolling shutter to acquire the plurality of images in determining the indicator of optical flow.

12. The navigation system of claim 1, wherein the at least one processing device is configured to select an operational mode from among at least a first mode and a second mode, where the first mode is associated with a wider image capture field of view than the second mode, based on the range information from the radar sensor.

13. A vehicle comprising a navigation system, the navigation system comprising:
at least one image capture device configured to acquire a plurality of images of an environment of the vehicle;
a radar sensor to detect an object in the environment of the vehicle and to provide and output including range information indicative of at least one of a range or range rate between the vehicle and the object; and
at least one processing device programmed to:
receive the plurality of images from the at least one image capture device;
receive the output from the radar sensor;
determine, for each of a plurality of image segments in a first image, from among the plurality of images, and corresponding image segments in a second image, from

among the plurality of images, an indicator of optical flow;
determine a value indicative of an expected optical inflation associated with the object in the plurality of images based at least upon the range information in the output of the radar sensor;
determine one or more areas of the second image where the value indicative of an expected optical inflation associated with the object substantially matches the indicator of optical flow;
identify a target object region, including at least a subset of the determined one or more areas of the second image; and
cause a system response based on the identified target object region.

14. The vehicle of claim 13, wherein the at least one processing device is further programmed to determine an angular position of the object relative to the vehicle based on a determined location of the target object region.

15. The vehicle of claim 13, wherein the at least one processing device is further programmed to determine a location of the object relative to the vehicle based on the identified target object region.

16. The vehicle of claim 13, wherein the system response includes at least one of braking or a change in current travel direction of the vehicle if the at least one processing device determines that the current travel direction may result in a collision with the object.

17. The vehicle of claim 13, wherein the at least one processing device is further programmed to identify edges of the object based on the identified target object region.

18. The vehicle of claim 13, wherein the at least one processing device is further programmed to determine a height of the object based on the identified target object region.

19. The vehicle of claim 13, wherein the output of the radar sensor includes an indicator of a time to contact between the vehicle and the object.

20. The vehicle of claim 13, wherein the processing device is further configured to identify a light source in at least one of the plurality of images, determine whether the light source is a traffic light based on the indicator of optical flow and the output of the radar sensor, and cause automatic application of brakes of the vehicle if the traffic light is determined to be yellow or red.

21. A method for navigating a vehicle based on radar detection and visual image data, the method comprising:
acquiring a plurality of images of an environment of the vehicle;
detecting an object in the environment of the vehicle using a radar sensor and providing an output of the radar sensor including range information indicative of at least one of a range or range rate between the object and the vehicle;
receiving, at one or more processing devices, the plurality of images from the at least one image capture device and the output from the radar sensor;
determining an indicator of optical flow, for each of a plurality of image segments in a first image, from among the plurality of images, and corresponding image segments in a second image, from among the plurality of images;
identifying a target object region having a substantially common focus of expansion,

based the range information provided by the radar sensor and the determined indicators of optical flow; and
causing a system response based on the identified target object region.

22. A navigation system for a vehicle, the system comprising:
at least one image capture device configured to acquire a plurality of images of an environment of a vehicle;
a radar sensor to detect an object in the environment of the vehicle and to provide and output including range information indicative of at least one of a range or range rate between the vehicle and the object; and
at least one processing device programmed to:
receive the plurality of images from the at least one image capture device;
receive the output from the radar sensor;
determine, for each of a plurality of image segments in a first image, from among the plurality of images, and corresponding image segments in a second image, from among the plurality of images, an indicator of optical flow;
identify a target object region having a substantially common focus of expansion, based on the range information provided by the radar sensor and the determined indicators of optical flow; and
cause a system response based on the identified target object region.

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Veröffentlichungsnummer:
(11) Publication number: **EP 3 108 264 A0**
(11) Numéro de publication:

Internationale Anmeldung veröffentlicht durch die Weltorganisation für geistiges Eigentum unter der Nummer:

WO 2015/125022 (Art. 153(3) EPÜ).

International application published by the World Intellectual Property Organization under number:

WO 2015/125022 (Art. 153(3) EPC).

Demande internationale publiée par l'Organisation Mondiale de la Propriété Intellectuelle sous le numéro:

WO 2015/125022 (art. 153(3) CBE).

Espacenet

Bibliographic data: EP3143607 (A1) — 2017-03-22

SYSTEMS AND METHODS FOR CURB DETECTION AND PEDESTRIAN HAZARD ASSESSMENT

Inventor(s): SPRINGER OFER [IL] ± (SPRINGER, Ofer)

Applicant(s): MOBILEYE VISION TECHNOLOGIES LTD [IL] ± (Mobileye Vision Technologies Ltd)

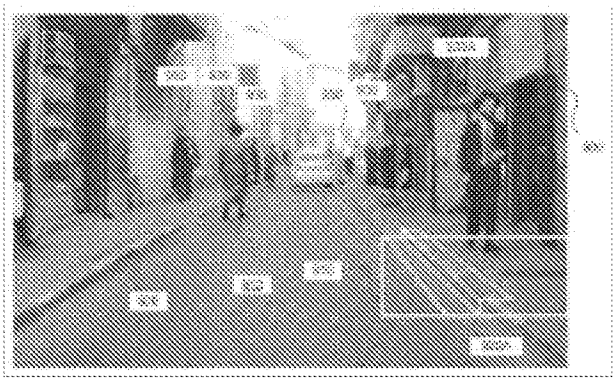
Classification: - international: *G06K9/00; G08G1/16*
- cooperative: **B60W30/09 (US)**; **B60W50/14 (US)**; **G06K9/00798 (EP, CN, US)**; **G06K9/00805 (EP, CN, US)**; **G06K9/4604 (US)**; **G06K9/6267 (US)**; **G06T3/0093 (US)**; **G06T7/13 (EP, US)**; **G06T7/97 (EP, US)**; **G08G1/166 (EP, CN, US)**; **G08G1/167 (EP, CN, US)**; **H04N7/183 (US)**; G06T2207/10021 (US); G06T2207/30241 (US); G06T2207/30261 (US)

Application number: EP20150772006 20150514 Global Dossier

Priority number(s): US201461993050P 20140514 ; WO2015IB01542 20150514

Also published as: CN106663193 (A) CN106663193 (B) CN112580456 (A) US10303958 (B2) US10685246 (B2) more

Abstract not available for EP3143607 (A1)
Abstract of corresponding document: US2015332114 (A1)



A detection system for a vehicle is provided. The detection system may include at least one image capture device configured to acquire a plurality of images of an area forward of the vehicle, the area including a curb separating a road surface from an off-road surface and a data interface. The detection system may also include at least one processing device programmed to receive the plurality of images via the data interface, and determine a plurality of curb edge line candidates in

the plurality of images. The at least one processing device may be further programmed to identify at least one edge line candidate as an edge line of the curb.

IPR2025-00944
Tesla EX1002 Page 1384

Espacenet

Description: EP3143607 (A1) — 2017-03-22

SYSTEMS AND METHODS FOR CURB DETECTION AND PEDESTRIAN HAZARD ASSESSMENT

Description not available for EP3143607 (A1)
Description of corresponding document: US2015332114 (A1)

A high quality text as facsimile in your desired language may be available amongst the following family members:

CN106663193 (A) US2015332114 (A1) WO2015177648 (A1) CN112580456 (A) US2016371550 (A1) US2018189579 (A1) US2019251375 (A1) US2020320315 (A1)

The EPO does not accept any responsibility for the accuracy of data and information originating from other authorities than the EPO; in particular, the EPO does not guarantee that they are complete, up-to-date or fit for specific purposes.

CROSS REFERENCE TO RELATED APPLICATION

[0001] This application claims the benefit of priority of U.S. Provisional Patent Application No. 61/993,050, filed on May 14, 2014. The foregoing application is incorporated herein by reference in its entirety.

BACKGROUND

[0002] I. Technical Field

[0003] The present disclosure relates generally to autonomous vehicle navigation and, more specifically, to systems and methods that use cameras to detect a curb associated with a roadway and assess potential pedestrian hazards.

[0004] II. Background Information

[0005] As technology continues to advance, the goal of a fully autonomous vehicle that is capable of navigating on roadways is on the horizon. Primarily, an autonomous vehicle may be able to identify its environment and navigate without input from a human operator. Autonomous vehicles may also take into account a variety of factors and make appropriate decisions based on those factors to safely and accurately reach an intended destination. For example, various objects—such as other vehicles and pedestrians—are encountered when a vehicle typically travels a roadway. Autonomous

driving systems may recognize these objects in a vehicle's environment and take appropriate and timely action to avoid collisions. Additionally, autonomous driving systems may identify other indicators—such as traffic signals, traffic signs, and lane markings—that regulate vehicle movement (e.g., when the vehicle must stop and may go, a speed at which the vehicle must not exceed, where the vehicle must be positioned on the roadway, etc.). Autonomous driving systems may need to determine when a vehicle should change lanes, turn at intersections, change roadways, etc. As is evident from these examples, many factors may need to be addressed in order to provide an autonomous vehicle that is capable of navigating safely and accurately.

SUMMARY

[0006] Embodiments consistent with the present disclosure provide systems and methods for autonomous vehicle navigation. The disclosed embodiments may use cameras to provide autonomous vehicle navigation features. For example, consistent with the disclosed embodiments, the disclosed systems may include one, two, or more cameras that monitor the environment of a vehicle and cause a navigational response based on an analysis of images captured by one or more of the cameras.

[0007] Consistent with a disclosed embodiment, a detection system for a vehicle is provided. The detection system may include at least one image capture device programmed to acquire a plurality of images of an area forward of the vehicle, the area including a curb separating a road surface from an off-road surface and a data interface. The detection system may also include at least one processing device configured to receive the plurality of images via the data interface, and determine a plurality of curb edge line candidates in the plurality of images. The at least one processing device may be further programmed to identify at least one edge line candidate as an edge line of the curb.

[0008] Consistent with another embodiment, a vehicle is provided. The vehicle may include a body, at least one image capture device configured to acquire a plurality of images of an area forward of the vehicle, the area including a curb separating a road surface from an off-road surface, and a data interface. The vehicle may also include at least one processing device programmed to receive the plurality of images via the data interface, and determine a plurality of curb edge line candidates in the plurality of images. The at least one processing device may be further programmed to identify at least one curb edge line candidate as an edge line of the curb, and identify regions in the plurality of images that correspond to the road surface and the off-road surface based on the at least one curb edge line.

[0009] Consistent with another disclosed embodiment, a method for curb detection is disclosed. The method may include acquiring, via at least one image capture device, a plurality of images of an area forward of a vehicle, the area including a curb separating a road surface from an off-road surface, and determining a plurality of curb edge line candidates in the plurality of images. The method may also include determining a height of a step function associated with each curb edge line candidate, and identifying at least one curb edge line candidate as an edge line of the curb based on the determined heights.

[0010] Consistent with other disclosed embodiments, non-transitory computer-readable storage media may store program instructions, which are executed by at least one processing device and perform any of the methods described herein.

IPR2025-00944
Tesla EX1002 Page 1386

[0011] The foregoing general description and the following detailed description are exemplary and explanatory only and are not restrictive of the claims.

BRIEF DESCRIPTION OF THE DRAWINGS

[0012] The accompanying drawings, which are incorporated in and constitute a part of this disclosure, illustrate various disclosed embodiments. In the drawings:

[0013] FIG. 1 is a diagrammatic representation of an exemplary system consistent with the disclosed embodiments.

[0014] FIG. 2A is a diagrammatic side view representation of an exemplary vehicle including a system consistent with the disclosed embodiments.

[0015] FIG. 2B is a diagrammatic top view representation of the vehicle and system shown in FIG. 2A consistent with the disclosed embodiments.

[0016] FIG. 2C is a diagrammatic top view representation of another embodiment of a vehicle including a system consistent with the disclosed embodiments.

[0017] FIG. 2D is a diagrammatic top view representation of yet another embodiment of a vehicle including a system consistent with the disclosed embodiments.

[0018] FIG. 2E is a diagrammatic top view representation of yet another embodiment of a vehicle including a system consistent with the disclosed embodiments.

[0019] FIG. 2F is a diagrammatic representation of exemplary vehicle control systems consistent with the disclosed embodiments.

[0020] FIG. 3A is a diagrammatic representation of an interior of a vehicle including a rearview mirror and a user interface for a vehicle imaging system consistent with the disclosed embodiments.

[0021] FIG. 3B is an illustration of an example of a camera mount that is configured to be positioned behind a rearview mirror and against a vehicle windshield consistent with the disclosed embodiments.

[0022] FIG. 3C is an illustration of the camera mount shown in FIG. 3B from a different perspective consistent with the disclosed embodiments.

[0023] FIG. 3D is an illustration of an example of a camera mount that is configured to be positioned behind a rearview mirror and against a vehicle windshield consistent with the disclosed embodiments.

[0024] FIG. 4 is an exemplary block diagram of a memory configured to store instructions for performing one or more operations consistent with the disclosed embodiments.

[0025] FIG. 5A is a flowchart showing an exemplary process for causing one or more navigational responses based on monocular image analysis consistent with disclosed embodiments.

Tesla EX1002 Page 1387 IPR2025-00944

[0026] FIG. 5B is a flowchart showing an exemplary process for detecting one or more vehicles and/or pedestrians in a set of images consistent with the disclosed embodiments.

[0027] FIG. 5C is a flowchart showing an exemplary process for detecting road marks and/or lane geometry information in a set of images consistent with the disclosed embodiments.

[0028] FIG. 5D is a flowchart showing an exemplary process for detecting traffic lights in a set of images consistent with the disclosed embodiments.

[0029] FIG. 5E is a flowchart showing an exemplary process for causing one or more navigational responses based on a vehicle path consistent with the disclosed embodiments.

[0030] FIG. 5F is a flowchart showing an exemplary process for determining whether a leading vehicle is changing lanes consistent with the disclosed embodiments.

[0031] FIG. 6 is a flowchart showing an exemplary process for causing one or more navigational responses based on stereo image analysis consistent with the disclosed embodiments.

[0032] FIG. 7 is a flowchart showing an exemplary process for causing one or more navigational responses based on an analysis of three sets of images consistent with the disclosed embodiments.

[0033] FIG. 8 is an block diagram of an exemplary memory configured to store instructions for performing one or more operations consistent with disclosed embodiments.

[0034] FIG. 9 is an illustration of an image of an environment of a vehicle consistent with disclosed embodiments.

[0035] FIGS. 10A and 10B are illustrations of images that include a curb consistent with disclosed embodiments.

[0036] FIG. 11 is an illustration of an exemplary modeled curb consistent with disclosed embodiments.

[0037] FIG. 12 is a flowchart of an exemplary process for identifying a curb within an image or video sequence consistent with disclosed embodiments.

[0038] FIG. 13 is a flowchart of an exemplary process for assessing hazards presented by pedestrians consistent with disclosed embodiments.

DETAILED DESCRIPTION

[0039] The following detailed description refers to the accompanying drawings. Wherever possible, the same reference numbers are used in the drawings and the following description to refer to the same or similar parts. While several illustrative embodiments are described herein, modifications, adaptations and other

implementations are possible. For example, substitutions, additions or modifications may be made to the components illustrated in the drawings, and the illustrative methods described herein may be modified by substituting, reordering, removing, or adding steps to the disclosed methods. Accordingly, the following detailed description is not limited to the disclosed embodiments and examples. Instead, the proper scope is defined by the appended claims.

[0040] FIG. 1 is a block diagram representation of a system 100 consistent with the exemplary disclosed embodiments. System 100 may include various components depending on the requirements of a particular implementation. In some embodiments, system 100 may include a processing unit 110, an image acquisition unit 120, a position sensor 130, one or more memory units 140, 150, a map database 160, and a user interface 170. Processing unit 110 may include one or more processing devices. In some embodiments, processing unit 110 may include an applications processor 180, an image processor 190, or any other suitable processing device. Similarly, image acquisition unit 120 may include any number of image acquisition devices and components depending on the requirements of a particular application. In some embodiments, image acquisition unit 120 may include one or more image capture devices (e.g., cameras), such as image capture device 122, image capture device 124, and image capture device 126. System 100 may also include a data interface 128 communicatively connecting processing device 110 to image acquisition device 120. For example, data interface 128 may include any wired and/or wireless link or links for transmitting image data acquired by image accusation device 120 to processing unit 110.

[0041] Both applications processor 180 and image processor 190 may include various types of processing devices. For example, either or both of applications processor 180 and image processor 190 may include a microprocessor, preprocessors (such as an image preprocessor), graphics processors, a central processing unit (CPU), support circuits, digital signal processors, integrated circuits, memory, or any other types of devices suitable for running applications and for image processing and analysis. In some embodiments, applications processor 180 and/or image processor 190 may include any type of single or multi-core processor, mobile device microcontroller, central processing unit, etc. Various processing devices may be used, including, for example, processors available from manufacturers such as Intel®, AMD®, etc. and may include various architectures (e.g., x86 processor, ARM®, etc.).

[0042] In some embodiments, applications processor 180 and/or image processor 190 may include any of the EyeQ series of processor chips available from Mobileye®. These processor designs each include multiple processing units with local memory and instruction sets. Such processors may include video inputs for receiving image data from multiple image sensors and may also include video out capabilities. In one example, the EyeQ2® uses 90 nm-micron technology operating at 332 Mhz. The EyeQ2® architecture consists of two floating point, hyper-thread 32-bit RISC CPUs (MIPS32® 34K® cores), five Vision Computing Engines (VCE), three Vector Microcode Processors (VMP®), Denali 64-bit Mobile DDR Controller, 128-bit internal Sonics Interconnect, dual 16-bit Video input and 18-bit Video output controllers, 16 channels DMA and several peripherals. The MIPS34K CPU manages the five VCEs, three VMP™ and the DMA, the second MIPS34K CPU and the multi-channel DMA as well as the other peripherals. The five VCEs, three VMP® and the MIPS34K CPU can perform intensive vision computations required by multi-function bundle applications. In another example, the EyeQ3®, which is a third generation processor and is six times more

Tesla EX1002 Page 1389 IPR2025-00944

powerful that the EyeQ2®, may be used in the disclosed embodiments.

[0043] Any of the processing devices disclosed herein may be configured to perform certain functions. Configuring a processing device, such as any of the described EyeQ processors or other controller or microprocessor, to perform certain functions may include programming of computer executable instructions and making those instructions available to the processing device for execution during operation of the processing device. In some embodiments, configuring a processing device may include programming the processing device directly with architectural instructions. In other embodiments, configuring a processing device may include storing executable instructions on a memory that is accessible to the processing device during operation. For example, the processing device may access the memory to obtain and execute the stored instructions during operation.

[0044] While FIG. 1 depicts two separate processing devices included in processing unit 110, more or fewer processing devices may be used. For example, in some embodiments, a single processing device may be used to accomplish the tasks of applications processor 180 and image processor 190. In other embodiments, these tasks may be performed by more than two processing devices.

[0045] Processing unit 110 may comprise various types of devices. For example, processing unit 110 may include various devices, such as a controller, an image preprocessor, a central processing unit (CPU), support circuits, digital signal processors, integrated circuits, memory, or any other types of devices for image processing and analysis. The image preprocessor may include a video processor for capturing, digitizing and processing the imagery from the image sensors. The CPU may comprise any number of microcontrollers or microprocessors. The support circuits may be any number of circuits generally well known in the art, including cache, power supply, clock and input-output circuits. The memory may store software that, when executed by the processor, controls the operation of the system. The memory may include databases and image processing software. The memory may comprise any number of random access memories, read only memories, flash memories, disk drives, optical storage, tape storage, removable storage and other types of storage. In one instance, the memory may be separate from the processing unit 110. In another instance, the memory may be integrated into the processing unit 110.

[0046] Each memory 140, 150 may include software instructions that when executed by a processor (e.g., applications processor 180 and/or image processor 190), may control operation of various aspects of system 100. These memory units may include various databases and image processing software. The memory units may include random access memory, read only memory, flash memory, disk drives, optical storage, tape storage, removable storage and/or any other types of storage. In some embodiments, memory units 140, 150 may be separate from the applications processor 180 and/or image processor 190. In other embodiments, these memory units may be integrated into applications processor 180 and/or image processor 190.

[0047] Position sensor 130 may include any type of device suitable for determining a location associated with at least one component of system 100. In some embodiments, position sensor 130 may include a GPS receiver. Such receivers can determine a user position and velocity by processing signals broadcasted by global positioning system satellites. Position information from position sensor 130 may be made available to applications processor 180 and/or image processor 190.

IPR2025-00944
Tesla EX1002 Page 1390

IPR2025-00944
Tesla EX1002 Page 1391

[0048] User interface 170 may include any device suitable for providing information to or for receiving inputs from one or more users of system 100. In some embodiments, user interface 170 may include user input devices, including, for example, a touchscreen, microphone, keyboard, pointer devices, track wheels, cameras, knobs, buttons, etc. With such input devices, a user may be able to provide information inputs or commands to system 100 by typing instructions or information, providing voice commands, selecting menu options on a screen using buttons, pointers, or eye-tracking capabilities, or through any other suitable techniques for communicating information to system 100.

[0049] User interface 170 may be equipped with one or more processing devices configured to provide and receive information to or from a user and process that information for use by, for example, applications processor 180. In some embodiments, such processing devices may execute instructions for recognizing and tracking eye movements, receiving and interpreting voice commands, recognizing and interpreting touches and/or gestures made on a touchscreen, responding to keyboard entries or menu selections, etc. In some embodiments, user interface 170 may include a display, speaker, tactile device, and/or any other devices for providing output information to a user.

[0050] Map database 160 may include any type of database for storing map data useful to system 100. In some embodiments, map database 160 may include data relating to the position, in a reference coordinate system, of various items, including roads, water features, geographic features, businesses, points of interest, restaurants, gas stations, etc. Map database 160 may store not only the locations of such items, but also descriptors relating to those items, including, for example, names associated with any of the stored features. In some embodiments, map database 160 may be physically located with other components of system 100. Alternatively or additionally, map database 160 or a portion thereof may be located remotely with respect to other components of system 100 (e.g., processing unit 110). In such embodiments, information from map database 160 may be downloaded over a wired or wireless data connection to a network (e.g., over a cellular network and/or the Internet, etc.).

[0051] Image capture devices 122, 124, and 126 may each include any type of device suitable for capturing at least one image from an environment. Moreover, any number of image capture devices may be used to acquire images for input to the image processor. Some embodiments may include only a single image capture device, while other embodiments may include two, three, or even four or more image capture devices. Image capture devices 122, 124, and 126 will be further described with reference to FIGS. 2B-2E, below.

[0052] System 100, or various components thereof, may be incorporated into various different platforms. In some embodiments, system 100 may be included on a vehicle 200, as shown in FIG. 2A. For example, vehicle 200 may be equipped with a processing unit 110 and any of the other components of system 100, as described above relative to FIG. 1. While in some embodiments vehicle 200 may be equipped with only a single image capture device (e.g., camera), in other embodiments, such as those discussed in connection with FIGS. 2B-2E, multiple image capture devices may be used. For example, either of image capture devices 122 and 124 of vehicle 200, as shown in FIG. 2A, may be part of an ADAS (Advanced Driver Assistance Systems) imaging set.

[0053] The image capture devices included on vehicle 200 as part of the image acquisition unit 120 may be positioned at any suitable location. In some embodiments, as shown in FIGS. 2A-2E and 3A-3C, image capture device 122 may be located in the vicinity of the rearview mirror. This position may provide a line of sight similar to that of the driver of vehicle 200, which may aid in determining what is and is not visible to the driver. Image capture device 122 may be positioned at any location near the rearview mirror, but placing image capture device 122 on the driver side of the mirror may further aid in obtaining images representative of the driver's field of view and/or line of sight.

[0054] Other locations for the image capture devices of image acquisition unit 120 may also be used. For example, image capture device 124 may be located on or in a bumper of vehicle 200. Such a location may be especially suitable for image capture devices having a wide field of view. The line of sight of bumper-located image capture devices can be different from that of the driver and, therefore, the bumper image capture device and driver may not always see the same objects. The image capture devices (e.g., image capture devices 122, 124, and 126) may also be located in other locations. For example, the image capture devices may be located on or in one or both of the side mirrors of vehicle 200, on the roof of vehicle 200, on the hood of vehicle 200, on the trunk of vehicle 200, on the sides of vehicle 200, mounted on, positioned behind, or positioned in front of any of the windows of vehicle 200, and mounted in or near light figures on the front and/or back of vehicle 200, etc.

[0055] In addition to image capture devices, vehicle 200 may include various other components of system 100. For example, processing unit 110 may be included on vehicle 200 either integrated with or separate from an engine control unit (ECU) of the vehicle. Vehicle 200 may also be equipped with a position sensor 130, such as a GPS receiver and may also include a map database 160 and memory units 140 and 150.

[0056] FIG. 2A is a diagrammatic side view representation of an exemplary vehicle imaging system consistent with the disclosed embodiments. FIG. 2B is a diagrammatic top view illustration of the embodiment shown in FIG. 2A. As illustrated in FIG. 2B, the disclosed embodiments may include a vehicle 200 including in its body a system 100 with a first image capture device 122 positioned in the vicinity of the rearview mirror and/or near the driver of vehicle 200, a second image capture device 124 positioned on or in a bumper region (e.g., one of bumper regions 210) of vehicle 200, and a processing unit 110.

[0057] As illustrated in FIG. 2C, image capture devices 122 and 124 may both be positioned in the vicinity of the rearview mirror and/or near the driver of vehicle 200. Additionally, while two image capture devices 122 and 124 are shown in FIGS. 2B and 2C, it should be understood that other embodiments may include more than two image capture devices. For example, in the embodiments shown in FIGS. 2D and 2E, first, second, and third image capture devices 122, 124, and 126, are included in the system 100 of vehicle 200.

[0058] As illustrated in FIG. 2D, image capture device 122 may be positioned in the vicinity of the rearview mirror and/or near the driver of vehicle 200, and image capture devices 124 and 126 may be positioned on or in a bumper region (e.g., one of bumper regions 210) of vehicle 200. And as shown in FIG. 2E, image capture devices 122, 124, and 126 may be positioned in the vicinity of the rearview mirror and/or near the driver seat of vehicle 200. The disclosed embodiments are not limited to any particular

IPR2025-00944
Tesla EX1002 Page 1392

number and configuration of the image capture devices, and the image capture devices may be positioned in any appropriate location within and/or on vehicle 200.

[0059] It is to be understood that the disclosed embodiments are not limited to vehicles and could be applied in other contexts. It is also to be understood that disclosed embodiments are not limited to a particular type of vehicle 200 and may be applicable to all types of vehicles including automobiles, trucks, trailers, and other types of vehicles.

[0060] The first image capture device 122 may include any suitable type of image capture device. Image capture device 122 may include an optical axis. In one instance, the image capture device 122 may include an Aptina M9V024 WVGA sensor with a global shutter. In other embodiments, image capture device 122 may provide a resolution of 1280×960 pixels and may include a rolling shutter. Image capture device 122 may include various optical elements. In some embodiments one or more lenses may be included, for example, to provide a desired focal length and field of view for the image capture device. In some embodiments, image capture device 122 may be associated with a 6 mm lens or a 12 mm lens. In some embodiments, image capture device 122 may be configured to capture images having a desired field-of-view (FOV) 202, as illustrated in FIG. 2D. For example, image capture device 122 may be configured to have a regular FOV, such as within a range of 40 degrees to 56 degrees, including a 46 degree FOV, 50 degree FOV, 52 degree FOV, or greater. Alternatively, image capture device 122 may be configured to have a narrow FOV in the range of 23 to 40 degrees, such as a 28 degree FOV or 36 degree FOV. In addition, image capture device 122 may be configured to have a wide FOV in the range of 100 to 180 degrees. In some embodiments, image capture device 122 may include a wide angle bumper camera or one with up to a 180 degree FOV.

[0061] The first image capture device 122 may acquire a plurality of first images relative to a scene associated with the vehicle 200. Each of the plurality of first images may be acquired as a series of image scan lines, which may be captured using a rolling shutter. Each scan line may include a plurality of pixels.

[0062] The first image capture device 122 may have a scan rate associated with acquisition of each of the first series of image scan lines. The scan rate may refer to a rate at which an image sensor can acquire image data associated with each pixel included in a particular scan line.

[0063] Image capture devices 122, 124, and 126 may contain any suitable type and number of image sensors, including CCD sensors or CMOS sensors, for example. In one embodiment, a CMOS image sensor may be employed along with a rolling shutter, such that each pixel in a row is read one at a time, and scanning of the rows proceeds on a row-by-row basis until an entire image frame has been captured. In some embodiments, the rows may be captured sequentially from top to bottom relative to the frame.

[0064] The use of a rolling shutter may result in pixels in different rows being exposed and captured at different times, which may cause skew and other image artifacts in the captured image frame. On the other hand, when the image capture device 122 is configured to operate with a global or synchronous shutter, all of the pixels may be exposed for the same amount of time and during a common exposure period. As a result, the image data in a frame collected from a system employing a global shutter

represents a snapshot of the entire FOV (such as FOV 202) at a particular time. In contrast, in a rolling shutter application, each row in a frame is exposed and data is capture at different times. Thus, moving objects may appear distorted in an image capture device having a rolling shutter. This phenomenon will be described in greater detail below.

[0065] The second image capture device 124 and the third image capturing device 126 may be any type of image capture device. Like the first image capture device 122, each of image capture devices 124 and 126 may include an optical axis. In one embodiment, each of image capture devices 124 and 126 may include an Aptina M9V024 WVGA sensor with a global shutter. Alternatively, each of image capture devices 124 and 126 may include a rolling shutter. Like image capture device 122, image capture devices 124 and 126 may be configured to include various lenses and optical elements. In some embodiments, lenses associated with image capture devices 124 and 126 may provide FOVs (such as FOVs 204 and 206) that are the same as, or narrower than, a FOV (such as FOV 202) associated with image capture device 122. For example, image capture devices 124 and 126 may have FOVs of 40 degrees, 30 degrees, 26 degrees, 23 degrees, 20 degrees, or less.

[0066] Image capture devices 124 and 126 may acquire a plurality of second and third images relative to a scene associated with the vehicle 200. Each of the plurality of second and third images may be acquired as a second and third series of image scan lines, which may be captured using a rolling shutter. Each scan line or row may have a plurality of pixels. Image capture devices 124 and 126 may have second and third scan rates associated with acquisition of each of image scan lines included in the second and third series.

[0067] Each image capture device 122, 124, and 126 may be positioned at any suitable position and orientation relative to vehicle 200. The relative positioning of the image capture devices 122, 124, and 126 may be selected to aid in fusing together the information acquired from the image capture devices. For example, in some embodiments, a FOV (such as FOV 204) associated with image capture device 124 may overlap partially or fully with a FOV (such as FOV 202) associated with image capture device 122 and a FOV (such as FOV 206) associated with image capture device 126.

[0068] Image capture devices 122, 124, and 126 may be located on vehicle 200 at any suitable relative heights. In one instance, there may be a height difference between the image capture devices 122, 124, and 126, which may provide sufficient parallax information to enable stereo analysis. For example, as shown in FIG. 2A, the two image capture devices 122 and 124 are at different heights. There may also be a lateral displacement difference between image capture devices 122, 124, and 126, giving additional parallax information for stereo analysis by processing unit 110, for example. The difference in the lateral displacement may be denoted by dx, as shown in FIGS. 2C and 2D. In some embodiments, fore or aft displacement (e.g., range displacement) may exist between image capture devices 122, 124, and 126. For example, image capture device 122 may be located 0.5 to 2 meters or more behind image capture device 124 and/or image capture device 126. This type of displacement may enable one of the image capture devices to cover potential blind spots of the other image capture device(s).

[0069] Image capture devices 122 may have any suitable resolution capability (e.g.,

Tesla EX1002 Page 1394 IPR2025-00944

number of pixels associated with the image sensor), and the resolution of the image sensor(s) associated with the image capture device 122 may be higher, lower, or the same as the resolution of the image sensor(s) associated with image capture devices 124 and 126. In some embodiments, the image sensor(s) associated with image capture device 122 and/or image capture devices 124 and 126 may have a resolution of 640×480, 1024×768, 1280×960, or any other suitable resolution.

[0070] The frame rate (e.g., the rate at which an image capture device acquires a set of pixel data of one image frame before moving on to capture pixel data associated with the next image frame) may be controllable. The frame rate associated with image capture device 122 may be higher, lower, or the same as the frame rate associated with image capture devices 124 and 126. The frame rate associated with image capture devices 122, 124, and 126 may depend on a variety of factors that may affect the timing of the frame rate. For example, one or more of image capture devices 122, 124, and 126 may include a selectable pixel delay period imposed before or after acquisition of image data associated with one or more pixels of an image sensor in image capture device 122, 124, and/or 126. Generally, image data corresponding to each pixel may be acquired according to a clock rate for the device (e.g., one pixel per clock cycle). Additionally, in embodiments including a rolling shutter, one or more of image capture devices 122, 124, and 126 may include a selectable horizontal blanking period imposed before or after acquisition of image data associated with a row of pixels of an image sensor in image capture device 122, 124, and/or 126. Further, one or more of image capture devices 122, 124, and/or 126 may include a selectable vertical blanking period imposed before or after acquisition of image data associated with an image frame of image capture device 122, 124, and 126.

[0071] These timing controls may enable synchronization of frame rates associated with image capture devices 122, 124, and 126, even where the line scan rates of each are different. Additionally, as will be discussed in greater detail below, these selectable timing controls, among other factors (e.g., image sensor resolution, maximum line scan rates, etc.) may enable synchronization of image capture from an area where the FOV of image capture device 122 overlaps with one or more FOVs of image capture devices 124 and 126, even where the field of view of image capture device 122 is different from the FOVs of image capture devices 124 and 126.

[0072] Frame rate timing in image capture device 122, 124, and 126 may depend on the resolution of the associated image sensors. For example, assuming similar line scan rates for both devices, if one device includes an image sensor having a resolution of 640×480 and another device includes an image sensor with a resolution of 1280×960, then more time will be required to acquire a frame of image data from the sensor having the higher resolution.

[0073] Another factor that may affect the timing of image data acquisition in image capture devices 122, 124, and 126 is the maximum line scan rate. For example, acquisition of a row of image data from an image sensor included in image capture device 122, 124, and 126 will require some minimum amount of time. Assuming no pixel delay periods are added, this minimum amount of time for acquisition of a row of image data will be related to the maximum line scan rate for a particular device. Devices that offer higher maximum line scan rates have the potential to provide higher frame rates than devices with lower maximum line scan rates. In some embodiments, one or more of image capture devices 124 and 126 may have a maximum line scan rate that is higher than a maximum line scan rate associated with image capture device

Tesla EX1002 Page 1395 IPR2025-00944

122. In some embodiments, the maximum line scan rate of image capture device 124 and/or 126 may be 1.25, 1.5, 1.75, or 2 times or more than a maximum line scan rate of image capture device 122.

[0074] In another embodiment, image capture devices 122, 124, and 126 may have the same maximum line scan rate, but image capture device 122 may be operated at a scan rate less than or equal to its maximum scan rate. The system may be configured such that one or more of image capture devices 124 and 126 operate at a line scan rate that is equal to the line scan rate of image capture device 122. In other instances, the system may be configured such that the line scan rate of image capture device 124 and/or image capture device 126 may be 1.25, 1.5, 1.75, or 2 times or more than the line scan rate of image capture device 122.

[0075] In some embodiments, image capture devices 122, 124, and 126 may be asymmetric. That is, they may include cameras having different fields of view (FOV) and focal lengths. The fields of view of image capture devices 122, 124, and 126 may include any desired area relative to an environment of vehicle 200, for example. In some embodiments, one or more of image capture devices 122, 124, and 126 may be configured to acquire image data from an environment in front of vehicle 200, behind vehicle 200, to the sides of vehicle 200, or combinations thereof.

[0076] Further, the focal length associated with each image capture device 122, 124, and/or 126 may be selectable (e.g., by inclusion of appropriate lenses etc.) such that each device acquires images of objects at a desired distance range relative to vehicle 200. For example, in some embodiments image capture devices 122, 124, and 126 may acquire images of close-up objects within a few meters from the vehicle. Image capture devices 122, 124, and 126 may also be configured to acquire images of objects at ranges more distant from the vehicle (e.g., 25 m, 50 m, 100 m, 150 m, or more). Further, the focal lengths of image capture devices 122, 124, and 126 may be selected such that one image capture device (e.g., image capture device 122) can acquire images of objects relatively close to the vehicle (e.g., within 10 m or within 20 m) while the other image capture devices (e.g., image capture devices 124 and 126) can acquire images of more distant objects (e.g., greater than 20 m, 50 m, 100 m, 150 m, etc.) from vehicle 200.

[0077] According to some embodiments, the FOV of one or more image capture devices 122, 124, and 126 may have a wide angle. For example, it may be advantageous to have a FOV of 140 degrees, especially for image capture devices 122, 124, and 126 that may be used to capture images of the area in the vicinity of vehicle 200. For example, image capture device 122 may be used to capture images of the area to the right or left of vehicle 200 and, in such embodiments, it may be desirable for image capture device 122 to have a wide FOV (e.g., at least 140 degrees).

[0078] The field of view associated with each of image capture devices 122, 124, and 126 may depend on the respective focal lengths. For example, as the focal length increases, the corresponding field of view decreases.

[0079] Image capture devices 122, 124, and 126 may be configured to have any suitable fields of view. In one particular example, image capture device 122 may have a horizontal FOV of 46 degrees, image capture device 124 may have a horizontal FOV of 23 degrees, and image capture device 126 may have a horizontal FOV in between 23

Tesla EX1002 Page 1396 IPR2025-00944

and 46 degrees. In another instance, image capture device 122 may have a horizontal FOV of 52 degrees, image capture device 124 may have a horizontal FOV of 26 degrees, and image capture device 126 may have a horizontal FOV in between 26 and 52 degrees. In some embodiments, a ratio of the FOV of image capture device 122 to the FOVs of image capture device 124 and/or image capture device 126 may vary from 1.5 to 2.0. In other embodiments, this ratio may vary between 1.25 and 2.25.

[0080] System 100 may be configured so that a field of view of image capture device 122 overlaps, at least partially or fully, with a field of view of image capture device 124 and/or image capture device 126. In some embodiments, system 100 may be configured such that the fields of view of image capture devices 124 and 126, for example, fall within (e.g., are narrower than) and share a common center with the field of view of image capture device 122. In other embodiments, the image capture devices 122, 124, and 126 may capture adjacent FOVs or may have partial overlap in their FOVs. In some embodiments, the fields of view of image capture devices 122, 124, and 126 may be aligned such that a center of the narrower FOV image capture devices 124 and/or 126 may be located in a lower half of the field of view of the wider FOV device 122.

[0081] FIG. 2F is a diagrammatic representation of exemplary vehicle control systems, consistent with the disclosed embodiments. As indicated in FIG. 2F, vehicle 200 may include throttling system 220, braking system 230, and steering system 240. System 100 may provide inputs (e.g., control signals) to one or more of throttling system 220, braking system 230, and steering system 240 over one or more data links (e.g., any wired and/or wireless link or links for transmitting data). For example, based on analysis of images acquired by image capture devices 122, 124, and/or 126, system 100 may provide control signals to one or more of throttling system 220, braking system 230, and steering system 240 to navigate vehicle 200 (e.g., by causing an acceleration, a turn, a lane shift, etc.). Further, system 100 may receive inputs from one or more of throttling system 220, braking system 230, and steering system 24 indicating operating conditions of vehicle 200 (e.g., speed, whether vehicle 200 is braking and/or turning, etc.). Further details are provided in connection with FIGS. 4-7, below.

[0082] As shown in FIG. 3A, vehicle 200 may also include a user interface 170 for interacting with a driver or a passenger of vehicle 200. For example, user interface 170 in a vehicle application may include a touch screen 320, knobs 330, buttons 340, and a microphone 350. A driver or passenger of vehicle 200 may also use handles (e.g., located on or near the steering column of vehicle 200 including, for example, turn signal handles), buttons (e.g., located on the steering wheel of vehicle 200), and the like, to interact with system 100. In some embodiments, microphone 350 may be positioned adjacent to a rearview mirror 310. Similarly, in some embodiments, image capture device 122 may be located near rearview mirror 310. In some embodiments, user interface 170 may also include one or more speakers 360 (e.g., speakers of a vehicle audio system). For example, system 100 may provide various notifications (e.g., alerts) via speakers 360.

[0083] FIGS. 3B-3D are illustrations of an exemplary camera mount 370 configured to be positioned behind a rearview mirror (e.g., rearview mirror 310) and against a vehicle windshield, consistent with disclosed embodiments. As shown in FIG. 3B, camera mount 370 may include image capture devices 122, 124, and 126. Image capture devices 124 and 126 may be positioned behind a glare shield 380, which may be flush against the vehicle windshield and include a composition of film and/or anti-reflective

IPR2025-00944
Tesla EX1002 Page 1397

materials. For example, glare shield 380 may be positioned such that it aligns against a vehicle windshield having a matching slope. In some embodiments, each of image capture devices 122, 124, and 126 may be positioned behind glare shield 380, as depicted, for example, in FIG. 3D. The disclosed embodiments are not limited to any particular configuration of image capture devices 122, 124, and 126, camera mount 370, and glare shield 380. FIG. 3C is an illustration of camera mount 370 shown in FIG. 3B from a front perspective.

[0084] As will be appreciated by a person skilled in the art having the benefit of this disclosure, numerous variations and/or modifications may be made to the foregoing disclosed embodiments. For example, not all components are essential for the operation of system 100. Further, any component may be located in any appropriate part of system 100 and the components may be rearranged into a variety of configurations while providing the functionality of the disclosed embodiments. Therefore, the foregoing configurations are examples and, regardless of the configurations discussed above, system 100 can provide a wide range of functionality to analyze the surroundings of vehicle 200 and navigate vehicle 200 in response to the analysis.

[0085] As discussed below in further detail and consistent with various disclosed embodiments, system 100 may provide a variety of features related to autonomous driving and/or driver assist technology. For example, system 100 may analyze image data, position data (e.g., GPS location information), map data, speed data, and/or data from sensors included in vehicle 200. System 100 may collect the data for analysis from, for example, image acquisition unit 120, position sensor 130, and other sensors. Further, system 100 may analyze the collected data to determine whether or not vehicle 200 should take a certain action, and then automatically take the determined action without human intervention. For example, when vehicle 200 navigates without human intervention, system 100 may automatically control the braking, acceleration, and/or steering of vehicle 200 (e.g., by sending control signals to one or more of throttling system 220, braking system 230, and steering system 240). Further, system 100 may analyze the collected data and issue warnings and/or alerts to vehicle occupants based on the analysis of the collected data. Additional details regarding the various embodiments that are provided by system 100 are provided below.

[0086] Forward-Facing Multi-Imaging System

[0087] As discussed above, system 100 may provide drive assist functionality that uses a multi-camera system. The multi-camera system may use one or more cameras facing in the forward direction of a vehicle. In other embodiments, the multi-camera system may include one or more cameras facing to the side of a vehicle or to the rear of the vehicle. In one embodiment, for example, system 100 may use a two-camera imaging system, where a first camera and a second camera (e.g., image capture devices 122 and 124) may be positioned at the front and/or the sides of a vehicle (e.g., vehicle 200). The first camera may have a field of view that is greater than, less than, or partially overlapping with, the field of view of the second camera. In addition, the first camera may be connected to a first image processor to perform monocular image analysis of images provided by the first camera, and the second camera may be connected to a second image processor to perform monocular image analysis of images provided by the second camera. The outputs (e.g., processed information) of the first and second image processors may be combined. In some embodiments, the second image processor may receive images from both the first camera and second camera to

IPR2025-00944
Tesla EX1002 Page 1399

perform stereo analysis. In another embodiment, system 100 may use a three-camera imaging system where each of the cameras has a different field of view. Such a system may, therefore, make decisions based on information derived from objects located at varying distances both forward and to the sides of the vehicle. References to monocular image analysis may refer to instances where image analysis is performed based on images captured from a single point of view (e.g., from a single camera). Stereo image analysis may refer to instances where image analysis is performed based on two or more images captured with one or more variations of an image capture parameter. For example, captured images suitable for performing stereo image analysis may include images captured: from two or more different positions, from different fields of view, using different focal lengths, along with parallax information, etc.

[0088] For example, in one embodiment, system 100 may implement a three camera configuration using image capture devices 122-126. In such a configuration, image capture device 122 may provide a narrow field of view (e.g., 34 degrees, or other values selected from a range of about 20 to 45 degrees, etc.), image capture device 124 may provide a wide field of view (e.g., 150 degrees or other values selected from a range of about 100 to about 180 degrees), and image capture device 126 may provide an intermediate field of view (e.g., 46 degrees or other values selected from a range of about 35 to about 60 degrees). In some embodiments, image capture device 126 may act as a main or primary camera. Image capture devices 122-126 may be positioned behind rearview mirror 310 and positioned substantially side-by-side (e.g., 6 cm apart). Further, in some embodiments, as discussed above, one or more of image capture devices 122-126 may be mounted behind glare shield 380 that is flush with the windshield of vehicle 200. Such shielding may act to minimize the impact of any reflections from inside the car on image capture devices 122-126.

[0089] In another embodiment, as discussed above in connection with FIGS. 3B and 3C, the wide field of view camera (e.g., image capture device 124 in the above example) may be mounted lower than the narrow and main field of view cameras (e.g., image devices 122 and 126 in the above example). This configuration may provide a free line of sight from the wide field of view camera. To reduce reflections, the cameras may be mounted close to the windshield of vehicle 200, and may include polarizers on the cameras to damp reflected light.

[0090] A three camera system may provide certain performance characteristics. For example, some embodiments may include an ability to validate the detection of objects by one camera based on detection results from another camera. In the three camera configuration discussed above, processing unit 110 may include, for example, three processing devices (e.g., three EyeQ series of processor chips, as discussed above), with each processing device dedicated to processing images captured by one or more of image capture devices 122-126.

[0091] In a three camera system, a first processing device may receive images from both the main camera and the narrow field of view camera, and perform vision processing of the narrow FOV camera to, for example, detect other vehicles, pedestrians, lane marks, traffic signs, traffic lights, and other road objects. Further, the first processing device may calculate a disparity of pixels between the images from the main camera and the narrow camera and create a 3D reconstruction of the environment of vehicle 200. The first processing device may then combine the 3D reconstruction with 3D map data or with 3D information calculated based on information from another camera.

[0092] The second processing device may receive images from main camera and perform vision processing to detect other vehicles, pedestrians, lane marks, traffic signs, traffic lights, and other road objects. Additionally, the second processing device may calculate a camera displacement and, based on the displacement, calculate a disparity of pixels between successive images and create a 3D reconstruction of the scene (e.g., a structure from motion). The second processing device may send the structure from motion based 3D reconstruction to the first processing device to be combined with the stereo 3D images.

[0093] The third processing device may receive images from the wide FOV camera and process the images to detect vehicles, pedestrians, lane marks, traffic signs, traffic lights, and other road objects. The third processing device may further execute additional processing instructions to analyze images to identify objects moving in the image, such as vehicles changing lanes, pedestrians, etc.

[0094] In some embodiments, having streams of image-based information captured and processed independently may provide an opportunity for providing redundancy in the system. Such redundancy may include, for example, using a first image capture device and the images processed from that device to validate and/or supplement information obtained by capturing and processing image information from at least a second image capture device.

[0095] In some embodiments, system 100 may use two image capture devices (e.g., image capture devices 122 and 124) in providing navigation assistance for vehicle 200 and use a third image capture device (e.g., image capture device 126) to provide redundancy and validate the analysis of data received from the other two image capture devices. For example, in such a configuration, image capture devices 122 and 124 may provide images for stereo analysis by system 100 for navigating vehicle 200, while image capture device 126 may provide images for monocular analysis by system 100 to provide redundancy and validation of information obtained based on images captured from image capture device 122 and/or image capture device 124. That is, image capture device 126 (and a corresponding processing device) may be considered to provide a redundant sub-system for providing a check on the analysis derived from image capture devices 122 and 124 (e.g., to provide an automatic emergency braking (AEB) system).

[0096] One of skill in the art will recognize that the above camera configurations, camera placements, number of cameras, camera locations, etc., are examples only. These components and others described relative to the overall system may be assembled and used in a variety of different configurations without departing from the scope of the disclosed embodiments. Further details regarding usage of a multi-camera system to provide driver assist and/or autonomous vehicle functionality follow below.

[0097] FIG. 4 is an exemplary functional block diagram of memory 140 and/or 150, which may be stored/programmed with instructions for performing one or more operations consistent with the disclosed embodiments. Although the following refers to memory 140, one of skill in the art will recognize that instructions may be stored in memory 140 and/or 150.

[0098] As shown in FIG. 4, memory 140 may store a monocular image analysis module 402, a stereo image analysis module 404, a velocity and acceleration module 406, and

a navigational response module 408. The disclosed embodiments are not limited to any particular configuration of memory 140. Further, application processor 180 and/or image processor 190 may execute the instructions stored in any of modules 402-408 included in memory 140. One of skill in the art will understand that references in the following discussions to processing unit 110 may refer to application processor 180 and image processor 190 individually or collectively. Accordingly, steps of any of the following processes may be performed by one or more processing devices.

[0099] In one embodiment, monocular image analysis module 402 may store instructions (such as computer vision software) which, when executed by processing unit 110, performs monocular image analysis of a set of images acquired by one of image capture devices 122, 124, and 126. In some embodiments, processing unit 110 may combine information from a set of images with additional sensory information (e.g., information from radar) to perform the monocular image analysis. As described in connection with FIGS. 5A-5D below, monocular image analysis module 402 may include instructions for detecting a set of features within the set of images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, hazardous objects, and any other feature associated with an environment of a vehicle. Based on the analysis, system 100 (e.g., via processing unit 110) may cause one or more navigational responses in vehicle 200, such as a turn, a lane shift, a change in acceleration, and the like, as discussed below in connection with navigational response module 408.

[0100] In one embodiment, stereo image analysis module 404 may store instructions (such as computer vision software) which, when executed by processing unit 110, performs stereo image analysis of first and second sets of images acquired by a combination of image capture devices selected from any of image capture devices 122, 124, and 126. In some embodiments, processing unit 110 may combine information from the first and second sets of images with additional sensory information (e.g., information from radar) to perform the stereo image analysis. For example, stereo image analysis module 404 may include instructions for performing stereo image analysis based on a first set of images acquired by image capture device 124 and a second set of images acquired by image capture device 126. As described in connection with FIG. 6 below, stereo image analysis module 404 may include instructions for detecting a set of features within the first and second sets of images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, hazardous objects, and the like. Based on the analysis, processing unit 110 may cause one or more navigational responses in vehicle 200, such as a turn, a lane shift, a change in acceleration, and the like, as discussed below in connection with navigational response module 408.

[0101] In one embodiment, velocity and acceleration module 406 may store software configured to analyze data received from one or more computing and electromechanical devices in vehicle 200 that are configured to cause a change in velocity and/or acceleration of vehicle 200. For example, processing unit 110 may execute instructions associated with velocity and acceleration module 406 to calculate a target speed for vehicle 200 based on data derived from execution of monocular image analysis module 402 and/or stereo image analysis module 404. Such data may include, for example, a target position, velocity, and/or acceleration, the position and/or speed of vehicle 200 relative to a nearby vehicle, pedestrian, or road object, position information for vehicle 200 relative to lane markings of the road, and the like. In addition, processing unit 110 may calculate a target speed for vehicle 200 based on

IPR2025-00944
Tesla EX1002 Page 1401

sensory input (e.g., information from radar) and input from other systems of vehicle 200, such as throttling system 220, braking system 230, and/or steering system 240 of vehicle 200. Based on the calculated target speed, processing unit 110 may transmit electronic signals to throttling system 220, braking system 230, and/or steering system 240 of vehicle 200 to trigger a change in velocity and/or acceleration by, for example, physically depressing the brake or easing up off the accelerator of vehicle 200.

[0102] In one embodiment, navigational response module 408 may store software executable by processing unit 110 to determine a desired navigational response based on data derived from execution of monocular image analysis module 402 and/or stereo image analysis module 404. Such data may include position and speed information associated with nearby vehicles, pedestrians, and road objects, target position information for vehicle 200, and the like. Additionally, in some embodiments, the navigational response may be based (partially or fully) on map data, a predetermined position of vehicle 200, and/or a relative velocity or a relative acceleration between vehicle 200 and one or more objects detected from execution of monocular image analysis module 402 and/or stereo image analysis module 404. Navigational response module 408 may also determine a desired navigational response based on sensory input (e.g., information from radar) and inputs from other systems of vehicle 200, such as throttling system 220, braking system 230, and steering system 240 of vehicle 200. Based on the desired navigational response, processing unit 110 may transmit electronic signals to throttling system 220, braking system 230, and steering system 240 of vehicle 200 to trigger a desired navigational response by, for example, turning the steering wheel of vehicle 200 to achieve a rotation of a predetermined angle. In some embodiments, processing unit 110 may use the output of navigational response module 408 (e.g., the desired navigational response) as an input to execution of velocity and acceleration module 406 for calculating a change in speed of vehicle 200.

[0103] FIG. 5A is a flowchart showing an exemplary process 500A for causing one or more navigational responses based on monocular image analysis, consistent with disclosed embodiments. At step 510, processing unit 110 may receive a plurality of images via data interface 128 between processing unit 110 and image acquisition unit 120. For instance, a camera included in image acquisition unit 120 (such as image capture device 122 having field of view 202) may capture a plurality of images of an area forward of vehicle 200 (or to the sides or rear of a vehicle, for example) and transmit them over a data connection (e.g., digital, wired, USB, wireless, Bluetooth, etc.) to processing unit 110. Processing unit 110 may execute monocular image analysis module 402 to analyze the plurality of images at step 520, as described in further detail in connection with FIGS. 5B-5D below. By performing the analysis, processing unit 110 may detect a set of features within the set of images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, and the like.

[0104] Processing unit 110 may also execute monocular image analysis module 402 to detect various road hazards at step 520, such as, for example, parts of a truck tire, fallen road signs, loose cargo, small animals, and the like. Road hazards may vary in structure, shape, size, and color, which may make detection of such hazards more challenging. In some embodiments, processing unit 110 may execute monocular image analysis module 402 to perform multi-frame analysis on the plurality of images to detect road hazards. For example, processing unit 110 may estimate camera motion between consecutive image frames and calculate the disparities in pixels between the frames to construct a 3D-map of the road. Processing unit 110 may then use the 3D-map to

IPR2025-00944
Tesla EX1002 Page 1402

detect the road surface, as well as hazards existing above the road surface.

[0105] At step 530, processing unit 110 may execute navigational response module 408 to cause one or more navigational responses in vehicle 200 based on the analysis performed at step 520 and the techniques as described above in connection with FIG. 4. Navigational responses may include, for example, a turn, a lane shift, a change in acceleration, and the like. In some embodiments, processing unit 110 may use data derived from execution of velocity and acceleration module 406 to cause the one or more navigational responses. Additionally, multiple navigational responses may occur simultaneously, in sequence, or any combination thereof. For instance, processing unit 110 may cause vehicle 200 to shift one lane over and then accelerate by, for example, sequentially transmitting control signals to steering system 240 and throttling system 220 of vehicle 200. Alternatively, processing unit 110 may cause vehicle 200 to brake while at the same time shifting lanes by, for example, simultaneously transmitting control signals to braking system 230 and steering system 240 of vehicle 200.

[0106] FIG. 5B is a flowchart showing an exemplary process 500B for detecting one or more vehicles and/or pedestrians in a set of images, consistent with disclosed embodiments. Processing unit 110 may execute monocular image analysis module 402 to implement process 500B. At step 540, processing unit 110 may determine a set of candidate objects representing possible vehicles and/or pedestrians. For example, processing unit 110 may scan one or more images, compare the images to one or more predetermined patterns, and identify within each image possible locations that may contain objects of interest (e.g., vehicles, pedestrians, or portions thereof). The predetermined patterns may be designed in such a way to achieve a high rate of "false hits" and a low rate of "misses." For example, processing unit 110 may use a low threshold of similarity to predetermined patterns for identifying candidate objects as possible vehicles or pedestrians. Doing so may allow processing unit 110 to reduce the probability of missing (e.g., not identifying) a candidate object representing a vehicle or pedestrian.

[0107] At step 542, processing unit 110 may filter the set of candidate objects to exclude certain candidates (e.g., irrelevant or less relevant objects) based on classification criteria. Such criteria may be derived from various properties associated with object types stored in a database (e.g., a database stored in memory 140). Properties may include object shape, dimensions, texture, position (e.g., relative to vehicle 200), and the like. Thus, processing unit 110 may use one or more sets of criteria to reject false candidates from the set of candidate objects.

[0108] At step 544, processing unit 110 may analyze multiple frames of images to determine whether objects in the set of candidate objects represent vehicles and/or pedestrians. For example, processing unit 110 may track a detected candidate object across consecutive frames and accumulate frame-by-frame data associated with the detected object (e.g., size, position relative to vehicle 200, etc.). Additionally, processing unit 110 may estimate parameters for the detected object and compare the object's frame-by-frame position data to a predicted position.

[0109] At step 546, processing unit 110 may construct a set of measurements for the detected objects. Such measurements may include, for example, position, velocity, and acceleration values (relative to vehicle 200) associated with the detected objects. In some embodiments, processing unit 110 may construct the measurements based on estimation techniques using a series of time-based observations such as Kalman filters

or linear quadratic estimation (LQE), and/or based on available modeling data for different object types (e.g., cars, trucks, pedestrians, bicycles, road signs, etc.). The Kalman filters may be based on a measurement of an object's scale, where the scale measurement is proportional to a time to collision (e.g., the amount of time for vehicle 200 to reach the object). Thus, by performing steps 540-546, processing unit 110 may identify vehicles and pedestrians appearing within the set of captured images and derive information (e.g., position, speed, size) associated with the vehicles and pedestrians. Based on the identification and the derived information, processing unit 110 may cause one or more navigational responses in vehicle 200, as described in connection with FIG. 5A, above.

[0110] At step 548, processing unit 110 may perform an optical flow analysis of one or more images to reduce the probabilities of detecting a "false hit" and missing a candidate object that represents a vehicle or pedestrian. The optical flow analysis may refer to, for example, analyzing motion patterns relative to vehicle 200 in the one or more images associated with other vehicles and pedestrians, and that are distinct from road surface motion. Processing unit 110 may calculate the motion of candidate objects by observing the different positions of the objects across multiple image frames, which are captured at different times. Processing unit 110 may use the position and time values as inputs into mathematical models for calculating the motion of the candidate objects. Thus, optical flow analysis may provide another method of detecting vehicles and pedestrians that are nearby vehicle 200. Processing unit 110 may perform optical flow analysis in combination with steps 540-546 to provide redundancy for detecting vehicles and pedestrians and increase the reliability of system 100.

[0111] FIG. 5C is a flowchart showing an exemplary process 500C for detecting road marks and/or lane geometry information in a set of images, consistent with disclosed embodiments. Processing unit 110 may execute monocular image analysis module 402 to implement process 500C. At step 550, processing unit 110 may detect a set of objects by scanning one or more images. To detect segments of lane markings, lane geometry information, and other pertinent road marks, processing unit 110 may filter the set of objects to exclude those determined to be irrelevant (e.g., minor potholes, small rocks, etc). At step 552, processing unit 110 may group together the segments detected in step 550 belonging to the same road mark or lane mark. Based on the grouping, processing unit 110 may develop a model to represent the detected segments, such as a mathematical model.

[0112] At step 554, processing unit 110 may construct a set of measurements associated with the detected segments. In some embodiments, processing unit 110 may create a projection of the detected segments from the image plane onto the real-world plane. The projection may be characterized using a 3rd-degree polynomial having coefficients corresponding to physical properties such as the position, slope, curvature, and curvature derivative of the detected road. In generating the projection, processing unit 110 may take into account changes in the road surface, as well as pitch and roll rates associated with vehicle 200. In addition, processing unit 110 may model the road elevation by analyzing position and motion cues present on the road surface. Further, processing unit 110 may estimate the pitch and roll rates associated with vehicle 200 by tracking a set of feature points in the one or more images.

[0113] At step 556, processing unit 110 may perform multi-frame analysis by, for example, tracking the detected segments across consecutive image frames and accumulating frame-by-frame data associated with detected segments. As processing

IPR2025-00944
Tesla EX1002 Page 1404

unit 110 performs multi-frame analysis, the set of measurements constructed at step 554 may become more reliable and associated with an increasingly higher confidence level. Thus, by performing steps 550-556, processing unit 110 may identify road marks appearing within the set of captured images and derive lane geometry information. Based on the identification and the derived information, processing unit 110 may cause one or more navigational responses in vehicle 200, as described in connection with FIG. 5A, above.

[0114] At step 558, processing unit 110 may consider additional sources of information to further develop a safety model for vehicle 200 in the context of its surroundings. Processing unit 110 may use the safety model to define a context in which system 100 may execute autonomous control of vehicle 200 in a safe manner. To develop the safety model, in some embodiments, processing unit 110 may consider the position and motion of other vehicles, the detected road edges and barriers, and/or general road shape descriptions extracted from map data (such as data from map database 160). By considering additional sources of information, processing unit 110 may provide redundancy for detecting road marks and lane geometry and increase the reliability of system 100.

[0115] FIG. 5D is a flowchart showing an exemplary process 500D for detecting traffic lights in a set of images, consistent with disclosed embodiments. Processing unit 110 may execute monocular image analysis module 402 to implement process 500D. At step 560, processing unit 110 may scan the set of images and identify objects appearing at locations in the images likely to contain traffic lights. For example, processing unit 110 may filter the identified objects to construct a set of candidate objects, excluding those objects unlikely to correspond to traffic lights. The filtering may be done based on various properties associated with traffic lights, such as shape, dimensions, texture, position (e.g., relative to vehicle 200), and the like. Such properties may be based on multiple examples of traffic lights and traffic control signals and stored in a database. In some embodiments, processing unit 110 may perform multi-frame analysis on the set of candidate objects reflecting possible traffic lights. For example, processing unit 110 may track the candidate objects across consecutive image frames, estimate the real-world position of the candidate objects, and filter out those objects that are moving (which are unlikely to be traffic lights). In some embodiments, processing unit 110 may perform color analysis on the candidate objects and identify the relative position of the detected colors appearing inside possible traffic lights.

[0116] At step 562, processing unit 110 may analyze the geometry of a junction. The analysis may be based on any combination of: (i) the number of lanes detected on either side of vehicle 200, (ii) markings (such as arrow marks) detected on the road, and (iii) descriptions of the junction extracted from map data (such as data from map database 160). Processing unit 110 may conduct the analysis using information derived from execution of monocular analysis module 402. In addition, Processing unit 110 may determine a correspondence between the traffic lights detected at step 560 and the lanes appearing near vehicle 200.

[0117] As vehicle 200 approaches the junction, at step 564, processing unit 110 may update the confidence level associated with the analyzed junction geometry and the detected traffic lights. For instance, the number of traffic lights estimated to appear at the junction as compared with the number actually appearing at the junction may impact the confidence level. Thus, based on the confidence level, processing unit 110

IPR2025-00944
Tesla EX1002 Page 1405

may delegate control to the driver of vehicle 200 in order to improve safety conditions. By performing steps 560-564, processing unit 110 may identify traffic lights appearing within the set of captured images and analyze junction geometry information. Based on the identification and the analysis, processing unit 110 may cause one or more navigational responses in vehicle 200, as described in connection with FIG. 5A, above.

[0118] FIG. 5E is a flowchart showing an exemplary process 500E for causing one or more navigational responses in vehicle 200 based on a vehicle path, consistent with the disclosed embodiments. At step 570, processing unit 110 may construct an initial vehicle path associated with vehicle 200. The vehicle path may be represented using a set of points expressed in coordinates (x, z), and the distance di between two points in the set of points may fall in the range of 1 to 5 meters. In one embodiment, processing unit 110 may construct the initial vehicle path using two polynomials, such as left and right road polynomials. Processing unit 110 may calculate the geometric midpoint between the two polynomials and offset each point included in the resultant vehicle path by a predetermined offset (e.g., a smart lane offset), if any (an offset of zero may correspond to travel in the middle of a lane). The offset may be in a direction perpendicular to a segment between any two points in the vehicle path. In another embodiment, processing unit 110 may use one polynomial and an estimated lane width to offset each point of the vehicle path by half the estimated lane width plus a predetermined offset (e.g., a smart lane offset).

[0119] At step 572, processing unit 110 may update the vehicle path constructed at step 570. Processing unit 110 may reconstruct the vehicle path constructed at step 570 using a higher resolution, such that the distance dk between two points in the set of points representing the vehicle path is less than the distance di described above. For example, the distance dk may fall in the range of 0.1 to 0.3 meters. Processing unit 110 may reconstruct the vehicle path using a parabolic spline algorithm, which may yield a cumulative distance vector S corresponding to the total length of the vehicle path (i.e., based on the set of points representing the vehicle path).

[0120] At step 574, processing unit 110 may determine a look-ahead point (expressed in coordinates as (xl, zl)) based on the updated vehicle path constructed at step 572. Processing unit 110 may extract the look-ahead point from the cumulative distance vector S, and the look-ahead point may be associated with a look-ahead distance and look-ahead time. The look-ahead distance, which may have a lower bound ranging from 10 to 20 meters, may be calculated as the product of the speed of vehicle 200 and the look-ahead time. For example, as the speed of vehicle 200 decreases, the look-ahead distance may also decrease (e.g., until it reaches the lower bound). The look-ahead time, which may range from 0.5 to 1.5 seconds, may be inversely proportional to the gain of one or more control loops associated with causing a navigational response in vehicle 200, such as the heading error tracking control loop. For example, the gain of the heading error tracking control loop may depend on the bandwidth of a yaw rate loop, a steering actuator loop, car lateral dynamics, and the like. Thus, the higher the gain of the heading error tracking control loop, the lower the look-ahead time.

[0121] At step 576, processing unit 110 may determine a heading error and yaw rate command based on the look-ahead point determined at step 574. Processing unit 110 may determine the heading error by calculating the arctangent of the look-ahead point, e.g., arctan (xl/zl). Processing unit 110 may determine the yaw rate command as the product of the heading error and a high-level control gain. The high-level control gain may be equal to: (2/look-ahead time), if the look-ahead distance is not at the lower

Tesla EX1002 Page 1406 IPR2025-00944

bound. Otherwise, the high-level control gain may be equal to: (2*speed of vehicle 200/look-ahead distance).

[0122] FIG. 5F is a flowchart showing an exemplary process 500F for determining whether a leading vehicle is changing lanes, consistent with the disclosed embodiments. At step 580, processing unit 110 may determine navigation information associated with a leading vehicle (e.g., a vehicle traveling ahead of vehicle 200). For example, processing unit 110 may determine the position, velocity (e.g., direction and speed), and/or acceleration of the leading vehicle, using the techniques described in connection with FIGS. 5A and 5B, above. Processing unit 110 may also determine one or more road polynomials, a look-ahead point (associated with vehicle 200), and/or a snail trail (e.g., a set of points describing a path taken by the leading vehicle), using the techniques described in connection with FIG. 5E, above.

[0123] At step 582, processing unit 110 may analyze the navigation information determined at step 580. In one embodiment, processing unit 110 may calculate the distance between a snail trail and a road polynomial (e.g., along the trail). If the variance of this distance along the trail exceeds a predetermined threshold (for example, 0.1 to 0.2 meters on a straight road, 0.3 to 0.4 meters on a moderately curvy road, and 0.5 to 0.6 meters on a road with sharp curves), processing unit 110 may determine that the leading vehicle is likely changing lanes. In the case where multiple vehicles are detected traveling ahead of vehicle 200, processing unit 110 may compare the snail trails associated with each vehicle. Based on the comparison, processing unit 110 may determine that a vehicle whose snail trail does not match with the snail trails of the other vehicles is likely changing lanes. Processing unit 110 may additionally compare the curvature of the snail trail (associated with the leading vehicle) with the expected curvature of the road segment in which the leading vehicle is traveling. The expected curvature may be extracted from map data (e.g., data from map database 160), from road polynomials, from other vehicles' snail trails, from prior knowledge about the road, and the like. If the difference in curvature of the snail trail and the expected curvature of the road segment exceeds a predetermined threshold, processing unit 110 may determine that the leading vehicle is likely changing lanes.

[0124] In another embodiment, processing unit 110 may compare the leading vehicle's instantaneous position with the look-ahead point (associated with vehicle 200) over a specific period of time (e.g., 0.5 to 1.5 seconds). If the distance between the leading vehicle's instantaneous position and the look-ahead point varies during the specific period of time, and the cumulative sum of variation exceeds a predetermined threshold (for example, 0.3 to 0.4 meters on a straight road, 0.7 to 0.8 meters on a moderately curvy road, and 1.3 to 1.7 meters on a road with sharp curves), processing unit 110 may determine that the leading vehicle is likely changing lanes. In another embodiment, processing unit 110 may analyze the geometry of the snail trail by comparing the lateral distance traveled along the trail with the expected curvature of the snail trail. The expected radius of curvature may be determined according to the calculation: (δz<2>+δx<2>)/2/(δx), where δx represents the lateral distance traveled and δz represents the longitudinal distance traveled. If the difference between the lateral distance traveled and the expected curvature exceeds a predetermined threshold (e.g., 500 to 700 meters), processing unit 110 may determine that the leading vehicle is likely changing lanes. In another embodiment, processing unit 110 may analyze the position of the leading vehicle. If the position of the leading vehicle obscures a road polynomial (e.g., the leading vehicle is overlaid on top of the road polynomial), then processing unit 110 may determine that the leading vehicle is likely

Tesla EX1002 Page 1407 IPR2025-00944

changing lanes. In the case where the position of the leading vehicle is such that, another vehicle is detected ahead of the leading vehicle and the snail trails of the two vehicles are not parallel, processing unit 110 may determine that the (closer) leading vehicle is likely changing lanes.

[0125] At step 584, processing unit 110 may determine whether or not leading vehicle 200 is changing lanes based on the analysis performed at step 582. For example, processing unit 110 may make the determination based on a weighted average of the individual analyses performed at step 582. Under such a scheme, for example, a decision by processing unit 110 that the leading vehicle is likely changing lanes based on a particular type of analysis may be assigned a value of "1" (and "0" to represent a determination that the leading vehicle is not likely changing lanes). Different analyses performed at step 582 may be assigned different weights, and the disclosed embodiments are not limited to any particular combination of analyses and weights.

[0126] FIG. 6 is a flowchart showing an exemplary process 600 for causing one or more navigational responses based on stereo image analysis, consistent with disclosed embodiments. At step 610, processing unit 110 may receive a first and second plurality of images via data interface 128. For example, cameras included in image acquisition unit 120 (such as image capture devices 122 and 124 having fields of view 202 and 204) may capture a first and second plurality of images of an area forward of vehicle 200 and transmit them over a digital connection (e.g., USB, wireless, Bluetooth, etc.) to processing unit 110. In some embodiments, processing unit 110 may receive the first and second plurality of images via two or more data interfaces. The disclosed embodiments are not limited to any particular data interface configurations or protocols.

[0127] At step 620, processing unit 110 may execute stereo image analysis module 404 to perform stereo image analysis of the first and second plurality of images to create a 3D map of the road in front of the vehicle and detect features within the images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, road hazards, and the like. Stereo image analysis may be performed in a manner similar to the steps described in connection with FIGS. 5A-5D, above. For example, processing unit 110 may execute stereo image analysis module 404 to detect candidate objects (e.g., vehicles, pedestrians, road marks, traffic lights, road hazards, etc.) within the first and second plurality of images, filter out a subset of the candidate objects based on various criteria, and perform multi-frame analysis, construct measurements, and determine a confidence level for the remaining candidate objects. In performing the steps above, processing unit 110 may consider information from both the first and second plurality of images, rather than information from one set of images alone. For example, processing unit 110 may analyze the differences in pixel-level data (or other data subsets from among the two streams of captured images) for a candidate object appearing in both the first and second plurality of images. As another example, processing unit 110 may estimate a position and/or velocity of a candidate object (e.g., relative to vehicle 200) by observing that the object appears in one of the plurality of images but not the other or relative to other differences that may exist relative to objects appearing if the two image streams. For example, position, velocity, and/or acceleration relative to vehicle 200 may be determined based on trajectories, positions, movement characteristics, etc. of features associated with an object appearing in one or both of the image streams.

[0128] At step 630, processing unit 110 may execute navigational response module 408 to cause one or more navigational responses in vehicle 200 based on the analysis

IPR2025-00944
Tesla EX1002 Page 1408

performed at step 620 and the techniques as described above in connection with FIG. 4. Navigational responses may include, for example, a turn, a lane shift, a change in acceleration, a change in velocity, braking, and the like. In some embodiments, processing unit 110 may use data derived from execution of velocity and acceleration module 406 to cause the one or more navigational responses. Additionally, multiple navigational responses may occur simultaneously, in sequence, or any combination thereof.

[0129] FIG. 7 is a flowchart showing an exemplary process 700 for causing one or more navigational responses based on an analysis of three sets of images, consistent with disclosed embodiments. At step 710, processing unit 110 may receive a first, second, and third plurality of images via data interface 128. For instance, cameras included in image acquisition unit 120 (such as image capture devices 122, 124, and 126 having fields of view 202, 204, and 206) may capture a first, second, and third plurality of images of an area forward and/or to the side of vehicle 200 and transmit them over a digital connection (e.g., USB, wireless, Bluetooth, etc.) to processing unit 110. In some embodiments, processing unit 110 may receive the first, second, and third plurality of images via three or more data interfaces. For example, each of image capture devices 122, 124, 126 may have an associated data interface for communicating data to processing unit 110. The disclosed embodiments are not limited to any particular data interface configurations or protocols.

[0130] At step 720, processing unit 110 may analyze the first, second, and third plurality of images to detect features within the images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, road hazards, and the like. The analysis may be performed in a manner similar to the steps described in connection with FIGS. 5A-5D and 6, above. For instance, processing unit 110 may perform monocular image analysis (e.g., via execution of monocular image analysis module 402 and based on the steps described in connection with FIGS. 5A-5D, above) on each of the first, second, and third plurality of images. Alternatively, processing unit 110 may perform stereo image analysis (e.g., via execution of stereo image analysis module 404 and based on the steps described in connection with FIG. 6, above) on the first and second plurality of images, the second and third plurality of images, and/or the first and third plurality of images. The processed information corresponding to the analysis of the first, second, and/or third plurality of images may be combined. In some embodiments, processing unit 110 may perform a combination of monocular and stereo image analyses. For example, processing unit 110 may perform monocular image analysis (e.g., via execution of monocular image analysis module 402) on the first plurality of images and stereo image analysis (e.g., via execution of stereo image analysis module 404) on the second and third plurality of images. The configuration of image capture devices 122, 124, and 126—including their respective locations and fields of view 202, 204, and 206—may influence the types of analyses conducted on the first, second, and third plurality of images. The disclosed embodiments are not limited to a particular configuration of image capture devices 122, 124, and 126, or the types of analyses conducted on the first, second, and third plurality of images.

[0131] In some embodiments, processing unit 110 may perform testing on system 100 based on the images acquired and analyzed at steps 710 and 720. Such testing may provide an indicator of the overall performance of system 100 for certain configurations of image capture devices 122, 124, and 126. For example, processing unit 110 may determine the proportion of "false hits" (e.g., cases where system 100 incorrectly determined the presence of a vehicle or pedestrian) and "misses."

Tesla EX1002 Page 1409 IPR2025-00944

[0132] At step 730, processing unit 110 may cause one or more navigational responses in vehicle 200 based on information derived from two of the first, second, and third plurality of images. Selection of two of the first, second, and third plurality of images may depend on various factors, such as, for example, the number, types, and sizes of objects detected in each of the plurality of images. Processing unit 110 may also make the selection based on image quality and resolution, the effective field of view reflected in the images, the number of captured frames, the extent to which one or more objects of interest actually appear in the frames (e.g., the percentage of frames in which an object appears, the proportion of the object that appears in each such frame, etc.), and the like.

[0133] In some embodiments, processing unit 110 may select information derived from two of the first, second, and third plurality of images by determining the extent to which information derived from one image source is consistent with information derived from other image sources. For example, processing unit 110 may combine the processed information derived from each of image capture devices 122, 124, and 126 (whether by monocular analysis, stereo analysis, or any combination of the two) and determine visual indicators (e.g., lane markings, a detected vehicle and its location and/or path, a detected traffic light, etc.) that are consistent across the images captured from each of image capture devices 122, 124, and 126. Processing unit 110 may also exclude information that is inconsistent across the captured images (e.g., a vehicle changing lanes, a lane model indicating a vehicle that is too close to vehicle 200, etc.). Thus, processing unit 110 may select information derived from two of the first, second, and third plurality of images based on the determinations of consistent and inconsistent information.

[0134] Navigational responses may include, for example, a turn, a lane shift, a change in acceleration, and the like. Processing unit 110 may cause the one or more navigational responses based on the analysis performed at step 720 and the techniques as described above in connection with FIG. 4. Processing unit 110 may also use data derived from execution of velocity and acceleration module 406 to cause the one or more navigational responses. In some embodiments, processing unit 110 may cause the one or more navigational responses based on a relative position, relative velocity, and/or relative acceleration between vehicle 200 and an object detected within any of the first, second, and third plurality of images. Multiple navigational responses may occur simultaneously, in sequence, or any combination thereof.

[0135] Curb Detection and Pedestrian Hazard Assessment

[0136] As described herein, system 100 may include technical features and capabilities that allow for the detection of certain elements that assist system 100 with staying on a road and avoiding collisions. In one example, system 100 is configured to detect road boundaries that define a roadway and/or pedestrians that are in the vicinity of the roadway. Knowledge of a physical road boundary allows system 100 to either alert a driver or control the vehicle in such a way as to avoid driving off of the road and/or colliding with an object or pedestrian. For example, identifying a road boundary helps to provide physical protection to pedestrians on a sidewalk adjacent to the road by demarcating the region in which pedestrians are likely to be present.

[0137] In an exemplary embodiment, system 100 may be configured to detect a curb as a road boundary. A curb in most situations is a physical boundary that separates the

road surface from an off-road surface (e.g., sidewalk, path, median, building, etc.). Keeping a vehicle from crossing a curb helps to keep the vehicle on the road and pedestrians on the sidewalk safe. However, pedestrians situated on the road surface are at a greater risk of being hit by moving vehicles. Therefore, an ability to classify pedestrians as either on-curb or off-curb, in addition to an estimation of the vehicle and pedestrian trajectory provides system 100 with an ability to assess a hazard posed by the vehicle to pedestrians. Disclosed embodiments include curb detection and pedestrian hazard assessment features to provide these and other functionalities to system 100.

[0138] To implement the curb detection and pedestrian hazard assessment functionality, processing unit 110 may receive images captured by at least one of image capture devices 122, 124, 126 and perform one or more image analysis processes. For example, processing unit 110 may perform a curb identification process to identify a curb in the captured images and a pedestrian identification process to determine a location of pedestrians relative to a road surface.

[0139] In some embodiments, memory 140 and/or 150 may store instructions programmed such that, upon execution by a processing device, curb detection and/or pedestrian hazard assessment functions are provided. As shown in FIG. 8, memory 140 and/or 150 may store a curb identification module 810, a curb registration module 820, a pedestrian identification module 830, and a database 840. Curb identification module 810 may store instructions for identifying elements of an image that correspond to a curb. Curb registration module 820 may store instructions for modeling a curb and/or determining the properties of each identified curb. Pedestrian identification module 830 may store instructions for detecting and classifying a pedestrian in one or more captured images. Database 840 may be configured to store data associated with curb detection and/or pedestrian hazard assessment functions. Further, curb identification module 810, curb registration module 820, and pedestrian identification module 830 may store instructions executable by one or more processors (e.g., processing unit 110), alone or in various combinations with each other. For example, curb identification module 810, curb registration module 820, and pedestrian identification module 830 may be configured to interact with each other and/or other modules of system 100 to perform functions consistent with disclosed embodiments.

[0140] Database 840 may include one or more memory devices that store information and are accessed and/or managed through a computing device, such as processing unit 110. In some embodiments, database 840 may be located in memory 140 or 150, as shown in FIG. 8. In other embodiments, database 840 may be located remotely from memory 140 or 150, and be accessible to other components of system 100 (e.g., processing unit 120) via one or more wireless connections (e.g., a wireless network). While one database 840 is shown, it should be understood that several separate and/or interconnected databases may make up database 840. Database 840 may include computing components (e.g., database management system, database server, etc.) configured to receive and process requests for data stored in memory devices associated with database 840 and to provide data from database 840 (e.g., to processing unit 110).

[0141] In some embodiments, database 840 may be configured to store data associated with providing curb detection and/or pedestrian hazard assessment functions. For example, database 840 may store data, such as images, parts of images, maps, algorithms, sets of values, or the like, which may allow processing unit

IPR2025-00944
Tesla EX1002 Page 1411

110 to identify information detected in an image. For instance, database 840 may store information that allows system 100 to identify a curb and/or pedestrians, such as stored images of curbs and/or images of pedestrians.

[0142] FIG. 9 depicts an exemplary image 900 captured by one of image capture devices 122, 124, 126. Image 900 includes an environment that may generally represent a setting that is seen as a vehicle travels on a road. For example, image 900 includes a road 910 defined between opposing curbs 920, and a plurality of pedestrians 930. In an exemplary embodiment, system 100 may be configured to identify and/or differentiate between the various features of image 900. Knowledge of these features allows system 100 to make decisions that help to promote safe operation of the vehicle. For example, identification of curbs 920 provides information that helps to keep the vehicle within the physical boundaries of road 910, away from pedestrians 930 on the sidewalk. Further, identification of curbs 920 allows system 100 to identify pedestrians that are on a surface of road 910, and thus, identify potential pedestrian hazards, which may require further action to avoid a collision.

[0143] The entirety of image 900 may represent a FOV of one or more of image capture devices 122, 124, 126. In some embodiments, system 100 may be configured to identify a region of interest (ROI) 950 within the FOV. In an exemplary embodiment, system 100 may be a detection system configured to process a selected ROI 950 and/or a full FOV associated with image 900 (and/or a plurality of images of which image 900 is one) to identify a curb 920.

[0144] In one embodiment, system 100 (e.g., curb identification module 810) may be configured to detect linear image edges to serve as curb edge line candidates. Curb identification module 810 may analyze an image to identify one or more edges lines of a curb, based on the identified curb edge line candidates. For example, curb identification module 810 may identify a bottom edge of a curb, which may be a transition line between the curb and the road surface in the image, by analyzing the edge line candidates to determine which, if any, include characteristics that match that of a curb bottom edge.

[0145] In some embodiments, system 100 (e.g., curb registration module 820) may be configured to determine whether the curb edge line candidate is part of a curb (e.g., the candidate would be correctly identified as an edge line of a curb) by determining whether a motion field of pixels in the vicinity of the curb edge line candidate is consistent with a three-dimensional structure of a step having an appropriate height. For example, system 100 may model a curb as a sharp step function and, using a plurality of captured images (e.g., a previous and current image), compute what the properties of the step are that would best model the motion field of pixels between the captured image frames. FIG. 9 illustrates an example of an identified curb 920A in the vicinity of the closest-in-path-pedestrian 930A.

[0146] In an exemplary embodiment, curb identification module 810 may identify curb edge line candidates by using an image analysis method. For example, system 100 may use a method that utilizes the Hough transform, which is an image processing technique that may be used to identify lines in an image. In another example, system 100 may use the Canny edge detection technique to identify edges present in an image using image gradient analysis. Other image analysis methods may be used, such as other methods that group image pixels by intensity, contrast, or other parameter to identify lines and/or curves present an image.

[0147] Curb edges, including the curb bottom edge, commonly appear as an elongated linear edge in the intensity pattern of an image, thus allowing them to be identified using one of the above described image analysis methods. However, this is also true for non-curb objects and road marks containing long linear features appearing in the image. Therefore, an elongated linear edge found in an image is not necessarily an edge of a curb. Curb registration module 820 may perform a curb detection process to analyze curb edge line candidates to identify a curb edge lines (and therefore the curb).

[0148] FIG. 10A depicts an exemplary image 1000, including linear edges 1010 identified by curb identification module 810 as potential edges of a curb 1020. Some of linear edges 1010 may correspond to edge lines of curb 1020 (e.g., top and/or bottom edges of the curb), while some other linear edges 1010 may correspond to non-curb objects, such as cracks, seams, markings, etc. Curb registration module 820 may analyze image 1000 and linear edges 1010 to determine the location of and/or model curb 1020. FIG. 10B depicts image 1000 with curb 1020 having been determined and modeled according to disclosed embodiments.

[0149] Curb registration module 820 may be configured to store instructions associated with one or more algorithms that allow curb 1020 to be identified and modeled as shown in FIG. 10B. For example, curb registration module 820 may include a curb analysis algorithm (CAA) that may allow system 100 to identify a curb within an image and/or plurality of images in a video sequence. In an exemplary embodiment, processing unit 110 may use the CAA to determine whether a curb edge line candidate corresponds to a curb edge line (e.g., a curb bottom edge or a curb top edge).

[0150] FIG. 11 depicts an example of a modeled step function 1100 which may be used to model a curb that corresponds to a real curb 1105, including at least some parameters that may be used by system 100 when executing the CAA. As shown in FIG. 11, system 110 may analyze a region of interest 1110 across multiple captured image frames to model step function 1100, while a mask region 1120 is not considered. This may allow processing unit 110 to focus only on pixels around a curb edge line candidate. In an exemplary embodiment, in order to determine whether a curb edge line candidate corresponds to a curb edge line, a height (hereinafter "Zcurb") of step function 1100 (i.e., a height of an associated modeled curb) may be calculated. For example, Zcurb may be calculated for a given curb edge line candidate detected in one image of a video sequence, the previous frame, and known, assumed, or estimated parameters (e.g., the camera's calibration and ego-motion parameters).

[0151] In one example, curb registration module 820 may utilize a parametric image warp function {tilde over (W)} (x; y; p) and an error minimization function E(p) to calculate Zcurb for a given curb edge line candidate. The image warp function {tilde over (W)} (x; y; p) may be a piecewise homography warp whereby each homography operates on different portions of a region of interest 1110 to match pixels in three-dimensional space between captured frame images. For example, the image warp function W(x; y; p) may be defined as:

[mathematical formula]

[0152] Each homography H1, H2, and H3 may correspond to a portion of region of interest 1120, as shown in FIG. 11. In an exemplary embodiment, H1 may represent a region inside of the bottom edge of the curb 1105 (e.g., the road surface), H2 may

IPR2025-00944
Tesla EX1002 Page 1413

represent a region between the bottom edge of the curb 1105 and the top edge of the curb 1105 (e.g., the height of the curb), and H3 may represent a region that extends away from the top of the curb 1105 and the road surface (e.g., the sidewalk), all within region of interest 1120. System 100 may use the image warp function {tilde over (W)}(x; y; p) to follow pixels between a plurality of images, such as a current image Ia and a previous image Ib, to determine the step function associated with the curb edge line candidate.

[0153] In an exemplary embodiment, each homography H1, H2, and H3 may be calculated based on a homography determination equation, such as:

[mathematical formula]

[0154] In general, the parameters that define each homography may be known, assumed, or estimated based on information associated with vehicle 200 and an image capture device 122, 124, 126. For example, K may be an image capturing device parameter which may take into account focal lengths, image centers, and/or other camera settings, and may be defined as:

[mathematical formula]

[0155] The inter-frame ego-motion parameters may be found using:

tab=(ΔX,ΔY,ΔZ)<T >

Δθab=(Δα,Δβ,Δγ)<T >

[0156] Rab=Rz(Δγ)Ry(Δβ)Rx(Δα) may be the camera rotation matrix composed of pitch (x), yaw (y), and roll (z) rotations (for camera pitch α, yaw β, and roll γ angles relative to the road plane), between frames Ia and Ib in the camera coordinate system of frame Ia.

[0157] Regarding the surface normals ni:

n1=n3=Rcalib{circumflex over (Z)}=Rcalib(0,0,1)<T >

n2=−u×n1

[0158] In these equations, u=(X2−X1)/‖(X2−X1)‖, u being a unit vector in the direction of the curb's bottom edge, and Rcalib=Rz(γ)Ry(β)Rx(α), Rcalib being a road calibration matrix.

[0159] Regarding the distances to the three planes di:

d1=Zcam

d2=|X1<T>n2|

d3=Zcam−Zcurb

[0160] In the above equations, the world coordinates X1 and X2 of the image coordinates {right arrow over (x)}1 and {right arrow over (x)}2 projected onto the road plane may be used. The image coordinates {right arrow over (x)}3 and {right arrow over

IPR2025-00944
Tesla EX1002 Page 1414

(x)}4 of the points corresponding to {right arrow over (x)}1 and {right arrow over (x)}2 on the upper part of the step function 1100 may define the line separating regions H2 and H3 within the region of interest 1110. These points may thus be defined:

[mathematical formula][mathematical formula][mathematical formula]

[0161] In these equations, x may be used to denote the three-dimensional homogenous image coordinates of the corresponding two-dimensional image point x.

[0162] The above equations may be used to determine the homographies H1, H2, and H3 using the image warp function {tilde over (W)} (x; y; p). Since the curb height Zcurb for step function 1100 is an unknown parameter, the image warp function {tilde over (W)} (x; y; p) may be iteratively solved until an appropriate Zcurb is determined. The appropriate Zcurb may correspond to a result of the image warp equation {tilde over (W)} (x; y; p) that is within a satisfactory error when warping the coordinates of the region of interest 1120 between captured image frames. This may be determined using the error minimization function E(p), where E(p) is defined as:

[mathematical formula]

[0163] The error minimization function E(p) may be used to minimize a difference between a current frame and the warp of a previous frame using known information, assumptions, and or estimated information. For example, system 100 may use the brightness constancy assumption when using the error minimization function E(p). The error minimization function E(p) allows one captured image frame to be compared to another captured image frame by warping the coordinates of one frame to the coordinates of the other frame. Minimizing the error of the warp allows system 100 to provide a more accurate calculation of Zcurb (e.g., because a low error value may indicate that the difference between captured image frame is accurately considered).

[0164] For each curb edge line candidate, system 100 may use the image warp function {tilde over (W)} (x; y; p) to identify the parameter vector p. The parameter vector p includes at least Zcurb, depending on whether other parameters are known, assumed, or estimated (e.g., additional degrees of freedom may be built in to allow some parameters to be unknown). Based on the parameter vector p, system 100 may determine a curb height Zcurb that is within an error threshold determined based on the error minimization function E(p). This Zcurb value may represent a value of a height of a step function modeled to a given curb edge line candidate. System 100 may compare the calculated Zcurb value to an expected Zcurb value (e.g., an average height of a curb that may be expected to be found on the road) to determine whether the calculated Zcurb value is within a threshold difference. If the calculated Zcurb value is within the threshold difference, system 100 may determine that the curb edge line candidate corresponds to a curb edge line.

[0165] System 100 may repeat the process of using the CAA (which may include the equations described above) to determine whether one or more identified curb edge line candidate corresponds to an curb edge line. System 100 may model each curb edge line candidate as both a curb bottom edge and a curb top edge to determine whether the curb edge line candidate matches either edge of a curb. Based on these processes, system 100 may determine the location of one or more curbs within an image or video sequence of images.

Tesla EX1002 Page 1415 IPR2025-00944

[0166] After a curb has been identified, pedestrian identification module 830 may perform one or more processes to identify a pedestrian relative to the curb. For example, pedestrian identification module 830 may assess whether an identified pedestrian is a potential hazard. System 100 may use the information related to the curb and the pedestrians to provide warnings to a driver and/or control the vehicle as it travels on the road.

[0167] FIG. 12 is a flowchart showing an exemplary process 1200 for identifying a curb within an image and/or video sequence. In an exemplary embodiment, system 100 may perform process 1200 to determine the location of a curb such that system 100 may determine a boundary of a road surface. Knowledge of the boundary of the road surface may allow system 100 to promote safe operation of a vehicle, through helping to maintain the vehicle on the road surface and out of collisions with objects and/or pedestrians. Further, knowledge of the boundary of the road surface may allow system 100 to assess whether any pedestrians or other objects are on the road surface and therefor hazards for which some action to prevent a collision should be taken.

[0168] In step 1202, processing unit 110 may receive a plurality of images of a vehicle environment. For example, one or more of image capture devices 122, 124, 126 may acquire a plurality of images of an area forward of vehicle 200. The area forward of the vehicle may include various features, such as those depicted in FIG. 9.

[0169] In step 1204, processing unit 110 may identify curb edge line candidates. For example, processing unit 110 (e.g., curb identification module 810) may use an algorithm or equation to identify linear edges within a capture images. These linear edges may represent curb edge candidates that may be further processed to determine whether they correctly represent edges of an curb present in the image or video sequence. In one example, processing unit 110 may use an edge detection method, such as using the Hough transform or by fitting lines to long connected components of the Canny edge response, to identify curb edge line candidates.

[0170] In step 1206, processing unit 110 may analyze each curb edge line candidate. For example, processing unit 110 (e.g., curb registration module 820) may use an algorithm or equation to iteratively process curb edge line candidates to determine whether any of them correspond to a curb edge line. For example, processing unit may use the CAA to model each curb edge line candidate as a step function and determine whether a corresponding height of that step function corresponds to what would be expected of a real curb.

[0171] In step 1208, processing unit 110 may identify at least one of the curb edge line candidates as an edge line of a curb. For example, processing unit 110 (e.g., curb registration module 820) may select a curb edge line candidate that is determined to be an edge line of a curb and store or otherwise identify that edge line candidate as a curb edge line. The curb edge line may be identified as a curb top edge or a curb bottom edge.

[0172] In addition, processing unit 110 may identify regions that correspond to the road surface and the off-road surface based on the identified edge line of the curb. For example, processing unit 110 may use determined edge lines of the curb to identify the curb throughout an image and/or video sequence and may identify a first region as a road surface and a second region as an off-road surface, throughout the image and/or each image of a video sequence. In this way, processing unit 110 may identify a

Tesla EX1002 Page 1416 IPR2025-00944

physical boundary of a road on which vehicle 200 travels. In some embodiments, processing unit 110 may store location and/or coordinate information associated with the identified curb edge line or curb.

[0173] In step 1210, processing unit 110 may use the identified curb edge line or curb to control vehicle 200. For example, processing unit 110 may provide location and/or coordinate information to a vehicle control module that makes decisions regarding speed, direction, acceleration, etc., of vehicle 200. In some instances, processing unit 110 may use the modeled curb to assess whether there are hazards that need to be considered. For example, processing unit 110 may use a modeled curb to assess pedestrian hazards, as explained in more detail below.

[0174] FIG. 13 is a flowchart showing an exemplary process 1300 for assessing hazards presented by pedestrians or other objects that may be located within a roadway. In an exemplary embodiment, system 100 may perform process 1300 to determine the location on one or more pedestrians in a FOV of an image capturing device 122, 124, 126. For example, system 100 may determine that one or more pedestrians is located on a road surface of a road (e.g., and therefore not off-road) and determine one or more actions that should be taken to prevent a collision with the pedestrian.

[0175] In step 1310, processing unit 110 may identify pedestrians within an image and/or images of a video sequence. For example, processing unit 110 (e.g., pedestrian identification module 830) may process one or more images to identify objects of interest within the images. Processing unit 110 may further analyze the objects of interest to determine whether they correspond to a pedestrian (or other object which may be treated as a pedestrian, such as an animal, bicyclist, etc.). In one example, processing unit may compare an image to stored images to match characteristics of an image with information that is known to correspond to a pedestrian (e.g., a figure in the shape of a person). In another example, processing unit may compare multiple images of a video sequence to identify objects that are moving in a manner consistent with a pedestrian.

[0176] In step 1320, processing unit 110 may classify pedestrians. For example, processing unit 110 (e.g., pedestrian identification module 830) may classify pedestrians as either on the road surface or on the off-road surface. In one example, processing unit 110 may compare the location and/or trajectory of an identified pedestrian with the location of an identified and modeled curb (e.g., a curb modeled using process 1200). For instance, processing unit may determine whether the pedestrian is within a region that is determined to be the road surface (e.g., inside of a bottom edge of the curb) or within a region that is determined to be the off-road surface (e.g., outside of a top edge of the curb).

[0177] In step 1330, processing unit 110 may assess whether a pedestrian presents a hazard to vehicle 200. In one example, processing unit 110 may determine that a pedestrian on the road surface is a hazard. In another example, processing unit 110 may compare a trajectory of vehicle 200 to a trajectory of an identified pedestrian on the road surface to determine whether there is a possibility of a collision between vehicle and the pedestrian.

[0178] In step 1340, processing unit 110 may alert a driver and/or control vehicle 200 to prevent a collision. In one example, processing unit 110 may alert a driver to a

Tesla EX1002 Page 1417 IPR2025-00944

pedestrian on a road surface whenever a pedestrian is classified as being on the road surface. In another example, processing unit 110 may alert a driver only if processing unit 110 determines that a collision with the pedestrian is likely (e.g., above a risk threshold, such as based on a predicted closest distance between vehicle 200 and the pedestrian at some point in time). In some embodiments, processing unit 110 may use the assessment of a pedestrian hazard to automatically control travel of vehicle 200. For example, processing unit 110 may modify a path (e.g., steering direction), speed, acceleration, etc., of vehicle 200 in order to avoid a collision with a pedestrian.

[0179] System 100 may use processes 1200 and/or 1300 to identify features of an environment of vehicle 200, and use those features to operate/control vehicle 200. In particular, as discussed in the examples above, system 100 may identify a road boundary and determine whether there are any pedestrian hazards based on their location with respect to the road boundary. It should be understood, however, that these are examples, and that other similar processes may be implemented. For example, other road boundaries besides curbs, such as railings, medians, shoulders, etc., may be determined, and other hazards besides pedestrians, such as other vehicles, debris, road imperfections, etc., may be identified and assessed based on their location and/or position with respect to a road boundary.

[0180] The foregoing description has been presented for purposes of illustration. It is not exhaustive and is not limited to the precise forms or embodiments disclosed. Modifications and adaptations will be apparent to those skilled in the art from consideration of the specification and practice of the disclosed embodiments. Additionally, although aspects of the disclosed embodiments are described as being stored in memory, one skilled in the art will appreciate that these aspects can also be stored on other types of computer readable media, such as secondary storage devices, for example, hard disks or CD ROM, or other forms of RAM or ROM, USB media, DVD, Blu-ray, or other optical drive media.

[0181] Computer programs based on the written description and disclosed methods are within the skill of an experienced developer. The various programs or program modules can be created using any of the techniques known to one skilled in the art or can be designed in connection with existing software. For example, program sections or program modules can be designed in or by means of .Net Framework, .Net Compact Framework (and related languages, such as Visual Basic, C, etc.), Java, C++, Objective-C, HTML, HTML/AJAX combinations, XML, or HTML with included Java applets.

[0182] Moreover, while illustrative embodiments have been described herein, the scope of any and all embodiments having equivalent elements, modifications, omissions, combinations (e.g., of aspects across various embodiments), adaptations and/or alterations as would be appreciated by those skilled in the art based on the present disclosure. The limitations in the claims are to be interpreted broadly based on the language employed in the claims and not limited to examples described in the present specification or during the prosecution of the application. The examples are to be construed as non-exclusive. Furthermore, the steps of the disclosed methods may be modified in any manner, including by reordering steps and/or inserting or deleting steps. It is intended, therefore, that the specification and examples be considered as illustrative only, with a true scope and spirit being indicated by the following claims and their full scope of equivalents.

IPR2025-00944
Tesla EX1002 Page 1418

IPR2025-00944
Tesla EX1002 Page 1419

Espacenet

Claims: EP3143607 (A1) — 2017-03-22

SYSTEMS AND METHODS FOR CURB DETECTION AND PEDESTRIAN HAZARD ASSESSMENT

Claims not available for EP3143607 (A1)
Claims of corresponding document: US2015332114 (A1)

A high quality text as facsimile in your desired language may be available amongst the following family members:

CN106663193 (A) US9443163 (B2) WO2015177648 (A1) CN112580456 (A) US2016371550 (A1) US2018189579 (A1) US2019251375 (A1) US2020320315 (A1)

- Original claims
- Claims tree

The EPO does not accept any responsibility for the accuracy of data and information originating from other authorities than the EPO; in particular, the EPO does not guarantee that they are complete, up-to-date or fit for specific purposes.

1. What is claimed is: 1. A detection system for a vehicle, the system comprising:
at least one image capture device configured to acquire a plurality of images of an area forward of the vehicle, the area including a curb separating a road surface from an off-road surface;
a data interface; and
at least one processing device programmed to:
receive the plurality of images via the data interface;
determine a plurality of curb edge line candidates in the plurality of images; and
identify at least one curb edge line candidate as an edge line of the curb.

2. The detection system of claim 1, wherein the edge line of the curb is a curb bottom edge or a curb top edge.

3. The detection system of claim 1, wherein the at least one processing device is further programmed to determine a height of a step function associated with each curb edge line candidate.

4. The detection system of claim 3, wherein identifying the at least one curb edge line candidate as an edge line of the curb includes comparing the height of the step function to a threshold value.

5. The detection system of claim 3, wherein determining the curb height includes executing an image warp function between pixel coordinates of the plurality of images.

6. The detection system of claim 1, wherein the at least one processing device is further programmed to identify regions in the plurality of images that correspond to the road surface and the off-road surface based on the identified edge line of the curb.

7. The detection system of claim 6, wherein the at least one processing device is further programmed to:
identify a pedestrian in one or more of the plurality of images;
assess whether the pedestrian presents a hazard; and
perform an action to avoid a collision with the pedestrian, based on the assessment.

8. The detection system of claim 7, wherein assessing whether the pedestrian presents a hazard includes classifying the pedestrian based on the identified regions and at least one of a location or a trajectory of the pedestrian.

9. The detection system of claim 8, wherein classifying the pedestrian includes classifying the pedestrian as on the road surface or on the off-road surface.

10. The detection system of claim 7, wherein assessing whether the pedestrian presents a hazard includes comparing a trajectory of the pedestrian to a trajectory of a vehicle.

11. The detection system of claim 7, wherein the action to avoid a collision includes alerting a driver.

12. The detection system of claim 7, wherein the action to avoid the collision includes modifying at least one of a speed, acceleration, or direction of a vehicle.

13. A vehicle, comprising:
a body;
at least one image capture device configured to acquire a plurality of images of an area forward of the vehicle, the area including a curb separating a road surface from an off-road surface;
a data interface; and
at least one processing device programmed to:
receive the plurality of images via the data interface;
determine a plurality of curb edge line candidates in the plurality of images;
identify at least one curb edge line candidate as an edge line of the curb; and
identify regions in the plurality of images that correspond to the road surface and the off-road surface based on the identified edge line of the curb.

14. The vehicle of claim 13, wherein the at least one processing device is further programmed to:
identify a pedestrian in one or more of the plurality of images;
assess whether the pedestrian presents a hazard; and
perform an action to avoid a collision with the pedestrian, based on the assessment.

15. The vehicle of claim 14, wherein assessing whether the pedestrian presents a hazard includes classifying the pedestrian based on the identified regions and a location of the pedestrian.

16. The vehicle of claim 15, wherein classifying the pedestrian includes comparing a

IPR2025-00944
Tesla EX1002 Page 1421

location of the pedestrian to the identified regions.

17. The vehicle of claim 16, wherein classifying the pedestrian includes classifying the pedestrian as on the road surface or on the off-road surface based on the comparison.

18. The vehicle of claim 13, wherein assessing whether the pedestrian presents a hazard includes comparing a trajectory of the pedestrian to a trajectory of a vehicle.

19. A method for curb detection, the method comprising:
acquiring, via at least one image capture device, a plurality of images of an area forward of a vehicle, the area including a curb separating a road surface from an off-road surface;
determining a plurality of curb edge line candidates in the plurality of images;
determining a height of a step function associated with each curb edge line candidate; and
identifying at least one curb edge line candidate as an edge line of the curb based on the determined heights.

20. The method of claim 19, wherein determining the height of the step function includes executing an image warp function between pixel coordinates of the plurality of images.

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Veröffentlichungsnummer:
(11) Publication number: **EP 3 143 607 A0**
(11) Numéro de publication:

Internationale Anmeldung veröffentlicht durch die Weltorganisation für geistiges Eigentum unter der Nummer:

WO 2015/177648 (Art. 153(3) EPÜ).

International application published by the World Intellectual Property Organization under number:

WO 2015/177648 (Art. 153(3) EPC).

Demande internationale publiée par l'Organisation Mondiale de la Propriété Intellectuelle sous le numéro:

WO 2015/177648 (art. 153(3) CBE).

Espacenet

Bibliographic data: EP3146506 (A2) — 2017-03-29

SYSTEMS AND METHODS FOR BRAKING A VEHICLE BASED ON A DETECTED OBJECT

Inventor(s): DAGAN EREZ [IL] ± (DAGAN, Erez)

Applicant(s): MOBILEYE VISION TECHNOLOGIES LTD [IL] ± (Mobileye Vision Technologies Ltd)

Classification: **- international:** ***G06T7/20***
- cooperative: **B60T17/221 (US)**; **B60T7/22 (EP, CN, US)**; **B60T8/171 (US)**; **B60T8/172 (US)**; **G06K9/00791 (EP)**; **G06K9/00805 (EP)**; **G06T7/269 (EP, US)**; **G07C5/0808 (US)**; B60T2201/022 (EP, CN, US); B60T2270/40 (US); G06T2207/30252 (EP, CN, US); G06T2207/30261 (EP, CN, US)

Application number: EP20150770635 20150522 Global Dossier

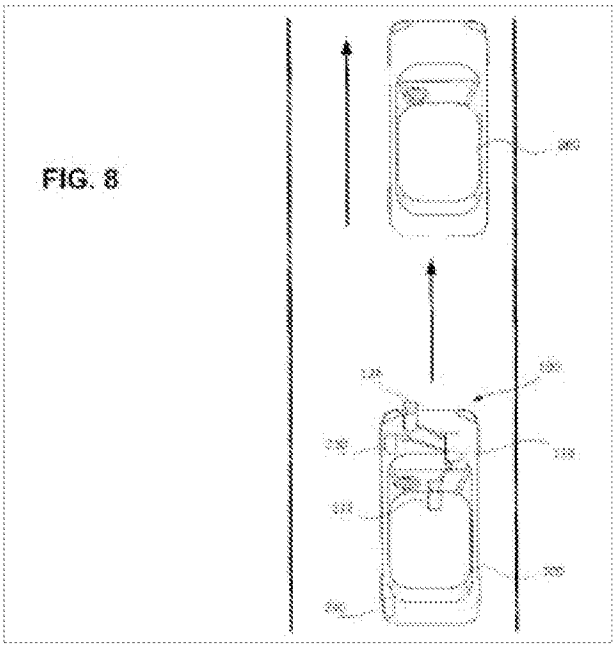
Priority number(s): US201462001772P 20140522 ; WO2015IB01457 20150522

Also published as: EP3146506 (B1) CN106536299 (A) CN106536299 (B) CN110171405 (A) CN110171405 (B) more

Abstract not available for EP3146506 (A2)
Abstract of corresponding document: WO2015177643 (A2)

Systems and methods are provided for detecting an object and causing a vehicle to brake based on the detection. In one implementation, an object detecting and braking system for a vehicle includes at least one image capture device configured to acquire a plurality of images of an area including an object in front of the vehicle. The system includes at least one processing device programmed to perform a first image analysis to determine a first estimated time-to-collision of the vehicle with the object, and to perform a second image analysis to determine a second estimated time-to-collision of the vehicle with the object. The processing device is also programmed to calculate a difference between the first estimated time-to-collision and the second estimated time-to-collision, to determine that the difference does not exceed a predetermined threshold, and to cause the vehicle to brake based on the determination that the difference does not exceed the predetermined threshold.

FIG. 8



Espacenet

Description: EP3146506 (A2) — 2017-03-29

SYSTEMS AND METHODS FOR BRAKING A VEHICLE BASED ON A DETECTED OBJECT

Description of EP3146506 (B1)

A high quality text as facsimile in your desired language may be available amongst the following family members:

CN106536299 (A) US2015336547 (A1) WO2015177643 (A2) CN110171405 (A) US2017203744 (A1) US2019084549 (A1)

The EPO does not accept any responsibility for the accuracy of data and information originating from other authorities than the EPO; in particular, the EPO does not guarantee that they are complete, up-to-date or fit for specific purposes.

BACKGROUND

I. Technical Field

[0001] The present disclosure relates generally to autonomous vehicle navigation and, more specifically, to systems and methods that use cameras to detect an object in front of the vehicle and that apply brakes based on the detected object.

II. Background Information

[0002] As technology continues to advance, the goal of a fully autonomous vehicle that is capable of navigating on roadways is on the horizon. Primarily, an autonomous vehicle may be able to identify its environment and navigate without input from a human operator. Autonomous vehicles may also take into account a variety of factors and make appropriate decisions based on those factors to safely and accurately reach an intended destination. For example, various objects-such as other vehicles and pedestrians-are encountered when a vehicle typically travels a roadway. Autonomous driving systems may recognize these objects in a vehicle's environment and take appropriate and timely action to avoid collisions. Additionally, autonomous driving systems may identify other indicators-such as traffic signals, traffic signs, and lane markings-that regulate vehicle movement (e.g., when the vehicle must stop and may go, a speed at which the vehicle must not exceed, where the vehicle must be positioned on the roadway, etc.). Autonomous driving systems may need to determine when a vehicle should change lanes, turn at intersections, change roadways, etc. As is

IPR2025-00944
Tesla EX1002 Page 1426

evident from these examples, many factors may need to be addressed in order to provide an autonomous vehicle that is capable of navigating safely and accurately.

[0003] US 2012/0314071 A1 discloses a method for preventing collisions between a vehicle and pedestrians using a camera mounted in a vehicle for capturing image frames of the environment of the vehicle in the field of view of the camera. When a candidate image of a pedestrian in the field of view of the camera is detected in an image frame, the optical flow between the image frames is analyzed to validate the pedestrian's detection. When the analyzed optical flow is consistent with an upright pedestrian, a collision warning signal is provided.

[0004] US 2006/0178830 A1 relates to collision warning methods and apparatus for vehicles. An apparatus for facilitating avoidance of a vehicle collision with an object comprises a vision sensor providing image data, and an image analyzer, operable to provide an estimated time to collision of the vehicle with the object. The estimated time to collision is determined from a rate of expansion of an image element corresponding to the object within the image data, and the apparatus provides a warning if the estimated time to collision is less than or approximately equal to a predetermined time.

[0005] US 5,559,695 A discloses an apparatus and method for real-time, adaptive self-calibration of a visual time-to-contact sensor mounted on a vehicle to predict the time-to-contact with stationary and moving obstacles in a moving vehicle's immediate path. The focus of expansion is actively centered in the image sequence, using optical flow, and the pan and tilt of the sensor are continuously calibrated to point in the direction of the focus of expansion in order to accurately predict the time-to-contact with stationary and moving obstacles in a moving vehicle's immediate path, and enable warning or evasive action.

[0006] " A study on the reliability improvement for applying GNSS to railroad systems" by Kyung-Ho Shin , et al. investigates reliability of application of Global Navigation Satellite (GNSS) systems in the railway field and proposes a dual-mode GNSS configuration method for controlling trains with improved reliability.

SUMMARY

[0007] Embodiments consistent with the present disclosure provide systems and methods for autonomous vehicle navigation. The disclosed embodiments may use cameras to provide autonomous vehicle navigation features. For example, consistent with the disclosed embodiments, the disclosed systems may include one, two, or more cameras that monitor the environment of a vehicle and cause a navigational response based on an analysis of images captured by one or more of the cameras.

[0008] In accordance with aspects of the invention there are provided an object detecting and braking system for a vehicle, a vehicle, and a method for detecting an object and braking a vehicle, as set out in claims 1, 7, and 8 respectively.

[0009] Consistent with a disclosed embodiment, an object detecting and braking system for a vehicle is provided, according to claim 1.

[0010] Consistent with another disclosed embodiment, a vehicle is provided, according to claim 7.

Tesla EX1002 Page 1427 IPR2025-00944

[0011] Consistent with yet another disclosed embodiment, a method for detecting an object and braking a vehicle is provided, according to claim 8.

[0012] Consistent with another disclosed example, a method is provided for assessing an overall system failure rate associated with a braking decisioning system for a vehicle. The method includes determining, using a processing device, a first failure rate associated with a textural analysis based sub-system. The textural analysis based sub-system is configured to make a decision to brake based on a change in texture between at least two images of an area in front of a vehicle. The method further includes determining, using the processing device, a second failure rate associated with a structural analysis based sub-system. The structural analysis based sub-system is configured to make a decision to brake based on optical flow information derived from at least two images of the area in front of the vehicle. The method also includes determining, using the processing device, the overall system failure rate based on the first failure rate and the second failure rate.

[0013] Consistent with other disclosed examples, non-transitory computer-readable storage media may store program instructions, which are executed by at least one processing device and perform any of the methods described herein.

[0014] The foregoing general description and the following detailed description are exemplary and explanatory only. The invention is defined solely by the appended claims.

BRIEF DESCRIPTION OF THE DRAWINGS

[0015] The accompanying drawings, which are incorporated in and constitute a part of this disclosure, illustrate various disclosed embodiments. In the drawings:
FIG. 1 is a diagrammatic representation of an exemplary system consistent with the disclosed embodiments.
FIG. 2A is a diagrammatic side view representation of an exemplary vehicle including a system consistent with the disclosed embodiments.
FIG. 2B is a diagrammatic top view representation of the vehicle and system shown in FIG. 2A consistent with the disclosed embodiments.
FIG. 2C is a diagrammatic top view representation of another embodiment of a vehicle including a system consistent with the disclosed embodiments.
FIG. 2D is a diagrammatic top view representation of yet another embodiment of a vehicle including a system consistent with the disclosed embodiments.
FIG. 2E is a diagrammatic top view representation of yet another embodiment of a vehicle including a system consistent with the disclosed embodiments.
FIG. 2F is a diagrammatic representation of exemplary vehicle control systems consistent with the disclosed embodiments.
FIG. 3A is a diagrammatic representation of an interior of a vehicle including a rearview mirror and a user interface for a vehicle imaging system consistent with the disclosed embodiments.
FIG. 3B is an illustration of an example of a camera mount that is configured to be positioned behind a rearview mirror and against a vehicle windshield consistent with the disclosed embodiments.
FIG. 3C is an illustration of the camera mount shown in FIG. 3B from a different perspective consistent with the disclosed embodiments.
FIG. 3D is an illustration of an example of a camera mount that is configured to be positioned behind a rearview mirror and against a vehicle windshield consistent with

IPR2025-00944
Tesla EX1002 Page 1428

the disclosed embodiments.
FIG. 4 is an exemplary block diagram of a memory configured to store instructions for performing one or more operations consistent with the disclosed embodiments.
FIG. 5A is a flowchart showing an exemplary process for causing one or more navigational responses based on monocular image analysis consistent with disclosed embodiments.
FIG. 5B is a flowchart showing an exemplary process for detecting one or more vehicles and/or pedestrians in a set of images consistent with the disclosed embodiments.
FIG. 5C is a flowchart showing an exemplary process for detecting road marks and/or lane geometry information in a set of images consistent with the disclosed embodiments.
FIG. 5D is a flowchart showing an exemplary process for detecting traffic lights in a set of images consistent with the disclosed embodiments.
FIG. 5E is a flowchart showing an exemplary process for causing one or more navigational responses based on a vehicle path consistent with the disclosed embodiments.
FIG. 5F is a flowchart showing an exemplary process for determining whether a leading vehicle is changing lanes consistent with the disclosed embodiments.
FIG. 6 is a flowchart showing an exemplary process for causing one or more navigational responses based on stereo image analysis consistent with the disclosed embodiments.
FIG. 7 is a flowchart showing an exemplary process for causing one or more navigational responses based on an analysis of three sets of images consistent with the disclosed embodiments.
FIG. 8 is a diagrammatic representation of an exemplary vehicle including an object detecting and braking system traveling on a road behind another vehicle consistent with the disclosed embodiments.
FIG. 9 is a diagrammatic representation of an exemplary vehicle including an object detecting and braking system traveling on a road including an oncoming vehicle consistent with the disclosed embodiments.
FIG. 10 is an exemplary block diagram of a memory that may store instructions for performing one or more operations for detecting an object in front of the vehicle and causing the vehicle to brake based on the detection consistent with the disclosed embodiments.
FIG. 11 is a flowchart showing an exemplary process for detecting an object in front of the vehicle and causing the vehicle to brake based on the detection consistent with the disclosed embodiments.
FIG. 12 is a flowchart showing an exemplary first image analysis for determining a first estimated time-to-collision consistent with the disclosed embodiments.
FIG. 13 is a flowchart showing an exemplary second image analysis for determining a second estimated time-to-collision consistent with the disclosed embodiments.
FIGS. 14A-14C show an image with an optical flow field and a distribution of a second estimated time-to-collision in the optical flow field consistent with the disclosed embodiments.
FIGS. 15A and 15B show exemplary focuses of expansion when an object travels in different directions consistent with the disclosed embodiments.
FIG. 16 shows an exemplary optical flow field and a focus of expansion consistent with the disclosed embodiments.
FIGS. 17A and 17B show an image of an object and a linear relationship between distances from points in an optical flow field to the focus of expansion and magnitudes of movement of the points consistent with the disclosed embodiments.

IPR2025-00944
Tesla EX1002 Page 1429

FIGS. 18A and 18B show an image of an object and a non-linear relationship between distances from points in an optical flow field to the focus of expansion and magnitudes of movement of the points consistent with the disclosed embodiments.
FIG. 19 is a flowchart showing an exemplary process for triggering a warning based on probabilities indicating that a time-to-collision is smaller than one or more thresholds consistent with the disclosed embodiments.
FIG. 20 is a flowchart showing an exemplary process for determining the first time-to-collision (TTC1) that may be implemented in the first image analysis shown in FIG. 11 consistent with the disclosed embodiments.
FIG. 21 is a flowchart showing an exemplary process for performing a system validation using two modalities consistent with the disclosed embodiments.

DETAILED DESCRIPTION

[0016] The following detailed description refers to the accompanying drawings. Wherever possible, the same reference numbers are used in the drawings and the following description to refer to the same or similar parts. While several illustrative embodiments are described herein, modifications, adaptations and other implementations are possible. For example, substitutions, additions or modifications may be made to the components illustrated in the drawings, and the illustrative methods described herein may be modified by substituting, reordering, removing, or adding steps to the disclosed methods. Accordingly, the following detailed description is not limited to the disclosed embodiments and examples. Instead, the proper scope is defined by the appended claims.

[0017] FIG. 1 is a block diagram representation of a system 100 consistent with the exemplary disclosed embodiments. System 100 may include various components depending on the requirements of a particular implementation. In some embodiments, system 100 may include a processing unit 110, an image acquisition unit 120, a position sensor 130, one or more memory units 140, 150, a map database 160, and a user interface 170. Processing unit 110 may include one or more processing devices. In some embodiments, processing unit 110 may include an applications processor 180, an image processor 190, or any other suitable processing device. Similarly, image acquisition unit 120 may include any number of image acquisition devices and components depending on the requirements of a particular application. In some embodiments, image acquisition unit 120 may include one or more image capture devices (e.g., cameras), such as image capture device 122, image capture device 124, and image capture device 126. System 100 may also include a data interface 128 communicatively connecting processing device 110 to image acquisition device 120. For example, data interface 128 may include any wired and/or wireless link or links for transmitting image data acquired by image accusation device 120 to processing unit 110.

[0018] Both applications processor 180 and image processor 190 may include various types of processing devices. For example, either or both of applications processor 180 and image processor 190 may include a microprocessor, preprocessors (such as an image preprocessor), graphics processors, a central processing unit (CPU), support circuits, digital signal processors, integrated circuits, memory, or any other types of devices suitable for running applications and for image processing and analysis. In some embodiments, applications processor 180 and/or image processor 190 may include any type of single or multi-core processor, mobile device microcontroller, central processing unit, etc. Various processing devices may be used, including, for

Tesla EX1002 Page 1430 IPR2025-00944

example, processors available from manufacturers such as Intel®, AMD®, etc. and may include various architectures (e.g., x86 processor, ARM®, etc.).

[0019] In some embodiments, applications processor 180 and/or image processor 190 may include any of the EyeQ series of processor chips available from Mobileye®. These processor designs each include multiple processing units with local memory and instruction sets. Such processors may include video inputs for receiving image data from multiple image sensors and may also include video out capabilities. In one example, the EyeQ2® uses 90nm-micron technology operating at 332Mhz. The EyeQ2® architecture consists of two floating point, hyper-thread 32-bit RISC CPUs (MIPS32® 34K® cores), five Vision Computing Engines (VCE), three Vector Microcode Processors (VMP®), Denali 64-bit Mobile DDR Controller, 128-bit internal Sonics Interconnect, dual 16-bit Video input and 18-bit Video output controllers, 16 channels DMA and several peripherals. The MIPS34K CPU manages the five VCEs, three VMP™ and the DMA, the second MIPS34K CPU and the multi-channel DMA as well as the other peripherals. The five VCEs, three VMP® and the MIPS34K CPU can perform intensive vision computations required by multi-function bundle applications. In another example, the EyeQ3®, which is a third generation processor and is six times more powerful that the EyeQ2®, may be used in the disclosed embodiments.

[0020] Any of the processing devices disclosed herein may be configured to perform certain functions. Configuring a processing device, such as any of the described EyeQ processors or other controller or microprocessor, to perform certain functions may include programming of computer executable instructions and making those instructions available to the processing device for execution during operation of the processing device. In some embodiments, configuring a processing device may include programming the processing device directly with architectural instructions. In other embodiments, configuring a processing device may include storing executable instructions on a memory that is accessible to the processing device during operation. For example, the processing device may access the memory to obtain and execute the stored instructions during operation.

[0021] While FIG. 1 depicts two separate processing devices included in processing unit 110, more or fewer processing devices may be used. For example, in some embodiments, a single processing device may be used to accomplish the tasks of applications processor 180 and image processor 190. In other embodiments, these tasks may be performed by more than two processing devices.

[0022] Processing unit 110 may comprise various types of devices. For example, processing unit 110 may include various devices, such as a controller, an image preprocessor, a central processing unit (CPU), support circuits, digital signal processors, integrated circuits, memory, or any other types of devices for image processing and analysis. The image preprocessor may include a video processor for capturing, digitizing and processing the imagery from the image sensors. The CPU may comprise any number of microcontrollers or microprocessors. The support circuits may be any number of circuits generally well known in the art, including cache, power supply, clock and input-output circuits. The memory may store software that, when executed by the processor, controls the operation of the system. The memory may include databases and image processing software. The memory may comprise any number of random access memories, read only memories, flash memories, disk drives, optical storage, tape storage, removable storage and other types of storage. In one instance, the memory may be separate from the processing unit 110. In another

IPR2025-00944
Tesla EX1002 Page 1431

IPR2025-00944
Tesla EX1002 Page 1432

instance, the memory may be integrated into the processing unit 110.

[0023] Each memory 140, 150 may include software instructions that when executed by a processor (e.g., applications processor 180 and/or image processor 190), may control operation of various aspects of system 100. These memory units may include various databases and image processing software. The memory units may include random access memory, read only memory, flash memory, disk drives, optical storage, tape storage, removable storage and/or any other types of storage. In some embodiments, memory units 140, 150 may be separate from the applications processor 180 and/or image processor 190. In other embodiments, these memory units may be integrated into applications processor 180 and/or image processor 190.

[0024] Position sensor 130 may include any type of device suitable for determining a location associated with at least one component of system 100. In some embodiments, position sensor 130 may include a GPS receiver. Such receivers can determine a user position and velocity by processing signals broadcasted by global positioning system satellites. Position information from position sensor 130 may be made available to applications processor 180 and/or image processor 190.

[0025] User interface 170 may include any device suitable for providing information to or for receiving inputs from one or more users of system 100. In some embodiments, user interface 170 may include user input devices, including, for example, a touchscreen, microphone, keyboard, pointer devices, track wheels, cameras, knobs, buttons, etc. With such input devices, a user may be able to provide information inputs or commands to system 100 by typing instructions or information, providing voice commands, selecting menu options on a screen using buttons, pointers, or eye-tracking capabilities, or through any other suitable techniques for communicating information to system 100.

[0026] User interface 170 may be equipped with one or more processing devices configured to provide and receive information to or from a user and process that information for use by, for example, applications processor 180. In some embodiments, such processing devices may execute instructions for recognizing and tracking eye movements, receiving and interpreting voice commands, recognizing and interpreting touches and/or gestures made on a touchscreen, responding to keyboard entries or menu selections, etc. In some embodiments, user interface 170 may include a display, speaker, tactile device, and/or any other devices for providing output information to a user.

[0027] Map database 160 may include any type of database for storing map data useful to system 100. In some embodiments, map database 160 may include data relating to the position, in a reference coordinate system, of various items, including roads, water features, geographic features, businesses, points of interest, restaurants, gas stations, etc. Map database 160 may store not only the locations of such items, but also descriptors relating to those items, including, for example, names associated with any of the stored features. In some embodiments, map database 160 may be physically located with other components of system 100. Alternatively or additionally, map database 160 or a portion thereof may be located remotely with respect to other components of system 100 (e.g., processing unit 110). In such embodiments, information from map database 160 may be downloaded over a wired or wireless data connection to a network (e.g., over a cellular network and/or the Internet, etc.).

[0028] Image capture devices 122, 124, and 126 may each include any type of device suitable for capturing at least one image from an environment. Moreover, any number of image capture devices may be used to acquire images for input to the image processor. Some embodiments may include only a single image capture device, while other embodiments may include two, three, or even four or more image capture devices. Image capture devices 122, 124, and 126 will be further described with reference to FIGS. 2B-2E , below.

[0029] System 100, or various components thereof, may be incorporated into various different platforms. In some embodiments, system 100 may be included on a vehicle 200, as shown in FIG. 2A . For example, vehicle 200 may be equipped with a processing unit 110 and any of the other components of system 100, as described above relative to FIG. 1 . While in some embodiments vehicle 200 may be equipped with only a single image capture device (e.g., camera), in other embodiments, such as those discussed in connection with FIGS. 2B-2E , multiple image capture devices may be used. For example, either of image capture devices 122 and 1 24 of vehicle 200, as shown in FIG. 2A , may be part of an ADAS (Advanced Driver Assistance Systems) imaging set.

[0030] The image capture devices included on vehicle 200 as part of the image acquisition unit 120 may be positioned at any suitable location. In some embodiments, as shown in FIGS. 2A-2E and 3A-3C , image capture device 122 may be located in the vicinity of the rearview mirror. This position may provide a line of sight similar to that of the driver of vehicle 200, which may aid in determining what is and is not visible to the driver. Image capture device 122 may be positioned at any location near the rearview mirror, but placing image capture device 122 on the driver side of the mirror may further aid in obtaining images representative of the driver's field of view and/or line of sight.

[0031] Other locations for the image capture devices of image acquisition unit 120 may also be used. For example, image capture device 124 may be located on or in a bumper of vehicle 200. Such a location may be especially suitable for image capture devices having a wide field of view. The line of sight of bumper-located image capture devices can be different from that of the driver and, therefore, the bumper image capture device and driver may not always see the same objects. The image capture devices (e.g., image capture devices 122, 124, and 126) may also be located in other locations. For example, the image capture devices may be located on or in one or both of the side mirrors of vehicle 200, on the roof of vehicle 200, on the hood of vehicle 200, on the trunk of vehicle 200, on the sides of vehicle 200, mounted on, positioned behind, or positioned in front of any of the windows of vehicle 200, and mounted in or near light figures on the front and/or back of vehicle 200, etc.

[0032] In addition to image capture devices, vehicle 200 may include various other components of system 100. For example, processing unit 110 may be included on vehicle 200 either integrated with or separate from an engine control unit (ECU) of the vehicle. Vehicle 200 may also be equipped with a position sensor 130, such as a GPS receiver and may also include a map database 160 and memory units 140 and 150.

[0033] FIG. 2A is a diagrammatic side view representation of an exemplary vehicle imaging system consistent with the disclosed embodiments. FIG. 2B is a diagrammatic top view illustration of the embodiment shown in FIG. 2A . As illustrated in FIG. 2B , the disclosed embodiments may include a vehicle 200 including in its body a system 100 with a first image capture device 122 positioned in the vicinity of the rearview mirror

IPR2025-00944
Tesla EX1002 Page 1433

and/or near the driver of vehicle 200, a second image capture device 124 positioned on or in a bumper region (e.g., one of bumper regions 210) of vehicle 200, and a processing unit 110.

[0034] As illustrated in FIG. 2C , image capture devices 122 and 124 may both be positioned in the vicinity of the rearview mirror and/or near the driver of vehicle 200. Additionally, while two image capture devices 122 and 124 are shown in FIGS. 2B and 2C , it should be understood that other embodiments may include more than two image capture devices. For example, in the embodiments shown in FIGS. 2D and 2E , first, second, and third image capture devices 122, 124, and 126, are included in the system 100 of vehicle 200.

[0035] As illustrated in FIG. 2D , image capture device 122 may be positioned in the vicinity of the rearview mirror and/or near the driver of vehicle 200, and image capture devices 124 and 126 may be positioned on or in a bumper region (e.g., one of bumper regions 210) of vehicle 200. And as shown in FIG. 2E , image capture devices 122, 124, and 126 may be positioned in the vicinity of the rearview mirror and/or near the driver seat of vehicle 200. The disclosed embodiments are not limited to any particular number and configuration of the image capture devices, and the image capture devices may be positioned in any appropriate location within and/or on vehicle 200.

[0036] It is to be understood that the disclosed embodiments are not limited to vehicles and could be applied in other contexts. It is also to be understood that disclosed embodiments are not limited to a particular type of vehicle 200 and may be applicable to all types of vehicles including automobiles, trucks, trailers, and other types of vehicles.

[0037] The first image capture device 122 may include any suitable type of image capture device. Image capture device 122 may include an optical axis. In one instance, the image capture device 122 may include an Aptina M9V024 WVGA sensor with a global shutter. In other embodiments, image capture device 122 may provide a resolution of 1280x960 pixels and may include a rolling shutter. Image capture device 122 may include various optical elements. In some embodiments one or more lenses may be included, for example, to provide a desired focal length and field of view for the image capture device. In some embodiments, image capture device 122 may be associated with a 6mm lens or a 12mm lens. In some embodiments, image capture device 122 may be configured to capture images having a desired field-of-view (FOV) 202, as illustrated in FIG. 2D . For example, image capture device 122 may be configured to have a regular FOV, such as within a range of 40 degrees to 56 degrees, including a 46 degree FOV, 50 degree FOV, 52 degree FOV, or greater. Alternatively, image capture device 122 may be configured to have a narrow FOV in the range of 23 to 40 degrees, such as a 28 degree FOV or 36 degree FOV. In addition, image capture device 122 may be configured to have a wide FOV in the range of 100 to 180 degrees. In some embodiments, image capture device 122 may include a wide angle bumper camera or one with up to a 180 degree FOV.

[0038] The first image capture device 122 may acquire a plurality of first images relative to a scene associated with the vehicle 200. Each of the plurality of first images may be acquired as a series of image scan lines, which may be captured using a rolling shutter. Each scan line may include a plurality of pixels.

[0039] The first image capture device 122 may have a scan rate associated with

Tesla EX1002 Page 1434 IPR2025-00944

acquisition of each of the first series of image scan lines. The scan rate may refer to a rate at which an image sensor can acquire image data associated with each pixel included in a particular scan line.

[0040] Image capture devices 122, 124, and 126 may contain any suitable type and number of image sensors, including CCD sensors or CMOS sensors, for example. In one embodiment, a CMOS image sensor may be employed along with a rolling shutter, such that each pixel in a row is read one at a time, and scanning of the rows proceeds on a row-by-row basis until an entire image frame has been captured. In some embodiments, the rows may be captured sequentially from top to bottom relative to the frame.

[0041] The use of a rolling shutter may result in pixels in different rows being exposed and captured at different times, which may cause skew and other image artifacts in the captured image frame. On the other hand, when the image capture device 122 is configured to operate with a global or synchronous shutter, all of the pixels may be exposed for the same amount of time and during a common exposure period. As a result, the image data in a frame collected from a system employing a global shutter represents a snapshot of the entire FOV (such as FOV 202) at a particular time. In contrast, in a rolling shutter application, each row in a frame is exposed and data is capture at different times. Thus, moving objects may appear distorted in an image capture device having a rolling shutter. This phenomenon will be described in greater detail below.

[0042] The second image capture device 124 and the third image capturing device 126 may be any type of image capture device. Like the first image capture device 122, each of image capture devices 124 and 126 may include an optical axis. In one embodiment, each of image capture devices 124 and 1 26 may include an Aptina M9V024 WVGA sensor with a global shutter. Alternatively, each of image capture devices 124 and 126 may include a rolling shutter. Like image capture device 122, image capture devices 124 and 126 may be configured to include various lenses and optical elements. In some embodiments, lenses associated with image capture devices 124 and 126 may provide FOVs (such as FOVs 204 and 206) that are the same as, or narrower than, a FOV (such as FOV 202) associated with image capture device 122. For example, image capture devices 124 and 126 may have FOVs of 40 degrees, 30 degrees, 26 degrees, 23 degrees, 20 degrees, or less.

[0043] Image capture devices 124 and 126 may acquire a plurality of second and third images relative to a scene associated with the vehicle 200. Each of the plurality of second and third images may be acquired as a second and third series of image scan lines, which may be captured using a rolling shutter. Each scan line or row may have a plurality of pixels. Image capture devices 124 and 126 may have second and third scan rates associated with acquisition of each of image scan lines included in the second and third series.

[0044] Each image capture device 122, 124, and 126 may be positioned at any suitable position and orientation relative to vehicle 200. The relative positioning of the image capture devices 122, 124, and 126 may be selected to aid in fusing together the information acquired from the image capture devices. For example, in some embodiments, a FOV (such as FOV 204) associated with image capture device 124 may overlap partially or fully with a FOV (such as FOV 202) associated with image capture device 122 and a FOV (such as FOV 206) associated with image capture

IPR2025-00944
Tesla EX1002 Page 1435

device 126.

[0045] Image capture devices 122, 124, and 126 may be located on vehicle 200 at any suitable relative heights. In one instance, there may be a height difference between the image capture devices 122, 124, and 126, which may provide sufficient parallax information to enable stereo analysis. For example, as shown in FIG. 2A , the two image capture devices 122 and 124 are at different heights. There may also be a lateral displacement difference between image capture devices 122, 124, and 126, giving additional parallax information for stereo analysis by processing unit 110, for example. The difference in the lateral displacement may be denoted by dx, as shown in FIGS. 2C and 2D . In some embodiments, fore or aft displacement (e.g., range displacement) may exist between image capture devices 122, 124, and 126. For example, image capture device 122 may be located 0.5 to 2 meters or more behind image capture device 124 and/or image capture device 126. This type of displacement may enable one of the image capture devices to cover potential blind spots of the other image capture device(s).

[0046] Image capture devices 122 may have any suitable resolution capability (e.g., number of pixels associated with the image sensor), and the resolution of the image sensor(s) associated with the image capture device 122 may be higher, lower, or the same as the resolution of the image sensor(s) associated with image capture devices 124 and 126. In some embodiments, the image sensor(s) associated with image capture device 122 and/or image capture devices 124 and 126 may have a resolution of 640 x 480, 1024 x 768, 1280 x 960, or any other suitable resolution.

[0047] The frame rate (e.g., the rate at which an image capture device acquires a set of pixel data of one image frame before moving on to capture pixel data associated with the next image frame) may be controllable. The frame rate associated with image capture device 122 may be higher, lower, or the same as the frame rate associated with image capture devices 124 and 126. The frame rate associated with image capture devices 122, 124, and 126 may depend on a variety of factors that may affect the timing of the frame rate. For example, one or more of image capture devices 122, 124, and 126 may include a selectable pixel delay period imposed before or after acquisition of image data associated with one or more pixels of an image sensor in image capture device 122, 124, and/or 126. Generally, image data corresponding to each pixel may be acquired according to a clock rate for the device (e.g., one pixel per clock cycle). Additionally, in embodiments including a rolling shutter, one or more of image capture devices 122, 124, and 126 may include a selectable horizontal blanking period imposed before or after acquisition of image data associated with a row of pixels of an image sensor in image capture device 122, 124, and/or 126. Further, one or more of image capture devices 122, 124, and/or 126 may include a selectable vertical blanking period imposed before or after acquisition of image data associated with an image frame of image capture device 122, 124, and 126.

[0048] These timing controls may enable synchronization of frame rates associated with image capture devices 122, 124, and 126, even where the line scan rates of each are different. Additionally, as will be discussed in greater detail below, these selectable timing controls, among other factors (e.g., image sensor resolution, maximum line scan rates, etc.) may enable synchronization of image capture from an area where the FOV of image capture device 122 overlaps with one or more FOVs of image capture devices 124 and 126, even where the field of view of image capture device 122 is different from the FOVs of image capture devices 124 and 126.

IPR2025-00944
Tesla EX1002 Page 1436

[0049] Frame rate timing in image capture device 122, 124, and 126 may depend on the resolution of the associated image sensors. For example, assuming similar line scan rates for both devices, if one device includes an image sensor having a resolution of 640 x 480 and another device includes an image sensor with a resolution of 1280 x 960, then more time will be required to acquire a frame of image data from the sensor having the higher resolution.

[0050] Another factor that may affect the timing of image data acquisition in image capture devices 122, 124, and 126 is the maximum line scan rate. For example, acquisition of a row of image data from an image sensor included in image capture device 122, 124, and 126 will require some minimum amount of time. Assuming no pixel delay periods are added, this minimum amount of time for acquisition of a row of image data will be related to the maximum line scan rate for a particular device. Devices that offer higher maximum line scan rates have the potential to provide higher frame rates than devices with lower maximum line scan rates. In some embodiments, one or more of image capture devices 124 and 126 may have a maximum line scan rate that is higher than a maximum line scan rate associated with image capture device 122. In some embodiments, the maximum line scan rate of image capture device 124 and/or 126 may be 1.25, 1.5, 1.75, or 2 times or more than a maximum line scan rate of image capture device 122.

[0051] In another embodiment, image capture devices 122, 124, and 126 may have the same maximum line scan rate, but image capture device 122 may be operated at a scan rate less than or equal to its maximum scan rate. The system may be configured such that one or more of image capture devices 124 and 126 operate at a line scan rate that is equal to the line scan rate of image capture device 122. In other instances, the system may be configured such that the line scan rate of image capture device 124 and/or image capture device 126 may be 1.25, 1.5, 1.75, or 2 times or more than the line scan rate of image capture device 122.

[0052] In some embodiments, image capture devices 122, 124, and 126 may be asymmetric. That is, they may include cameras having different fields of view (FOV) and focal lengths. The fields of view of image capture devices 122, 124, and 126 may include any desired area relative to an environment of vehicle 200, for example. In some embodiments, one or more of image capture devices 122, 124, and 126 may be configured to acquire image data from an environment in front of vehicle 200, behind vehicle 200, to the sides of vehicle 200, or combinations thereof.

[0053] Further, the focal length associated with each image capture device 122, 124, and/or 126 may be selectable (e.g., by inclusion of appropriate lenses etc.) such that each device acquires images of objects at a desired distance range relative to vehicle 200. For example, in some embodiments image capture devices 122, 124, and 126 may acquire images of close-up objects within a few meters from the vehicle. Image capture devices 122, 124, and 126 may also be configured to acquire images of objects at ranges more distant from the vehicle (e.g., 25 m, 50 m, 100 m, 150 m, or more). Further, the focal lengths of image capture devices 122, 124, and 126 may be selected such that one image capture device (e.g., image capture device 122) can acquire images of objects relatively close to the vehicle (e.g., within 10 m or within 20 m) while the other image capture devices (e.g., image capture devices 124 and 126) can acquire images of more distant objects (e.g., greater than 20 m, 50 m, 100 m, 150 m, etc.) from vehicle 200.

Tesla EX1002 Page 1437 IPR2025-00944

[0054] According to some embodiments, the FOV of one or more image capture devices 122, 124, and 126 may have a wide angle. For example, it may be advantageous to have a FOV of 140 degrees, especially for image capture devices 122, 124, and 126 that may be used to capture images of the area in the vicinity of vehicle 200. For example, image capture device 122 may be used to capture images of the area to the right or left of vehicle 200 and, in such embodiments, it may be desirable for image capture device 122 to have a wide FOV (e.g., at least 140 degrees).

[0055] The field of view associated with each of image capture devices 122, 124, and 126 may depend on the respective focal lengths. For example, as the focal length increases, the corresponding field of view decreases.

[0056] Image capture devices 122, 124, and 126 may be configured to have any suitable fields of view. In one particular example, image capture device 122 may have a horizontal FOV of 46 degrees, image capture device 124 may have a horizontal FOV of 23 degrees, and image capture device 126 may have a horizontal FOV in between 23 and 46 degrees. In another instance, image capture device 122 may have a horizontal FOV of 52 degrees, image capture device 124 may have a horizontal FOV of 26 degrees, and image capture device 126 may have a horizontal FOV in between 26 and 52 degrees. In some embodiments, a ratio of the FOV of image capture device 122 to the FOVs of image capture device 124 and/or image capture device 126 may vary from 1.5 to 2.0. In other embodiments, this ratio may vary between 1.25 and 2.25.

[0057] System 100 may be configured so that a field of view of image capture device 122 overlaps, at least partially or fully, with a field of view of image capture device 124 and/or image capture device 126. In some embodiments, system 100 may be configured such that the fields of view of image capture devices 124 and 126, for example, fall within (e.g., are narrower than) and share a common center with the field of view of image capture device 122. In other embodiments, the image capture devices 122, 124, and 126 may capture adjacent FOVs or may have partial overlap in their FOVs. In some embodiments, the fields of view of image capture devices 122, 124, and 126 may be aligned such that a center of the narrower FOV image capture devices 124 and/or 126 may be located in a lower half of the field of view of the wider FOV device 122.

[0058] FIG. 2F is a diagrammatic representation of exemplary vehicle control systems, consistent with the disclosed embodiments. As indicated in FIG. 2F , vehicle 200 may include throttling system 220, braking system 230, and steering system 240. System 100 may provide inputs (e.g., control signals) to one or more of throttling system 220, braking system 230, and steering system 240 over one or more data links (e.g., any wired and/or wireless link or links for transmitting data). For example, based on analysis of images acquired by image capture devices 122, 124, and/or 126, system 100 may provide control signals to one or more of throttling system 220, braking system 230, and steering system 240 to navigate vehicle 200 (e.g., by causing an acceleration, a turn, a lane shift, etc.). Further, system 100 may receive inputs from one or more of throttling system 220, braking system 230, and steering system 24 indicating operating conditions of vehicle 200 (e.g., speed, whether vehicle 200 is braking and/or turning, etc.). Further details are provided in connection with FIGS. 4-7 , below.

[0059] As shown in FIG. 3A , vehicle 200 may also include a user interface 170 for

interacting with a driver or a passenger of vehicle 200. For example, user interface 170 in a vehicle application may include a touch screen 320, knobs 330, buttons 340, and a microphone 350. A driver or passenger of vehicle 200 may also use handles (e.g., located on or near the steering column of vehicle 200 including, for example, turn signal handles), buttons (e.g., located on the steerin' wheel of vehicle 200), and the like, to interact with system 100. In some embodiments, microphone 350 may be positioned adjacent to a rearview mirror 310. Similarly, in some embodiments, image capture device 122 may be located near rearview mirror 310. In some embodiments, user interface 170 may also include one or more speakers 360 (e.g., speakers of a vehicle audio system). For example, system 100 may provide various notifications (e.g., alerts) via speakers 360.

[0060] FIGS. 3B-3D are illustrations of an exemplary camera mount 370 configured to be positioned behind a rearview mirror (e.g., rearview mirror 310) and against a vehicle windshield, consistent with disclosed embodiments. As shown in FIG. 3B , camera mount 370 may include image capture devices 122, 124, and 126. Image capture devices 124 and 126 may be positioned behind a glare shield 380, which may be flush against the vehicle windshield and include a composition of film and/or anti-reflective materials. For example, glare shield 380 may be positioned such that it aligns against a vehicle windshield having a matching slope. In some embodiments, each of image capture devices 122, 124, and 126 may be positioned behind glare shield 380, as depicted, for example, in FIG. 3D . The disclosed embodiments are not limited to any particular configuration of image capture devices 122, 124, and 126, camera mount 370, and glare shield 380. FIG. 3C is an illustration of camera mount 370 shown in FIG. 3B from a front perspective.

[0061] As will be appreciated by a person skilled in the art having the benefit of this disclosure, numerous variations and/or modifications may be made to the foregoing disclosed embodiments. For example, not all components are essential for the operation of system 100. Further, any component may be located in any appropriate part of system 100 and the components may be rearranged into a variety of configurations while providing the functionality of the disclosed embodiments. Therefore, the foregoing configurations are examples and, regardless of the configurations discussed above, system 100 can provide a wide range of functionality to analyze the surroundings of vehicle 200 and navigate vehicle 200 in response to the analysis.

[0062] As discussed below in further detail and consistent with various disclosed embodiments, system 100 may provide a variety of features related to autonomous driving and/or driver assist technology. For example, system 100 may analyze image data, position data (e.g., GPS location information), map data, speed data, and/or data from sensors included in vehicle 200. System 100 may collect the data for analysis from, for example, image acquisition unit 120, position sensor 130, and other sensors. Further, system 100 may analyze the collected data to determine whether or not vehicle 200 should take a certain action, and then automatically take the determined action without human intervention. For example, when vehicle 200 navigates without human intervention, system 100 may automatically control the braking, acceleration, and/or steering of vehicle 200 (e.g., by sending control signals to one or more of throttling system 220, braking system 230, and steering system 240). Further, system 100 may analyze the collected data and issue warnings and/or alerts to vehicle occupants based on the analysis of the collected data. Additional details regarding the various embodiments that are provided by system 100 are provided below.

IPR2025-00944
Tesla EX1002 Page 1439

Forward-Facing Multi-Imaging System

[0063] As discussed above, system 100 may provide drive assist functionality that uses a multi-camera system. The multi-camera system may use one or more cameras facing in the forward direction of a vehicle. In other embodiments, the multi-camera system may include one or more cameras facing to the side of a vehicle or to the rear of the vehicle. In one embodiment, for example, system 100 may use a two-camera imaging system, where a first camera and a second camera (e.g., image capture devices 122 and 124) may be positioned at the front and/or the sides of a vehicle (e.g., vehicle 200). The first camera may have a field of view that is greater than, less than, or partially overlapping with, the field of view of the second camera. In addition, the first camera may be connected to a first image processor to perform monocular image analysis of images provided by the first camera, and the second camera may be connected to a second image processor to perform monocular image analysis of images provided by the second camera. The outputs (e.g., processed information) of the first and second image processors may be combined. In some embodiments, the second image processor may receive images from both the first camera and second camera to perform stereo analysis. In another embodiment, system 100 may use a three-camera imaging system where each of the cameras has a different field of view. Such a system may, therefore, make decisions based on information derived from objects located at varying distances both forward and to the sides of the vehicle. References to monocular image analysis may refer to instances where image analysis is performed based on images captured from a single point of view (e.g., from a single camera). Stereo image analysis may refer to instances where image analysis is performed based on two or more images captured with one or more variations of an image capture parameter. For example, captured images suitable for performing stereo image analysis may include images captured: from two or more different positions, from different fields of view, using different focal lengths, along with parallax information, etc.

[0064] For example, in one embodiment, system 100 may implement a three camera configuration using image capture devices 122-126. In such a configuration, image capture device 122 may provide a narrow field of view (e.g., 34 degrees, or other values selected from a range of about 20 to 45 degrees, etc.), image capture device 124 may provide a wide field of view (e.g., 150 degrees or other values selected from a range of about 100 to about 180 degrees), and image capture device 126 may provide an intermediate field of view (e.g., 46 degrees or other values selected from a range of about 35 to about 60 degrees). In some embodiments, image capture device 126 may act as a main or primary camera. Image capture devices 122-126 may be positioned behind rearview mirror 310 and positioned substantially side-by-side (e.g., 6 cm apart). Further, in some embodiments, as discussed above, one or more of image capture devices 122-126 may be mounted behind glare shield 380 that is flush with the windshield of vehicle 200. Such shielding may act to minimize the impact of any reflections from inside the car on image capture devices 122-126.

[0065] In another embodiment, as discussed above in connection with FIGS. 3B and 3C , the wide field of view camera (e.g., image capture device 124 in the above example) may be mounted lower than the narrow and main field of view cameras (e.g., image devices 122 and 126 in the above example). This configuration may provide a free line of sight from the wide field of view camera. To reduce reflections, the cameras may be mounted close to the windshield of vehicle 200, and may include polarizers on the cameras to damp reflected light.

IPR2025-00944
Tesla EX1002 Page 1440

[0066] A three camera system may provide certain performance characteristics. For example, some embodiments may include an ability to validate the detection of objects by one camera based on detection results from another camera. In the three camera configuration discussed above, processing unit 110 may include, for example, three processing devices (e.g., three EyeQ series of processor chips, as discussed above), with each processing device dedicated to processing images captured by one or more of image capture devices 122-126.

[0067] In a three camera system, a first processing device may receive images from both the main camera and the narrow field of view camera, and perform vision processing of the narrow FOV camera to, for example, detect other vehicles, pedestrians, lane marks, traffic signs, traffic lights, and other road objects. Further, the first processing device may calculate a disparity of pixels between the images from the main camera and the narrow camera and create a 3D reconstruction of the environment of vehicle 200. The first processing device may then combine the 3D reconstruction with 3D map data or with 3D information calculated based on information from another camera.

[0068] The second processing device may receive images from main camera and perform vision processing to detect other vehicles, pedestrians, lane marks, traffic signs, traffic lights, and other road objects. Additionally, the second processing device may calculate a camera displacement and, based on the displacement, calculate a disparity of pixels between successive images and create a 3D reconstruction of the scene (e.g., a structure from motion). The second processing device may send the structure from motion based 3D reconstruction to the first processing device to be combined with the stereo 3D images.

[0069] The third processing device may receive images from the wide FOV camera and process the images to detect vehicles, pedestrians, lane marks, traffic signs, traffic lights, and other road objects. The third processing device may further execute additional processing instructions to analyze images to identify objects moving in the image, such as vehicles changing lanes, pedestrians, etc.

[0070] In some embodiments, having streams of image-based information captured and processed independently may provide an opportunity for providing redundancy in the system. Such redundancy may include, for example, using a first image capture device and the images processed from that device to validate and/or supplement information obtained by capturing and processing image information from at least a second image capture device.

[0071] In some embodiments, system 100 may use two image capture devices (e.g., image capture devices 122 and 124) in providing navigation assistance for vehicle 200 and use a third image capture device (e.g., image capture device 126) to provide redundancy and validate the analysis of data received from the other two image capture devices. For example, in such a configuration, image capture devices 122 and 124 may provide images for stereo analysis by system 100 for navigating vehicle 200, while image capture device 126 may provide images for monocular analysis by system 100 to provide redundancy and validation of information obtained based on images captured from image capture device 122 and/or image capture device 124. That is, image capture device 126 (and a corresponding processing device) may be considered to provide a redundant sub-system for providing a check on the analysis derived from

Tesla EX1002 Page 1441 IPR2025-00944

image capture devices 122 and 124 (e.g., to provide an automatic emergency braking (AEB) system).

[0072] One of skill in the art will recognize that the above camera configurations, camera placements, number of cameras, camera locations, etc., are examples only. These components and others described relative to the overall system may be assembled and used in a variety of different configurations without departing from the scope of the disclosed embodiments. Further details regarding usage of a multi-camera system to provide driver assist and/or autonomous vehicle functionality follow below.

[0073] FIG. 4 is an exemplary functional block diagram of memory 140 and/or 150, which may be stored/programmed with instructions for performing one or more operations consistent with the disclosed embodiments. Although the following refers to memory 140, one of skill in the art will recognize that instructions may be stored in memory 140 and/or 150.

[0074] As shown in FIG. 4 , memory 140 may store a monocular image analysis module 402, a stereo image analysis module 404, a velocity and acceleration module 406, and a navigational response module 408. The disclosed embodiments are not limited to any particular configuration of memory 140. Further, application processor 180 and/or image processor 190 may execute the instructions stored in any of modules 402-408 included in memory 140. One of skill in the art will understand that references in the following discussions to processing unit 110 may refer to application processor 180 and image processor 190 individually or collectively. Accordingly, steps of any of the following processes may be performed by one or more processing devices.

[0075] In one embodiment, monocular image analysis module 402 may store instructions (such as computer vision software) which, when executed by processing unit 110, performs monocular image analysis of a set of images acquired by one of image capture devices 122, 124, and 126. In some embodiments, processing unit 110 may combine information from a set of images with additional sensory information (e.g., information from radar) to perform the monocular image analysis. As described in connection with FIGS. 5A-5D below, monocular image analysis module 402 may include instructions for detecting a set of features within the set of images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, hazardous objects, and any other feature associated with an environment of a vehicle. Based on the analysis, system 100 (e.g., via processing unit 110) may cause one or more navigational responses in vehicle 200, such as a turn, a lane shift, a change in acceleration, and the like, as discussed below in connection with navigational response module 408.

[0076] In one embodiment, stereo image analysis module 404 may store instructions (such as computer vision software) which, when executed by processing unit 110, performs stereo image analysis of first and second sets of images acquired by a combination of image capture devices selected from any of image capture devices 122, 124, and 126. In some embodiments, processing unit 110 may combine information from the first and second sets of images with additional sensory information (e.g., information from radar) to perform the stereo image analysis. For example, stereo image analysis module 404 may include instructions for performing stereo image analysis based on a first set of images acquired by image capture device 124 and a second set of images acquired by image capture device 126. As described in connection with FIG. 6 below, stereo image analysis module 404 may include

IPR2025-00944
Tesla EX1002 Page 1442

instructions for detecting a set of features within the first and second sets of images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, hazardous objects, and the like. Based on the analysis, processing unit 110 may cause one or more navigational responses in vehicle 200, such as a turn, a lane shift, a change in acceleration, and the like, as discussed below in connection with navigational response module 408.

[0077] In one embodiment, velocity and acceleration module 406 may store software configured to analyze data received from one or more computing and electromechanical devices in vehicle 200 that are configured to cause a change in velocity and/or acceleration of vehicle 200. For example, processing unit 110 may execute instructions associated with velocity and acceleration module 406 to calculate a target speed for vehicle 200 based on data derived from execution of monocular image analysis module 402 and/or stereo image analysis module 404. Such data may include, for example, a target position, velocity, and/or acceleration, the position and/or speed of vehicle 200 relative to a nearby vehicle, pedestrian, or road object, position information for vehicle 200 relative to lane markings of the road, and the like. In addition, processing unit 110 may calculate a target speed for vehicle 200 based on sensory input (e.g., information from radar) and input from other systems of vehicle 200, such as throttling system 220, braking system 230, and/or steering system 240 of vehicle 200. Based on the calculated target speed, processing unit 110 may transmit electronic signals to throttling system 220, braking system 230, and/or steering system 240 of vehicle 200 to trigger a change in velocity and/or acceleration by, for example, physically depressing the brake or easing up off the accelerator of vehicle 200.

[0078] In one embodiment, navigational response module 408 may store software executable by processing unit 110 to determine a desired navigational response based on data derived from execution of monocular image analysis module 402 and/or stereo image analysis module 404. Such data may include position and speed information associated with nearby vehicles, pedestrians, and road objects, target position information for vehicle 200, and the like. Additionally, in some embodiments, the navigational response may be based (partially or fully) on map data, a predetermined position of vehicle 200, and/or a relative velocity or a relative acceleration between vehicle 200 and one or more objects detected from execution of monocular image analysis module 402 and/or stereo image analysis module 404. Navigational response module 408 may also determine a desired navigational response based on sensory input (e.g., information from radar) and inputs from other systems of vehicle 200, such as throttling system 220, braking system 230, and steering system 240 of vehicle 200. Based on the desired navigational response, processing unit 110 may transmit electronic signals to throttling system 220, braking system 230, and steering system 240 of vehicle 200 to trigger a desired navigational response by, for example, turning the steering wheel of vehicle 200 to achieve a rotation of a predetermined angle. In some embodiments, processing unit 110 may use the output of navigational response module 408 (e.g., the desired navigational response) as an input to execution of velocity and acceleration module 406 for calculating a change in speed of vehicle 200.

[0079] FIG. 5A is a flowchart showing an exemplary process 500A for causing one or more navigational responses based on monocular image analysis, consistent with disclosed embodiments. At step 510, processing unit 110 may receive a plurality of images via data interface 128 between processing unit 110 and image acquisition unit 120. For instance, a camera included in image acquisition unit 120 (such as image capture device 122 having field of view 202) may capture a plurality of images of an

Tesla EX1002 Page 1443 IPR2025-00944

area forward of vehicle 200 (or to the sides or rear of a vehicle, for example) and transmit them over a data connection (e.g., digital, wired, USB, wireless, Bluetooth, etc.) to processing unit 1 10. Processing unit 110 may execute monocular image analysis module 402 to analyze the plurality of images at step 520, as described in further detail in connection with FIGS. 5B-5D below. By performing the analysis, processing unit 1 10 may detect a set of features within the set of images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, and the like.

[0080] Processing 3unit 110 may also execute monocular image analysis module 402 to detect various road hazards at step 520, such as, for example, parts of a truck tire, fallen road signs, loose cargo, small animals, and the like. Road hazards may vary in structure, shape, size, and color, which may make detection of such hazards more challenging. In some embodiments, processing unit 110 may execute monocular image analysis module 402 to perform multi-frame analysis on the plurality of images to detect road hazards. For example, processing unit 110 may estimate camera motion between consecutive image frames and calculate the disparities in pixels between the frames to construct a 3D-map of the road. Processing unit 110 may then use the 3D-map to detect the road surface, as well as hazards existing above the road surface.

[0081] At step 530, processing unit 110 may execute navigational response module 408 to cause one or more navigational responses in vehicle 200 based on the analysis performed at step 520 and the techniques as described above in connection with FIG. 4 . Navigational responses may include, for example, a turn, a lane shift, a change in acceleration, and the like. In some embodiments, processing unit 110 may use data derived from execution of velocity and acceleration module 406 to cause the one or more navigational responses. Additionally, multiple navigational responses may occur simultaneously, in sequence, or any combination thereof. For instance, processing unit 110 may cause vehicle 200 to shift one lane over and then accelerate by, for example, sequentially transmitting control signals to steering system 240 and throttling system 220 of vehicle 200. Alternatively, processing unit 110 may cause vehicle 200 to brake while at the same time shifting lanes by, for example, simultaneously transmitting control signals to braking system 230 and steering system 240 of vehicle 200.

[0082] FIG. 5B is a flowchart showing an exemplary process 500B for detecting one or more vehicles and/or pedestrians in a set of images, consistent with disclosed embodiments. Processing unit 110 may execute monocular image analysis module 402 to implement process 500B. At step 540, processing unit 1 10 may determine a set of candidate objects representing possible vehicles and/or pedestrians. For example, processing unit 110 may scan one or more images, compare the images to one or more predetermined patterns, and identify within each image possible locations that may contain objects of interest (e.g., vehicles, pedestrians, or portions thereof). The predetermined patterns may be designed in such a way to achieve a high rate of "false hits" and a low rate of "misses." For example, processing unit 110 may use a low threshold of similarity to predetermined patterns for identifying candidate objects as possible vehicles or pedestrians. Doing so may allow processing unit 110 to reduce the probability of missing (e.g., not identifying) a candidate object representing a vehicle or pedestrian.

[0083] At step 542, processing unit 1 10 may filter the set of candidate objects to exclude certain candidates (e.g., irrelevant or less relevant objects) based on classification criteria. Such criteria may be derived from various properties associated

IPR2025-00944
Tesla EX1002 Page 1444

with object types stored in a database (e.g., a database stored in memory 140). Properties may include object shape, dimensions, texture, position (e.g., relative to vehicle 200), and the like. Thus, processing unit 110 may use one or more sets of criteria to reject false candidates from the set of candidate objects.

[0084] At step 544, processing unit 110 may analyze multiple frames of images to determine whether objects in the set of candidate objects represent vehicles and/or pedestrians. For example, processing unit 110 may track a detected candidate object across consecutive frames and accumulate frame-by-frame data associated with the detected object (e.g., size, position relative to vehicle 200, etc.). Additionally, processing unit 110 may estimate parameters for the detected object and compare the object's frame-by-frame position data to a predicted position.

[0085] At step 546, processing unit 110 may construct a set of measurements for the detected objects. Such measurements may include, for example, position, velocity, and acceleration values (relative to vehicle 200) associated with the detected objects. In some embodiments, processing unit 110 may construct the measurements based on estimation techniques using a series of time-based observations such as Kalman filters or linear quadratic estimation (LQE), and/or based on available modeling data for different object types (e.g., cars, trucks, pedestrians, bicycles, road signs, etc.). The Kalman filters may be based on a measurement of an object's scale, where the scale measurement is proportional to a time to collision (e.g., the amount of time for vehicle 200 to reach the object). Thus, by performing steps 540-546, processing unit 110 may identify vehicles and pedestrians appearing within the set of captured images and derive information (e.g., position, speed, size) associated with the vehicles and pedestrians. Based on the identification and the derived information, processing unit 110 may cause one or more navigational responses in vehicle 200, as described in connection with FIG. 5A , above.

[0086] At step 548, processing unit 110 may perform an optical flow analysis of one or more images to reduce the probabilities of detecting a "false hit" and missing a candidate object that represents a vehicle or pedestrian. The optical flow analysis may refer to, for example, analyzing motion patterns relative to vehicle 200 in the one or more images associated with other vehicles and pedestrians, and that are distinct from road surface motion. Processing unit 1 10 may calculate the motion of candidate objects by observing the different positions of the objects across multiple image frames, which are captured at different times. Processing unit 110 may use the position and time values as inputs into mathematical models for calculating the motion of the candidate objects. Thus, optical flow analysis may provide another method of detecting vehicles and pedestrians that are nearby vehicle 200. Processing unit 110 may perform optical flow analysis in combination with steps 540-546 to provide redundancy for detecting vehicles and pedestrians and increase the reliability of system 100.

[0087] FIG. 5C is a flowchart showing an exemplary process 500C for detecting road marks and/or lane geometry information in a set of images, consistent with disclosed embodiments. Processing unit 110 may execute monocular image analysis module 402 to implement process 500C. At step 550, processing unit 1 10 may detect a set of objects by scanning one or more images. To detect segments of lane markings, lane geometry information, and other pertinent road marks, processing unit 110 may filter the set of objects to exclude those determined to be irrelevant (e.g., minor potholes, small rocks, etc.). At step 552, processing unit 110 may group together the segments detected in step 550 belonging to the same road mark or lane mark. Based on the

grouping, processing unit 110 may develop a model to represent the detected segments, such as a mathematical model.

[0088] At step 554, processing unit 110 may construct a set of measurements associated with the detected segments. In some embodiments, processing unit 110 may create a projection of the detected segments from the image plane onto the real-world plane. The projection may be characterized using a 3rd-degree polynomial having coefficients corresponding to physical properties such as the position, slope, curvature, and curvature derivative of the detected road. In generating the projection, processing unit 110 may take into account changes in the road surface, as well as pitch and roll rates associated with vehicle 200. In addition, processing unit 110 may model the road elevation by analyzing position and motion cues present on the road surface. Further, processing unit 110 may estimate the pitch and roll rates associated with vehicle 200 by tracking a set of feature points in the one or more images.

[0089] At step 556, processing unit 110 may perform multi-frame analysis by, for example, tracking the detected segments across consecutive image frames and accumulating frame-by-frame data associated with detected segments. As processing unit 110 performs multi-frame analysis, the set of measurements constructed at step 554 may become more reliable and associated with an increasingly higher confidence level. Thus, by performing steps 550-556, processing unit 110 may identify road marks appearing within the set of captured images and derive lane geometry information. Based on the identification and the derived information, processing unit 110 may cause one or more navigational responses in vehicle 200, as described in connection with FIG. 5A , above.

[0090] At step 558, processing unit 110 may consider additional sources of information to further develop a safety model for vehicle 200 in the context of its surroundings. Processing unit 110 may use the safety model to define a context in which system 100 may execute autonomous control of vehicle 200 in a safe manner. To develop the safety model, in some embodiments, processing unit 110 may consider the position and motion of other vehicles, the detected road edges and barriers, and/or general road shape descriptions extracted from map data (such as data from map database 160). By considering additional sources of information, processing unit 110 may provide redundancy for detecting road marks and lane geometry and increase the reliability of system 100.

[0091] FIG. 5D is a flowchart showing an exemplary process 500D for detecting traffic lights in a set of images, consistent with disclosed embodiments. Processing unit 1 10 may execute monocular image analysis module 402 to implement process 500D. At step 560, processing unit 110 may scan the set of images and identify objects appearing at locations in the images likely to contain traffic lights. For example, processing unit 110 may filter the identified objects to construct a set of candidate objects, excluding those objects unlikely to correspond to traffic lights. The filtering may be done based on various properties associated with traffic lights, such as shape, dimensions, texture, position (e.g., relative to vehicle 200), and the like. Such properties may be based on multiple examples of traffic lights and traffic control signals and stored in a database. In some embodiments, processing unit 110 may perform multi-frame analysis on the set of candidate objects reflecting possible traffic lights. For example, processing unit 110 may track the candidate objects across consecutive image frames, estimate the real-world position of the candidate objects, and filter out those objects that are moving (which are unlikely to be traffic lights). In some

Tesla EX1002 Page 1446 IPR2025-00944

embodiments, processing unit 110 may perform color analysis on the candidate objects and identify the relative position of the detected colors appearing inside possible traffic lights.

[0092] At step 562, processing unit 110 may analyze the geometry of a junction. The analysis may be based on any combination of: (i) the number of lanes detected on either side of vehicle 200, (ii) markings (such as arrow marks) detected on the road, and (iii) descriptions of the junction extracted from map data (such as data from map database 160). Processing unit 110 may conduct the analysis using information derived from execution of monocular analysis module 402. In addition, Processing unit 110 may determine a correspondence between the traffic lights detected at step 560 and the lanes appearing near vehicle 200.

[0093] As vehicle 200 approaches the junction, at step 564, processing unit 110 may update the confidence level associated with the analyzed junction geometry and the detected traffic lights. For instance, the number of traffic lights estimated to appear at the junction as compared with the number actually appearing at the junction may impact the confidence level. Thus, based on the confidence level, processing unit 110 may delegate control to the driver of vehicle 200 in order to improve safety conditions. By performing steps 560-564, processing unit 110 may identify traffic lights appearing within the set of captured images and analyze junction geometry information. Based on the identification and the analysis, processing unit 1 10 may cause one or more navigational responses in vehicle 200, as described in connection with FIG. 5A , above.

[0094] FIG. 5E is a flowchart showing an exemplary process 500E for causing one or more navigational responses in vehicle 200 based on a vehicle path, consistent with the disclosed embodiments. At step 570, processing unit 110 may construct an initial vehicle path associated with vehicle 200. The vehicle path may be represented using a set of points expressed in coordinates (x, z), and the distance di between two points in the set of points may fall in the range of 1 to 5 meters. In one embodiment, processing unit 110 may construct the initial vehicle path using two polynomials, such as left and right road polynomials. Processing unit 110 may calculate the geometric midpoint between the two polynomials and offset each point included in the resultant vehicle path by a predetermined offset (e.g., a smart lane offset), if any (an offset of zero may correspond to travel in the middle of a lane). The offset may be in a direction perpendicular to a segment between any two points in the vehicle path. In another embodiment, processing unit 110 may use one polynomial and an estimated lane width to offset each point of the vehicle path by half the estimated lane width plus a predetermined offset (e.g., a smart lane offset).

[0095] At step 572, processing unit 110 may update the vehicle path constructed at step 570. Processing unit 1 10 may reconstruct the vehicle path constructed at step 570 using a higher resolution, such that the distance dk between two points in the set of points representing the vehicle path is less than the distance di described above. For example, the distance dk may fall in the range of 0.1 to 0.3 meters. Processing unit 110 may reconstruct the vehicle path using a parabolic spline algorithm, which may yield a cumulative distance vector S corresponding to the total length of the vehicle path (i.e., based on the set of points representing the vehicle path).

[0096] At step 574, processing unit 110 may determine a look-ahead point (expressed in coordinates as (xl, zl)) based on the updated vehicle path constructed at step 572. Processing unit 1 10 may extract the look-ahead point from the cumulative distance

vector S, and the look-ahead point may be associated with a look-ahead distance and look-ahead time. The look-ahead distance, which may have a lower bound ranging from 1 0 to 20 meters, may be calculated as the product of the speed of vehicle 200 and the look-ahead time. For example, as the speed of vehicle 200 decreases, the look-ahead distance may also decrease (e.g., until it reaches the lower bound). The look-ahead time, which may range from 0.5 to 1.5 seconds, may be inversely proportional to the gain of one or more control loops associated with causing a navigational response in vehicle 200, such as the heading error tracking control loop. For example, the gain of the heading error tracking control loop may depend on the bandwidth of a yaw rate loop, a steering actuator loop, car lateral dynamics, and the like. Thus, the higher the gain of the heading error tracking control loop, the lower the look-ahead time.

[0097] At step 576, processing unit 110 may determine a heading error and yaw rate command based on the look-ahead point determined at step 574. Processing unit 110 may determine the heading error by calculating the arctangent of the look-ahead point, e.g., arctan (xl/zl). Processing unit 110 may determine the yaw rate command as the product of the heading error and a high-level control gain. The high-level control gain may be equal to: (2 / look-ahead time), if the look-ahead distance is not at the lower bound. Otherwise, the high-level control gain may be equal to: (2 <*> speed of vehicle 200 / look-ahead distance).

[0098] FIG. 5F is a flowchart showing an exemplary process 500F for determining whether a leading vehicle is changing lanes, consistent with the disclosed embodiments. At step 580, processing unit 110 may determine navigation information associated with a leading vehicle (e.g., a vehicle traveling ahead of vehicle 200). For example, processing unit 110 may determine the position, velocity (e.g., direction and speed), and/or acceleration of the leading vehicle, using the techniques described in connection with FIGS. 5A and 5B , above. Processing unit 110 may also determine one or more road polynomials, a look-ahead point (associated with vehicle 200), and/or a snail trail (e.g., a set of points describing a path taken by the leading vehicle), using the techniques described in connection with FIG. 5E , above.

[0099] At step 582, processing unit 110 may analyze the navigation information determined at step 580. In one embodiment, processing unit 110 may calculate the distance between a snail trail and a road polynomial (e.g., along the trail). If the variance of this distance along the trail exceeds a predetermined threshold (for example, 0.1 to 0.2 meters on a straight road, 0.3 to 0.4 meters on a moderately curvy road, and 0.5 to 0.6 meters on a road with sharp curves), processing unit 110 may determine that the leading vehicle is likely changing lanes. In the case where multiple vehicles are detected traveling ahead of vehicle 200, processing unit 110 may compare the snail trails associated with each vehicle. Based on the comparison, processing unit 110 may determine that a vehicle whose snail trail does not match with the snail trails of the other vehicles is likely changing lanes. Processing unit 110 may additionally compare the curvature of the snail trail (associated with the leading vehicle) with the expected curvature of the road segment in which the leading vehicle is traveling. The expected curvature may be extracted from map data (e.g., data from map database 160), from road polynomials, from other vehicles' snail trails, from prior knowledge about the road, and the like. If the difference in curvature of the snail trail and the expected curvature of the road segment exceeds a predetermined threshold, processing unit 110 may determine that the leading vehicle is likely changing lanes.

IPR2025-00944
Tesla EX1002 Page 1448

[0100] In another embodiment, processing unit 110 may compare the leading vehicle's instantaneous position with the look-ahead point (associated with vehicle 200) over a specific period of time (e.g., 0.5 to 1.5 seconds). If the distance between the leading vehicle's instantaneous position and the look-ahead point varies during the specific period of time, and the cumulative sum of variation exceeds a predetermined threshold (for example, 0.3 to 0.4 meters on a straight road, 0.7 to 0.8 meters on a moderately curvy road, and 1.3 to 1.7 meters on a road with sharp curves), processing unit 110 may determine that the leading vehicle is likely changing lanes. In another embodiment, processing unit 110 may analyze the geometry of the snail trail by comparing the lateral distance traveled along the trail with the expected curvature of the snail trail. The expected radius of curvature may be determined according to the calculation: (δz<2> + δx<2>) / 2 / (δx), where δx represents the lateral distance traveled and δz represents the longitudinal distance traveled. If the difference between the lateral distance traveled and the expected curvature exceeds a predetermined threshold (e.g., 500 to 700 meters), processing unit 110 may determine that the leading vehicle is likely changing lanes. In another embodiment, processing unit 110 may analyze the position of the leading vehicle. If the position of the leading vehicle obscures a road polynomial (e.g., the leading vehicle is overlaid on top of the road polynomial), then processing unit 110 may determine that the leading vehicle is likely changing lanes. In the case where the position of the leading vehicle is such that, another vehicle is detected ahead of the leading vehicle and the snail trails of the two vehicles are not parallel, processing unit 110 may determine that the (closer) leading vehicle is likely changing lanes.

[0101] At step 584, processing unit 110 may determine whether or not leading vehicle 200 is changing lanes based on the analysis performed at step 582. For example, processing unit 110 may make the determination based on a weighted average of the individual analyses performed at step 582. Under such a scheme, for example, a decision by processing unit 110 that the leading vehicle is likely changing lanes based on a particular type of analysis may be assigned a value of "1" (and "0" to represent a determination that the leading vehicle is not likely changing lanes). Different analyses performed at step 582 may be assigned different weights, and the disclosed embodiments are not limited to any particular combination of analyses and weights.

[0102] FIG. 6 is a flowchart showing an exemplary process 600 for causing one or more navigational responses based on stereo image analysis, consistent with disclosed embodiments. At step 610, processing unit 110 may receive a first and second plurality of images via data interface 128. For example, cameras included in image acquisition unit 120 (such as image capture devices 122 and 124 having fields of view 202 and 204) may capture a first and second plurality of images of an area forward of vehicle 200 and transmit them over a digital connection (e.g., USB, wireless, Bluetooth, etc.) to processing unit 110. In some embodiments, processing unit 110 may receive the first and second plurality of images via two or more data interfaces. The disclosed embodiments are not limited to any particular data interface configurations or protocols.

[0103] At step 620, processing unit 110 may execute stereo image analysis module 404 to perform stereo image analysis of the first and second plurality of images to create a 3D map of the road in front of the vehicle and detect features within the images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, road hazards, and the like. Stereo image analysis may be performed in a manner similar to the steps described in connection with FIGS. 5A-5D , above. For example, processing unit 110 may execute stereo image analysis module 404 to detect

Tesla EX1002 Page 1449 IPR2025-00944

candidate objects (e.g., vehicles, pedestrians, road marks, traffic lights, road hazards, etc.) within the first and second plurality of images, filter out a subset of the candidate objects based on various criteria, and perform multi-frame analysis, construct measurements, and determine a confidence level for the remaining candidate objects. In performing the steps above, processing unit 110 may consider information from both the first and second plurality of images, rather than information from one set of images alone. For example, processing unit 110 may analyze the differences in pixel-level data (or other data subsets from among the two streams of captured images) for a candidate object appearing in both the first and second plurality of images. As another example, processing unit 110 may estimate a position and/or velocity of a candidate object (e.g., relative to vehicle 200) by observing that the object appears in one of the plurality of images but not the other or relative to other differences that may exist relative to objects appearing if the two image streams. For example, position, velocity, and/or acceleration relative to vehicle 200 may be determined based on trajectories, positions, movement characteristics, etc. of features associated with an object appearing in one or both of the image streams.

[0104] At step 630, processing unit 110 may execute navigational response module 408 to cause one or more navigational responses in vehicle 200 based on the analysis performed at step 620 and the techniques as described above in connection with FIG. 4 . Navigational responses may include, for example, a turn, a lane shift, a change in acceleration, a change in velocity, braking, and the like. In some embodiments, processing unit 1 10 may use data derived from execution of velocity and acceleration module 406 to cause the one or more navigational responses. Additionally, multiple navigational responses may occur simultaneously, in sequence, or any combination thereof.

[0105] FIG. 7 is a flowchart showing an exemplary process 700 for causing one or more navigational responses based on an analysis of three sets of images, consistent with disclosed embodiments. At step 710, processing unit 110 may receive a first, second, and third plurality of images via data interface 128. For instance, cameras included in image acquisition unit 120 (such as image capture devices 122, 124, and 126 having fields of view 202, 204, and 206) may capture a first, second, and third plurality of images of an area forward and/or to the side of vehicle 200 and transmit them over a digital connection (e.g., USB, wireless, Bluetooth, etc.) to processing unit 110. In some embodiments, processing unit 110 may receive the first, second, and third plurality of images via three or more data interfaces. For example, each of image capture devices 122, 124, 126 may have an associated data interface for communicating data to processing unit 110. The disclosed embodiments are not limited to any particular data interface configurations or protocols.

[0106] At step 720, processing unit 110 may analyze the first, second, and third plurality of images to detect features within the images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, road hazards, and the like. The analysis may be performed in a manner similar to the steps described in connection with FIGS. 5A-5D and 6 , above. For instance, processing unit 110 may perform monocular image analysis (e.g., via execution of monocular image analysis module 402 and based on the steps described in connection with FIGS. 5A-5D , above) on each of the first, second, and third plurality of images. Alternatively, processing unit 110 may perform stereo image analysis (e.g., via execution of stereo image analysis module 404 and based on the steps described in connection with FIG. 6 , above) on the first and second plurality of images, the second and third plurality of images, and/or

IPR2025-00944
Tesla EX1002 Page 1450

the first and third plurality of images. The processed information corresponding to the analysis of the first, second, and/or third plurality of images may be combined. In some embodiments, processing unit 110 may perform a combination of monocular and stereo image analyses. For example, processing unit 110 may perform monocular image analysis (e.g., via execution of monocular image analysis module 402) on the first plurality of images and stereo image analysis (e.g., via execution of stereo image analysis module 404) on the second and third plurality of images. The configuration of image capture devices 122, 124, and 126-including their respective locations and fields of view 202, 204, and 206-may influence the types of analyses conducted on the first, second, and third plurality of images. The disclosed embodiments are not limited to a particular configuration of image capture devices 122, 124, and 126, or the types of analyses conducted on the first, second, and third plurality of images.

[0107] In some embodiments, processing unit 110 may perform testing on system 100 based on the images acquired and analyzed at steps 710 and 720. Such testing may provide an indicator of the overall performance of system 100 for certain configurations of image capture devices 122, 124, and 126. For example, processing unit 1 10 may determine the proportion of "false hits" (e.g., cases where system 100 incorrectly determined the presence of a vehicle or pedestrian) and "misses."

[0108] At step 730, processing unit 110 may cause one or more navigational responses in vehicle 200 based on information derived from two of the first, second, and third plurality of images. Selection of two of the first, second, and third plurality of images may depend on various factors, such as, for example, the number, types, and sizes of objects detected in each of the plurality of images. Processing unit 110 may also make the selection based on image quality and resolution, the effective field of view reflected in the images, the number of captured frames, the extent to which one or more objects of interest actually appear in the frames (e.g., the percentage of frames in which an object appears, the proportion of the object that appears in each such frame, etc.), and the like.

[0109] In some embodiments, processing unit 110 may select information derived from two of the first, second, and third plurality of images by determining the extent to which information derived from one image source is consistent with information derived from other image sources. For example, processing unit 1 10 may combine the processed information derived from each of image capture devices 122, 124, and 126 (whether by monocular analysis, stereo analysis, or any combination of the two) and determine visual indicators (e.g., lane markings, a detected vehicle and its location and/or path, a detected traffic light, etc.) that are consistent across the images captured from each of image capture devices 122, 124, and 126. Processing unit 110 may also exclude information that is inconsistent across the captured images (e.g., a vehicle changing lanes, a lane model indicating a vehicle that is too close to vehicle 200, etc.). Thus, processing unit 110 may select information derived from two of the first, second, and third plurality of images based on the determinations of consistent and inconsistent information.

[0110] Navigational responses may include, for example, a turn, a lane shift, a change in acceleration, and the like. Processing unit 1 10 may cause the one or more navigational responses based on the analysis performed at step 720 and the techniques as described above in connection with FIG. 4 . Processing unit 110 may also use data derived from execution of velocity and acceleration module 406 to cause the one or more navigational responses. In some embodiments, processing unit 110

IPR2025-00944
Tesla EX1002 Page 1451

may cause the one or more navigational responses based on a relative position, relative velocity, and/or relative acceleration between vehicle 200 and an object detected within any of the first, second, and third plurality of images. Multiple navigational responses may occur simultaneously, in sequence, or any combination thereof.

Object Detecting and Braking System

[0111] In some embodiments, vehicle 200 may include an object detecting and braking system configured to detect an object in front of vehicle 200 and cause vehicle 200 to brake, e.g., using braking system 230, based on the detection of the object. For example, the object detecting and braking system may cause vehicle 200 to brake automatically, or autonomously, without or with limited driver intervention. The automatic or autonomous object detecting and braking system may or may not allow a driver to intervene to take over control of braking system 230.

[0112] The object detecting and braking system may include a plurality of other systems discussed above. For example, the object detecting and braking system may include system 100, braking system 230, and one or more of image capture devices 122, 124, and 126. The one or more of image capture devices 122, 124, and 126 may acquire one or more images of an environment or an area in front of vehicle 200, which may include an object, such as another vehicle, a pedestrian, a road surface, etc. Processing unit 110 may process the images based on at least two independent analyses to determine at least two times-to-collision (TTCs) of vehicle 200 with the object. In some embodiments, the two TTCs of vehicle 200 with the object may be determined separately and independently of each other. A "time-to-collision" (TTC) may also be referred to as a "time-to-contact." The time-to-collision is an estimated time that would lapse before vehicle 200 collides with another object. Processing unit 110 may determine that a difference between the two times-to-collision does not exceed a predetermined threshold. Based on this determination, processing unit 110 may cause vehicle 200 to brake using, e.g., braking system 230. The automatic or autonomous braking may assist the driver to avoid collision with the object in front of vehicle 200, thereby improving driving safety. By causing vehicle 200 to brake when the difference between the two TTCs does not exceed the predetermined threshold, false braking may be prevented or reduced.

[0113] In some embodiments, the object detecting and braking system may use two independent channels of information inherent to monocular video sensing. The first channel of information may be referred to as a "texture" or "appearance information" channel. System 100 may perform an analysis based on the texture or appearance information. In this channel, system 100 may perform pattern recognition over the pixels intensities in each image separately among different images (e.g., two sequential images or frames). System 100 may estimate the change of a footprint size (e.g., a dimension length of a feature of an object in the image) by comparing image data of the two sequential images. The second channel of information may be referred to as a "structural" or "optical flow information" channel. System 100 may perform an analysis based on the structural or optical flow information. For example, system 100 may process two or more images (e.g., sequentially acquired images or frames) to obtain an optical flow field. The optical flow field may provide information regarding the movement of trackable textures between the two or more images or frames. The analysis of the optical flow field may indicate a low time-to-collision upright object in the images. The two analyses may also be referred to as two modalities, methods, or

IPR2025-00944
Tesla EX1002 Page 1452

IPR2025-00944
Tesla EX1002 Page 1453

models.

[0114] In some embodiments, the structure (or optical flow) based analysis may serve to validate a result of a texture based analysis. For example, the texture based analysis may be performed on a plurality of images (e.g., at least two images), and the structure based analysis may be performed based on the same two or more images used in the texture based analysis after the texture based analysis is performed. In such embodiments, the structure based analysis may validate a result of the texture based analysis. In some embodiments, the texture based analysis may be performed after the structure based analysis has been performed, and may validate a result of the structure based analysis.

[0115] In some embodiments, the texture based analysis may use a first group of at least two images, and the structure based analysis may use a second group of at least two images. The first group and the second group may be the same or different. In some embodiments, the first group of at least two images may be selected from odd numbered images among a plurality of images acquired by at least one of image capture devices 122-126, and the second group of at least two images may be selected from even numbered images among the same plurality of images acquired by at least of image capture devices 122-126.

[0116] In some embodiments, the first group of at least two images may be selected from a plurality of images that are acquired at a first time interval and the second group of at least two images may be selected from the same plurality of images that are acquired at a second time interval, which may be different from the first time interval. In some embodiments, a first image capture device (or a first image sensor) may be designated for capturing images for the texture based analysis, and a second image capture device (or a second image sensor) may be designated for capturing images for structure based analysis.

[0117] The texture based analysis and the structure based analysis may complement each other, and the combination of their results that may be used by processing unit 110 to make a braking decision. For example, the texture based analysis may result in a first estimate of time-to-collision, and the structure based analysis may also result in a second estimate of the time-to-collision. Processing unit 110 may cause braking system 230 to brake vehicle 200 when the difference between the first and second estimates is smaller than a threshold. One of the texture based analysis and the structure based analysis may provide a redundancy for the other, thereby improve accuracy of braking, for example, autonomous braking in emergency driving situation, by preventing or reducing the occurrence of false braking.

[0118] In some embodiments, when the structure based analysis serves as a validator of the texture based analysis, the texture based analysis may detect an object (e.g., a car or a pedestrian) in front of vehicle 200, and may determine that the time-to-collision is low (e.g., lower than a predetermined threshold). The texture based analysis may reach this determination by calculating the time-to-collision based on a change of a footprint size through time in at least two sequentially acquired images. This determination from the texture based analysis may then trigger the structure based analysis. In some embodiments, a determination that the time-to-collision is low from the texture based analysis may serve as a precondition for releasing a braking decision (e.g., an autonomous emergency braking decision) to vehicle systems. In some embodiments, braking (e.g., autonomous emergency braking) may not be performed

until the structure based analysis confirms the determination made via the texture based analysis.

[0119] In some embodiments, system 100 may require that the estimated time-to-collision (TTC) determined by the structure based analysis is identical to or substantially identical to (up to an error tolerance) to the estimated time-to-collision (TTC) determined by the texture based analysis. For example, as discussed below, processing unit 10 may determine that a difference between the two TTCs determined by the texture based analysis and the structure based analysis does not exceed a predetermined threshold, which may be the error tolerance. In some embodiments, system 100 may take into account the divergence of the object's focus of expansion from that of the static environment. If the focus of expansion of the object diverges (e.g., moves) away from the focus of expansion of the static environment, this may be a strong indication that the object is a low TTC upright object. An example is discussed in greater detail with respect to FIGS. 15A and 15B . This requirement is optional because it may only apply in the case where the target object has a velocity component that is perpendicular to the driving direction of vehicle 200 (e.g., target object traveling in a direction that is not parallel to the driving direction of vehicle 200).

[0120] In some embodiments, processing unit 110 may use certain criteria for deciding whether structure based analysis supports texture based analysis in making a braking (e.g., autonomous emergency braking) decision. For example, such criteria may include a requirement that the optical flow fits a model of low time-to-collision upright object (LTU object), and not an open-road (or an open road surface).

[0121] FIG. 8 is a diagrammatic representation of an exemplary vehicle 200 including an object detecting and braking system (or referred to as an emergency braking decisioning system). The object detecting and braking system may include system 100 and braking system 230. The object detecting and braking system may also include at least one of image capture devices 122 and 124 (and 126, which is not shown in FIG. 8). The object detecting and braking system may include processing unit 110, which may be part of system 100. At least one of image capture devices 122 and 1 24 may acquire one or a plurality of images of an area including an object in front of vehicle 200. The object may be another vehicle, such as a vehicle 800 that may be traveling in front of vehicle 200. As shown in FIG. 8 , in some embodiments, vehicle 800 may be traveling in a direction that is parallel to a traveling direction of vehicle 200. A parallel direction may include a situation in which vehicle 800 is traveling in the same direction as and in alignment or substantially in alignment with vehicle 200 (e.g , vehicle 800 traveling in front of vehicle 200), as shown in FIG. 8 . A parallel direction may also include a situation in which vehicle 800 is traveling in an opposite direction relative to vehicle 200 and in alignment with vehicle 200 (e.g., vehicle 800 moving toward vehicle 200 in a direction that is 180 degrees, or substantially 180 degrees, opposite to the traveling direction of vehicle 200). A parallel direction may also include a situation in which vehicle 800 is not traveling in alignment with vehicle 200, but is traveling in a direction parallel to the traveling direction of vehicle 200 (e.g., vehicle 800 traveling on a front side of vehicle 200 at an offset distance).

[0122] In some embodiments, the object may be a pedestrian who may be standing in the road, or walking or running across the road. The area included in the images may include other objects, such as a road surface, a house, a wall, a sign, etc., which may form a static environment. When an object is not traveling in a direction parallel to the traveling direction of vehicle 200, the traveling direction of the object may transverse

Tesla EX1002 Page 1454 IPR2025-00944

the traveling direction of vehicle 200 with an angle that is neither zero (or substantially zero) nor 180 degrees (or substantially 180 degrees). The images may include both moving and/or static objects.

[0123] Processing unit 1 1 0 may communicate with braking system 230 and at least one of image capture devices 122 and 124. For example, processing unit 110 may receive images from at least one of image capture devices 122 and 124, and may process the images using various methods or modalities to detect one or more times-to-collision of vehicle 200 with object 800. Although two image capture devices 122 and 124 are shown in FIG. 8 , the use of additional image capture devices (e.g., 3, 4, 5, etc., image capture devices) is consistent with the disclosed embodiments.

[0124] FIG. 9 is a diagrammatic representation of an exemplary vehicle 200 including an object detecting and braking system that may be configured to detect an object not traveling in a direction parallel to the traveling direction of vehicle 200. The object may be a vehicle 900 or a pedestrian (not shown). The object detecting and braking system may detect vehicle 900 traveling toward vehicle 200 based on one or more images acquired by at least one of image capture devices 122 and 1 24. In addition, the object detecting and braking system may determine at least two times-to-collision of vehicle 200 with vehicle 900 based on the at least one or more images. When a difference between the at least two times-to-collision does not exceed the predetermined threshold, processing unit 110 may cause vehicle 200 to brake in order to prevent collision with vehicle 900.

[0125] FIG. 10 is an exemplary block diagram of memory 140 or 150 that may store instructions for performing one or more operations for detecting an object in front of vehicle 200 and causing vehicle 200 to brake based on the detection. As shown in FIG. 10 , memory 140 or 150 may store one or more modules for performing the operations, as described herein. For example, memory 140 or 150 may store a first image analysis module 1002, a second image analysis module 1003, and a braking module 1004. Application processor 180 and/or image processor 190 may execute the instructions stored in any of modules 1002-1004 included in memory 140 or 150. One of skill in the art will understand that references in the following discussions to processing unit 110 may refer to application processor 180 and image processor 190 individually or collectively. Accordingly, steps of any of the following processes may be performed by one or more processing devices.

[0126] First image analysis module 1002 may store instructions which, when executed by processing unit 1 10, may perform a first image analysis to calculate or generate a first estimated time-to-collision (TTC1) of vehicle 200 with an object (e.g., vehicle 800 or 900). First image analysis module 1002 may process at least two images acquired by at least one of image capture devices 122 and 124 using, for example, an appearance based detection method, which may calculate TTC1 based on a change in the appearance (e.g., a size) of the object in the at least two images. An exemplary first image analysis is discussed below in greater detail with respect to FIG. 12 .

[0127] Second image analysis module 1002 may store instructions which, when executed by processing unit 110, may perform a second image analysis to calculate or generate a second estimated time-to-collision (TTC2) of vehicle 200 with the object (e.g., vehicle 800 or 900). Second image analysis module 1003 may process at least two images acquired by at least one of image capture devices 122 and 124 using, for example, an optical flow based method, which may calculate TTC2 based on optical

IPR2025-00944
Tesla EX1002 Page 1455

flow information derived from the at least two images. An exemplary second image analysis is discussed below in greater detail with respect to FIG. 13 .

[0128] Braking module 1004 may store instructions which, when executed by processing unit 110, may cause vehicle 200 to brake using braking system 230. For example, when processing unit 110 determines that a difference between TTC1 and TTC2 does not exceed the predetermined threshold, braking module 1004 may cause braking system 230 to apply braking to reduce the speed of vehicle 200 or stop vehicle 200, thereby preventing collision with the object.

[0129] FIG. 11 is a flowchart showing an exemplary process or method 1100 for detecting an object in front of vehicle 200 and causing vehicle 200 to brake based on the detection. Method 1100 may include acquiring a plurality of images of an area in front of vehicle 200 (step 1110). For example, at least one of image capture devices 122 and 124 may acquire one or more images of an area in front of vehicle 200. The area may include an object, such as another vehicle or a pedestrian. The area may include other static objects, such as a sign, a house, a mountain, a road surface, etc., which may form a static environment.

[0130] Method 1100 may include performing the first image analysis based on at least a first image and a second image to generate a first time-to-collision (TTC1) (step 1120). For example, processing unit 110 may perform the first image analysis based on at least the first image and the second image. More than two images may be used in the first image analysis to generate TTC1. In some embodiments, the first image analysis may include an appearance based detection method. The appearance based detection may detect TTC1 based on a change in appearance (e.g., a change in size) of an object in the first image and the second image. The appearance based detection technique is further discussed below in connection with FIG. 12 .

[0131] FIG. 12 is a flowchart showing an exemplary first image analysis 1200 for determining the first estimated time-to-collision (TTC1). The first image analysis 1200 may include the appearance or texture based detection method (hence, the first image analysis may be referred to as the texture based analysis or textural analysis). The appearance based detection method may include determining a change in the appearance of an object or a feature of the object in at least two images (e.g., the first image and the second image), and may calculate TTC1 based on the change in the appearance. In some embodiments, the change in the appearance may be a change in a size of the object (or the feature of the object).

[0132] In the exemplary first image analysis 1200 shown in FIG. 12 , processing unit 1 10 may determine a first size of an object (or a feature of the object) appearing in the first image (step 1210). Processing unit 110 may determine a second size of the object (or the feature of the object) appearing in the second image (step 1220). The second size may be larger than the first size as vehicle 200 approaches the object. Processing unit 110 may determine the first size and the second size based on the dimensions (e.g., pixel sizes) of the object appearing in the first image and the second image. Processing unit 110 may determine a change in size (or relative scale) of the object (or the feature of the object) (step 1230). For example, processing unit 110 may subtract the first size from the second size to determine a difference in size. In some embodiments, processing unit 110 may divide the difference in size by the first size to determine a relative difference in size. The change in size (or relative scale) may refer to the relative difference in size, which may be a percentage. Processing unit 110 may

IPR2025-00944
Tesla EX1002 Page 1456

determine TTC1 based on the change in size and a first predetermined time interval between times of acquisition of the first image and the second image (step 1240). The first predetermined time interval may depend on settings of the camera (e.g., the camera included in the first image capture device 122) that acquires the first image and the second image, and may be known once the camera settings are fixed. The first predetermined time interval may be any suitable interval, such as, for example, 0.05 second, 0.1 second, 0.2 second, etc. For example, the first predetermined time interval may be represented by dtl, the relative change in size may be represented by S0, and processing unit 110 may calculate TTC1 from TTC1 = dt1 / S0.

[0133] Referring back to FIG. 11 , method 1100 may include performing a second image analysis based on at least a third image and a fourth image to generate a second estimated time-to-collision (TTC2) (step 1130). For example, processing unit 110 may derive optical flow information from the third image and the fourth image, and may calculate TTC2 based on the optical flow information. The second image analysis may be referred to as the structure based analysis or structural analysis. In some embodiments, more than two images may be used in the second image analysis to derive the optical flow information.

[0134] FIG. 13 is a flowchart showing an exemplary second image analysis 1300 for determining the second estimated time-to-collision (TTC2). The second image analysis 1300 may include obtaining or deriving an optical flow field from at least the third image and the fourth image (step 1310). In some embodiments, processing unit 110 may derive the optical flow field from the third image and the fourth image, and/or additional images acquired by at least one of image capture devices 122 and 124. The optical flow field may be a vector field in which each point may be represented by a vector showing a magnitude and a direction of movement from a first point in, e.g., the third image, to a corresponding second point in, e.g., the fourth image.

[0135] As shown in FIG. 13 , the second image analysis 1300 may include determining a focus of expansion based on the optical flow field (step 1320). The focus of expansion may be a converging point of vectors in the optical flow field, which is discussed below in greater detail with respect to FIG. 16 . The second image analysis 1300 may also include determining a rate of expansion relative to the focus of expansion for each point in the optical flow field (step 1330). For example, for each point in the optical flow field, processing unit 110 may calculate the rate of expansion based on a magnitude of movement of the point, and a distance from the point to the focus of expansion. For example, the magnitude of movement of any point may be represented by d, the distance from the point to the focus of expansion may be represented by r, and processing unit 110 may calculate the rate of expansion from Rexp = d / r. For each point in the optical flow field, the magnitude of movement d may be calculated from the magnitude of the vector associated with that point. The magnitude of movement r may indicate how much a first point in, e.g., the third image, has moved to reach a second point in, e.g., the fourth image.

[0136] As shown in FIG. 13 , the second image analysis 1300 may include determining the second estimated time-to-collision (TTC2) based on the rate of expansion and a second predetermined time interval. The second predetermined time interval may be the difference between times of acquisition of the third image and the fourth image. The second predetermined time interval dt2 may depend on the settings (e.g., the frame rate) of the camera, and may or may not be the same as the first predetermined time interval used in the first image analysis. The second predetermined time interval may

IPR2025-00944
Tesla EX1002 Page 1457

be any suitable interval, such as, for example, 0.05 second, 0.1 second, 0.2 second, etc. In some embodiments, TTC2 may be inversely proportional to the rate of expansion d/r, and may be directly proportional to the second predetermined time interval dt2. For example, processing unit 110 may calculate TTC2 from TTC2 = dt2 / (d/r).

[0137] The first image analysis 1200 and second image analysis 1300 shown in FIGS. 12 and 13 may be performed independently and separately. The first image and second image used in the first image analysis 1200 may be the same as or different from the third image and fourth image used in the second image analysis 1300. In some embodiments, the first image may be the same as the third image, and the second image may be the same as the fourth image. In some embodiments, the first image may be different from the third image, and the second image may be different from the fourth image. In some embodiments, images of alternating sequences may be used for the first and second analyses. For example, from the same set of images, odd numbered images may be used for the first image analysis, and even numbered images may be used for the second image analysis. In some embodiments, at least one image capture device (e.g., image capture device 122) may be configured to capture images (e.g., the first and second images) for use in the first image analysis to generate the first estimated time-to-collision. At least one other image capture device (e.g., image capture device 124) may be configured to capture images (e.g., the third and fourth images) for use in the second image analysis to generate the second estimated time-to-collision. In some embodiments, when at least one image capture device is used for capturing a plurality of images for the first image analysis and the second image analysis, the at least one image capture device may include a first image sensor configured to capture images (e.g., the first and 34 second images) for use in the first image analysis, and a second image sensor that is separate and independent of the first image sensor configured to capture images (e.g., the third and fourth images) for use in the second image analysis.

[0138] Referring back to FIG. 11 , method 1100 may include calculating a difference between the first estimated time-to-collision (TTC1) and the second estimated time-to-collision (TTC2) (step 1140). In some embodiments, processing unit 110 may subtract TTC1 from TTC2 to result in a difference. Processing unit 110 may determine whether the difference exceed a predetermined threshold (step 1150). The predetermined threshold may be any suitable threshold, such as, for example, 0.05 second, 0.1 second, 0.2 second, etc. When processing unit 1 10 determines that the difference exceeds the predetermined threshold (Yes, step 1150), processing unit 110 may repeat steps 1110-1150. When processing unit 110 determines that the difference does not exceed the predetermined threshold (No, step 1150), processing unit 110 may cause vehicle 200 to brake using, e.g., braking system 230 (step 1160).

[0139] FIGS. 14A-14C show an image with an optical flow field and a distribution of the second estimated time-to-collision (TTC2) in the optical flow field. FIG. 14A shows an image 1400 showing a vehicle 1410 traveling on a road 1420. For illustrative purposes, other objects included in image 1400 are not shown. FIG. 14B shows an optical flow field 1430 derived based on at least two images. For example, optical flow field 1430 may be derived from image 1400 and at least one second image (not shown) of vehicle 1410 acquired at the second predetermined time interval dt2 apart from the time when image 1400 was acquired. A plurality of image points may be selected in the image 1400 or the at least one second image, and vectors connecting the corresponding points of these images may be generated by processing unit 110. The vectors may

Tesla EX1002 Page 1458 IPR2025-00944

form optical flow field 1430, which is schematically shown in FIG. 14B , as a region surrounding vehicle 1410.

[0140] Although optical flow field 1430 is shown within a region surrounding vehicle 1410 for illustrative purposes, optical flow field 1430 may be derived for a smaller region (e.g., focusing on the window of vehicle 1430) or a greater region (e.g., the entire area shown in image 1400). Each point in optical flow field 1430 may be associated with a vector represented in FIG. 14B by an arrow having a length and pointing at a direction. The length of the arrow may represent a magnitude of the vector. For example, the magnitude of the vector may indicate a magnitude of movement of a first point in image 1400 to a second point in the at least one second image. The direction of the arrow may indicate the direction of movement between each pair of corresponding points in these two images.

[0141] After performing the second image analysis 1300 discussed above, processing unit 110 may determine the second estimated time-to-collision (TTC2) for each point in the optical flow field 1430. FIG. 14C shows an exemplary distribution of TTC2 over the entire area shown in image 1400. As shown in FIG. 14C , points in different regions may be associated with different values of TTC2. The different patterns and/or shades shown in FIG. 14C are shown to schematically illustrate different values of TTC2. For a region 1440 substantially focusing on the body of vehicle 1410, all or most of the points within region 1440 may be associated with substantially the same value of TTC2. In other words, for an upright moving object like vehicle 1410, all or most of the points on that upright object may have the same or substantially the same times-to-collision.

[0142] For road surfaces, however, all or most of the points may not have the same or substantially the same times-to-collision. For example, points in a first portion 1450 of road 1420 may be associated with one time-to-collision, and points in a second portion 1460 of road 1420 may be associated with another time-to-collision. These times-to-collision may be different, as shown with different patterns/shades in FIG. 14C . In some embodiments, there may be a discontinuity between the times-to-collision for the first portion 1450 and second portion 1460 of road 1420 (e.g., the changes between these times-to-collision may not be gradual). Although the first and second portions 1450 and 1460 appear as upright objects in a single image, the optical flow field generated from at least two images may reveal additional information (e.g., time-to-collision) that may distinguish a true upright object (which may or may not be moving) and a road surface, which is not upright. This aspect is further discussed with respect to FIGS. 17A-18B .

[0143] As shown in FIG. 14C , points in regions 1470, 1480, and 1490 may be associated with different times-to-collision, as indicated by the different patterns/shading shown in the figure. Regions 1470 and 1480 may each represent a portion of road 1420, and the change in time-to-collision values in regions 1470 and 1480 and first portion 1450 may be gradual (e.g., continuous). For example, region 1490 may represent a portion of the sky, which may be associated with a time-to-collision that is different from those associated with regions 1470 and/or 1480.

[0144] FIGS. 15A and 15B show exemplary focuses of expansion in an optical flow field. FIG. 15A shows an object in an image traveling in the same direction (or substantially the same direction) as and in alignment (or substantially in alignment) with vehicle 200 (e.g., in the same line of movement in front of vehicle 200, toward or away from vehicle 200). As shown in FIG. 15A , a focus of expansion of the object may

Tesla EX1002 Page 1459 IPR2025-00944

overlap completely or substantially with a focus of expansion of the static environment. FIG. 15B shows an object not traveling in a direction parallel to the traveling direction of vehicle 200, e.g., when the object is traveling laterally in a direction that transverses the traveling direction of vehicle 200. As shown in FIG. 15B , a focus of expansion of the object may not overlap the focus of expansion of the static environment.

[0145] For example, FIG. 15A shows an environment 1500 of an area in front of vehicle 200. Consistent with disclosed embodiments, an image capture device (e.g., one of image capture devices 122-126) may acquire a first image 1510 of environment 1500 that includes an object located within the first image 1510. The area surrounding the object may represent a static environment (e.g., roads, buildings, mountains, sky, etc.). FIG. 15A also shows a second image 1520 that an image capture device (e.g., one of image capture devices 122-126) may acquire of the object after the second predetermined time interval dt2. Image 1520 may be acquired as vehicle 200 moves toward the object and, therefore, may be larger than image 1510, indicating an expansion.

[0146] FIG. 15A further shows a focus of expansion 1530. The focus of expansion 1530 may be a focus of expansion of the static environment, as indicated by two dashed lines converging to focus of expansion point 1530. The location of focus of expansion 1530 of the static environment may depend on the camera, and may not change over time as the camera continuously acquires a plurality of images of the object or area for deriving optical flow information. When the object is traveling in the same (or substantially the same) direction as and in (or substantially in) alignment with vehicle 200, e.g., in front of vehicle 200 and in the same or opposite traveling direction as vehicle 200, the focus of expansion 1530 may also constitute a focus of expansion of the object. In other words, the focus of expansion of the object traveling in the same direction as and in alignment with vehicle 200 may completely or substantially overlap (or completely or substantially converge with) the focus of expansion of the static environment. Thus, processing unit 110 may detect one focus of expansion from the optical flow field derived from a plurality of images. FIG. 15A also shows, for illustrative purpose, the distance r from certain points to the focus of expansion.

[0147] FIG. 15B shows an environment 1540 of an area in front of vehicle 200, and a first image 1550 and a second image 1560 of an object (the object being included in the first image 1550 and second image 1560) acquired by the camera one after another at the second predetermined time interval dt2. For example, as vehicle 200 moves toward the object, the image of the object may expand, e.g., from image 1550 to image 1560. FIG. 15B further shows a first focus of expansion 1570 that may be calculated from the optical flow field derived from image 1550 and image 1560 of the object. For illustrative purposes, FIG. 15B shows the distances r from certain points to the first focus of expansion 1570. FIG. 15B also shows a second focus of expansion 1580, which may be a focus of expansion of the static environment. The location of the second focus of expansion 1580 may depend on the camera, and may not change as the camera continuously acquires images of the object or area for deriving the optical flow information. In the example shown in FIG. 15B , the first focus of expansion 1570 may not overlap the second focus of expansion 1580. The detection of two or at least two non-overlapping focuses of expansion may indicate that there is at least one object in the image 1540 that is traveling laterally, e.g., in a direction that is not parallel to the traveling direction of vehicle 200, or in a direction that transverses the traveling direction of vehicle 200 at an angle that is neither zero (or substantially zero) nor 1 80 degrees (or substantially 180 degrees).

Tesla EX1002 Page 1460 IPR2025-00944

[0148] FIG. 16 shows an exemplary optical flow field and a focus of expansion. The following discussion of FIG. 16 illustrates how an optical flow field may be derived and how a focus of expansion may be determined. For example, vehicle 1600 is an object included in an image acquired by, e.g., image capture device 122 installed on vehicle 200, and vehicle 200 is traveling toward vehicle 1600. Image capture device 122 may sequentially and/or continuously acquire a plurality of images of an area in front of vehicle 200 as vehicle 200 approaches vehicle 1600. Any two or more images may be selected from the plurality of images to derive an optical flow field. In FIG. 16 , the first image is represented by vehicle 1600 shown in solid lines, and the second image is represented by vehicle 1610 shown in dashed lines. For a plurality of selected points on the first image, e.g., points (x1, y1), (x2, y2), and (x3, y3), processing unit 110 may determine their corresponding points, e.g., (u1, v1), (u2, v2), and (u3, v3) in the second image. Vectors may be determined for each pair of points (x1, y1) and (u1, v1), (x2, y2) and (u2, v2), and (x3, y3) and (u3, v3). The vectors may indicate a magnitude and direction of expansion for the points in the first image. The length of each vector (represented by a solid arrow in FIG. 16) may indicate the distance or magnitude of movement of from one point in the first image to a corresponding point in the second image, and the direction of the vector may indicate the expansion direction or direction of movement. The converging point 1620 of the vectors may represent the focus of expansion. In some embodiments, the location of the converging point 1620 may be calculated using a suitable mathematical model based on the coordinates of the vectors. For example, the location of converging point 1620 may be calculated from:
SS ∗ x i − FOE x = u i
image available on "Original document"
SS ∗ y i − FOE y = v iImage available on "Original document"

[0149] Where SS may be a constant, (xi, yi) may be coordinates of the i-th point within a first image, and (ui, vi) may be coordinates of the corresponding i-th point within the second image. FOEx and FOEy may be coordinates representing the location of converging point 1620 (i.e., focus of expansion).

[0150] FIGS. 17A and 17B show an image of an object and a linear relationship between distances from points in an optical flow field to the focus of expansion and magnitudes of movement of the points. FIGS. 17A and 17B show that a linear relationship between distances from points to the focus of expansion and magnitudes of movement of the points may be used to determine that the object is an upright object (e.g., a low time-to-collision upright object).

[0151] FIG. 17A shows an image including an object, i.e., a vehicle 1700. The image of vehicle 1700 may include an image of two car windows 1710 and 1720, which appear to be darker in the image of vehicle 1700. A plurality of points 1730 may be selected, most of which may be from a region in the image of vehicle 1700. An optical flow field may be derived from at least two images of vehicle 1700 acquired one after another with the second predetermined time interval dt2. Processing unit 110 may determine the magnitudes of movement d and the distances r from the points to the focus of expansion based on the optical flow field. Processing unit 110 may determine a relationship between the magnitudes d and distances r. In some embodiments, the relationship may be a linear relationship, as shown in FIG. 17B , indicating that d/r (e.g., slope of the plotted line) is the same for all or substantially all of the points in the optical flow field. As discussed above, d/r may be inversely proportional to the second

Tesla EX1002 Page 1461 IPR2025-00944

estimated time-to-collision (TTC2). Thus, a linear relationship between d and r may indicate a constant or substantially constant value of d/r, which in turn may indicate a constant or substantially constant TTC2 for all or substantially all of the points in the optical flow field. The constant or same TTC2 may indicate that the object within the optical flow field is an upright object, or a low time-to-collision upright object.

[0152] To illustrate the linear relationship between magnitudes of movement d and distances r, FIG. 17B shows a plot of the vertical components (y components) related to the magnitudes of movement d (e.g., the magnitude of movement in the y direction, "dy" in FIG. 17B) versus vertical components of distances r (e.g., the distance to the focus of expansion in the y direction, "y" in FIG. 17B). Although not shown in FIG. 17B , a plot of the horizontal components, e.g., magnitudes in the x direction and distances in the x direction, may also indicate a linear relationship, which in turn may indicate a constant or substantially constant time-to-collision for all or substantially all of the selected points in the optical flow field. In addition, a plot of distances r and magnitudes d may also indicate a linear relationship, which in turn may indicate a constant (or same) or substantially constant (or same) time-to-collision for all or substantially all of the selected points in the optical flow field.

[0153] When a non-linear relationship between the magnitudes of movement d and the distances r to the focus of expansion is detected or determined, the non-linear relationship may indicate that the second estimated times-to-collision are different for different points included in the optical flow field, which in turn may indicate that the image may include a road surface, rather than an upright object. As discussed above, the second estimated times-to-collision may be determined based on the optical flow information derived from at least two images. FIGS. 18A and 18B show an image of an object and a non-linear relationship between distances r from points in an optical flow field to the focus of expansion and magnitudes d of movement of the points. FIG. 18A shows an image 1800 of a house 1810 having two garages 1820 and 1830 with garage doors opened, and a driveway 1840 (e.g., a road surface) leading up to the garages 1820 and 1830. Because the garage doors are opened, garages 1820 and 1830 may appear darker on the image, similar to the car windows 1710 and 1720 shown in FIG. 17A . A region 1850 of the image may be selected for image processing. A plurality of points may be selected within region 1850, and at least two images including regions corresponding to region 1850 may be acquired to derive an optical flow field. Based on the optical flow field having a plurality of vectors, processing unit 110 may determine a focus of expansion, the distances r of the points to the focus of expansion, and the magnitudes d of movement of the points included in the optical flow field. Processing unit 110 may determine a relationship between the distances r and the magnitudes d. In the situation shown in FIG. 18A , processing unit 110 may determine that the relationship between the distances r and the magnitudes d is non-linear. The non-linear relationship between the distances r and the magnitudes d may indicate that the times-to-collision for the points are different for different points. This in turn may indicate that selected region 1850 includes an image of a road surface, rather than an upright object like vehicle 1700 shown in FIG. 17A .

[0154] To illustrate the non-linear relationship between magnitudes of movement d and distances r, FIG. 18B shows a plot of the vertical components (y components) related to the magnitudes d of movement (e.g., the magnitude of movement in the y direction, "dy" in FIG. 18B) versus vertical components of distances r (e.g., the distance to the focus of expansion in the y direction, "y" in FIG. 18B). Although not shown in FIG. 18B , a plot of the horizontal components, e.g., magnitudes in the x direction and

Tesla EX1002 Page 1462 IPR2025-00944

distances in the x direction, may also indicate a nonlinear relationship, and hence, different times-to-collision for the selected points in the optical flow field. In addition, a plot of distances r and magnitudes d may also indicate a nonlinear relationship, and hence, different times-to-collision for the selected points in the optical flow field. Here, the image in the region 1850 may appear similar to the image of vehicle 1700 shown in FIG. 17A in a single image. The optical flow field derived from a plurality of images acquired at the second predetermined time interval dt2 may reveal additional information, e.g., time-to-collision information, about the objects included in the selected region. The time-to-collision information may be used to determine whether the object is an upright object or a road surface. Processing unit 110 may combine this determination with the calculated TTC2 to determine whether or not to cause vehicle 200 to brake.

[0155] FIG. 19 is a flowchart showing an exemplary process 1900 for triggering a warning based on one or more probabilities indicating that a time-to-collision is smaller than one or more thresholds. Instead of determining the time-to-collision as discussed above, processing unit 110 may determine, based on at least two images acquired by one or more image capture devices 122, 124, and 126, a probability that the time-to-collision is smaller than a predetermined threshold. For example, in process 1900, one or more image capture devices 122-126 may acquire a plurality of images of an area in front of vehicle 200 (step 1910). In some embodiments, one image capture device having one camera may be used. In some embodiments, two or more image capture devices have two or more cameras may be used. For example, one camera may have a 45-degree field of view of the area in front of vehicle 200, and another camera may have a 30-degree field of view.

[0156] In some embodiments, processing unit 110 may perform a third image analysis based on at least two images to determine a first probability (P1) that a time-to-collision (TTC) is smaller than a first predetermined threshold T1 (i.e., TTC<T1) (step 1920). The first predetermined threshold T1 may be any suitable value, such as, for example, 0.1 second, 0.2 second, 05. second, etc. Processing unit 110 may determine a second probability (P2) that the time-to-collision (TTC) is smaller than a second predetermined threshold T2 (i.e., TTC < T2) (step 1930). The second predetermined threshold T2 may be any suitable value, such as, for example, 0.1 second, 0.2 second, 05. second, etc. The second predetermined threshold T2 may or may not be the same as the first predetermined threshold T1.

[0157] Processing unit 110 may trigger a warning if the first probability P1 is greater than a first predetermined probability threshold Tpl, or if the second probability P2 is greater than a second predetermined probability threshold Tp2. Tp1 and Tp2 may be any suitable value, such as, for example, 0.5, 0.8, 0.9, etc. Tp1 and Tp2 may or may not be the same. In some embodiments, Tp2 may be smaller than Tp1. Processing unit 110 may compare the first probability P1 with the first predetermined probability threshold Tp1 to determine whether the first probability P1 is greater than the first predetermined probability threshold Tp1. Processing unit 110 may also compare the second probability P2 with the second predetermined probability threshold Tp2 to determine whether the second probability P2 is greater than the second predetermined probability threshold Tp2. The warning may include an audio alert, a video alert, a vibrational alert, or a combination thereof. The warning may alert the driver so that the driver may take suitable navigational action, such as applying brakes, steering the wheel, etc.

IPR2025-00944
Tesla EX1002 Page 1463

[0158] FIG. 20 is a flowchart showing an exemplary process 2000 for determining the first time-to-collision (TTC1) that may be implemented in the first image analysis shown in FIG. 11 . Process 2000 may be used in the first image analysis as an alternative texture based TTC1 calculation method replacing the process 1200 shown in FIG. 12 . Processing unit 110 may process at least two images that are acquired at a predetermined time interval DT. Processing unit 110 may determine a first dimension length y1 of an object in a first image (step 2010). The first dimension length of the object may be a length of a feature of the object in the vertical direction, horizontal direction, or in any direction on the first image. For example, when the object is a vehicle, the feature may be a vertical side of a license plate on the vehicle. Processing unit 110 may determine a second dimension length y2 of the same object (or the same feature of the same object) in a second image, which may be acquired after the first image at a time that is DT after the time the first image is acquired (step 2020). Processing unit 110 may determine a scaling factor Sf based on the first dimension length and the second dimension length. For example, processing unit 110 may calculate Sf based on Sf = y1 / y2. In some embodiments, processing unit may track the horizontal and vertical coordinates (dx, dy) of the object in the first and second image, and may calculate Sf based on the coordinates (dx, dy), for example, Sf = dy1 / dy2.

[0159] Processing unit 110 may determine a time-to-collision (which may be the first time-to-collision shown in FIG. 11) based on the scaling factor and the time interval DT. In some embodiments, processing unit 110 may calculate TTC1 from TTC1 = DT / (Sf-1).

[0160] Using two, i.e., texture based and structure based, sensing modalities, methods, or models, as described above, have several benefits or advantages. For example, an immediate result of using two modalities in a braking decisioning system is a reduction of the probability of system failure may be achieved, as compared to braking decisioning systems using only one of the modalities (or only one information channel).

[0161] Another benefit relates to system validation. Using at least two modalities (or channels of information, or two independent and/or separate analyses) enable the vehicle to assess the overall system failure rate (e.g., mean-time-between-failures or MTBF). The overall system failure rate may be determined by assessing each modality's failure rate separately (e.g., by assessing the failure rate of the texture based analysis and the failure rate of the structure based analysis separately). The two modalities may be implemented in two sub-systems: a textural analysis based sub-system and a structural analysis based sub-system. The failure rates of the two modalities may be multiplied. A result of the multiplication may be used as an indicator of or may be used to compute the overall system failure rate. In some embodiments, the MTBF may be estimated based on the result of the multiplication, even if it is significantly higher than the extent (in time) of available validation data.

[0162] For example, in one scenario, a record of only 1000 hours of data is available, and the combined systems MTBF is 1000000 hours. In 1/100 hours the texture based analysis may trigger a false braking decision. In 1/10000 of the cases the texture based analysis is checking the hypothesized decision (e.g., the hypothesized decision being a decision to brake because of a low TTC upright object in front of vehicle 200). In such cases, the texture based analysis may falsely approve the hypothesized decision. By direct assessment of the combined system (or combined analyses of the texture based analysis and the structure based analysis), vehicle systems may only achieve 0 failure

IPR2025-00944
Tesla EX1002 Page 1464

IPR2025-00944
Tesla EX1002 Page 1465

over 1000 hours, and may not be able to assess the MTBF. By taking two independent channels of information into account in the braking decisioning for vehicle 200, e.g., using the texture based analysis and the structure based analysis, a vehicle system may be able to assess the failure rate of each sub-system (e.g., textural analysis based sub-system and structural analysis based sub-system) separately. Such failure rates may be multiplied, and a result of the multiplication may be used to assess the MTBF. For example, in some embodiments, the textural analysis based sub-system may be keep running alone and it may fail 10 times over 1000 hours. If the structural analysis based sub-system is falsely provoked 1000000 times, as if, for example, the textural analysis based sub-system reached a decision that there is a need to issue an autonomous emergency braking on an open road, the structural analysis based sub-system may fail once in 10000 provocations. Based on these observations, processing unit 110 may deduce (or determine) that the total system mean-time-between failures (MTBFs) is indeed 1000000 hours. This estimate may be reached using only 1000 hours of recorded data. Thus, by using the disclosed two modalities, system validation may be more efficiently performed based on only limited hours of recorded data.

[0163] FIG. 21 is a flowchart showing an exemplary process 2100 for performing system validation using two modalities. Process 2100 summarizes the system validation discussed above. Processing unit 110 may determine a first failure rate associated with the textural analysis based sub-system (step 2110). For example, processing unit 110 may determine the first failure rate based on the total number of times the textural analysis based sub-system makes a false braking decision (e.g., falsely detected a low TTC upright object) among a total number of system running hours. As discussed above, the textural analysis based sub-system may make a decision to brake based on detection of a change in a texture (e.g., a change in footprint size) between at least two images of an area in front of vehicle 200. Processing unit 110 may determine a second failure rate associated with the structural analysis based sub-system (step 2120). For example, processing unit 110 may determine the second failure rate based on the total number of times the structural analysis based sub-system makes a false braking decision (e.g., falsely detected a low TTC upright object) among the same total number of system running hours or a different total number of system running hours. Processing unit 110 may determine an overall system failure rate based on the first failure rate and the second failure rate (step 2130). For example, processing unit 110 may multiply the first failure rate and the second failure rate, and use a result of the multiplication as an indicator of overall system mean-time-between-failures (MTBF).

[0164] The foregoing description has been presented for purposes of illustration. It is not exhaustive and is not limited to the precise forms or embodiments disclosed. Modifications and adaptations will be apparent to those skilled in the art from consider! tion of the specification and practice of the disclosed embodiments. Additionally, although aspects of the disclosed embodiments are described as being stored in memory, one skilled in the art will appreciate that these aspects can also be stored on other types of computer readable media, such as secondary storage devices, for example, hard disks or CD ROM, or other forms of RAM or ROM, USB media, DVD, Blu-ray, or other optical drive media.

[0165] Computer programs based on the written description and disclosed methods are within the skill of an experienced developer. The various programs or program modules can be created using any of the techniques known to one skilled in the art or can be designed in connection with existing software. For example, program sections or

program modules can be designed in or by means of .Net Framework, .Net Compact Framework (and related languages, such as Visual Basic, C, etc.), Java, C++, Objective-C, HTML, HTML/AJAX combinations, XML, or HTML with included Java applets.

[0166] Moreover, while illustrative embodiments have been described herein, the scope of any and all embodiments having equivalent elements, modifications, omissions, combinations (e.g., of aspects across various embodiments), adaptations and/or alterations as would be appreciated by those skilled in the art based on the present disclosure. The limitations in the claims are to be interpreted broadly based on the language employed in the claims and not limited to examples described in the present specification or during the prosecution of the application. The examples are to be construed as non-exclusive. Furthermore, the steps of the disclosed methods may be modified in any manner, including by reordering steps and/or inserting or deleting steps. It is intended, therefore, that the specification and examples be considered as illustrative only, with a true scope being indicated by the following claims.

Espacenet

Claims: EP3146506 (A2) — 2017-03-29

SYSTEMS AND METHODS FOR BRAKING A VEHICLE BASED ON A DETECTED OBJECT

Claims of EP3146506 (B1)

A high quality text as facsimile in your desired language may be available amongst the following family members:

CN106536299 (A) US2015336547 (A1) WO2015177643 (A2) CN110171405 (A) US2017203744 (A1) US2019084549 (A1)

- Original claims
- Claims tree

The EPO does not accept any responsibility for the accuracy of data and information originating from other authorities than the EPO; in particular, the EPO does not guarantee that they are complete, up-to-date or fit for specific purposes.

1. An object detecting and braking system for a vehicle (200), comprising:
at least one image capture device (122, 124, 126) configured to acquire a plurality of images of an area including an object in front of the vehicle (200);
a data interface (128); and
at least one processing device (110) programmed to:
receive (1110) the plurality of images via the data interface (128);
perform (1120) a first image analysis based on a time interval between a first image and a second image to determine a first estimated time-to-collision of the vehicle (200) with the object, the first image analysis including:
determining (1210) a first size of the object appearing in the first image;
determining (1220) a second size of the object appearing in the second image;
determining (1230) a change in size of the object based on the first size and the second size; and
determining (1240) the first estimated time-to-collision based on the change in size of the object and a predetermined time interval between times of
acquisition of the first image and the second image;
characterized in that the at least one processing device (110) is further programmed to:
perform (1130) a second image analysis based on a time interval between a third image and a fourth image to determine a second estimated time-to-collision of the vehicle (200) with the object, the second image analysis including:
determining (1320) a focus of expansion based on the third image and the fourth image, the focus of expansion being a converging point of vectors in a vector field determined from the third image and the fourth image, wherein each point in the vector

field is represented by a vector indicating a magnitude and a direction of movement from a first point in the third image to a corresponding second point in the fourth image;
determining (1330) a rate of expansion relative to the focus of expansion for each point in the vector field by dividing a magnitude of movement of a vector associated with the point by a distance from the point to the focus of expansion; and
determining for a region substantially focusing on the object the second estimated time-to-collision based on the rate of expansion and a time interval between times of acquisition of the third image and the fourth image;
calculate (1140) a difference between the first estimated time-to-collision and the second estimated time-to-collision;
determine (1150) that the difference does not exceed a predetermined threshold; and
cause (1160) the vehicle (200) to brake based on the determination that the difference does not exceed the predetermined threshold.

2. The object detecting and braking system of claim 1, wherein the second image analysis further includes:
determining a relationship between magnitudes of movement and distances to the focus of expansion for a plurality of points in the vector field; and
determining that the object represented by the plurality of points is a low time-to-collision upright object based on a determination that the relationship is a linear relationship.

3. The object detecting and braking system of claim 1, wherein the second image analysis further includes:
determining a relationship between magnitudes of motion and to the focus of expansion. for a plurality of points in the vector field; and
determining that at least a portion of the object is a road surface based on a determination that the relationship is a non-linear relationship.

4. The object detecting and braking system of claim 1, wherein the second image analysis further includes:
detecting a first focus of expansion that is a first converging point of a first plurality of vectors in the vector field;
detecting a second focus of expansion that is a second converging point of a second plurality of vectors in the vector field; and
determining that the third image and the fourth image include a moving object that moves in a direction that is not parallel to a travel direction of the vehicle.

5. The object detecting and braking system of claim 1, wherein the second image analysis further includes:
determining that the third image and the fourth image represent an environment that is static relative to a road surface on which the vehicle is traveling.

6. The object detecting and braking system of claim 1, wherein the at least one image capture device includes:
a first image sensor configured to capture the first image and the second image to be used in the first image analysis; and
a second image sensor independent of the first image sensor and configured to capture the third image and the fourth image to be used in the second image analysis.

7. A vehicle (200), comprising:
a body; and

Tesla EX1002 Page 1468 IPR2025-00944

the object detecting and braking system of any of the preceding claims.

8. A method for detecting an object and braking a vehicle (200), comprising:
acquiring (1110), via at least one image capture device (122, 124, 126), a plurality of images of an area including an object in front of the vehicle (200);
receiving, via a processing device (110), the plurality of images;
performing (1120), via the processing device (110), a first image analysis based on a time interval between a first image and a second image to determine a first estimated time-to-collision of the vehicle (200) with the object, wherein performing the first image analysis includes:
determining (1210) a first size of the object appearing in the first image;
determining (1220) a second size of the object appearing in the second image;
determining (1230) a change in size of the object based on the first size and the second size; and
determining (1240) the first estimated time-to-collision based on the change in size of the object and a predetermined time interval between times of acquisition of the first image and the second image;
characterized by
performing (1130), via the processing device (110), a second image analysis based on a time interval between a third image and a fourth image to determine a second estimated time-to-collision of the vehicle with the object, wherein performing the second image analysis includes:
determining (1320) a focus of expansion based on the third image and the fourth image, the focus of expansion being a converging point of vectors in a vector field determined from the third image and the fourth image, wherein each point in the vector field is represented by a vector indicating a magnitude and a direction of movement from a first point in the third image to a corresponding second point in the fourth image;
determining (1330) a rate of expansion relative to the focus of expansion for each point in the vector field by dividing a magnitude of movement of a vector associated with the point by a distance from the point to the focus of expansion; and
determining for a region substantially focusing on the object the second estimated time-to-collision based on the rate of expansion and the time interval between times of acquisition of the third image and the fourth image;
calculating (1140), via the processing device (110), a difference between the first estimated time-to-collision and the second estimated time-to-collision;
determining (1150), via the processing device (110), that the difference does not exceed a predetermined threshold; and
causing (1160), via the processing device (110), the vehicle (200) to brake based on the determination that the difference does not exceed the predetermined threshold.

9. The method of claim 8, wherein performing the second image analysis further includes:
determining a relationship between magnitudes of movement and distances to the focus of expansion for a plurality of points in the vector field; and
determining that the object represented by the plurality of points is a low time-to-collision upright object based on a determination that the relationship is a linear relationship.

10. The method of claim 8, wherein performing the second image analysis further includes:
determining a relationship between magnitudes of movement and distances to the focus of expansion for a plurality of points in the vector field; and

determining that at least a portion of the object is a road surface based on a determination that the relationship is a non-linear relationship.

11. The method of claim 8, wherein performing the second image analysis further includes:
detecting a first focus of expansion that is a first converging point of a first plurality of vectors in the vector field;
detecting a second focus of expansion that is a second converging point of a second plurality of vectors in the vector field; and
determining that the third image and the fourth image include a moving object that moves in a direction that is not parallel to a travel direction of the vehicle.

12. The method of claim 8, wherein performing the second image analysis further includes:
determining that the third image and the fourth image represent an environment that is static relative to a road surface on which the vehicle is traveling.

13. The method of claim 8, wherein acquiring (1110) the plurality of images comprises:
capturing, by a first image sensor included in the at least one image capture device, the first image and the second image to be used in the first image analysis; and
capturing, by a second image sensor included in the at least one image capture device and being independent of the first image sensor, the third image and the fourth image to be used in the second image analysis.

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Veröffentlichungsnummer:
(11) Publication number: **EP 3 146 506 A0**
(11) Numéro de publication:

Internationale Anmeldung veröffentlicht durch die Weltorganisation für geistiges Eigentum unter der Nummer:

WO 2015/177643 (Art. 153(3) EPÜ).

International application published by the World Intellectual Property Organization under number:

WO 2015/177643 (Art. 153(3) EPC).

Demande internationale publiée par l'Organisation Mondiale de la Propriété Intellectuelle sous le numéro:

WO 2015/177643 (art. 153(3) CBE).

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 3 146 506 B1

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention of the grant of the patent:
11.12.2019 Bulletin 2019/50

(21) Application number: **15770635.9**

(22) Date of filing: **22.05.2015**

(51) Int Cl.:
G06T 7/20 *(2017.01)*

(86) International application number:
PCT/IB2015/001457

(87) International publication number:
WO 2015/177643 (26.11.2015 Gazette 2015/47)

(54) **SYSTEMS AND METHODS FOR BRAKING A VEHICLE BASED ON A DETECTED OBJECT**

SYSTEME UND VERFAHREN ZUM ABBREMSEN EINES FAHRZEUGS AUF BASIS EINES ERFASSTEN OBJEKTS

SYSTÈMES ET PROCÉDÉS DE FREINAGE D'UN VÉHICULE SUR LA BASE D'UN OBJET DÉTECTÉ

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR

(30) Priority: **22.05.2014 US 201462001772 P**

(43) Date of publication of application:
29.03.2017 Bulletin 2017/13

(73) Proprietor: **Mobileye Vision Technologies Ltd.**
91450 Jerusalem (IL)

(72) Inventor: **DAGAN, Erez**
Tel Aviv (IL)

(74) Representative: **Finnegan Europe LLP**
1 London Bridge
London SE1 9BG (GB)

(56) References cited:
WO-A1-2013/107525 **US-A- 5 559 695**
US-A1- 2006 178 830 **US-A1- 2012 314 071**

- **Kyung-Ho Shin ET AL: "25-11-2010", 8th World Congress on Railway Research, May 18-22, 2008, 18 May 2008 (2008-05-18), pages 1-7, XP055227717, Seoul, Korea Retrieved from the Internet: URL:https://web.archive.org/web/20101125041611/http://www.uic.org/cdrom/2008/11_wcrr2008/pdf/O.1.4.5.3.pdf [retrieved on 2015-11-11]**

Note: Within nine months of the publication of the mention of the grant of the European patent in the European Patent Bulletin, any person may give notice to the European Patent Office of opposition to that patent, in accordance with the Implementing Regulations. Notice of opposition shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

EP 3 146 506 B1

Printed by Jouve, 75001 PARIS (FR)

Description

BACKGROUND

I. Technical Field

[0001] The present disclosure relates generally to autonomous vehicle navigation and, more specifically, to systems and methods that use cameras to detect an object in front of the vehicle and that apply brakes based on the detected object.

II. Background Information

[0002] As technology continues to advance, the goal of a fully autonomous vehicle that is capable of navigating on roadways is on the horizon. Primarily, an autonomous vehicle may be able to identify its environment and navigate without input from a human operator. Autonomous vehicles may also take into account a variety of factors and make appropriate decisions based on those factors to safely and accurately reach an intended destination. For example, various objects-such as other vehicles and pedestrians-are encountered when a vehicle typically travels a roadway. Autonomous driving systems may recognize these objects in a vehicle's environment and take appropriate and timely action to avoid collisions. Additionally, autonomous driving systems may identify other indicators-such as traffic signals, traffic signs, and lane markings-that regulate vehicle movement (e.g., when the vehicle must stop and may go, a speed at which the vehicle must not exceed, where the vehicle must be positioned on the roadway, etc.). Autonomous driving systems may need to determine when a vehicle should change lanes, turn at intersections, change roadways, etc. As is evident from these examples, many factors may need to be addressed in order to provide an autonomous vehicle that is capable of navigating safely and accurately.

[0003] US 2012/0314071 A1 discloses a method for preventing collisions between a vehicle and pedestrians using a camera mounted in a vehicle for capturing image frames of the environment of the vehicle in the field of view of the camera. When a candidate image of a pedestrian in the field of view of the camera is detected in an image frame, the optical flow between the image frames is analyzed to validate the pedestrian's detection. When the analyzed optical flow is consistent with an upright pedestrian, a collision warning signal is provided.

[0004] US 2006/0178830 A1 relates to collision warning methods and apparatus for vehicles. An apparatus for facilitating avoidance of a vehicle collision with an object comprises a vision sensor providing image data, and an image analyzer, operable to provide an estimated time to collision of the vehicle with the object. The estimated time to collision is determined from a rate of expansion of an image element corresponding to the object within the image data, and the apparatus provides a warning if the estimated time to collision is less than or approximately equal to a predetermined time.

[0005] US 5,559,695 A discloses an apparatus and method for real-time, adaptive self-calibration of a visual time-to-contact sensor mounted on a vehicle to predict the time-to-contact with stationary and moving obstacles in a moving vehicle's immediate path. The focus of expansion is actively centered in the image sequence, using optical flow, and the pan and tilt of the sensor are continuously calibrated to point in the direction of the focus of expansion in order to accurately predict the time-to-contact with stationary and moving obstacles in a moving vehicle's immediate path, and enable warning or evasive action.

[0006] "A study on the reliability improvement for applying GNSS to railroad systems" by Kyung-Ho Shin, et al. investigates reliability of application of Global Navigation Satellite (GNSS) systems in the railway field and proposes a dual-mode GNSS configuration method for controlling trains with improved reliability.

SUMMARY

[0007] Embodiments consistent with the present disclosure provide systems and methods for autonomous vehicle navigation. The disclosed embodiments may use cameras to provide autonomous vehicle navigation features. For example, consistent with the disclosed embodiments, the disclosed systems may include one, two, or more cameras that monitor the environment of a vehicle and cause a navigational response based on an analysis of images captured by one or more of the cameras.

[0008] In accordance with aspects of the invention there are provided an object detecting and braking system for a vehicle, a vehicle, and a method for detecting an object and braking a vehicle, as set out in claims 1, 7, and 8 respectively.

[0009] Consistent with a disclosed embodiment, an object detecting and braking system for a vehicle is provided, according to claim 1.

[0010] Consistent with another disclosed embodiment, a vehicle is provided, according to claim 7.

[0011] Consistent with yet another disclosed embodiment, a method for detecting an object and braking a vehicle is provided, according to claim 8.

[0012] Consistent with another disclosed example, a method is provided for assessing an overall system failure rate associated with a braking decisioning system for a vehicle. The method includes determining, using a processing device, a first failure rate associated with a textural analysis based sub-system. The textural analysis based sub-system is configured to make a decision to brake based on a change in texture between at least two images of an area in front of a vehicle. The method further includes determining, using the processing device, a second failure rate associated with a structural analysis based sub-system. The structural analysis based sub-system is configured to make a decision to

IPR2025-00944
Tesla EX1002 Page 1473

brake based on optical flow information derived from at least two images of the area in front of the vehicle. The method also includes determining, using the processing device, the overall system failure rate based on the first failure rate and the second failure rate.

[0013] Consistent with other disclosed examples, non-transitory computer-readable storage media may store program instructions, which are executed by at least one processing device and perform any of the methods described herein.

[0014] The foregoing general description and the following detailed description are exemplary and explanatory only. The invention is defined solely by the appended claims.

BRIEF DESCRIPTION OF THE DRAWINGS

[0015] The accompanying drawings, which are incorporated in and constitute a part of this disclosure, illustrate various disclosed embodiments. In the drawings:

FIG. 1 is a diagrammatic representation of an exemplary system consistent with the disclosed embodiments.

FIG. 2A is a diagrammatic side view representation of an exemplary vehicle including a system consistent with the disclosed embodiments.

FIG. 2B is a diagrammatic top view representation of the vehicle and system shown in FIG. 2A consistent with the disclosed embodiments.

FIG. 2C is a diagrammatic top view representation of another embodiment of a vehicle including a system consistent with the disclosed embodiments.

FIG. 2D is a diagrammatic top view representation of yet another embodiment of a vehicle including a system consistent with the disclosed embodiments.

FIG. 2E is a diagrammatic top view representation of yet another embodiment of a vehicle including a system consistent with the disclosed embodiments.

FIG. 2F is a diagrammatic representation of exemplary vehicle control systems consistent with the disclosed embodiments.

FIG. 3A is a diagrammatic representation of an interior of a vehicle including a rearview mirror and a user interface for a vehicle imaging system consistent with the disclosed embodiments.

FIG. 3B is an illustration of an example of a camera mount that is configured to be positioned behind a rearview mirror and against a vehicle windshield consistent with the disclosed embodiments.

FIG. 3C is an illustration of the camera mount shown in FIG. 3B from a different perspective consistent with the disclosed embodiments.

FIG. 3D is an illustration of an example of a camera mount that is configured to be positioned behind a rearview mirror and against a vehicle windshield consistent with the disclosed embodiments.

FIG. 4 is an exemplary block diagram of a memory configured to store instructions for performing one or more operations consistent with the disclosed embodiments.

FIG. 5A is a flowchart showing an exemplary process for causing one or more navigational responses based on monocular image analysis consistent with disclosed embodiments.

FIG. 5B is a flowchart showing an exemplary process for detecting one or more vehicles and/or pedestrians in a set of images consistent with the disclosed embodiments.

FIG. 5C is a flowchart showing an exemplary process for detecting road marks and/or lane geometry information in a set of images consistent with the disclosed embodiments.

FIG. 5D is a flowchart showing an exemplary process for detecting traffic lights in a set of images consistent with the disclosed embodiments.

FIG. 5E is a flowchart showing an exemplary process for causing one or more navigational responses based on a vehicle path consistent with the disclosed embodiments.

FIG. 5F is a flowchart showing an exemplary process for determining whether a leading vehicle is changing lanes consistent with the disclosed embodiments.

FIG. 6 is a flowchart showing an exemplary process for causing one or more navigational responses based on stereo image analysis consistent with the disclosed embodiments.

FIG. 7 is a flowchart showing an exemplary process for causing one or more navigational responses based on an analysis of three sets of images consistent with the disclosed embodiments.

FIG. 8 is a diagrammatic representation of an exemplary vehicle including an object detecting and braking system traveling on a road behind another vehicle consistent with the disclosed embodiments.

FIG. 9 is a diagrammatic representation of an exemplary vehicle including an object detecting and braking system traveling on a road including an oncoming vehicle consistent with the disclosed embodiments.

FIG. 10 is an exemplary block diagram of a memory that may store instructions for performing one or more operations for detecting an object in front of the vehicle and causing the vehicle to brake based on the detection consistent with the disclosed embodiments.

FIG. 11 is a flowchart showing an exemplary process for detecting an object in front of the vehicle and causing the vehicle to brake based on the detection consistent with the disclosed embodiments.

FIG. 12 is a flowchart showing an exemplary first image analysis for determining a first estimated time-to-collision consistent with the disclosed embodiments.

FIG. 13 is a flowchart showing an exemplary second image analysis for determining a second estimated

IPR2025-00944
Tesla EX1002 Page 1474

time-to-collision consistent with the disclosed embodiments.

FIGS. 14A-14C show an image with an optical flow field and a distribution of a second estimated time-to-collision in the optical flow field consistent with the disclosed embodiments.

FIGS. 15A and 15B show exemplary focuses of expansion when an object travels in different directions consistent with the disclosed embodiments.

FIG. 16 shows an exemplary optical flow field and a focus of expansion consistent with the disclosed embodiments.

FIGS. 17A and 17B show an image of an object and a linear relationship between distances from points in an optical flow field to the focus of expansion and magnitudes of movement of the points consistent with the disclosed embodiments.

FIGS. 18A and 18B show an image of an object and a non-linear relationship between distances from points in an optical flow field to the focus of expansion and magnitudes of movement of the points consistent with the disclosed embodiments.

FIG. 19 is a flowchart showing an exemplary process for triggering a warning based on probabilities indicating that a time-to-collision is smaller than one or more thresholds consistent with the disclosed embodiments.

FIG. 20 is a flowchart showing an exemplary process for determining the first time-to-collision (TTC1) that may be implemented in the first image analysis shown in FIG. 11 consistent with the disclosed embodiments.

FIG. 21 is a flowchart showing an exemplary process for performing a system validation using two modalities consistent with the disclosed embodiments.

DETAILED DESCRIPTION

[0016] The following detailed description refers to the accompanying drawings. Wherever possible, the same reference numbers are used in the drawings and the following description to refer to the same or similar parts. While several illustrative embodiments are described herein, modifications, adaptations and other implementations are possible. For example, substitutions, additions or modifications may be made to the components illustrated in the drawings, and the illustrative methods described herein may be modified by substituting, reordering, removing, or adding steps to the disclosed methods. Accordingly, the following detailed description is not limited to the disclosed embodiments and examples. Instead, the proper scope is defined by the appended claims.

[0017] FIG. 1 is a block diagram representation of a system 100 consistent with the exemplary disclosed embodiments. System 100 may include various components depending on the requirements of a particular implementation. In some embodiments, system 100 may include a processing unit 110, an image acquisition unit 120, a position sensor 130, one or more memory units 140, 150, a map database 160, and a user interface 170. Processing unit 110 may include one or more processing devices. In some embodiments, processing unit 110 may include an applications processor 180, an image processor 190, or any other suitable processing device. Similarly, image acquisition unit 120 may include any number of image acquisition devices and components depending on the requirements of a particular application. In some embodiments, image acquisition unit 120 may include one or more image capture devices (e.g., cameras), such as image capture device 122, image capture device 124, and image capture device 126. System 100 may also include a data interface 128 communicatively connecting processing device 110 to image acquisition device 120. For example, data interface 128 may include any wired and/or wireless link or links for transmitting image data acquired by image accusation device 120 to processing unit 110.

[0018] Both applications processor 180 and image processor 190 may include various types of processing devices. For example, either or both of applications processor 180 and image processor 190 may include a microprocessor, preprocessors (such as an image preprocessor), graphics processors, a central processing unit (CPU), support circuits, digital signal processors, integrated circuits, memory, or any other types of devices suitable for running applications and for image processing and analysis. In some embodiments, applications processor 180 and/or image processor 190 may include any type of single or multi-core processor, mobile device microcontroller, central processing unit, etc. Various processing devices may be used, including, for example, processors available from manufacturers such as Intel®, AMD®, etc. and may include various architectures (e.g., x86 processor, ARM®, etc.).

[0019] In some embodiments, applications processor 180 and/or image processor 190 may include any of the EyeQ series of processor chips available from Mobileye®. These processor designs each include multiple processing units with local memory and instruction sets. Such processors may include video inputs for receiving image data from multiple image sensors and may also include video out capabilities. In one example, the EyeQ2® uses 90nm-micron technology operating at 332Mhz. The EyeQ2® architecture consists of two floating point, hyper-thread 32-bit RISC CPUs (MIPS32® 34K® cores), five Vision Computing Engines (VCE), three Vector Microcode Processors (VMP®), Denali 64-bit Mobile DDR Controller, 128-bit internal Sonics Interconnect, dual 16-bit Video input and 18-bit Video output controllers, 16 channels DMA and several peripherals. The MIPS34K CPU manages the five VCEs, three VMP™ and the DMA, the second MIPS34K CPU and the multi-channel DMA as well as the other peripherals. The five VCEs, three VMP® and the MIPS34K CPU can perform intensive vision computations required by multi-

IPR2025-00944
Tesla EX1002 Page 1475

function bundle applications. In another example, the EyeQ3®, which is a third generation processor and is six times more powerful that the EyeQ2®, may be used in the disclosed embodiments.

[0020] Any of the processing devices disclosed herein may be configured to perform certain functions. Configuring a processing device, such as any of the described EyeQ processors or other controller or microprocessor, to perform certain functions may include programming of computer executable instructions and making those instructions available to the processing device for execution during operation of the processing device. In some embodiments, configuring a processing device may include programming the processing device directly with architectural instructions. In other embodiments, configuring a processing device may include storing executable instructions on a memory that is accessible to the processing device during operation. For example, the processing device may access the memory to obtain and execute the stored instructions during operation.

[0021] While FIG. 1 depicts two separate processing devices included in processing unit 110, more or fewer processing devices may be used. For example, in some embodiments, a single processing device may be used to accomplish the tasks of applications processor 180 and image processor 190. In other embodiments, these tasks may be performed by more than two processing devices.

[0022] Processing unit 110 may comprise various types of devices. For example, processing unit 110 may include various devices, such as a controller, an image preprocessor, a central processing unit (CPU), support circuits, digital signal processors, integrated circuits, memory, or any other types of devices for image processing and analysis. The image preprocessor may include a video processor for capturing, digitizing and processing the imagery from the image sensors. The CPU may comprise any number of microcontrollers or microprocessors. The support circuits may be any number of circuits generally well known in the art, including cache, power supply, clock and input-output circuits. The memory may store software that, when executed by the processor, controls the operation of the system. The memory may include databases and image processing software. The memory may comprise any number of random access memories, read only memories, flash memories, disk drives, optical storage, tape storage, removable storage and other types of storage. In one instance, the memory may be separate from the processing unit 110. In another instance, the memory may be integrated into the processing unit 110.

[0023] Each memory 140, 150 may include software instructions that when executed by a processor (e.g., applications processor 180 and/or image processor 190), may control operation of various aspects of system 100. These memory units may include various databases and image processing software. The memory units may include random access memory, read only memory, flash memory, disk drives, optical storage, tape storage, removable storage and/or any other types of storage. In some embodiments, memory units 140, 150 may be separate from the applications processor 180 and/or image processor 190. In other embodiments, these memory units may be integrated into applications processor 180 and/or image processor 190.

[0024] Position sensor 130 may include any type of device suitable for determining a location associated with at least one component of system 100. In some embodiments, position sensor 130 may include a GPS receiver. Such receivers can determine a user position and velocity by processing signals broadcasted by global positioning system satellites. Position information from position sensor 130 may be made available to applications processor 180 and/or image processor 190.

[0025] User interface 170 may include any device suitable for providing information to or for receiving inputs from one or more users of system 100. In some embodiments, user interface 170 may include user input devices, including, for example, a touchscreen, microphone, keyboard, pointer devices, track wheels, cameras, knobs, buttons, etc. With such input devices, a user may be able to provide information inputs or commands to system 100 by typing instructions or information, providing voice commands, selecting menu options on a screen using buttons, pointers, or eye-tracking capabilities, or through any other suitable techniques for communicating information to system 100.

[0026] User interface 170 may be equipped with one or more processing devices configured to provide and receive information to or from a user and process that information for use by, for example, applications processor 180. In some embodiments, such processing devices may execute instructions for recognizing and tracking eye movements, receiving and interpreting voice commands, recognizing and interpreting touches and/or gestures made on a touchscreen, responding to keyboard entries or menu selections, etc. In some embodiments, user interface 170 may include a display, speaker, tactile device, and/or any other devices for providing output information to a user.

[0027] Map database 160 may include any type of database for storing map data useful to system 100. In some embodiments, map database 160 may include data relating to the position, in a reference coordinate system, of various items, including roads, water features, geographic features, businesses, points of interest, restaurants, gas stations, etc. Map database 160 may store not only the locations of such items, but also descriptors relating to those items, including, for example, names associated with any of the stored features. In some embodiments, map database 160 may be physically located with other components of system 100. Alternatively or additionally, map database 160 or a portion thereof may be located remotely with respect to other components of system 100 (e.g., processing unit 110). In such embodiments, information from map database 160 may be

IPR2025-00944
Tesla EX1002 Page 1476

downloaded over a wired or wireless data connection to a network (e.g., over a cellular network and/or the Internet, etc.).

[0028] Image capture devices 122, 124, and 126 may each include any type of device suitable for capturing at least one image from an environment. Moreover, any number of image capture devices may be used to acquire images for input to the image processor. Some embodiments may include only a single image capture device, while other embodiments may include two, three, or even four or more image capture devices. Image capture devices 122, 124, and 126 will be further described with reference to FIGS. 2B-2E, below.

[0029] System 100, or various components thereof, may be incorporated into various different platforms. In some embodiments, system 100 may be included on a vehicle 200, as shown in FIG. 2A. For example, vehicle 200 may be equipped with a processing unit 110 and any of the other components of system 100, as described above relative to FIG. 1. While in some embodiments vehicle 200 may be equipped with only a single image capture device (e.g., camera), in other embodiments, such as those discussed in connection with FIGS. 2B-2E, multiple image capture devices may be used. For example, either of image capture devices 122 and 1 24 of vehicle 200, as shown in FIG. 2A, may be part of an ADAS (Advanced Driver Assistance Systems) imaging set.

[0030] The image capture devices included on vehicle 200 as part of the image acquisition unit 120 may be positioned at any suitable location. In some embodiments, as shown in FIGS. 2A-2E and 3A-3C, image capture device 122 may be located in the vicinity of the rearview mirror. This position may provide a line of sight similar to that of the driver of vehicle 200, which may aid in determining what is and is not visible to the driver. Image capture device 122 may be positioned at any location near the rearview mirror, but placing image capture device 122 on the driver side of the mirror may further aid in obtaining images representative of the driver's field of view and/or line of sight.

[0031] Other locations for the image capture devices of image acquisition unit 120 may also be used. For example, image capture device 124 may be located on or in a bumper of vehicle 200. Such a location may be especially suitable for image capture devices having a wide field of view. The line of sight of bumper-located image capture devices can be different from that of the driver and, therefore, the bumper image capture device and driver may not always see the same objects. The image capture devices (e.g., image capture devices 122, 124, and 126) may also be located in other locations. For example, the image capture devices may be located on or in one or both of the side mirrors of vehicle 200, on the roof of vehicle 200, on the hood of vehicle 200, on the trunk of vehicle 200, on the sides of vehicle 200, mounted on, positioned behind, or positioned in front of any of the windows of vehicle 200, and mounted in or near light figures on the front and/or back of vehicle 200, etc.

[0032] In addition to image capture devices, vehicle 200 may include various other components of system 100. For example, processing unit 110 may be included on vehicle 200 either integrated with or separate from an engine control unit (ECU) of the vehicle. Vehicle 200 may also be equipped with a position sensor 130, such as a GPS receiver and may also include a map database 160 and memory units 140 and 150.

[0033] FIG. 2A is a diagrammatic side view representation of an exemplary vehicle imaging system consistent with the disclosed embodiments. FIG. 2B is a diagrammatic top view illustration of the embodiment shown in FIG. 2A. As illustrated in FIG. 2B, the disclosed embodiments may include a vehicle 200 including in its body a system 100 with a first image capture device 122 positioned in the vicinity of the rearview mirror and/or near the driver of vehicle 200, a second image capture device 124 positioned on or in a bumper region (e.g., one of bumper regions 210) of vehicle 200, and a processing unit 110.

[0034] As illustrated in FIG. 2C, image capture devices 122 and 124 may both be positioned in the vicinity of the rearview mirror and/or near the driver of vehicle 200. Additionally, while two image capture devices 122 and 124 are shown in FIGS. 2B and 2C, it should be understood that other embodiments may include more than two image capture devices. For example, in the embodiments shown in FIGS. 2D and 2E, first, second, and third image capture devices 122, 124, and 126, are included in the system 100 of vehicle 200.

[0035] As illustrated in FIG. 2D, image capture device 122 may be positioned in the vicinity of the rearview mirror and/or near the driver of vehicle 200, and image capture devices 124 and 126 may be positioned on or in a bumper region (e.g., one of bumper regions 210) of vehicle 200. And as shown in FIG. 2E, image capture devices 122, 124, and 126 may be positioned in the vicinity of the rearview mirror and/or near the driver seat of vehicle 200. The disclosed embodiments are not limited to any particular number and configuration of the image capture devices, and the image capture devices may be positioned in any appropriate location within and/or on vehicle 200.

[0036] It is to be understood that the disclosed embodiments are not limited to vehicles and could be applied in other contexts. It is also to be understood that disclosed embodiments are not limited to a particular type of vehicle 200 and may be applicable to all types of vehicles including automobiles, trucks, trailers, and other types of vehicles.

[0037] The first image capture device 122 may include any suitable type of image capture device. Image capture device 122 may include an optical axis. In one instance, the image capture device 122 may include an Aptina M9V024 WVGA sensor with a global shutter. In other embodiments, image capture device 122 may provide a resolution of 1280x960 pixels and may include a rolling

IPR2025-00944
Tesla EX1002 Page 1477

shutter. Image capture device 122 may include various optical elements. In some embodiments one or more lenses may be included, for example, to provide a desired focal length and field of view for the image capture device. In some embodiments, image capture device 122 may be associated with a 6mm lens or a 12mm lens. In some embodiments, image capture device 122 may be configured to capture images having a desired field-of-view (FOV) 202, as illustrated in FIG. 2D. For example, image capture device 122 may be configured to have a regular FOV, such as within a range of 40 degrees to 56 degrees, including a 46 degree FOV, 50 degree FOV, 52 degree FOV, or greater. Alternatively, image capture device 122 may be configured to have a narrow FOV in the range of 23 to 40 degrees, such as a 28 degree FOV or 36 degree FOV. In addition, image capture device 122 may be configured to have a wide FOV in the range of 100 to 180 degrees. In some embodiments, image capture device 122 may include a wide angle bumper camera or one with up to a 180 degree FOV.

[0038] The first image capture device 122 may acquire a plurality of first images relative to a scene associated with the vehicle 200. Each of the plurality of first images may be acquired as a series of image scan lines, which may be captured using a rolling shutter. Each scan line may include a plurality of pixels.

[0039] The first image capture device 122 may have a scan rate associated with acquisition of each of the first series of image scan lines. The scan rate may refer to a rate at which an image sensor can acquire image data associated with each pixel included in a particular scan line.

[0040] Image capture devices 122, 124, and 126 may contain any suitable type and number of image sensors, including CCD sensors or CMOS sensors, for example. In one embodiment, a CMOS image sensor may be employed along with a rolling shutter, such that each pixel in a row is read one at a time, and scanning of the rows proceeds on a row-by-row basis until an entire image frame has been captured. In some embodiments, the rows may be captured sequentially from top to bottom relative to the frame.

[0041] The use of a rolling shutter may result in pixels in different rows being exposed and captured at different times, which may cause skew and other image artifacts in the captured image frame. On the other hand, when the image capture device 122 is configured to operate with a global or synchronous shutter, all of the pixels may be exposed for the same amount of time and during a common exposure period. As a result, the image data in a frame collected from a system employing a global shutter represents a snapshot of the entire FOV (such as FOV 202) at a particular time. In contrast, in a rolling shutter application, each row in a frame is exposed and data is capture at different times. Thus, moving objects may appear distorted in an image capture device having a rolling shutter. This phenomenon will be described in greater detail below.

[0042] The second image capture device 124 and the third image capturing device 126 may be any type of image capture device. Like the first image capture device 122, each of image capture devices 124 and 126 may include an optical axis. In one embodiment, each of image capture devices 124 and 1 26 may include an Aptina M9V024 WVGA sensor with a global shutter. Alternatively, each of image capture devices 124 and 126 may include a rolling shutter. Like image capture device 122, image capture devices 124 and 126 may be configured to include various lenses and optical elements. In some embodiments, lenses associated with image capture devices 124 and 126 may provide FOVs (such as FOVs 204 and 206) that are the same as, or narrower than, a FOV (such as FOV 202) associated with image capture device 122. For example, image capture devices 124 and 126 may have FOVs of 40 degrees, 30 degrees, 26 degrees, 23 degrees, 20 degrees, or less.

[0043] Image capture devices 124 and 126 may acquire a plurality of second and third images relative to a scene associated with the vehicle 200. Each of the plurality of second and third images may be acquired as a second and third series of image scan lines, which may be captured using a rolling shutter. Each scan line or row may have a plurality of pixels. Image capture devices 124 and 126 may have second and third scan rates associated with acquisition of each of image scan lines included in the second and third series.

[0044] Each image capture device 122, 124, and 126 may be positioned at any suitable position and orientation relative to vehicle 200. The relative positioning of the image capture devices 122, 124, and 126 may be selected to aid in fusing together the information acquired from the image capture devices. For example, in some embodiments, a FOV (such as FOV 204) associated with image capture device 124 may overlap partially or fully with a FOV (such as FOV 202) associated with image capture device 122 and a FOV (such as FOV 206) associated with image capture device 126.

[0045] Image capture devices 122, 124, and 126 may be located on vehicle 200 at any suitable relative heights. In one instance, there may be a height difference between the image capture devices 122, 124, and 126, which may provide sufficient parallax information to enable stereo analysis. For example, as shown in FIG. 2A, the two image capture devices 122 and 124 are at different heights. There may also be a lateral displacement difference between image capture devices 122, 124, and 126, giving additional parallax information for stereo analysis by processing unit 110, for example. The difference in the lateral displacement may be denoted by d_x, as shown in FIGS. 2C and 2D. In some embodiments, fore or aft displacement (e.g., range displacement) may exist between image capture devices 122, 124, and 126. For example, image capture device 122 may be located 0.5 to 2 meters or more behind image capture device 124 and/or image capture device 126. This type of displacement may enable one of the image capture devices to

IPR2025-00944
Tesla EX1002 Page 1478

cover potential blind spots of the other image capture device(s).

[0046] Image capture devices 122 may have any suitable resolution capability (e.g., number of pixels associated with the image sensor), and the resolution of the image sensor(s) associated with the image capture device 122 may be higher, lower, or the same as the resolution of the image sensor(s) associated with image capture devices 124 and 126. In some embodiments, the image sensor(s) associated with image capture device 122 and/or image capture devices 124 and 126 may have a resolution of 640 x 480, 1024 x 768, 1280 x 960, or any other suitable resolution.

[0047] The frame rate (e.g., the rate at which an image capture device acquires a set of pixel data of one image frame before moving on to capture pixel data associated with the next image frame) may be controllable. The frame rate associated with image capture device 122 may be higher, lower, or the same as the frame rate associated with image capture devices 124 and 126. The frame rate associated with image capture devices 122, 124, and 126 may depend on a variety of factors that may affect the timing of the frame rate. For example, one or more of image capture devices 122, 124, and 126 may include a selectable pixel delay period imposed before or after acquisition of image data associated with one or more pixels of an image sensor in image capture device 122, 124, and/or 126. Generally, image data corresponding to each pixel may be acquired according to a clock rate for the device (e.g., one pixel per clock cycle). Additionally, in embodiments including a rolling shutter, one or more of image capture devices 122, 124, and 126 may include a selectable horizontal blanking period imposed before or after acquisition of image data associated with a row of pixels of an image sensor in image capture device 122, 124, and/or 126. Further, one or more of image capture devices 122, 124, and/or 126 may include a selectable vertical blanking period imposed before or after acquisition of image data associated with an image frame of image capture device 122, 124, and 126.

[0048] These timing controls may enable synchronization of frame rates associated with image capture devices 122, 124, and 126, even where the line scan rates of each are different. Additionally, as will be discussed in greater detail below, these selectable timing controls, among other factors (e.g., image sensor resolution, maximum line scan rates, etc.) may enable synchronization of image capture from an area where the FOV of image capture device 122 overlaps with one or more FOVs of image capture devices 124 and 126, even where the field of view of image capture device 122 is different from the FOVs of image capture devices 124 and 126.

[0049] Frame rate timing in image capture device 122, 124, and 126 may depend on the resolution of the associated image sensors. For example, assuming similar line scan rates for both devices, if one device includes an image sensor having a resolution of 640 x 480 and another device includes an image sensor with a resolution of 1280 x 960, then more time will be required to acquire a frame of image data from the sensor having the higher resolution.

[0050] Another factor that may affect the timing of image data acquisition in image capture devices 122, 124, and 126 is the maximum line scan rate. For example, acquisition of a row of image data from an image sensor included in image capture device 122, 124, and 126 will require some minimum amount of time. Assuming no pixel delay periods are added, this minimum amount of time for acquisition of a row of image data will be related to the maximum line scan rate for a particular device. Devices that offer higher maximum line scan rates have the potential to provide higher frame rates than devices with lower maximum line scan rates. In some embodiments, one or more of image capture devices 124 and 126 may have a maximum line scan rate that is higher than a maximum line scan rate associated with image capture device 122. In some embodiments, the maximum line scan rate of image capture device 124 and/or 126 may be 1.25, 1.5, 1.75, or 2 times or more than a maximum line scan rate of image capture device 122.

[0051] In another embodiment, image capture devices 122, 124, and 126 may have the same maximum line scan rate, but image capture device 122 may be operated at a scan rate less than or equal to its maximum scan rate. The system may be configured such that one or more of image capture devices 124 and 126 operate at a line scan rate that is equal to the line scan rate of image capture device 122. In other instances, the system may be configured such that the line scan rate of image capture device 124 and/or image capture device 126 may be 1.25, 1.5, 1.75, or 2 times or more than the line scan rate of image capture device 122.

[0052] In some embodiments, image capture devices 122, 124, and 126 may be asymmetric. That is, they may include cameras having different fields of view (FOV) and focal lengths. The fields of view of image capture devices 122, 124, and 126 may include any desired area relative to an environment of vehicle 200, for example. In some embodiments, one or more of image capture devices 122, 124, and 126 may be configured to acquire image data from an environment in front of vehicle 200, behind vehicle 200, to the sides of vehicle 200, or combinations thereof.

[0053] Further, the focal length associated with each image capture device 122, 124, and/or 126 may be selectable (e.g., by inclusion of appropriate lenses etc.) such that each device acquires images of objects at a desired distance range relative to vehicle 200. For example, in some embodiments image capture devices 122, 124, and 126 may acquire images of close-up objects within a few meters from the vehicle. Image capture devices 122, 124, and 126 may also be configured to acquire images of objects at ranges more distant from the vehicle (e.g., 25 m, 50 m, 100 m, 150 m, or more). Further, the focal lengths of image capture devices 122, 124, and 126 may be selected such that one image cap-

IPR2025-00944
Tesla EX1002 Page 1479

ture device (e.g., image capture device 122) can acquire images of objects relatively close to the vehicle (e.g., within 10 m or within 20 m) while the other image capture devices (e.g., image capture devices 124 and 126) can acquire images of more distant objects (e.g., greater than 20 m, 50 m, 100 m, 150 m, etc.) from vehicle 200.

[0054] According to some embodiments, the FOV of one or more image capture devices 122, 124, and 126 may have a wide angle. For example, it may be advantageous to have a FOV of 140 degrees, especially for image capture devices 122, 124, and 126 that may be used to capture images of the area in the vicinity of vehicle 200. For example, image capture device 122 may be used to capture images of the area to the right or left of vehicle 200 and, in such embodiments, it may be desirable for image capture device 122 to have a wide FOV (e.g., at least 140 degrees).

[0055] The field of view associated with each of image capture devices 122, 124, and 126 may depend on the respective focal lengths. For example, as the focal length increases, the corresponding field of view decreases.

[0056] Image capture devices 122, 124, and 126 may be configured to have any suitable fields of view. In one particular example, image capture device 122 may have a horizontal FOV of 46 degrees, image capture device 124 may have a horizontal FOV of 23 degrees, and image capture device 126 may have a horizontal FOV in between 23 and 46 degrees. In another instance, image capture device 122 may have a horizontal FOV of 52 degrees, image capture device 124 may have a horizontal FOV of 26 degrees, and image capture device 126 may have a horizontal FOV in between 26 and 52 degrees. In some embodiments, a ratio of the FOV of image capture device 122 to the FOVs of image capture device 124 and/or image capture device 126 may vary from 1.5 to 2.0. In other embodiments, this ratio may vary between 1.25 and 2.25.

[0057] System 100 may be configured so that a field of view of image capture device 122 overlaps, at least partially or fully, with a field of view of image capture device 124 and/or image capture device 126. In some embodiments, system 100 may be configured such that the fields of view of image capture devices 124 and 126, for example, fall within (e.g., are narrower than) and share a common center with the field of view of image capture device 122. In other embodiments, the image capture devices 122, 124, and 126 may capture adjacent FOVs or may have partial overlap in their FOVs. In some embodiments, the fields of view of image capture devices 122, 124, and 126 may be aligned such that a center of the narrower FOV image capture devices 124 and/or 126 may be located in a lower half of the field of view of the wider FOV device 122.

[0058] FIG. 2F is a diagrammatic representation of exemplary vehicle control systems, consistent with the disclosed embodiments. As indicated in FIG. 2F, vehicle 200 may include throttling system 220, braking system 230, and steering system 240. System 100 may provide inputs (e.g., control signals) to one or more of throttling system 220, braking system 230, and steering system 240 over one or more data links (e.g., any wired and/or wireless link or links for transmitting data). For example, based on analysis of images acquired by image capture devices 122, 124, and/or 126, system 100 may provide control signals to one or more of throttling system 220, braking system 230, and steering system 240 to navigate vehicle 200 (e.g., by causing an acceleration, a turn, a lane shift, etc.). Further, system 100 may receive inputs from one or more of throttling system 220, braking system 230, and steering system 24 indicating operating conditions of vehicle 200 (e.g., speed, whether vehicle 200 is braking and/or turning, etc.). Further details are provided in connection with FIGS. 4-7, below.

[0059] As shown in FIG. 3A, vehicle 200 may also include a user interface 170 for interacting with a driver or a passenger of vehicle 200. For example, user interface 170 in a vehicle application may include a touch screen 320, knobs 330, buttons 340, and a microphone 350. A driver or passenger of vehicle 200 may also use handles (e.g., located on or near the steering column of vehicle 200 including, for example, turn signal handles), buttons (e.g., located on the steering wheel of vehicle 200), and the like, to interact with system 100. In some embodiments, microphone 350 may be positioned adjacent to a rearview mirror 310. Similarly, in some embodiments, image capture device 122 may be located near rearview mirror 310. In some embodiments, user interface 170 may also include one or more speakers 360 (e.g., speakers of a vehicle audio system). For example, system 100 may provide various notifications (e.g., alerts) via speakers 360.

[0060] FIGS. 3B-3D are illustrations of an exemplary camera mount 370 configured to be positioned behind a rearview mirror (e.g., rearview mirror 310) and against a vehicle windshield, consistent with disclosed embodiments. As shown in FIG. 3B, camera mount 370 may include image capture devices 122, 124, and 126. Image capture devices 124 and 126 may be positioned behind a glare shield 380, which may be flush against the vehicle windshield and include a composition of film and/or anti-reflective materials. For example, glare shield 380 may be positioned such that it aligns against a vehicle windshield having a matching slope. In some embodiments, each of image capture devices 122, 124, and 126 may be positioned behind glare shield 380, as depicted, for example, in FIG. 3D. The disclosed embodiments are not limited to any particular configuration of image capture devices 122, 124, and 126, camera mount 370, and glare shield 380. FIG. 3C is an illustration of camera mount 370 shown in FIG. 3B from a front perspective.

[0061] As will be appreciated by a person skilled in the art having the benefit of this disclosure, numerous variations and/or modifications may be made to the foregoing disclosed embodiments. For example, not all components are essential for the operation of system 100. Further, any component may be located in any appropriate

IPR2025-00944
Tesla EX1002 Page 1480

part of system 100 and the components may be rearranged into a variety of configurations while providing the functionality of the disclosed embodiments. Therefore, the foregoing configurations are examples and, regardless of the configurations discussed above, system 100 can provide a wide range of functionality to analyze the surroundings of vehicle 200 and navigate vehicle 200 in response to the analysis.

[0062] As discussed below in further detail and consistent with various disclosed embodiments, system 100 may provide a variety of features related to autonomous driving and/or driver assist technology. For example, system 100 may analyze image data, position data (e.g., GPS location information), map data, speed data, and/or data from sensors included in vehicle 200. System 100 may collect the data for analysis from, for example, image acquisition unit 120, position sensor 130, and other sensors. Further, system 100 may analyze the collected data to determine whether or not vehicle 200 should take a certain action, and then automatically take the determined action without human intervention. For example, when vehicle 200 navigates without human intervention, system 100 may automatically control the braking, acceleration, and/or steering of vehicle 200 (e.g., by sending control signals to one or more of throttling system 220, braking system 230, and steering system 240). Further, system 100 may analyze the collected data and issue warnings and/or alerts to vehicle occupants based on the analysis of the collected data. Additional details regarding the various embodiments that are provided by system 100 are provided below.

Forward-Facing Multi-Imaging System

[0063] As discussed above, system 100 may provide drive assist functionality that uses a multi-camera system. The multi-camera system may use one or more cameras facing in the forward direction of a vehicle. In other embodiments, the multi-camera system may include one or more cameras facing to the side of a vehicle or to the rear of the vehicle. In one embodiment, for example, system 100 may use a two-camera imaging system, where a first camera and a second camera (e.g., image capture devices 122 and 124) may be positioned at the front and/or the sides of a vehicle (e.g., vehicle 200). The first camera may have a field of view that is greater than, less than, or partially overlapping with, the field of view of the second camera. In addition, the first camera may be connected to a first image processor to perform monocular image analysis of images provided by the first camera, and the second camera may be connected to a second image processor to perform monocular image analysis of images provided by the second camera. The outputs (e.g., processed information) of the first and second image processors may be combined. In some embodiments, the second image processor may receive images from both the first camera and second camera to perform stereo analysis. In another embodiment, system 100 may use a three-camera imaging system where each of the cameras has a different field of view. Such a system may, therefore, make decisions based on information derived from objects located at varying distances both forward and to the sides of the vehicle. References to monocular image analysis may refer to instances where image analysis is performed based on images captured from a single point of view (e.g., from a single camera). Stereo image analysis may refer to instances where image analysis is performed based on two or more images captured with one or more variations of an image capture parameter. For example, captured images suitable for performing stereo image analysis may include images captured: from two or more different positions, from different fields of view, using different focal lengths, along with parallax information, etc.

[0064] For example, in one embodiment, system 100 may implement a three camera configuration using image capture devices 122-126. In such a configuration, image capture device 122 may provide a narrow field of view (e.g., 34 degrees, or other values selected from a range of about 20 to 45 degrees, etc.), image capture device 124 may provide a wide field of view (e.g., 150 degrees or other values selected from a range of about 100 to about 180 degrees), and image capture device 126 may provide an intermediate field of view (e.g., 46 degrees or other values selected from a range of about 35 to about 60 degrees). In some embodiments, image capture device 126 may act as a main or primary camera. Image capture devices 122-126 may be positioned behind rearview mirror 310 and positioned substantially side-by-side (e.g., 6 cm apart). Further, in some embodiments, as discussed above, one or more of image capture devices 122-126 may be mounted behind glare shield 380 that is flush with the windshield of vehicle 200. Such shielding may act to minimize the impact of any reflections from inside the car on image capture devices 122-126.

[0065] In another embodiment, as discussed above in connection with FIGS. 3B and 3C, the wide field of view camera (e.g., image capture device 124 in the above example) may be mounted lower than the narrow and main field of view cameras (e.g., image devices 122 and 126 in the above example). This configuration may provide a free line of sight from the wide field of view camera. To reduce reflections, the cameras may be mounted close to the windshield of vehicle 200, and may include polarizers on the cameras to damp reflected light.

[0066] A three camera system may provide certain performance characteristics. For example, some embodiments may include an ability to validate the detection of objects by one camera based on detection results from another camera. In the three camera configuration discussed above, processing unit 110 may include, for example, three processing devices (e.g., three EyeQ series of processor chips, as discussed above), with each processing device dedicated to processing images captured by one or more of image capture devices 122-126.

IPR2025-00944
Tesla EX1002 Page 1481

[0067] In a three camera system, a first processing device may receive images from both the main camera and the narrow field of view camera, and perform vision processing of the narrow FOV camera to, for example, detect other vehicles, pedestrians, lane marks, traffic signs, traffic lights, and other road objects. Further, the first processing device may calculate a disparity of pixels between the images from the main camera and the narrow camera and create a 3D reconstruction of the environment of vehicle 200. The first processing device may then combine the 3D reconstruction with 3D map data or with 3D information calculated based on information from another camera.

[0068] The second processing device may receive images from main camera and perform vision processing to detect other vehicles, pedestrians, lane marks, traffic signs, traffic lights, and other road objects. Additionally, the second processing device may calculate a camera displacement and, based on the displacement, calculate a disparity of pixels between successive images and create a 3D reconstruction of the scene (e.g., a structure from motion). The second processing device may send the structure from motion based 3D reconstruction to the first processing device to be combined with the stereo 3D images.

[0069] The third processing device may receive images from the wide FOV camera and process the images to detect vehicles, pedestrians, lane marks, traffic signs, traffic lights, and other road objects. The third processing device may further execute additional processing instructions to analyze images to identify objects moving in the image, such as vehicles changing lanes, pedestrians, etc.

[0070] In some embodiments, having streams of image-based information captured and processed independently may provide an opportunity for providing redundancy in the system. Such redundancy may include, for example, using a first image capture device and the images processed from that device to validate and/or supplement information obtained by capturing and processing image information from at least a second image capture device.

[0071] In some embodiments, system 100 may use two image capture devices (e.g., image capture devices 122 and 124) in providing navigation assistance for vehicle 200 and use a third image capture device (e.g., image capture device 126) to provide redundancy and validate the analysis of data received from the other two image capture devices. For example, in such a configuration, image capture devices 122 and 124 may provide images for stereo analysis by system 100 for navigating vehicle 200, while image capture device 126 may provide images for monocular analysis by system 100 to provide redundancy and validation of information obtained based on images captured from image capture device 122 and/or image capture device 124. That is, image capture device 126 (and a corresponding processing device) may be considered to provide a redundant sub-system for providing a check on the analysis derived from image capture devices 122 and 124 (e.g., to provide an automatic emergency braking (AEB) system).

[0072] One of skill in the art will recognize that the above camera configurations, camera placements, number of cameras, camera locations, etc., are examples only. These components and others described relative to the overall system may be assembled and used in a variety of different configurations without departing from the scope of the disclosed embodiments. Further details regarding usage of a multi-camera system to provide driver assist and/or autonomous vehicle functionality follow below.

[0073] FIG. 4 is an exemplary functional block diagram of memory 140 and/or 150, which may be stored/programmed with instructions for performing one or more operations consistent with the disclosed embodiments. Although the following refers to memory 140, one of skill in the art will recognize that instructions may be stored in memory 140 and/or 150.

[0074] As shown in FIG. 4, memory 140 may store a monocular image analysis module 402, a stereo image analysis module 404, a velocity and acceleration module 406, and a navigational response module 408. The disclosed embodiments are not limited to any particular configuration of memory 140. Further, application processor 180 and/or image processor 190 may execute the instructions stored in any of modules 402-408 included in memory 140. One of skill in the art will understand that references in the following discussions to processing unit 110 may refer to application processor 180 and image processor 190 individually or collectively. Accordingly, steps of any of the following processes may be performed by one or more processing devices.

[0075] In one embodiment, monocular image analysis module 402 may store instructions (such as computer vision software) which, when executed by processing unit 110, performs monocular image analysis of a set of images acquired by one of image capture devices 122, 124, and 126. In some embodiments, processing unit 110 may combine information from a set of images with additional sensory information (e.g., information from radar) to perform the monocular image analysis. As described in connection with FIGS. 5A-5D below, monocular image analysis module 402 may include instructions for detecting a set of features within the set of images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, hazardous objects, and any other feature associated with an environment of a vehicle. Based on the analysis, system 100 (e.g., via processing unit 110) may cause one or more navigational responses in vehicle 200, such as a turn, a lane shift, a change in acceleration, and the like, as discussed below in connection with navigational response module 408.

[0076] In one embodiment, stereo image analysis module 404 may store instructions (such as computer vision software) which, when executed by processing unit 110, performs stereo image analysis of first and second

IPR2025-00944
Tesla EX1002 Page 1482

sets of images acquired by a combination of image capture devices selected from any of image capture devices 122, 124, and 126. In some embodiments, processing unit 110 may combine information from the first and second sets of images with additional sensory information (e.g., information from radar) to perform the stereo image analysis. For example, stereo image analysis module 404 may include instructions for performing stereo image analysis based on a first set of images acquired by image capture device 124 and a second set of images acquired by image capture device 126. As described in connection with FIG. 6 below, stereo image analysis module 404 may include instructions for detecting a set of features within the first and second sets of images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, hazardous objects, and the like. Based on the analysis, processing unit 110 may cause one or more navigational responses in vehicle 200, such as a turn, a lane shift, a change in acceleration, and the like, as discussed below in connection with navigational response module 408.

[0077] In one embodiment, velocity and acceleration module 406 may store software configured to analyze data received from one or more computing and electromechanical devices in vehicle 200 that are configured to cause a change in velocity and/or acceleration of vehicle 200. For example, processing unit 110 may execute instructions associated with velocity and acceleration module 406 to calculate a target speed for vehicle 200 based on data derived from execution of monocular image analysis module 402 and/or stereo image analysis module 404. Such data may include, for example, a target position, velocity, and/or acceleration, the position and/or speed of vehicle 200 relative to a nearby vehicle, pedestrian, or road object, position information for vehicle 200 relative to lane markings of the road, and the like. In addition, processing unit 110 may calculate a target speed for vehicle 200 based on sensory input (e.g., information from radar) and input from other systems of vehicle 200, such as throttling system 220, braking system 230, and/or steering system 240 of vehicle 200. Based on the calculated target speed, processing unit 110 may transmit electronic signals to throttling system 220, braking system 230, and/or steering system 240 of vehicle 200 to trigger a change in velocity and/or acceleration by, for example, physically depressing the brake or easing up off the accelerator of vehicle 200.

[0078] In one embodiment, navigational response module 408 may store software executable by processing unit 110 to determine a desired navigational response based on data derived from execution of monocular image analysis module 402 and/or stereo image analysis module 404. Such data may include position and speed information associated with nearby vehicles, pedestrians, and road objects, target position information for vehicle 200, and the like. Additionally, in some embodiments, the navigational response may be based (partially or fully) on map data, a predetermined position of vehicle 200, and/or a relative velocity or a relative acceleration between vehicle 200 and one or more objects detected from execution of monocular image analysis module 402 and/or stereo image analysis module 404. Navigational response module 408 may also determine a desired navigational response based on sensory input (e.g., information from radar) and inputs from other systems of vehicle 200, such as throttling system 220, braking system 230, and steering system 240 of vehicle 200. Based on the desired navigational response, processing unit 110 may transmit electronic signals to throttling system 220, braking system 230, and steering system 240 of vehicle 200 to trigger a desired navigational response by, for example, turning the steering wheel of vehicle 200 to achieve a rotation of a predetermined angle. In some embodiments, processing unit 110 may use the output of navigational response module 408 (e.g., the desired navigational response) as an input to execution of velocity and acceleration module 406 for calculating a change in speed of vehicle 200.

[0079] FIG. 5A is a flowchart showing an exemplary process 500A for causing one or more navigational responses based on monocular image analysis, consistent with disclosed embodiments. At step 510, processing unit 110 may receive a plurality of images via data interface 128 between processing unit 110 and image acquisition unit 120. For instance, a camera included in image acquisition unit 120 (such as image capture device 122 having field of view 202) may capture a plurality of images of an area forward of vehicle 200 (or to the sides or rear of a vehicle, for example) and transmit them over a data connection (e.g., digital, wired, USB, wireless, Bluetooth, etc.) to processing unit 1 10. Processing unit 110 may execute monocular image analysis module 402 to analyze the plurality of images at step 520, as described in further detail in connection with FIGS. 5B-5D below. By performing the analysis, processing unit 1 10 may detect a set of features within the set of images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, and the like.

[0080] Processing 3unit 110 may also execute monocular image analysis module 402 to detect various road hazards at step 520, such as, for example, parts of a truck tire, fallen road signs, loose cargo, small animals, and the like. Road hazards may vary in structure, shape, size, and color, which may make detection of such hazards more challenging. In some embodiments, processing unit 110 may execute monocular image analysis module 402 to perform multi-frame analysis on the plurality of images to detect road hazards. For example, processing unit 110 may estimate camera motion between consecutive image frames and calculate the disparities in pixels between the frames to construct a 3D-map of the road. Processing unit 110 may then use the 3D-map to detect the road surface, as well as hazards existing above the road surface.

[0081] At step 530, processing unit 110 may execute navigational response module 408 to cause one or more

IPR2025-00944
Tesla EX1002 Page 1483

navigational responses in vehicle 200 based on the analysis performed at step 520 and the techniques as described above in connection with FIG. 4. Navigational responses may include, for example, a turn, a lane shift, a change in acceleration, and the like. In some embodiments, processing unit 110 may use data derived from execution of velocity and acceleration module 406 to cause the one or more navigational responses. Additionally, multiple navigational responses may occur simultaneously, in sequence, or any combination thereof. For instance, processing unit 110 may cause vehicle 200 to shift one lane over and then accelerate by, for example, sequentially transmitting control signals to steering system 240 and throttling system 220 of vehicle 200. Alternatively, processing unit 110 may cause vehicle 200 to brake while at the same time shifting lanes by, for example, simultaneously transmitting control signals to braking system 230 and steering system 240 of vehicle 200.

[0082] FIG. 5B is a flowchart showing an exemplary process 500B for detecting one or more vehicles and/or pedestrians in a set of images, consistent with disclosed embodiments. Processing unit 110 may execute monocular image analysis module 402 to implement process 500B. At step 540, processing unit 1 10 may determine a set of candidate objects representing possible vehicles and/or pedestrians. For example, processing unit 110 may scan one or more images, compare the images to one or more predetermined patterns, and identify within each image possible locations that may contain objects of interest (e.g., vehicles, pedestrians, or portions thereof). The predetermined patterns may be designed in such a way to achieve a high rate of "false hits" and a low rate of "misses." For example, processing unit 110 may use a low threshold of similarity to predetermined patterns for identifying candidate objects as possible vehicles or pedestrians. Doing so may allow processing unit 110 to reduce the probability of missing (e.g., not identifying) a candidate object representing a vehicle or pedestrian.

[0083] At step 542, processing unit 1 10 may filter the set of candidate objects to exclude certain candidates (e.g., irrelevant or less relevant objects) based on classification criteria. Such criteria may be derived from various properties associated with object types stored in a database (e.g., a database stored in memory 140). Properties may include object shape, dimensions, texture, position (e.g., relative to vehicle 200), and the like. Thus, processing unit 110 may use one or more sets of criteria to reject false candidates from the set of candidate objects.

[0084] At step 544, processing unit 110 may analyze multiple frames of images to determine whether objects in the set of candidate objects represent vehicles and/or pedestrians. For example, processing unit 110 may track a detected candidate object across consecutive frames and accumulate frame-by-frame data associated with the detected object (e.g., size, position relative to vehicle 200, etc.). Additionally, processing unit 110 may estimate parameters for the detected object and compare the object's frame-by-frame position data to a predicted position.

[0085] At step 546, processing unit 110 may construct a set of measurements for the detected objects. Such measurements may include, for example, position, velocity, and acceleration values (relative to vehicle 200) associated with the detected objects. In some embodiments, processing unit 110 may construct the measurements based on estimation techniques using a series of time-based observations such as Kalman filters or linear quadratic estimation (LQE), and/or based on available modeling data for different object types (e.g., cars, trucks, pedestrians, bicycles, road signs, etc.). The Kalman filters may be based on a measurement of an object's scale, where the scale measurement is proportional to a time to collision (e.g., the amount of time for vehicle 200 to reach the object). Thus, by performing steps 540-546, processing unit 110 may identify vehicles and pedestrians appearing within the set of captured images and derive information (e.g., position, speed, size) associated with the vehicles and pedestrians. Based on the identification and the derived information, processing unit 110 may cause one or more navigational responses in vehicle 200, as described in connection with FIG. 5A, above.

[0086] At step 548, processing unit 110 may perform an optical flow analysis of one or more images to reduce the probabilities of detecting a "false hit" and missing a candidate object that represents a vehicle or pedestrian. The optical flow analysis may refer to, for example, analyzing motion patterns relative to vehicle 200 in the one or more images associated with other vehicles and pedestrians, and that are distinct from road surface motion. Processing unit 1 10 may calculate the motion of candidate objects by observing the different positions of the objects across multiple image frames, which are captured at different times. Processing unit 110 may use the position and time values as inputs into mathematical models for calculating the motion of the candidate objects. Thus, optical flow analysis may provide another method of detecting vehicles and pedestrians that are nearby vehicle 200. Processing unit 110 may perform optical flow analysis in combination with steps 540-546 to provide redundancy for detecting vehicles and pedestrians and increase the reliability of system 100.

[0087] FIG. 5C is a flowchart showing an exemplary process 500C for detecting road marks and/or lane geometry information in a set of images, consistent with disclosed embodiments. Processing unit 110 may execute monocular image analysis module 402 to implement process 500C. At step 550, processing unit 1 10 may detect a set of objects by scanning one or more images. To detect segments of lane markings, lane geometry information, and other pertinent road marks, processing unit 110 may filter the set of objects to exclude those determined to be irrelevant (e.g., minor potholes, small rocks, etc.). At step 552, processing unit 110 may group together the segments detected in step 550 belonging to the same road mark or lane mark. Based on the grouping,

IPR2025-00944
Tesla EX1002 Page 1484

processing unit 110 may develop a model to represent the detected segments, such as a mathematical model.

[0088] At step 554, processing unit 110 may construct a set of measurements associated with the detected segments. In some embodiments, processing unit 110 may create a projection of the detected segments from the image plane onto the real-world plane. The projection may be characterized using a 3rd-degree polynomial having coefficients corresponding to physical properties such as the position, slope, curvature, and curvature derivative of the detected road. In generating the projection, processing unit 110 may take into account changes in the road surface, as well as pitch and roll rates associated with vehicle 200. In addition, processing unit 110 may model the road elevation by analyzing position and motion cues present on the road surface. Further, processing unit 110 may estimate the pitch and roll rates associated with vehicle 200 by tracking a set of feature points in the one or more images.

[0089] At step 556, processing unit 110 may perform multi-frame analysis by, for example, tracking the detected segments across consecutive image frames and accumulating frame-by-frame data associated with detected segments. As processing unit 110 performs multi-frame analysis, the set of measurements constructed at step 554 may become more reliable and associated with an increasingly higher confidence level. Thus, by performing steps 550-556, processing unit 110 may identify road marks appearing within the set of captured images and derive lane geometry information. Based on the identification and the derived information, processing unit 110 may cause one or more navigational responses in vehicle 200, as described in connection with FIG. 5A, above.

[0090] At step 558, processing unit 110 may consider additional sources of information to further develop a safety model for vehicle 200 in the context of its surroundings. Processing unit 110 may use the safety model to define a context in which system 100 may execute autonomous control of vehicle 200 in a safe manner. To develop the safety model, in some embodiments, processing unit 110 may consider the position and motion of other vehicles, the detected road edges and barriers, and/or general road shape descriptions extracted from map data (such as data from map database 160). By considering additional sources of information, processing unit 110 may provide redundancy for detecting road marks and lane geometry and increase the reliability of system 100.

[0091] FIG. 5D is a flowchart showing an exemplary process 500D for detecting traffic lights in a set of images, consistent with disclosed embodiments. Processing unit 1 10 may execute monocular image analysis module 402 to implement process 500D. At step 560, processing unit 110 may scan the set of images and identify objects appearing at locations in the images likely to contain traffic lights. For example, processing unit 110 may filter the identified objects to construct a set of candidate objects, excluding those objects unlikely to correspond to traffic lights. The filtering may be done based on various properties associated with traffic lights, such as shape, dimensions, texture, position (e.g., relative to vehicle 200), and the like. Such properties may be based on multiple examples of traffic lights and traffic control signals and stored in a database. In some embodiments, processing unit 110 may perform multi-frame analysis on the set of candidate objects reflecting possible traffic lights. For example, processing unit 110 may track the candidate objects across consecutive image frames, estimate the real-world position of the candidate objects, and filter out those objects that are moving (which are unlikely to be traffic lights). In some embodiments, processing unit 110 may perform color analysis on the candidate objects and identify the relative position of the detected colors appearing inside possible traffic lights.

[0092] At step 562, processing unit 110 may analyze the geometry of a junction. The analysis may be based on any combination of: (i) the number of lanes detected on either side of vehicle 200, (ii) markings (such as arrow marks) detected on the road, and (iii) descriptions of the junction extracted from map data (such as data from map database 160). Processing unit 110 may conduct the analysis using information derived from execution of monocular analysis module 402. In addition, Processing unit 110 may determine a correspondence between the traffic lights detected at step 560 and the lanes appearing near vehicle 200.

[0093] As vehicle 200 approaches the junction, at step 564, processing unit 110 may update the confidence level associated with the analyzed junction geometry and the detected traffic lights. For instance, the number of traffic lights estimated to appear at the junction as compared with the number actually appearing at the junction may impact the confidence level. Thus, based on the confidence level, processing unit 110 may delegate control to the driver of vehicle 200 in order to improve safety conditions. By performing steps 560-564, processing unit 110 may identify traffic lights appearing within the set of captured images and analyze junction geometry information. Based on the identification and the analysis, processing unit 1 10 may cause one or more navigational responses in vehicle 200, as described in connection with FIG. 5A, above.

[0094] FIG. 5E is a flowchart showing an exemplary process 500E for causing one or more navigational responses in vehicle 200 based on a vehicle path, consistent with the disclosed embodiments. At step 570, processing unit 110 may construct an initial vehicle path associated with vehicle 200. The vehicle path may be represented using a set of points expressed in coordinates (x, z), and the distance d_i between two points in the set of points may fall in the range of 1 to 5 meters. In one embodiment, processing unit 110 may construct the initial vehicle path using two polynomials, such as left and right road polynomials. Processing unit 110 may calculate the geometric midpoint between the two polynomials and offset each point included in the resultant ve-

IPR2025-00944
Tesla EX1002 Page 1485

hicle path by a predetermined offset (e.g., a smart lane offset), if any (an offset of zero may correspond to travel in the middle of a lane). The offset may be in a direction perpendicular to a segment between any two points in the vehicle path. In another embodiment, processing unit 110 may use one polynomial and an estimated lane width to offset each point of the vehicle path by half the estimated lane width plus a predetermined offset (e.g., a smart lane offset).

[0095] At step 572, processing unit 110 may update the vehicle path constructed at step 570. Processing unit 1 10 may reconstruct the vehicle path constructed at step 570 using a higher resolution, such that the distance d_k between two points in the set of points representing the vehicle path is less than the distance d_i described above. For example, the distance d_k may fall in the range of 0.1 to 0.3 meters. Processing unit 110 may reconstruct the vehicle path using a parabolic spline algorithm, which may yield a cumulative distance vector S corresponding to the total length of the vehicle path (i.e., based on the set of points representing the vehicle path).

[0096] At step 574, processing unit 110 may determine a look-ahead point (expressed in coordinates as (x_l, z_l)) based on the updated vehicle path constructed at step 572. Processing unit 1 10 may extract the look-ahead point from the cumulative distance vector S, and the look-ahead point may be associated with a look-ahead distance and look-ahead time. The look-ahead distance, which may have a lower bound ranging from 1 0 to 20 meters, may be calculated as the product of the speed of vehicle 200 and the look-ahead time. For example, as the speed of vehicle 200 decreases, the look-ahead distance may also decrease (e.g., until it reaches the lower bound). The look-ahead time, which may range from 0.5 to 1.5 seconds, may be inversely proportional to the gain of one or more control loops associated with causing a navigational response in vehicle 200, such as the heading error tracking control loop. For example, the gain of the heading error tracking control loop may depend on the bandwidth of a yaw rate loop, a steering actuator loop, car lateral dynamics, and the like. Thus, the higher the gain of the heading error tracking control loop, the lower the look-ahead time.

[0097] At step 576, processing unit 110 may determine a heading error and yaw rate command based on the look-ahead point determined at step 574. Processing unit 110 may determine the heading error by calculating the arctangent of the look-ahead point, e.g., arctan (x_l/z_l). Processing unit 110 may determine the yaw rate command as the product of the heading error and a high-level control gain. The high-level control gain may be equal to: (2 / look-ahead time), if the look-ahead distance is not at the lower bound. Otherwise, the high-level control gain may be equal to: (2 * speed of vehicle 200 / look-ahead distance).

[0098] FIG. 5F is a flowchart showing an exemplary process 500F for determining whether a leading vehicle is changing lanes, consistent with the disclosed embodiments. At step 580, processing unit 110 may determine navigation information associated with a leading vehicle (e.g., a vehicle traveling ahead of vehicle 200). For example, processing unit 110 may determine the position, velocity (e.g., direction and speed), and/or acceleration of the leading vehicle, using the techniques described in connection with FIGS. 5A and 5B, above. Processing unit 110 may also determine one or more road polynomials, a look-ahead point (associated with vehicle 200), and/or a snail trail (e.g., a set of points describing a path taken by the leading vehicle), using the techniques described in connection with FIG. 5E, above.

[0099] At step 582, processing unit 110 may analyze the navigation information determined at step 580. In one embodiment, processing unit 110 may calculate the distance between a snail trail and a road polynomial (e.g., along the trail). If the variance of this distance along the trail exceeds a predetermined threshold (for example, 0.1 to 0.2 meters on a straight road, 0.3 to 0.4 meters on a moderately curvy road, and 0.5 to 0.6 meters on a road with sharp curves), processing unit 110 may determine that the leading vehicle is likely changing lanes. In the case where multiple vehicles are detected traveling ahead of vehicle 200, processing unit 110 may compare the snail trails associated with each vehicle. Based on the comparison, processing unit 110 may determine that a vehicle whose snail trail does not match with the snail trails of the other vehicles is likely changing lanes. Processing unit 110 may additionally compare the curvature of the snail trail (associated with the leading vehicle) with the expected curvature of the road segment in which the leading vehicle is traveling. The expected curvature may be extracted from map data (e.g., data from map database 160), from road polynomials, from other vehicles' snail trails, from prior knowledge about the road, and the like. If the difference in curvature of the snail trail and the expected curvature of the road segment exceeds a predetermined threshold, processing unit 110 may determine that the leading vehicle is likely changing lanes.

[0100] In another embodiment, processing unit 110 may compare the leading vehicle's instantaneous position with the look-ahead point (associated with vehicle 200) over a specific period of time (e.g., 0.5 to 1.5 seconds). If the distance between the leading vehicle's instantaneous position and the look-ahead point varies during the specific period of time, and the cumulative sum of variation exceeds a predetermined threshold (for example, 0.3 to 0.4 meters on a straight road, 0.7 to 0.8 meters on a moderately curvy road, and 1.3 to 1.7 meters on a road with sharp curves), processing unit 110 may determine that the leading vehicle is likely changing lanes. In another embodiment, processing unit 110 may analyze the geometry of the snail trail by comparing the lateral distance traveled along the trail with the expected curvature of the snail trail. The expected radius of curvature may be determined according to the calculation: $(\delta_z^2 + \delta_x^2) / 2 / (\delta_x)$, where δ_x represents the lateral distance

IPR2025-00944
Tesla EX1002 Page 1486

traveled and δ_z represents the longitudinal distance traveled. If the difference between the lateral distance traveled and the expected curvature exceeds a predetermined threshold (e.g., 500 to 700 meters), processing unit 110 may determine that the leading vehicle is likely changing lanes. In another embodiment, processing unit 110 may analyze the position of the leading vehicle. If the position of the leading vehicle obscures a road polynomial (e.g., the leading vehicle is overlaid on top of the road polynomial), then processing unit 110 may determine that the leading vehicle is likely changing lanes. In the case where the position of the leading vehicle is such that, another vehicle is detected ahead of the leading vehicle and the snail trails of the two vehicles are not parallel, processing unit 110 may determine that the (closer) leading vehicle is likely changing lanes.

[0101] At step 584, processing unit 110 may determine whether or not leading vehicle 200 is changing lanes based on the analysis performed at step 582. For example, processing unit 110 may make the determination based on a weighted average of the individual analyses performed at step 582. Under such a scheme, for example, a decision by processing unit 110 that the leading vehicle is likely changing lanes based on a particular type of analysis may be assigned a value of "1" (and "0" to represent a determination that the leading vehicle is not likely changing lanes). Different analyses performed at step 582 may be assigned different weights, and the disclosed embodiments are not limited to any particular combination of analyses and weights.

[0102] FIG. 6 is a flowchart showing an exemplary process 600 for causing one or more navigational responses based on stereo image analysis, consistent with disclosed embodiments. At step 610, processing unit 110 may receive a first and second plurality of images via data interface 128. For example, cameras included in image acquisition unit 120 (such as image capture devices 122 and 124 having fields of view 202 and 204) may capture a first and second plurality of images of an area forward of vehicle 200 and transmit them over a digital connection (e.g., USB, wireless, Bluetooth, etc.) to processing unit 110. In some embodiments, processing unit 110 may receive the first and second plurality of images via two or more data interfaces. The disclosed embodiments are not limited to any particular data interface configurations or protocols.

[0103] At step 620, processing unit 110 may execute stereo image analysis module 404 to perform stereo image analysis of the first and second plurality of images to create a 3D map of the road in front of the vehicle and detect features within the images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, road hazards, and the like. Stereo image analysis may be performed in a manner similar to the steps described in connection with FIGS. 5A-5D, above. For example, processing unit 110 may execute stereo image analysis module 404 to detect candidate objects (e.g., vehicles, pedestrians, road marks, traffic lights, road hazards, etc.) within the first and second plurality of images, filter out a subset of the candidate objects based on various criteria, and perform multi-frame analysis, construct measurements, and determine a confidence level for the remaining candidate objects. In performing the steps above, processing unit 110 may consider information from both the first and second plurality of images, rather than information from one set of images alone. For example, processing unit 110 may analyze the differences in pixel-level data (or other data subsets from among the two streams of captured images) for a candidate object appearing in both the first and second plurality of images. As another example, processing unit 110 may estimate a position and/or velocity of a candidate object (e.g., relative to vehicle 200) by observing that the object appears in one of the plurality of images but not the other or relative to other differences that may exist relative to objects appearing if the two image streams. For example, position, velocity, and/or acceleration relative to vehicle 200 may be determined based on trajectories, positions, movement characteristics, etc. of features associated with an object appearing in one or both of the image streams.

[0104] At step 630, processing unit 110 may execute navigational response module 408 to cause one or more navigational responses in vehicle 200 based on the analysis performed at step 620 and the techniques as described above in connection with FIG. 4. Navigational responses may include, for example, a turn, a lane shift, a change in acceleration, a change in velocity, braking, and the like. In some embodiments, processing unit 1 10 may use data derived from execution of velocity and acceleration module 406 to cause the one or more navigational responses. Additionally, multiple navigational responses may occur simultaneously, in sequence, or any combination thereof.

[0105] FIG. 7 is a flowchart showing an exemplary process 700 for causing one or more navigational responses based on an analysis of three sets of images, consistent with disclosed embodiments. At step 710, processing unit 110 may receive a first, second, and third plurality of images via data interface 128. For instance, cameras included in image acquisition unit 120 (such as image capture devices 122, 124, and 126 having fields of view 202, 204, and 206) may capture a first, second, and third plurality of images of an area forward and/or to the side of vehicle 200 and transmit them over a digital connection (e.g., USB, wireless, Bluetooth, etc.) to processing unit 110. In some embodiments, processing unit 110 may receive the first, second, and third plurality of images via three or more data interfaces. For example, each of image capture devices 122, 124, 126 may have an associated data interface for communicating data to processing unit 110. The disclosed embodiments are not limited to any particular data interface configurations or protocols.

[0106] At step 720, processing unit 110 may analyze the first, second, and third plurality of images to detect

IPR2025-00944
Tesla EX1002 Page 1487

features within the images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, road hazards, and the like. The analysis may be performed in a manner similar to the steps described in connection with FIGS. 5A-5D and 6, above. For instance, processing unit 110 may perform monocular image analysis (e.g., via execution of monocular image analysis module 402 and based on the steps described in connection with FIGS. 5A-5D, above) on each of the first, second, and third plurality of images. Alternatively, processing unit 110 may perform stereo image analysis (e.g., via execution of stereo image analysis module 404 and based on the steps described in connection with FIG. 6, above) on the first and second plurality of images, the second and third plurality of images, and/or the first and third plurality of images. The processed information corresponding to the analysis of the first, second, and/or third plurality of images may be combined. In some embodiments, processing unit 110 may perform a combination of monocular and stereo image analyses. For example, processing unit 110 may perform monocular image analysis (e.g., via execution of monocular image analysis module 402) on the first plurality of images and stereo image analysis (e.g., via execution of stereo image analysis module 404) on the second and third plurality of images. The configuration of image capture devices 122, 124, and 126-including their respective locations and fields of view 202, 204, and 206-may influence the types of analyses conducted on the first, second, and third plurality of images. The disclosed embodiments are not limited to a particular configuration of image capture devices 122, 124, and 126, or the types of analyses conducted on the first, second, and third plurality of images.

[0107] In some embodiments, processing unit 110 may perform testing on system 100 based on the images acquired and analyzed at steps 710 and 720. Such testing may provide an indicator of the overall performance of system 100 for certain configurations of image capture devices 122, 124, and 126. For example, processing unit 1 10 may determine the proportion of "false hits" (e.g., cases where system 100 incorrectly determined the presence of a vehicle or pedestrian) and "misses."

[0108] At step 730, processing unit 110 may cause one or more navigational responses in vehicle 200 based on information derived from two of the first, second, and third plurality of images. Selection of two of the first, second, and third plurality of images may depend on various factors, such as, for example, the number, types, and sizes of objects detected in each of the plurality of images. Processing unit 110 may also make the selection based on image quality and resolution, the effective field of view reflected in the images, the number of captured frames, the extent to which one or more objects of interest actually appear in the frames (e.g., the percentage of frames in which an object appears, the proportion of the object that appears in each such frame, etc.), and the like.

[0109] In some embodiments, processing unit 110 may select information derived from two of the first, second, and third plurality of images by determining the extent to which information derived from one image source is consistent with information derived from other image sources. For example, processing unit 1 10 may combine the processed information derived from each of image capture devices 122, 124, and 126 (whether by monocular analysis, stereo analysis, or any combination of the two) and determine visual indicators (e.g., lane markings, a detected vehicle and its location and/or path, a detected traffic light, etc.) that are consistent across the images captured from each of image capture devices 122, 124, and 126. Processing unit 110 may also exclude information that is inconsistent across the captured images (e.g., a vehicle changing lanes, a lane model indicating a vehicle that is too close to vehicle 200, etc.). Thus, processing unit 110 may select information derived from two of the first, second, and third plurality of images based on the determinations of consistent and inconsistent information.

[0110] Navigational responses may include, for example, a turn, a lane shift, a change in acceleration, and the like. Processing unit 1 10 may cause the one or more navigational responses based on the analysis performed at step 720 and the techniques as described above in connection with FIG. 4. Processing unit 110 may also use data derived from execution of velocity and acceleration module 406 to cause the one or more navigational responses. In some embodiments, processing unit 110 may cause the one or more navigational responses based on a relative position, relative velocity, and/or relative acceleration between vehicle 200 and an object detected within any of the first, second, and third plurality of images. Multiple navigational responses may occur simultaneously, in sequence, or any combination thereof.

Object Detecting and Braking System

[0111] In some embodiments, vehicle 200 may include an object detecting and braking system configured to detect an object in front of vehicle 200 and cause vehicle 200 to brake, e.g., using braking system 230, based on the detection of the object. For example, the object detecting and braking system may cause vehicle 200 to brake automatically, or autonomously, without or with limited driver intervention. The automatic or autonomous object detecting and braking system may or may not allow a driver to intervene to take over control of braking system 230.

[0112] The object detecting and braking system may include a plurality of other systems discussed above. For example, the object detecting and braking system may include system 100, braking system 230, and one or more of image capture devices 122, 124, and 126. The one or more of image capture devices 122, 124, and 126 may acquire one or more images of an environment or an area in front of vehicle 200, which may include an object, such as another vehicle, a pedestrian, a road surface, etc. Processing unit 110 may process the images based on

IPR2025-00944
Tesla EX1002 Page 1488

at least two independent analyses to determine at least two times-to-collision (TTCs) of vehicle 200 with the object. In some embodiments, the two TTCs of vehicle 200 with the object may be determined separately and independently of each other. A "time-to-collision" (TTC) may also be referred to as a "time-to-contact." The time-to-collision is an estimated time that would lapse before vehicle 200 collides with another object. Processing unit 110 may determine that a difference between the two times-to-collision does not exceed a predetermined threshold. Based on this determination, processing unit 110 may cause vehicle 200 to brake using, e.g., braking system 230. The automatic or autonomous braking may assist the driver to avoid collision with the object in front of vehicle 200, thereby improving driving safety. By causing vehicle 200 to brake when the difference between the two TTCs does not exceed the predetermined threshold, false braking may be prevented or reduced.

[0113] In some embodiments, the object detecting and braking system may use two independent channels of information inherent to monocular video sensing. The first channel of information may be referred to as a "texture" or "appearance information" channel. System 100 may perform an analysis based on the texture or appearance information. In this channel, system 100 may perform pattern recognition over the pixels intensities in each image separately among different images (e.g., two sequential images or frames). System 100 may estimate the change of a footprint size (e.g., a dimension length of a feature of an object in the image) by comparing image data of the two sequential images. The second channel of information may be referred to as a "structural" or "optical flow information" channel. System 100 may perform an analysis based on the structural or optical flow information. For example, system 100 may process two or more images (e.g., sequentially acquired images or frames) to obtain an optical flow field. The optical flow field may provide information regarding the movement of trackable textures between the two or more images or frames. The analysis of the optical flow field may indicate a low time-to-collision upright object in the images. The two analyses may also be referred to as two modalities, methods, or models.

[0114] In some embodiments, the structure (or optical flow) based analysis may serve to validate a result of a texture based analysis. For example, the texture based analysis may be performed on a plurality of images (e.g., at least two images), and the structure based analysis may be performed based on the same two or more images used in the texture based analysis after the texture based analysis is performed. In such embodiments, the structure based analysis may validate a result of the texture based analysis. In some embodiments, the texture based analysis may be performed after the structure based analysis has been performed, and may validate a result of the structure based analysis.

[0115] In some embodiments, the texture based analysis may use a first group of at least two images, and the structure based analysis may use a second group of at least two images. The first group and the second group may be the same or different. In some embodiments, the first group of at least two images may be selected from odd numbered images among a plurality of images acquired by at least one of image capture devices 122-126, and the second group of at least two images may be selected from even numbered images among the same plurality of images acquired by at least of image capture devices 122-126.

[0116] In some embodiments, the first group of at least two images may be selected from a plurality of images that are acquired at a first time interval and the second group of at least two images may be selected from the same plurality of images that are acquired at a second time interval, which may be different from the first time interval. In some embodiments, a first image capture device (or a first image sensor) may be designated for capturing images for the texture based analysis, and a second image capture device (or a second image sensor) may be designated for capturing images for structure based analysis.

[0117] The texture based analysis and the structure based analysis may complement each other, and the combination of their results that may be used by processing unit 110 to make a braking decision. For example, the texture based analysis may result in a first estimate of time-to-collision, and the structure based analysis may also result in a second estimate of the time-to-collision. Processing unit 110 may cause braking system 230 to brake vehicle 200 when the difference between the first and second estimates is smaller than a threshold. One of the texture based analysis and the structure based analysis may provide a redundancy for the other, thereby improve accuracy of braking, for example, autonomous braking in emergency driving situation, by preventing or reducing the occurrence of false braking.

[0118] In some embodiments, when the structure based analysis serves as a validator of the texture based analysis, the texture based analysis may detect an object (e.g., a car or a pedestrian) in front of vehicle 200, and may determine that the time-to-collision is low (e.g., lower than a predetermined threshold). The texture based analysis may reach this determination by calculating the time-to-collision based on a change of a footprint size through time in at least two sequentially acquired images. This determination from the texture based analysis may then trigger the structure based analysis. In some embodiments, a determination that the time-to-collision is low from the texture based analysis may serve as a precondition for releasing a braking decision (e.g., an autonomous emergency braking decision) to vehicle systems. In some embodiments, braking (e.g., autonomous emergency braking) may not be performed until the structure based analysis confirms the determination made via the texture based analysis.

[0119] In some embodiments, system 100 may require that the estimated time-to-collision (TTC) determined by

IPR2025-00944
Tesla EX1002 Page 1489

the structure based analysis is identical to or substantially identical to (up to an error tolerance) to the estimated time-to-collision (TTC) determined by the texture based analysis. For example, as discussed below, processing unit 10 may determine that a difference between the two TTCs determined by the texture based analysis and the structure based analysis does not exceed a predetermined threshold, which may be the error tolerance. In some embodiments, system 100 may take into account the divergence of the object's focus of expansion from that of the static environment. If the focus of expansion of the object diverges (e.g., moves) away from the focus of expansion of the static environment, this may be a strong indication that the object is a low TTC upright object. An example is discussed in greater detail with respect to FIGS. 15A and 15B. This requirement is optional because it may only apply in the case where the target object has a velocity component that is perpendicular to the driving direction of vehicle 200 (e.g., target object traveling in a direction that is not parallel to the driving direction of vehicle 200).

[0120] In some embodiments, processing unit 110 may use certain criteria for deciding whether structure based analysis supports texture based analysis in making a braking (e.g., autonomous emergency braking) decision. For example, such criteria may include a requirement that the optical flow fits a model of low time-to-collision upright object (LTU object), and not an open-road (or an open road surface).

[0121] FIG. 8 is a diagrammatic representation of an exemplary vehicle 200 including an object detecting and braking system (or referred to as an emergency braking decisioning system). The object detecting and braking system may include system 100 and braking system 230. The object detecting and braking system may also include at least one of image capture devices 122 and 124 (and 126, which is not shown in FIG. 8). The object detecting and braking system may include processing unit 110, which may be part of system 100. At least one of image capture devices 122 and 1 24 may acquire one or a plurality of images of an area including an object in front of vehicle 200. The object may be another vehicle, such as a vehicle 800 that may be traveling in front of vehicle 200. As shown in FIG. 8, in some embodiments, vehicle 800 may be traveling in a direction that is parallel to a traveling direction of vehicle 200. A parallel direction may include a situation in which vehicle 800 is traveling in the same direction as and in alignment or substantially in alignment with vehicle 200 (e.g., vehicle 800 traveling in front of vehicle 200), as shown in FIG. 8. A parallel direction may also include a situation in which vehicle 800 is traveling in an opposite direction relative to vehicle 200 and in alignment with vehicle 200 (e.g., vehicle 800 moving toward vehicle 200 in a direction that is 180 degrees, or substantially 180 degrees, opposite to the traveling direction of vehicle 200). A parallel direction may also include a situation in which vehicle 800 is not traveling in alignment with vehicle 200, but is traveling in a direction parallel to the traveling direction of vehicle 200 (e.g., vehicle 800 traveling on a front side of vehicle 200 at an offset distance).

[0122] In some embodiments, the object may be a pedestrian who may be standing in the road, or walking or running across the road. The area included in the images may include other objects, such as a road surface, a house, a wall, a sign, etc., which may form a static environment. When an object is not traveling in a direction parallel to the traveling direction of vehicle 200, the traveling direction of the object may transverse the traveling direction of vehicle 200 with an angle that is neither zero (or substantially zero) nor 180 degrees (or substantially 180 degrees). The images may include both moving and/or static objects.

[0123] Processing unit 1 1 0 may communicate with braking system 230 and at least one of image capture devices 122 and 124. For example, processing unit 110 may receive images from at least one of image capture devices 122 and 124, and may process the images using various methods or modalities to detect one or more times-to-collision of vehicle 200 with object 800. Although two image capture devices 122 and 124 are shown in FIG. 8, the use of additional image capture devices (e.g., 3, 4, 5, etc., image capture devices) is consistent with the disclosed embodiments.

[0124] FIG. 9 is a diagrammatic representation of an exemplary vehicle 200 including an object detecting and braking system that may be configured to detect an object not traveling in a direction parallel to the traveling direction of vehicle 200. The object may be a vehicle 900 or a pedestrian (not shown). The object detecting and braking system may detect vehicle 900 traveling toward vehicle 200 based on one or more images acquired by at least one of image capture devices 122 and 1 24. In addition, the object detecting and braking system may determine at least two times-to-collision of vehicle 200 with vehicle 900 based on the at least one or more images. When a difference between the at least two times-to-collision does not exceed the predetermined threshold, processing unit 110 may cause vehicle 200 to brake in order to prevent collision with vehicle 900.

[0125] FIG. 10 is an exemplary block diagram of memory 140 or 150 that may store instructions for performing one or more operations for detecting an object in front of vehicle 200 and causing vehicle 200 to brake based on the detection. As shown in FIG. 10, memory 140 or 150 may store one or more modules for performing the operations, as described herein. For example, memory 140 or 150 may store a first image analysis module 1002, a second image analysis module 1003, and a braking module 1004. Application processor 180 and/or image processor 190 may execute the instructions stored in any of modules 1002-1004 included in memory 140 or 150. One of skill in the art will understand that references in the following discussions to processing unit 110 may refer to application processor 180 and image processor 190 individually or collectively. Accordingly, steps of any of

IPR2025-00944
Tesla EX1002 Page 1490

the following processes may be performed by one or more processing devices.

[0126] First image analysis module 1002 may store instructions which, when executed by processing unit 1 10, may perform a first image analysis to calculate or generate a first estimated time-to-collision (TTC1) of vehicle 200 with an object (e.g., vehicle 800 or 900). First image analysis module 1002 may process at least two images acquired by at least one of image capture devices 122 and 124 using, for example, an appearance based detection method, which may calculate TTC1 based on a change in the appearance (e.g., a size) of the object in the at least two images. An exemplary first image analysis is discussed below in greater detail with respect to FIG. 12.

[0127] Second image analysis module 1002 may store instructions which, when executed by processing unit 110, may perform a second image analysis to calculate or generate a second estimated time-to-collision (TTC2) of vehicle 200 with the object (e.g., vehicle 800 or 900). Second image analysis module 1003 may process at least two images acquired by at least one of image capture devices 122 and 124 using, for example, an optical flow based method, which may calculate TTC2 based on optical flow information derived from the at least two images. An exemplary second image analysis is discussed below in greater detail with respect to FIG. 13.

[0128] Braking module 1004 may store instructions which, when executed by processing unit 110, may cause vehicle 200 to brake using braking system 230. For example, when processing unit 110 determines that a difference between TTC1 and TTC2 does not exceed the predetermined threshold, braking module 1004 may cause braking system 230 to apply braking to reduce the speed of vehicle 200 or stop vehicle 200, thereby preventing collision with the object.

[0129] FIG. 11 is a flowchart showing an exemplary process or method 1100 for detecting an object in front of vehicle 200 and causing vehicle 200 to brake based on the detection. Method 1100 may include acquiring a plurality of images of an area in front of vehicle 200 (step 1110). For example, at least one of image capture devices 122 and 124 may acquire one or more images of an area in front of vehicle 200. The area may include an object, such as another vehicle or a pedestrian. The area may include other static objects, such as a sign, a house, a mountain, a road surface, etc., which may form a static environment.

[0130] Method 1100 may include performing the first image analysis based on at least a first image and a second image to generate a first time-to-collision (TTC1) (step 1120). For example, processing unit 110 may perform the first image analysis based on at least the first image and the second image. More than two images may be used in the first image analysis to generate TTC1. In some embodiments, the first image analysis may include an appearance based detection method. The appearance based detection may detect TTC1 based on a change in appearance (e.g., a change in size) of an object in the first image and the second image. The appearance based detection technique is further discussed below in connection with FIG. 12.

[0131] FIG. 12 is a flowchart showing an exemplary first image analysis 1200 for determining the first estimated time-to-collision (TTC1). The first image analysis 1200 may include the appearance or texture based detection method (hence, the first image analysis may be referred to as the texture based analysis or textural analysis). The appearance based detection method may include determining a change in the appearance of an object or a feature of the object in at least two images (e.g., the first image and the second image), and may calculate TTC1 based on the change in the appearance. In some embodiments, the change in the appearance may be a change in a size of the object (or the feature of the object).

[0132] In the exemplary first image analysis 1200 shown in FIG. 12, processing unit 1 10 may determine a first size of an object (or a feature of the object) appearing in the first image (step 1210). Processing unit 110 may determine a second size of the object (or the feature of the object) appearing in the second image (step 1220). The second size may be larger than the first size as vehicle 200 approaches the object. Processing unit 110 may determine the first size and the second size based on the dimensions (e.g., pixel sizes) of the object appearing in the first image and the second image. Processing unit 110 may determine a change in size (or relative scale) of the object (or the feature of the object) (step 1230). For example, processing unit 110 may subtract the first size from the second size to determine a difference in size. In some embodiments, processing unit 110 may divide the difference in size by the first size to determine a relative difference in size. The change in size (or relative scale) may refer to the relative difference in size, which may be a percentage. Processing unit 110 may determine TTC1 based on the change in size and a first predetermined time interval between times of acquisition of the first image and the second image (step 1240). The first predetermined time interval may depend on settings of the camera (e.g., the camera included in the first image capture device 122) that acquires the first image and the second image, and may be known once the camera settings are fixed. The first predetermined time interval may be any suitable interval, such as, for example, 0.05 second, 0.1 second, 0.2 second, etc. For example, the first predetermined time interval may be represented by dtl, the relative change in size may be represented by S0, and processing unit 110 may calculate TTC1 from TTC1 = dt1 / S0.

[0133] Referring back to FIG. 11, method 1100 may include performing a second image analysis based on at least a third image and a fourth image to generate a second estimated time-to-collision (TTC2) (step 1130). For example, processing unit 110 may derive optical flow information from the third image and the fourth image, and may calculate TTC2 based on the optical flow informa-

IPR2025-00944
Tesla EX1002 Page 1491

tion. The second image analysis may be referred to as the structure based analysis or structural analysis. In some embodiments, more than two images may be used in the second image analysis to derive the optical flow information.

[0134] FIG. 13 is a flowchart showing an exemplary second image analysis 1300 for determining the second estimated time-to-collision (TTC2). The second image analysis 1300 may include obtaining or deriving an optical flow field from at least the third image and the fourth image (step 1310). In some embodiments, processing unit 110 may derive the optical flow field from the third image and the fourth image, and/or additional images acquired by at least one of image capture devices 122 and 124. The optical flow field may be a vector field in which each point may be represented by a vector showing a magnitude and a direction of movement from a first point in, e.g., the third image, to a corresponding second point in, e.g., the fourth image.

[0135] As shown in FIG. 13, the second image analysis 1300 may include determining a focus of expansion based on the optical flow field (step 1320). The focus of expansion may be a converging point of vectors in the optical flow field, which is discussed below in greater detail with respect to FIG. 16. The second image analysis 1300 may also include determining a rate of expansion relative to the focus of expansion for each point in the optical flow field (step 1330). For example, for each point in the optical flow field, processing unit 110 may calculate the rate of expansion based on a magnitude of movement of the point, and a distance from the point to the focus of expansion. For example, the magnitude of movement of any point may be represented by d, the distance from the point to the focus of expansion may be represented by r, and processing unit 110 may calculate the rate of expansion from $R_{exp} = d / r$. For each point in the optical flow field, the magnitude of movement d may be calculated from the magnitude of the vector associated with that point. The magnitude of movement r may indicate how much a first point in, e.g., the third image, has moved to reach a second point in, e.g., the fourth image.

[0136] As shown in FIG. 13, the second image analysis 1300 may include determining the second estimated time-to-collision (TTC2) based on the rate of expansion and a second predetermined time interval. The second predetermined time interval may be the difference between times of acquisition of the third image and the fourth image. The second predetermined time interval dt2 may depend on the settings (e.g., the frame rate) of the camera, and may or may not be the same as the first predetermined time interval used in the first image analysis. The second predetermined time interval may be any suitable interval, such as, for example, 0.05 second, 0.1 second, 0.2 second, etc. In some embodiments, TTC2 may be inversely proportional to the rate of expansion d/r, and may be directly proportional to the second predetermined time interval dt2. For example, processing unit 110 may calculate TTC2 from TTC2 = dt2 / (d/r).

[0137] The first image analysis 1200 and second image analysis 1300 shown in FIGS. 12 and 13 may be performed independently and separately. The first image and second image used in the first image analysis 1200 may be the same as or different from the third image and fourth image used in the second image analysis 1300. In some embodiments, the first image may be the same as the third image, and the second image may be the same as the fourth image. In some embodiments, the first image may be different from the third image, and the second image may be different from the fourth image. In some embodiments, images of alternating sequences may be used for the first and second analyses. For example, from the same set of images, odd numbered images may be used for the first image analysis, and even numbered images may be used for the second image analysis. In some embodiments, at least one image capture device (e.g., image capture device 122) may be configured to capture images (e.g., the first and second images) for use in the first image analysis to generate the first estimated time-to-collision. At least one other image capture device (e.g., image capture device 124) may be configured to capture images (e.g., the third and fourth images) for use in the second image analysis to generate the second estimated time-to-collision. In some embodiments, when at least one image capture device is used for capturing a plurality of images for the first image analysis and the second image analysis, the at least one image capture device may include a first image sensor configured to capture images (e.g., the first and 34 second images) for use in the first image analysis, and a second image sensor that is separate and independent of the first image sensor configured to capture images (e.g., the third and fourth images) for use in the second image analysis.

[0138] Referring back to FIG. 11, method 1100 may include calculating a difference between the first estimated time-to-collision (TTC1) and the second estimated time-to-collision (TTC2) (step 1140). In some embodiments, processing unit 110 may subtract TTC1 from TTC2 to result in a difference. Processing unit 110 may determine whether the difference exceed a predetermined threshold (step 1150). The predetermined threshold may be any suitable threshold, such as, for example, 0.05 second, 0.1 second, 0.2 second, etc. When processing unit 1 10 determines that the difference exceeds the predetermined threshold (Yes, step 1150), processing unit 110 may repeat steps 1110-1150. When processing unit 110 determines that the difference does not exceed the predetermined threshold (No, step 1150), processing unit 110 may cause vehicle 200 to brake using, e.g., braking system 230 (step 1160).

[0139] FIGS. 14A-14C show an image with an optical flow field and a distribution of the second estimated time-to-collision (TTC2) in the optical flow field. FIG. 14A shows an image 1400 showing a vehicle 1410 traveling on a road 1420. For illustrative purposes, other objects included in image 1400 are not shown. FIG. 14B shows

IPR2025-00944
Tesla EX1002 Page 1492

an optical flow field 1430 derived based on at least two images. For example, optical flow field 1430 may be derived from image 1400 and at least one second image (not shown) of vehicle 1410 acquired at the second predetermined time interval dt2 apart from the time when image 1400 was acquired. A plurality of image points may be selected in the image 1400 or the at least one second image, and vectors connecting the corresponding points of these images may be generated by processing unit 110. The vectors may form optical flow field 1430, which is schematically shown in FIG. 14B, as a region surrounding vehicle 1410.

[0140] Although optical flow field 1430 is shown within a region surrounding vehicle 1410 for illustrative purposes, optical flow field 1430 may be derived for a smaller region (e.g., focusing on the window of vehicle 1430) or a greater region (e.g., the entire area shown in image 1400). Each point in optical flow field 1430 may be associated with a vector represented in FIG. 14B by an arrow having a length and pointing at a direction. The length of the arrow may represent a magnitude of the vector. For example, the magnitude of the vector may indicate a magnitude of movement of a first point in image 1400 to a second point in the at least one second image. The direction of the arrow may indicate the direction of movement between each pair of corresponding points in these two images.

[0141] After performing the second image analysis 1300 discussed above, processing unit 110 may determine the second estimated time-to-collision (TTC2) for each point in the optical flow field 1430. FIG. 14C shows an exemplary distribution of TTC2 over the entire area shown in image 1400. As shown in FIG. 14C, points in different regions may be associated with different values of TTC2. The different patterns and/or shades shown in FIG. 14C are shown to schematically illustrate different values of TTC2. For a region 1440 substantially focusing on the body of vehicle 1410, all or most of the points within region 1440 may be associated with substantially the same value of TTC2. In other words, for an upright moving object like vehicle 1410, all or most of the points on that upright object may have the same or substantially the same times-to-collision.

[0142] For road surfaces, however, all or most of the points may not have the same or substantially the same times-to-collision. For example, points in a first portion 1450 of road 1420 may be associated with one time-to-collision, and points in a second portion 1460 of road 1420 may be associated with another time-to-collision. These times-to-collision may be different, as shown with different patterns/shades in FIG. 14C. In some embodiments, there may be a discontinuity between the times-to-collision for the first portion 1450 and second portion 1460 of road 1420 (e.g., the changes between these times-to-collision may not be gradual). Although the first and second portions 1450 and 1460 appear as upright objects in a single image, the optical flow field generated from at least two images may reveal additional information (e.g., time-to-collision) that may distinguish a true upright object (which may or may not be moving) and a road surface, which is not upright. This aspect is further discussed with respect to FIGS. 17A-18B.

[0143] As shown in FIG. 14C, points in regions 1470, 1480, and 1490 may be associated with different times-to-collision, as indicated by the different patterns/shading shown in the figure. Regions 1470 and 1480 may each represent a portion of road 1420, and the change in time-to-collision values in regions 1470 and 1480 and first portion 1450 may be gradual (e.g., continuous). For example, region 1490 may represent a portion of the sky, which may be associated with a time-to-collision that is different from those associated with regions 1470 and/or 1480.

[0144] FIGS. 15A and 15B show exemplary focuses of expansion in an optical flow field. FIG. 15A shows an object in an image traveling in the same direction (or substantially the same direction) as and in alignment (or substantially in alignment) with vehicle 200 (e.g., in the same line of movement in front of vehicle 200, toward or away from vehicle 200). As shown in FIG. 15A, a focus of expansion of the object may overlap completely or substantially with a focus of expansion of the static environment. FIG. 15B shows an object not traveling in a direction parallel to the traveling direction of vehicle 200, e.g., when the object is traveling laterally in a direction that transverses the traveling direction of vehicle 200. As shown in FIG. 15B, a focus of expansion of the object may not overlap the focus of expansion of the static environment.

[0145] For example, FIG. 15A shows an environment 1500 of an area in front of vehicle 200. Consistent with disclosed embodiments, an image capture device (e.g., one of image capture devices 122-126) may acquire a first image 1510 of environment 1500 that includes an object located within the first image 1510. The area surrounding the object may represent a static environment (e.g., roads, buildings, mountains, sky, etc.). FIG. 15A also shows a second image 1520 that an image capture device (e.g., one of image capture devices 122-126) may acquire of the object after the second predetermined time interval dt2. Image 1520 may be acquired as vehicle 200 moves toward the object and, therefore, may be larger than image 1510, indicating an expansion.

[0146] FIG. 15A further shows a focus of expansion 1530. The focus of expansion 1530 may be a focus of expansion of the static environment, as indicated by two dashed lines converging to focus of expansion point 1530. The location of focus of expansion 1530 of the static environment may depend on the camera, and may not change over time as the camera continuously acquires a plurality of images of the object or area for deriving optical flow information. When the object is traveling in the same (or substantially the same) direction as and in (or substantially in) alignment with vehicle 200, e.g., in front of vehicle 200 and in the same or opposite traveling direction as vehicle 200, the focus of expansion 1530 may also constitute a focus of expansion of the object. In other words, the focus of expansion of the ob-

IPR2025-00944
Tesla EX1002 Page 1493

ject traveling in the same direction as and in alignment with vehicle 200 may completely or substantially overlap (or completely or substantially converge with) the focus of expansion of the static environment. Thus, processing unit 110 may detect one focus of expansion from the optical flow field derived from a plurality of images. FIG. 15A also shows, for illustrative purpose, the distance r from certain points to the focus of expansion.

[0147] FIG. 15B shows an environment 1540 of an area in front of vehicle 200, and a first image 1550 and a second image 1560 of an object (the object being included in the first image 1550 and second image 1560) acquired by the camera one after another at the second predetermined time interval dt2. For example, as vehicle 200 moves toward the object, the image of the object may expand, e.g., from image 1550 to image 1560. FIG. 15B further shows a first focus of expansion 1570 that may be calculated from the optical flow field derived from image 1550 and image 1560 of the object. For illustrative purposes, FIG. 15B shows the distances r from certain points to the first focus of expansion 1570. FIG. 15B also shows a second focus of expansion 1580, which may be a focus of expansion of the static environment. The location of the second focus of expansion 1580 may depend on the camera, and may not change as the camera continuously acquires images of the object or area for deriving the optical flow information. In the example shown in FIG. 15B, the first focus of expansion 1570 may not overlap the second focus of expansion 1580. The detection of two or at least two non-overlapping focuses of expansion may indicate that there is at least one object in the image 1540 that is traveling laterally, e.g., in a direction that is not parallel to the traveling direction of vehicle 200, or in a direction that transverses the traveling direction of vehicle 200 at an angle that is neither zero (or substantially zero) nor 1 80 degrees (or substantially 180 degrees).

[0148] FIG. 16 shows an exemplary optical flow field and a focus of expansion. The following discussion of FIG. 16 illustrates how an optical flow field may be derived and how a focus of expansion may be determined. For example, vehicle 1600 is an object included in an image acquired by, e.g., image capture device 122 installed on vehicle 200, and vehicle 200 is traveling toward vehicle 1600. Image capture device 122 may sequentially and/or continuously acquire a plurality of images of an area in front of vehicle 200 as vehicle 200 approaches vehicle 1600. Any two or more images may be selected from the plurality of images to derive an optical flow field. In FIG. 16, the first image is represented by vehicle 1600 shown in solid lines, and the second image is represented by vehicle 1610 shown in dashed lines. For a plurality of selected points on the first image, e.g., points (x1, y1), (x2, y2), and (x3, y3), processing unit 110 may determine their corresponding points, e.g., (u1, v1), (u2, v2), and (u3, v3) in the second image. Vectors may be determined for each pair of points (x1, y1) and (u1, v1), (x2, y2) and (u2, v2), and (x3, y3) and (u3, v3). The vectors may indicate a magnitude and direction of expansion for the points in the first image. The length of each vector (represented by a solid arrow in FIG. 16) may indicate the distance or magnitude of movement of from one point in the first image to a corresponding point in the second image, and the direction of the vector may indicate the expansion direction or direction of movement. The converging point 1620 of the vectors may represent the focus of expansion. In some embodiments, the location of the converging point 1620 may be calculated using a suitable mathematical model based on the coordinates of the vectors. For example, the location of converging point 1620 may be calculated from:

$$SS * (x_i - FOE_x) = u_i \quad (1)$$

$$SS * (y_i - FOE_y) = v_i \quad (2)$$

[0149] Where SS may be a constant, (x_i, y_i) may be coordinates of the i-th point within a first image, and (u_i, v_i) may be coordinates of the corresponding i-th point within the second image. FOE_x and FOE_y may be coordinates representing the location of converging point 1620 (i.e., focus of expansion).

[0150] FIGS. 17A and 17B show an image of an object and a linear relationship between distances from points in an optical flow field to the focus of expansion and magnitudes of movement of the points. FIGS. 17A and 17B show that a linear relationship between distances from points to the focus of expansion and magnitudes of movement of the points may be used to determine that the object is an upright object (e.g., a low time-to-collision upright object).

[0151] FIG. 17A shows an image including an object, i.e., a vehicle 1700. The image of vehicle 1700 may include an image of two car windows 1710 and 1720, which appear to be darker in the image of vehicle 1700. A plurality of points 1730 may be selected, most of which may be from a region in the image of vehicle 1700. An optical flow field may be derived from at least two images of vehicle 1700 acquired one after another with the second predetermined time interval dt2. Processing unit 110 may determine the magnitudes of movement d and the distances r from the points to the focus of expansion based on the optical flow field. Processing unit 110 may determine a relationship between the magnitudes d and distances r. In some embodiments, the relationship may be a linear relationship, as shown in FIG. 17B, indicating that d/r (e.g., slope of the plotted line) is the same for all or substantially all of the points in the optical flow field. As discussed above, d/r may be inversely proportional to the second estimated time-to-collision (TTC2). Thus, a linear relationship between d and r may indicate a constant or substantially constant value of d/r, which in turn may indicate a constant or substantially constant TTC2 for all or substantially all of the points in the optical flow

IPR2025-00944
Tesla EX1002 Page 1494

field. The constant or same TTC2 may indicate that the object within the optical flow field is an upright object, or a low time-to-collision upright object.

[0152] To illustrate the linear relationship between magnitudes of movement d and distances r, FIG. 17B shows a plot of the vertical components (y components) related to the magnitudes of movement d (e.g., the magnitude of movement in the y direction, "dy" in FIG. 17B) versus vertical components of distances r (e.g., the distance to the focus of expansion in the y direction, "y" in FIG. 17B). Although not shown in FIG. 17B, a plot of the horizontal components, e.g., magnitudes in the x direction and distances in the x direction, may also indicate a linear relationship, which in turn may indicate a constant or substantially constant time-to-collision for all or substantially all of the selected points in the optical flow field. In addition, a plot of distances r and magnitudes d may also indicate a linear relationship, which in turn may indicate a constant (or same) or substantially constant (or same) time-to-collision for all or substantially all of the selected points in the optical flow field.

[0153] When a non-linear relationship between the magnitudes of movement d and the distances r to the focus of expansion is detected or determined, the non-linear relationship may indicate that the second estimated times-to-collision are different for different points included in the optical flow field, which in turn may indicate that the image may include a road surface, rather than an upright object. As discussed above, the second estimated times-to-collision may be determined based on the optical flow information derived from at least two images. FIGS. 18A and 18B show an image of an object and a non-linear relationship between distances r from points in an optical flow field to the focus of expansion and magnitudes d of movement of the points. FIG. 18A shows an image 1800 of a house 1810 having two garages 1820 and 1830 with garage doors opened, and a driveway 1840 (e.g., a road surface) leading up to the garages 1820 and 1830. Because the garage doors are opened, garages 1820 and 1830 may appear darker on the image, similar to the car windows 1710 and 1720 shown in FIG. 17A. A region 1850 of the image may be selected for image processing. A plurality of points may be selected within region 1850, and at least two images including regions corresponding to region 1850 may be acquired to derive an optical flow field. Based on the optical flow field having a plurality of vectors, processing unit 110 may determine a focus of expansion, the distances r of the points to the focus of expansion, and the magnitudes d of movement of the points included in the optical flow field. Processing unit 110 may determine a relationship between the distances r and the magnitudes d. In the situation shown in FIG. 18A, processing unit 110 may determine that the relationship between the distances r and the magnitudes d is non-linear. The non-linear relationship between the distances r and the magnitudes d may indicate that the times-to-collision for the points are different for different points. This in turn may indicate that selected region 1850 includes an image of a road surface, rather than an upright object like vehicle 1700 shown in FIG. 17A.

[0154] To illustrate the non-linear relationship between magnitudes of movement d and distances r, FIG. 18B shows a plot of the vertical components (y components) related to the magnitudes d of movement (e.g., the magnitude of movement in the y direction, "dy" in FIG. 18B) versus vertical components of distances r (e.g., the distance to the focus of expansion in the y direction, "y" in FIG. 18B). Although not shown in FIG. 18B, a plot of the horizontal components, e.g., magnitudes in the x direction and distances in the x direction, may also indicate a nonlinear relationship, and hence, different times-to-collision for the selected points in the optical flow field. In addition, a plot of distances r and magnitudes d may also indicate a nonlinear relationship, and hence, different times-to-collision for the selected points in the optical flow field. Here, the image in the region 1850 may appear similar to the image of vehicle 1700 shown in FIG. 17A in a single image. The optical flow field derived from a plurality of images acquired at the second predetermined time interval dt2 may reveal additional information, e.g., time-to-collision information, about the objects included in the selected region. The time-to-collision information may be used to determine whether the object is an upright object or a road surface. Processing unit 110 may combine this determination with the calculated TTC2 to determine whether or not to cause vehicle 200 to brake.

[0155] FIG. 19 is a flowchart showing an exemplary process 1900 for triggering a warning based on one or more probabilities indicating that a time-to-collision is smaller than one or more thresholds. Instead of determining the time-to-collision as discussed above, processing unit 110 may determine, based on at least two images acquired by one or more image capture devices 122, 124, and 126, a probability that the time-to-collision is smaller than a predetermined threshold. For example, in process 1900, one or more image capture devices 122-126 may acquire a plurality of images of an area in front of vehicle 200 (step 1910). In some embodiments, one image capture device having one camera may be used. In some embodiments, two or more image capture devices have two or more cameras may be used. For example, one camera may have a 45-degree field of view of the area in front of vehicle 200, and another camera may have a 30-degree field of view.

[0156] In some embodiments, processing unit 110 may perform a third image analysis based on at least two images to determine a first probability (P1) that a time-to-collision (TTC) is smaller than a first predetermined threshold T1 (i.e., TTC<T1) (step 1920). The first predetermined threshold T1 may be any suitable value, such as, for example, 0.1 second, 0.2 second, 05. second, etc. Processing unit 110 may determine a second probability (P2) that the time-to-collision (TTC) is smaller than a second predetermined threshold T2 (i.e., TTC < T2) (step 1930). The second predetermined threshold T2 may be

IPR2025-00944
Tesla EX1002 Page 1495

any suitable value, such as, for example, 0.1 second, 0.2 second, 05. second, etc. The second predetermined threshold T2 may or may not be the same as the first predetermined threshold T1.

[0157] Processing unit 110 may trigger a warning if the first probability P1 is greater than a first predetermined probability threshold Tpl, or if the second probability P2 is greater than a second predetermined probability threshold Tp2. Tp1 and Tp2 may be any suitable value, such as, for example, 0.5, 0.8, 0.9, etc. Tp1 and Tp2 may or may not be the same. In some embodiments, Tp2 may be smaller than Tp1. Processing unit 110 may compare the first probability P1 with the first predetermined probability threshold Tp1 to determine whether the first probability P1 is greater than the first predetermined probability threshold Tp1. Processing unit 110 may also compare the second probability P2 with the second predetermined probability threshold Tp2 to determine whether the second probability P2 is greater than the second predetermined probability threshold Tp2. The warning may include an audio alert, a video alert, a vibrational alert, or a combination thereof. The warning may alert the driver so that the driver may take suitable navigational action, such as applying brakes, steering the wheel, etc.

[0158] FIG. 20 is a flowchart showing an exemplary process 2000 for determining the first time-to-collision (TTC1) that may be implemented in the first image analysis shown in FIG. 11. Process 2000 may be used in the first image analysis as an alternative texture based TTC1 calculation method replacing the process 1200 shown in FIG. 12. Processing unit 110 may process at least two images that are acquired at a predetermined time interval DT. Processing unit 110 may determine a first dimension length y1 of an object in a first image (step 2010). The first dimension length of the object may be a length of a feature of the object in the vertical direction, horizontal direction, or in any direction on the first image. For example, when the object is a vehicle, the feature may be a vertical side of a license plate on the vehicle. Processing unit 110 may determine a second dimension length y2 of the same object (or the same feature of the same object) in a second image, which may be acquired after the first image at a time that is DT after the time the first image is acquired (step 2020). Processing unit 110 may determine a scaling factor Sf based on the first dimension length and the second dimension length. For example, processing unit 110 may calculate Sf based on Sf = y1 / y2. In some embodiments, processing unit may track the horizontal and vertical coordinates (dx, dy) of the object in the first and second image, and may calculate Sf based on the coordinates (dx, dy), for example, Sf = dy1 / dy2.

[0159] Processing unit 110 may determine a time-to-collision (which may be the first time-to-collision shown in FIG. 11) based on the scaling factor and the time interval DT. In some embodiments, processing unit 110 may calculate TTC1 from TTC1 = DT / (Sf-1).

[0160] Using two, i.e., texture based and structure based, sensing modalities, methods, or models, as described above, have several benefits or advantages. For example, an immediate result of using two modalities in a braking decisioning system is a reduction of the probability of system failure may be achieved, as compared to braking decisioning systems using only one of the modalities (or only one information channel).

[0161] Another benefit relates to system validation. Using at least two modalities (or channels of information, or two independent and/or separate analyses) enable the vehicle to assess the overall system failure rate (e.g., mean-time-between-failures or MTBF). The overall system failure rate may be determined by assessing each modality's failure rate separately (e.g., by assessing the failure rate of the texture based analysis and the failure rate of the structure based analysis separately). The two modalities may be implemented in two sub-systems: a textural analysis based sub-system and a structural analysis based sub-system. The failure rates of the two modalities may be multiplied. A result of the multiplication may be used as an indicator of or may be used to compute the overall system failure rate. In some embodiments, the MTBF may be estimated based on the result of the multiplication, even if it is significantly higher than the extent (in time) of available validation data.

[0162] For example, in one scenario, a record of only 1000 hours of data is available, and the combined systems MTBF is 1000000 hours. In 1/100 hours the texture based analysis may trigger a false braking decision. In 1/10000 of the cases the texture based analysis is checking the hypothesized decision (e.g., the hypothesized decision being a decision to brake because of a low TTC upright object in front of vehicle 200). In such cases, the texture based analysis may falsely approve the hypothesized decision. By direct assessment of the combined system (or combined analyses of the texture based analysis and the structure based analysis), vehicle systems may only achieve 0 failure over 1000 hours, and may not be able to assess the MTBF. By taking two independent channels of information into account in the braking decisioning for vehicle 200, e.g., using the texture based analysis and the structure based analysis, a vehicle system may be able to assess the failure rate of each sub-system (e.g., textural analysis based sub-system and structural analysis based sub-system) separately. Such failure rates may be multiplied, and a result of the multiplication may be used to assess the MTBF. For example, in some embodiments, the textural analysis based sub-system may be keep running alone and it may fail 10 times over 1000 hours. If the structural analysis based sub-system is falsely provoked 1000000 times, as if, for example, the textural analysis based sub-system reached a decision that there is a need to issue an autonomous emergency braking on an open road, the structural analysis based sub-system may fail once in 10000 provocations. Based on these observations, processing unit 110 may deduce (or determine) that the total system mean-time-between failures (MTBFs) is indeed 1000000 hours. This estimate may be reached using only 1000 hours of recorded data.

IPR2025-00944
Tesla EX1002 Page 1496

Thus, by using the disclosed two modalities, system validation may be more efficiently performed based on only limited hours of recorded data.

[0163] FIG. 21 is a flowchart showing an exemplary process 2100 for performing system validation using two modalities. Process 2100 summarizes the system validation discussed above. Processing unit 110 may determine a first failure rate associated with the textural analysis based sub-system (step 2110). For example, processing unit 110 may determine the first failure rate based on the total number of times the textural analysis based sub-system makes a false braking decision (e.g., falsely detected a low TTC upright object) among a total number of system running hours. As discussed above, the textural analysis based sub-system may make a decision to brake based on detection of a change in a texture (e.g., a change in footprint size) between at least two images of an area in front of vehicle 200. Processing unit 110 may determine a second failure rate associated with the structural analysis based sub-system (step 2120). For example, processing unit 110 may determine the second failure rate based on the total number of times the structural analysis based sub-system makes a false braking decision (e.g., falsely detected a low TTC upright object) among the same total number of system running hours or a different total number of system running hours. Processing unit 110 may determine an overall system failure rate based on the first failure rate and the second failure rate (step 2130). For example, processing unit 110 may multiply the first failure rate and the second failure rate, and use a result of the multiplication as an indicator of overall system mean-time-between-failures (MTBF).

[0164] The foregoing description has been presented for purposes of illustration. It is not exhaustive and is not limited to the precise forms or embodiments disclosed. Modifications and adaptations will be apparent to those skilled in the art from consideration of the specification and practice of the disclosed embodiments. Additionally, although aspects of the disclosed embodiments are described as being stored in memory, one skilled in the art will appreciate that these aspects can also be stored on other types of computer readable media, such as secondary storage devices, for example, hard disks or CD ROM, or other forms of RAM or ROM, USB media, DVD, Blu-ray, or other optical drive media.

[0165] Computer programs based on the written description and disclosed methods are within the skill of an experienced developer. The various programs or program modules can be created using any of the techniques known to one skilled in the art or can be designed in connection with existing software. For example, program sections or program modules can be designed in or by means of .Net Framework, .Net Compact Framework (and related languages, such as Visual Basic, C, etc.), Java, C++, Objective-C, HTML, HTML/AJAX combinations, XML, or HTML with included Java applets.

[0166] Moreover, while illustrative embodiments have been described herein, the scope of any and all embodiments having equivalent elements, modifications, omissions, combinations (e.g., of aspects across various embodiments), adaptations and/or alterations as would be appreciated by those skilled in the art based on the present disclosure. The limitations in the claims are to be interpreted broadly based on the language employed in the claims and not limited to examples described in the present specification or during the prosecution of the application. The examples are to be construed as non-exclusive. Furthermore, the steps of the disclosed methods may be modified in any manner, including by reordering steps and/or inserting or deleting steps. It is intended, therefore, that the specification and examples be considered as illustrative only, with a true scope being indicated by the following claims.

Claims

1. An object detecting and braking system for a vehicle (200), comprising:

   at least one image capture device (122, 124, 126) configured to acquire a plurality of images of an area including an object in front of the vehicle (200);
   a data interface (128); and
   at least one processing device (110) programmed to:

   receive (1110) the plurality of images via the data interface (128);
   perform (1120) a first image analysis based on a time interval between a first image and a second image to determine a first estimated time-to-collision of the vehicle (200) with the object, the first image analysis including:

   determining (1210) a first size of the object appearing in the first image;
   determining (1220) a second size of the object appearing in the second image;
   determining (1230) a change in size of the object based on the first size and the second size; and
   determining (1240) the first estimated time-to-collision based on the change in size of the object and a predetermined time interval between times of

   acquisition of the first image and the second image;

   characterized in that the at least one processing device (110) is further programmed to:

   perform (1130) a second image analysis based on a time interval between a third image and a

IPR2025-00944
Tesla EX1002 Page 1497

fourth image to determine a second estimated time-to-collision of the vehicle (200) with the object, the second image analysis including:

determining (1320) a focus of expansion based on the third image and the fourth image, the focus of expansion being a converging point of vectors in a vector field determined from the third image and the fourth image, wherein each point in the vector field is represented by a vector indicating a magnitude and a direction of movement from a first point in the third image to a corresponding second point in the fourth image;
determining (1330) a rate of expansion relative to the focus of expansion for each point in the vector field by dividing a magnitude of movement of a vector associated with the point by a distance from the point to the focus of expansion; and
determining for a region substantially focusing on the object the second estimated time-to-collision based on the rate of expansion and a time interval between times of acquisition of the third image and the fourth image;

calculate (1140) a difference between the first estimated time-to-collision and the second estimated time-to-collision;
determine (1150) that the difference does not exceed a predetermined threshold; and
cause (1160) the vehicle (200) to brake based on the determination that the difference does not exceed the predetermined threshold.

2. The object detecting and braking system of claim 1, wherein the second image analysis further includes:

determining a relationship between magnitudes of movement and distances to the focus of expansion for a plurality of points in the vector field; and
determining that the object represented by the plurality of points is a low time-to-collision upright object based on a determination that the relationship is a linear relationship.

3. The object detecting and braking system of claim 1, wherein the second image analysis further includes:

determining a relationship between magnitudes of motion and to the focus of expansion. for a plurality of points in the vector field; and
determining that at least a portion of the object is a road surface based on a determination that the relationship is a non-linear relationship.

4. The object detecting and braking system of claim 1, wherein the second image analysis further includes:

detecting a first focus of expansion that is a first converging point of a first plurality of vectors in the vector field;
detecting a second focus of expansion that is a second converging point of a second plurality of vectors in the vector field; and
determining that the third image and the fourth image include a moving object that moves in a direction that is not parallel to a travel direction of the vehicle.

5. The object detecting and braking system of claim 1, wherein the second image analysis further includes: determining that the third image and the fourth image represent an environment that is static relative to a road surface on which the vehicle is traveling.

6. The object detecting and braking system of claim 1, wherein the at least one image capture device includes:

a first image sensor configured to capture the first image and the second image to be used in the first image analysis; and
a second image sensor independent of the first image sensor and configured to capture the third image and the fourth image to be used in the second image analysis.

7. A vehicle (200), comprising:

a body; and
the object detecting and braking system of any of the preceding claims.

8. A method for detecting an object and braking a vehicle (200), comprising:

acquiring (1110), via at least one image capture device (122, 124, 126), a plurality of images of an area including an object in front of the vehicle (200);
receiving, via a processing device (110), the plurality of images;
performing (1120), via the processing device (110), a first image analysis based on a time interval between a first image and a second image to determine a first estimated time-to-collision of the vehicle (200) with the object, wherein performing the first image analysis includes:

determining (1210) a first size of the object appearing in the first image;
determining (1220) a second size of the object appearing in the second image;

IPR2025-00944
Tesla EX1002 Page 1498

determining (1230) a change in size of the object based on the first size and the second size; and
determining (1240) the first estimated time-to-collision based on the change in size of the object and a predetermined time interval between times of acquisition of the first image and the second image;

characterized by

performing (1130), via the processing device (110), a second image analysis based on a time interval between a third image and a fourth image to determine a second estimated time-to-collision of the vehicle with the object, wherein performing the second image analysis includes:

determining (1320) a focus of expansion based on the third image and the fourth image, the focus of expansion being a converging point of vectors in a vector field determined from the third image and the fourth image, wherein each point in the vector field is represented by a vector indicating a magnitude and a direction of movement from a first point in the third image to a corresponding second point in the fourth image;
determining (1330) a rate of expansion relative to the focus of expansion for each point in the vector field by dividing a magnitude of movement of a vector associated with the point by a distance from the point to the focus of expansion; and
determining for a region substantially focusing on the object the second estimated time-to-collision based on the rate of expansion and the time interval between times of acquisition of the third image and the fourth image;

calculating (1140), via the processing device (110), a difference between the first estimated time-to-collision and the second estimated time-to-collision;
determining (1150), via the processing device (110), that the difference does not exceed a predetermined threshold; and
causing (1160), via the processing device (110), the vehicle (200) to brake based on the determination that the difference does not exceed the predetermined threshold.

9. The method of claim 8, wherein performing the second image analysis further includes:

determining a relationship between magnitudes of movement and distances to the focus of expansion for a plurality of points in the vector field; and
determining that the object represented by the plurality of points is a low time-to-collision upright object based on a determination that the relationship is a linear relationship.

10. The method of claim 8, wherein performing the second image analysis further includes:

determining a relationship between magnitudes of movement and distances to the focus of expansion for a plurality of points in the vector field; and
determining that at least a portion of the object is a road surface based on a determination that the relationship is a non-linear relationship.

11. The method of claim 8, wherein performing the second image analysis further includes:

detecting a first focus of expansion that is a first converging point of a first plurality of vectors in the vector field;
detecting a second focus of expansion that is a second converging point of a second plurality of vectors in the vector field; and
determining that the third image and the fourth image include a moving object that moves in a direction that is not parallel to a travel direction of the vehicle.

12. The method of claim 8, wherein performing the second image analysis further includes:
determining that the third image and the fourth image represent an environment that is static relative to a road surface on which the vehicle is traveling.

13. The method of claim 8, wherein acquiring (1110) the plurality of images comprises:

capturing, by a first image sensor included in the at least one image capture device, the first image and the second image to be used in the first image analysis; and
capturing, by a second image sensor included in the at least one image capture device and being independent of the first image sensor, the third image and the fourth image to be used in the second image analysis.

Patentansprüche

1. Objekterkennungs- und Bremssystem für ein Fahrzeug (200), Folgendes umfassend:

mindestens eine Bilderfassungsvorrichtung

IPR2025-00944
Tesla EX1002 Page 1499

(122, 124, 126), die dafür konfiguriert ist, eine Vielzahl von Bildern in einem Bereich aufzunehmen, der ein vor dem Fahrzeug (200) befindliches Objekt beinhaltet;
eine Datenschnittstelle (128); und
mindestens eine Verarbeitungsvorrichtung (110), die für Folgendes programmiert ist:

Empfangen (1110) der Vielzahl von Bildern über die Datenschnittstelle (128);
Durchführen (1120) einer ersten Bildanalyse basierend auf einem Zeitintervall zwischen einem ersten Bild und einem zweiten Bild, um eine erste geschätzte Zeit bis zur Kollision des Fahrzeugs (200) mit dem Objekt zu bestimmen, wobei die erste Bildanalyse Folgendes umfasst:

Bestimmen (1210) einer ersten Größe des Objekts, das auf dem ersten Bild erscheint;
Bestimmen (1220) einer zweiten Größe des Objekts, das auf dem zweiten Bild erscheint;
Bestimmen (1230) einer Größenänderung des Objekts basierend auf der ersten Größe und der zweiten Größe; und
Bestimmen (1240) der ersten geschätzten Zeit bis zum Zusammenstoß basierend auf der Änderung der Größe des Objekts und einem vorbestimmten Zeitintervall zwischen den Zeitpunkten der Aufnahme des ersten und des zweiten Bildes;
dadurch gekennzeichnet, dass die mindestens eine Verarbeitungsvorrichtung (110) ferner für Folgendes programmiert ist:
Durchführen (1130) einer zweiten Bildanalyse basierend auf einem Zeitintervall zwischen einem dritten und einem vierten Bild, um eine zweite geschätzte Zeit bis zum Zusammenstoß des Fahrzeugs (200) mit dem Objekt zu bestimmen, wobei die zweite Bildanalyse Folgendes umfasst:

Bestimmen (1320) eines Expansionsfokus anhand des dritten und vierten Bildes, wobei der Expansionsfokus ein konvergierender Punkt von Vektoren in einem aus dem dritten und dem vierten Bild bestimmten Vektorfeld ist, wobei jeder Punkt im Vektorfeld durch einen Vektor dargestellt wird, der eine Größe und eine Bewegungsrichtung von einem ersten Punkt im dritten Bild zu einem entsprechenden zweiten Punkt im vierten Bild anzeigt;
Bestimmen (1330) einer Expansionsrate in Bezug auf den Expansionsfokus für jeden Punkt im Vektorfeld durch Dividieren einer Bewegungsgröße eines dem Punkt zugeordneten Vektors durch eine Entfernung von dem Punkt zum Expansionsfokus; und
Bestimmen der zweiten geschätzten Zeit bis zum Zusammenstoß für einen Bereich, der im Wesentlichen auf das Objekt ausgerichtet ist, basierend auf der Expansionsrate und einem Zeitintervall zwischen den Zeitpunkten der Aufnahme des dritten und vierten Bildes;
Berechnen (1140) einer Differenz zwischen der ersten geschätzten Zeit bis zum Zusammenstoß und der zweiten geschätzten Zeit bis zum Zusammenstoß;
Bestimmen (1150), dass die Differenz einen vorbestimmten Schwellenwert nicht überschreitet;
und
Veranlassen (1160) eines Bremsvorgangs des Fahrzeugs (200) aufgrund der Bestimmung, dass die Differenz den vorbestimmten Schwellenwert nicht überschreitet.

2. Objekterkennungs- und Bremssystem nach Anspruch 1, wobei die zweite Bildanalyse ferner Folgendes umfasst:

Bestimmen einer Beziehung zwischen Bewegungsgrößen und Abständen zum Expansionsfokus für eine Vielzahl von Punkten im Vektorfeld; und
Bestimmen, dass das durch die Vielzahl von Punkten dargestellte Objekt ein aufrecht stehendes Objekt mit kurzer Zeit bis zum Zusammenstoß ist, aufgrund einer Bestimmung, dass die Beziehung eine lineare Beziehung ist.

3. Objekterkennungs- und Bremssystem nach Anspruch 1, worin die zweite Bildanalyse ferner Folgendes umfasst:

Bestimmen einer Beziehung zwischen Bewegungsgrößen und dem Expansionsfokus für eine Vielzahl von Punkten im Vektorfeld; und
Bestimmen, dass mindestens ein Teil des Objekts eine Straßenoberfläche ist, aufgrund einer

IPR2025-00944
Tesla EX1002 Page 1500

Bestimmung, dass die Beziehung eine nichtlineare Beziehung ist.

4. Objekterkennungs- und Bremssystem nach Anspruch 1, wobei die zweite Bildanalyse ferner Folgendes umfasst:

Erkennen eines ersten Expansionsfokus, der ein erster Konvergenzpunkt einer ersten Vielzahl von Vektoren im Vektorfeld ist;
Erkennen eines zweiten Expansionsfokus, der ein zweiter Konvergenzpunkt einer zweiten Vielzahl von Vektoren im Vektorfeld ist; und
Bestimmen, dass das dritte und vierte Bild ein sich bewegendes Objekt beinhalten, das sich in einer Richtung bewegt, die nicht parallel zu einer Fahrtrichtung des Fahrzeugs ist.

5. Objekterkennungs- und Bremssystem nach Anspruch 1, wobei die zweite Bildanalyse ferner Folgendes umfasst:
Bestimmen, dass das dritte und vierte Bild eine Umgebung darstellen, die in Bezug auf eine Straßenoberfläche, auf der das Fahrzeug fährt, statisch ist.

6. Objekterkennungs- und Bremssystem nach Anspruch 1, wobei die mindestens eine Bilderfassungsvorrichtung Folgendes umfasst:

einen ersten Bildsensor, der konfiguriert ist, um das erste und das zweite Bild zu erfassen, die bei der ersten Bildanalyse verwendet werden sollen; und
einen zweiten Bildsensor, der unabhängig vom ersten Bildsensor ist und dafür konfiguriert ist, das dritte und vierte Bild zu erfassen, die bei der zweiten Bildanalyse verwendet werden sollen.

7. Ein Fahrzeug (200), Folgendes umfassend:

eine Karosserie; und
das Objekterkennungs- und Bremssystem nach einem der vorhergehenden Ansprüche.

8. Verfahren zum Erfassen eines Objekts und zum Bremsen eines Fahrzeugs (200), Folgendes umfassend:

Aufnehmen (1110) einer Vielzahl von Bildern über mindestens eine Bilderfassungsvorrichtung (122, 124, 126), und zwar in einem Bereich, der ein vor dem Fahrzeug (200) befindliches Objekt beinhaltet;
Empfangen der Vielzahl von Bildern über eine Verarbeitungsvorrichtung (110);
Durchführen (1120) einer ersten Bildanalyse basierend auf einem Zeitintervall zwischen einem ersten und einem zweiten Bild mittels der Verarbeitungsvorrichtung (110), um eine erste geschätzte Zeit bis zum Zusammenstoß des Fahrzeugs (200) mit dem Objekt zu bestimmen, wobei das Durchführen der ersten Bildanalyse Folgendes umfasst:

Bestimmen (1210) einer ersten Größe des Objekts, das auf dem ersten Bild erscheint;
Bestimmen (1220) einer zweiten Größe des Objekts, das auf dem zweiten Bild erscheint;
Bestimmen (1230) einer Größenänderung des Objekts basierend auf der ersten Größe und der zweiten Größe; und
Bestimmen (1240) der ersten geschätzten Zeit bis zum Zusammenstoß basierend auf der Änderung der Größe des Objekts und einem vorbestimmten Zeitintervall zwischen den Zeitpunkten der Aufnahme des ersten und des zweiten Bildes;
gekennzeichnet durch das Durchführen (1130) einer zweiten Bildanalyse mittels der Verarbeitungsvorrichtung (110), basierend auf einem Zeitintervall zwischen einem dritten und einem vierten Bild, um eine zweite geschätzte Zeit bis zum Zusammenstoß des Fahrzeugs mit dem Objekt zu bestimmen, wobei das Durchführen der zweiten Bildanalyse Folgendes umfasst:

Bestimmen (1320) eines Expansionsfokus anhand des dritten und vierten Bildes, wobei der Expansionsfokus ein konvergierender Punkt von Vektoren in einem aus dem dritten und dem vierten Bild bestimmten Vektorfeld ist, wobei jeder Punkt im Vektorfeld durch einen Vektor dargestellt wird, der eine Größe und eine Bewegungsrichtung von einem ersten Punkt im dritten Bild zu einem entsprechenden zweiten Punkt im vierten Bild anzeigt;
Bestimmen (1330) einer Expansionsrate in Bezug auf den Expansionsfokus für jeden Punkt im Vektorfeld durch Dividieren einer Bewegungsgröße eines dem Punkt zugeordneten Vektors durch eine Entfernung von dem Punkt zum Expansionsfokus; und
Bestimmen der zweiten geschätzten Zeit bis zum Zusammenstoß für einen Bereich, der im Wesentlichen auf das Objekt ausgerichtet ist, basierend auf der Expansionsrate und dem Zeitintervall zwischen den Zeiten der Aufnahme des dritten und vierten Bildes;
Berechnen (1140) mittels der Verarbeitungsvorrichtung (110) einer Differenz

IPR2025-00944
Tesla EX1002 Page 1501

zwischen der ersten geschätzten Zeit bis zum Zusammenstoß sowie der zweiten geschätzten Zeit bis zum Zusammenstoß;
Bestimmen (1150) mittels der Verarbeitungsvorrichtung (110), dass die Differenz einen vorbestimmten Schwellenwert nicht überschreitet; und
Veranlassen (1160) eines Bremsvorgangs des Fahrzeugs (200) mittels der Verarbeitungsvorrichtung (110), aufgrund der Bestimmung, dass die Differenz den vorbestimmten Schwellenwert nicht überschreitet.

9. Verfahren nach Anspruch 8, wobei das Durchführen der zweiten Bildanalyse ferner Folgendes umfasst:

Bestimmen einer Beziehung zwischen Bewegungsgrößen und Abständen zum Expansionsfokus für eine Vielzahl von Punkten im Vektorfeld; und
Bestimmen, dass das durch die Vielzahl von Punkten dargestellte Objekt ein aufrecht stehendes Objekt mit kurzer Zeit bis zum Zusammenstoß ist, aufgrund einer Bestimmung, dass die Beziehung eine lineare Beziehung ist.

10. Verfahren nach Anspruch 8, wobei das Durchführen der zweiten Bildanalyse ferner Folgendes umfasst:

Bestimmen einer Beziehung zwischen Bewegungsgrößen und Abständen zum Expansionsfokus für eine Vielzahl von Punkten im Vektorfeld; und
Bestimmen, dass mindestens ein Teil des Objekts eine Straßenoberfläche ist, aufgrund einer Bestimmung, dass die Beziehung eine nichtlineare Beziehung ist.

11. Verfahren nach Anspruch 8, wobei das Durchführen der zweiten Bildanalyse ferner Folgendes umfasst:

Erkennen eines ersten Expansionsfokus, der ein erster Konvergenzpunkt einer ersten Vielzahl von Vektoren im Vektorfeld ist;
Erkennen eines zweiten Expansionsfokus, der ein zweiter Konvergenzpunkt einer zweiten Vielzahl von Vektoren im Vektorfeld ist; und
Bestimmen, dass das dritte und vierte Bild ein sich bewegendes Objekt beinhalten, das sich in einer Richtung bewegt, die nicht parallel zu einer Fahrtrichtung des Fahrzeugs ist.

12. Verfahren nach Anspruch 8, wobei das Durchführen der zweiten Bildanalyse ferner Folgendes umfasst:
Bestimmen, dass das dritte und vierte Bild eine Umgebung darstellen, die in Bezug auf eine Straßenoberfläche, auf der das Fahrzeug fährt, statisch ist.

13. Verfahren nach Anspruch 8, wobei das Aufnehmen (1110) der Vielzahl von Bildern Folgendes umfasst:

Erfassen, durch einen ersten Bildsensor, der in der mindestens einen Bilderfassungsvorrichtung vorhanden ist, des
ersten und des zweiten Bildes, die bei der ersten Bildanalyse verwendet werden sollen; und
Erfassen des dritten und des vierten Bildes, die bei der zweiten Bildanalyse verwendet werden sollen, durch einen zweiten Bildsensor, der in der mindestens einen Bilderfassungsvorrichtung vorhanden ist und vom ersten Bildsensor unabhängig ist.

Revendications

1. Système de détection d'objet et de freinage pour véhicule (200), comprenant :

au moins un dispositif de capture d'image (122, 124, 126) conçu pour acquérir une pluralité d'images d'une zone comprenant un objet situé devant le véhicule (200) ;
une interface de données (128) ; et
au moins un dispositif de traitement (110) programmé pour :

recevoir (1110) la pluralité d'images par l'intermédiaire de l'interface de données (128) ;
effectuer (1120) une première analyse d'image sur la base d'un intervalle de temps entre une première image et une deuxième image pour déterminer un premier temps avant collision estimé du véhicule (200) avec l'objet, la première analyse d'image consistant à :

déterminer (1210) une première taille de l'objet apparaissant dans la première image ;
déterminer (1220) une seconde taille de l'objet apparaissant dans la deuxième image ;
déterminer (1230) un changement de taille de l'objet sur la base de la première taille et de la seconde taille ; et
déterminer (1240) le premier temps avant collision estimé sur la base du changement de taille de l'objet et d'un intervalle de temps prédéterminé entre les temps d'acquisition de la première image et de la deuxième image ;

IPR2025-00944
Tesla EX1002 Page 1502

caractérisé en ce que l'au moins un dispositif de traitement (110) est en outre programmé pour :

effectuer (1130) une seconde analyse d'image sur la base d'un intervalle de temps entre une troisième image et une quatrième image pour déterminer un second temps avant collision estimé du véhicule (200) avec l'objet, la seconde analyse d'image consistant à :

déterminer (1320) un foyer d'expansion basé sur la troisième image et la quatrième image, le foyer d'expansion étant un point convergent de vecteurs dans un champ de vecteurs déterminé à partir de la troisième image et de la quatrième image, dans lequel chaque point du champ de vecteurs est représenté par un vecteur indiquant une amplitude et une direction de mouvement d'un premier point dans la troisième image à un second point correspondant dans la quatrième image ;
déterminer (1330) une vitesse d'expansion par rapport au foyer d'expansion pour chaque point dans le champ de vecteurs en divisant une amplitude de mouvement d'un vecteur associé au point par une distance du point au foyer d'expansion ;
et
déterminer pour une région se focalisant sensiblement sur l'objet, le second temps avant collision estimé sur la base de la vitesse d'expansion et d'un intervalle de temps entre les temps d'acquisition de la troisième image et de la quatrième image ;
calculer (1140) une différence entre le premier temps avant collision estimé et le second temps avant collision estimé ;
déterminer (1150) que la différence ne dépasse pas un seuil prédéterminé ;
et
provoquer (1160) le freinage du véhicule (200) sur la base de la détermination que la différence ne dépasse pas le seuil prédéterminé.

2. Système de détection d'objet et de freinage selon la revendication 1, dans lequel la seconde analyse d'image consiste en outre à :

déterminer une relation entre des amplitudes de mouvement et des distances au foyer d'expansion pour une pluralité de points dans le champ de vecteurs ; et
déterminer que l'objet représenté par la pluralité de points est un objet vertical à faible temps de collision, basé sur la détermination que la relation est une relation linéaire.

3. Système de détection d'objet et de freinage selon la revendication 1, dans lequel la seconde analyse d'image consiste en outre à :

déterminer une relation entre des amplitudes de mouvement et le foyer d'expansion pour une pluralité de points dans le champ de vecteurs ; et
déterminer qu'au moins une partie de l'objet est une surface de chaussée sur la base d'une détermination que la relation est une relation non linéaire.

4. Système de détection d'objet et de freinage selon la revendication 1, dans lequel la seconde analyse d'image consiste en outre à :

détecter un premier foyer d'expansion qui est un premier point convergent d'une première pluralité de vecteurs dans le champ de vecteurs ;
détecter un second foyer d'expansion qui est un second point convergent d'une seconde pluralité de vecteurs dans le champ de vecteurs ; et
déterminer que la troisième image et la quatrième image comprennent un objet en mouvement qui se déplace dans une direction qui n'est pas parallèle à une direction de déplacement du véhicule.

5. Système de détection d'objet et de freinage selon la revendication 1, dans lequel la seconde analyse d'image consiste en outre à :
déterminer que la troisième image et la quatrième image représentent un environnement qui est statique par rapport à une surface de chaussée sur laquelle le véhicule se déplace.

6. Système de détection d'objet et de freinage selon la revendication 1, dans lequel l'au moins un dispositif de capture d'image comprend :

un premier capteur d'image conçu pour capturer la première image et la deuxième image à utiliser dans la première analyse d'image ; et
un second capteur d'image indépendant du premier capteur d'image et conçu pour capturer la troisième image et la quatrième image à utiliser dans la seconde analyse d'image.

7. Véhicule (200), comprenant :

un corps ; et
le système de détection d'objet et de freinage selon l'une quelconque des revendications précédentes.

8. Procédé de détection d'un objet et de freinage d'un véhicule (200), consistant à :

IPR2025-00944
Tesla EX1002 Page 1503

acquérir (1110), par l'intermédiaire d'au moins un dispositif de capture d'image (122, 124, 126), une pluralité d'images d'une zone comprenant un objet situé devant le véhicule (200) ;
recevoir, par l'intermédiaire d'un dispositif de traitement (110), la pluralité d'images ;
effectuer (1120), par l'intermédiaire du dispositif de traitement (110), une première analyse d'image sur la base d'un intervalle de temps entre une première image et une deuxième image pour déterminer un premier temps avant collision estimé du véhicule (200) avec l'objet, dans lequel effectuer la première analyse d'image consiste à :

déterminer (1210) une première taille de l'objet apparaissant dans la première image ;
déterminer (1220) une seconde taille de l'objet apparaissant dans la deuxième image ;
déterminer (1230) un changement de taille de l'objet sur la base de la première taille et de la seconde taille ; et
déterminer (1240) le premier temps avant collision estimé sur la base du changement de taille de l'objet et d'un intervalle de temps prédéterminé entre les temps d'acquisition de la première image et de la deuxième image ;

caractérisé par le fait
d'effectuer (1130), par l'intermédiaire du dispositif de traitement (110), une seconde analyse d'image sur la base d'un intervalle de temps entre une troisième image et une quatrième image pour déterminer un second temps de collision estimé du véhicule avec l'objet, dans lequel effectuer la seconde analyse d'image consiste à :

déterminer (1320) un foyer d'expansion basé sur la troisième image et la quatrième image, le foyer d'expansion étant un point convergent de vecteurs dans un champ de vecteurs déterminé à partir de la troisième image et de la quatrième image, dans lequel chaque point du champ de vecteurs est représenté par un vecteur indiquant une amplitude et une direction de mouvement d'un premier point dans la troisième image à un second point correspondant dans la quatrième image ;
déterminer (1330) une vitesse d'expansion par rapport au foyer d'expansion pour chaque point dans le champ de vecteurs en divisant une amplitude de mouvement d'un vecteur associé au point par une distance du point au foyer d'expansion ; et
déterminer pour une région se focalisant sensiblement sur l'objet, le second temps avant collision estimé sur la base de la vitesse d'expansion et de l'intervalle de temps entre les temps d'acquisition de la troisième image et de la quatrième image ;

calculer (1140), par l'intermédiaire du dispositif de traitement (110), une différence entre le premier temps avant collision estimé et le second temps avant collision estimé ;
déterminer (1150), par l'intermédiaire du dispositif de traitement (110), que la différence ne dépasse pas un seuil prédéterminé ; et
provoquer (1160), par l'intermédiaire du dispositif de traitement (110), le freinage du véhicule (200) sur la base de la détermination que la différence ne dépasse pas le seuil prédéterminé.

9. Procédé selon la revendication 8, dans lequel le fait d'effectuer la seconde analyse d'image consiste en outre à :

déterminer une relation entre des amplitudes de mouvement et des distances au foyer d'expansion pour une pluralité de points dans le champ de vecteurs ; et
déterminer que l'objet représenté par la pluralité de points est un objet vertical à faible temps de collision, basé sur la détermination que la relation est une relation linéaire.

10. Procédé selon la revendication 8, dans lequel le fait d'effectuer la seconde analyse d'image consiste en outre à :

déterminer une relation entre des amplitudes de mouvement et des distances au foyer d'expansion pour une pluralité de points dans le champ de vecteurs ; et
déterminer qu'au moins une partie de l'objet est une surface de chaussée sur la base d'une détermination que la relation est une relation non linéaire.

11. Procédé selon la revendication 8, dans lequel le fait d'effectuer la seconde analyse d'image consiste en outre à :

détecter un premier foyer d'expansion qui est un premier point convergent d'une première pluralité de vecteurs dans le champ de vecteurs ;
détecter un second foyer d'expansion qui est un second point convergent d'une seconde pluralité de vecteurs dans le champ de vecteurs ; et
déterminer que la troisième image et la quatrième image comprennent un objet en mouvement qui se déplace dans une direction qui n'est pas

IPR2025-00944
Tesla EX1002 Page 1504

parallèle à une direction de déplacement du véhicule.

12. Procédé selon la revendication 8, dans lequel le fait d'effectuer la seconde analyse d'image consiste en outre à :
déterminer que la troisième image et la quatrième image représentent un environnement qui est statique par rapport à une surface de chaussée sur laquelle le véhicule se déplace.

13. Procédé selon la revendication 8, dans lequel l'acquisition (1110) de la pluralité d'images consiste en outre à :

capturer, par un premier capteur d'image inclus dans l'au moins un dispositif de capture d'image, la première image et la deuxième image à utiliser dans la première analyse d'image ; et
capturer, par un second capteur d'image inclus dans l'au moins un dispositif de capture d'image et étant indépendant du premier capteur d'image, la troisième image et la quatrième image à utiliser dans la seconde analyse d'image.

IPR2025-00944
Tesla EX1002 Page 1505

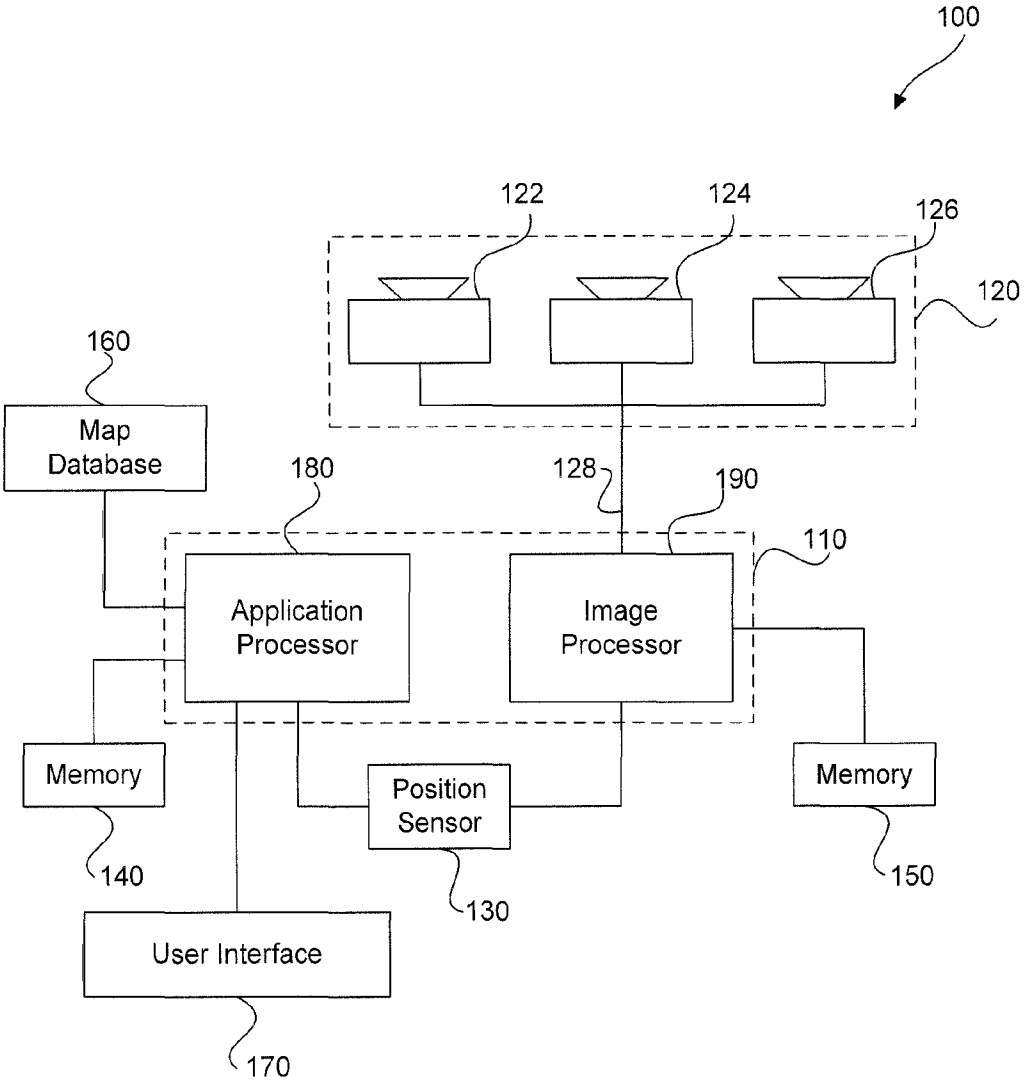


FIG. 1

IPR2025-00944
Tesla EX1002 Page 1506

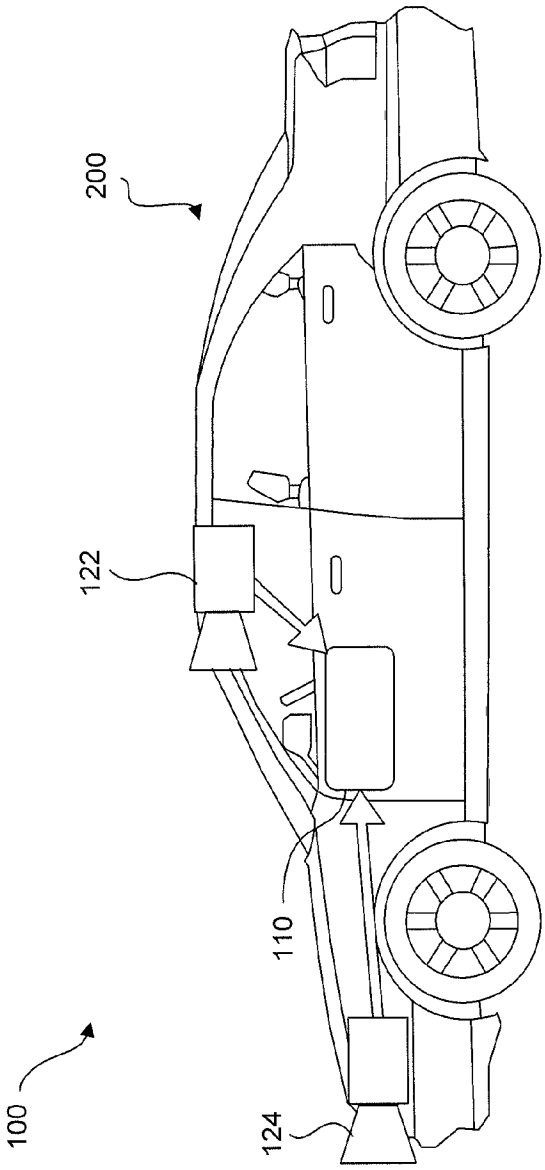


FIG. 2A

IPR2025-00944
Tesla EX1002 Page 1507

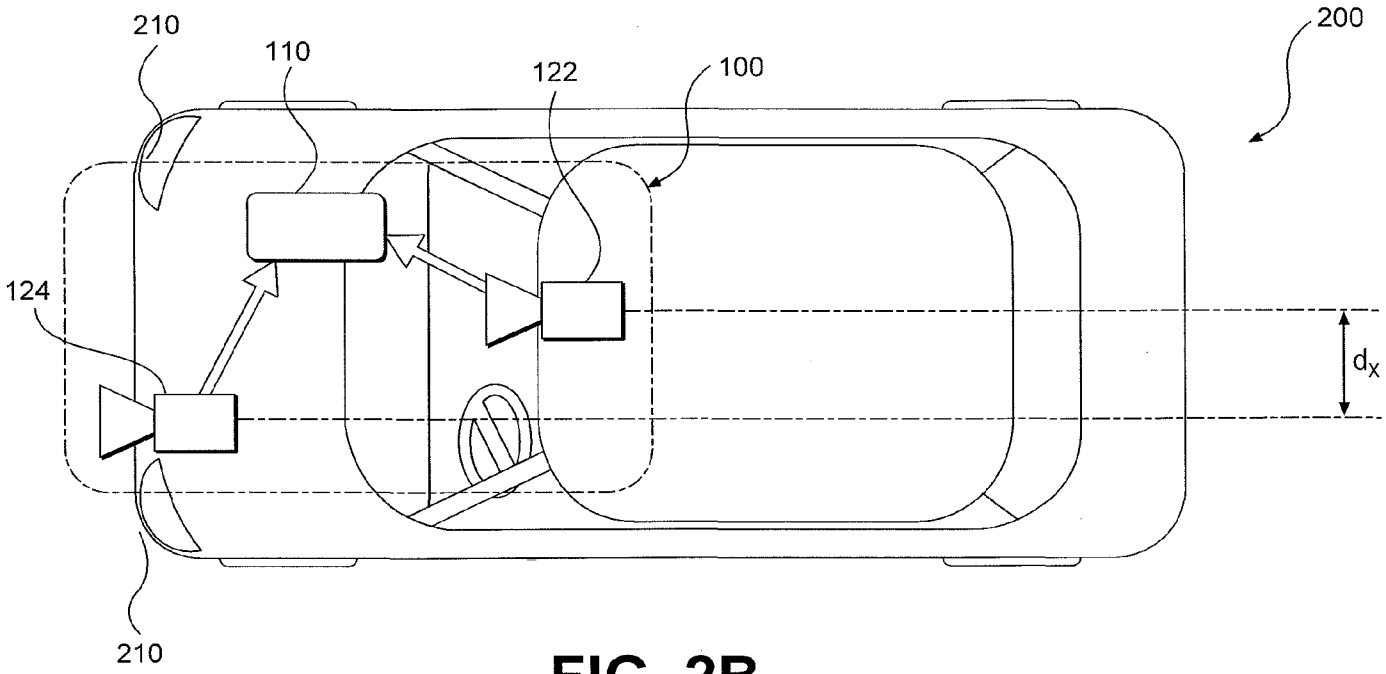


FIG. 2B

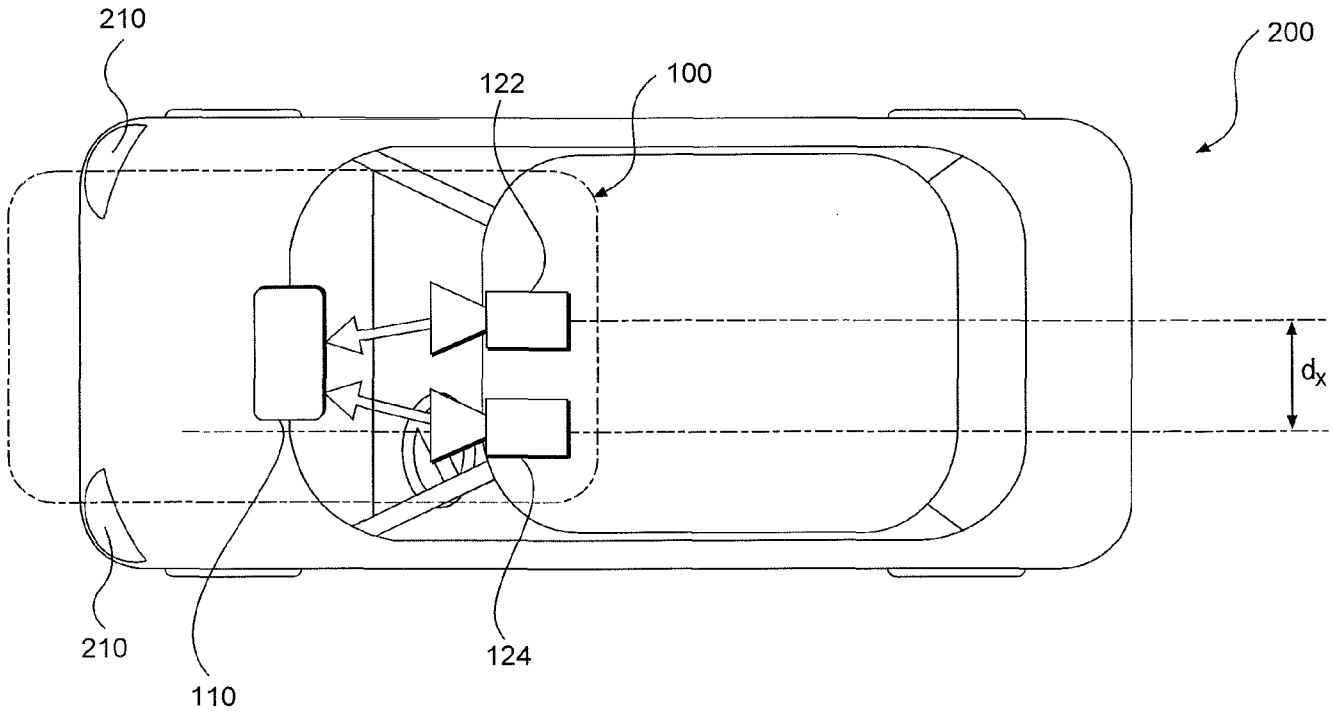


FIG. 2C

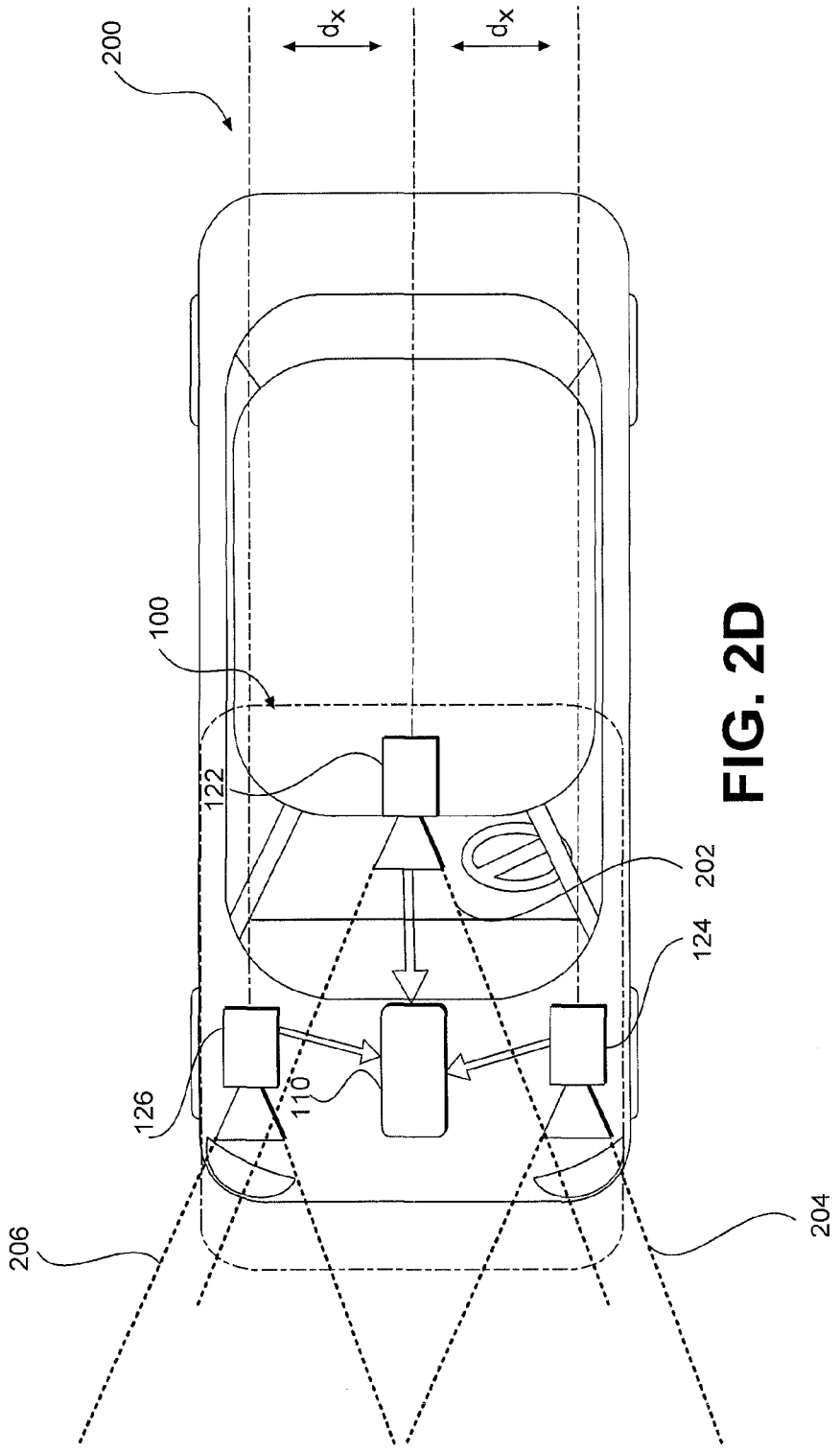


FIG. 2D

IPR2025-00944
Tesla EX1002 Page 1510

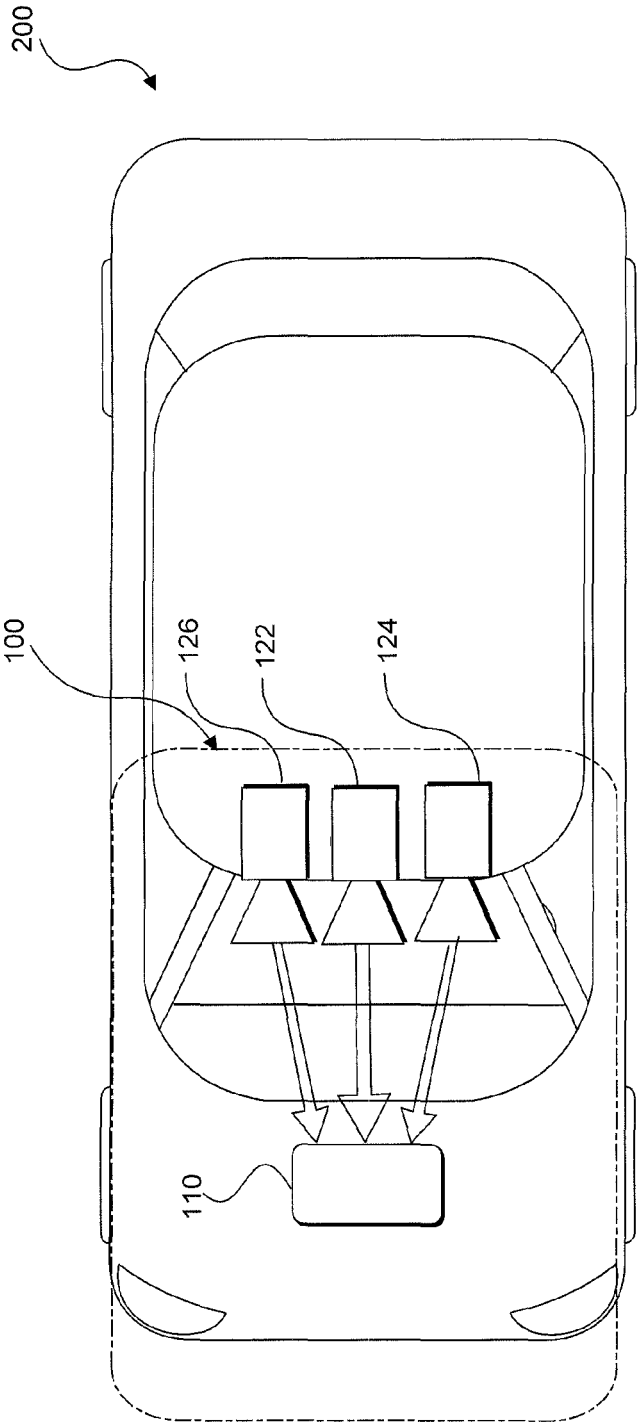


FIG. 2E

IPR2025-00944
Tesla EX1002 Page 1511

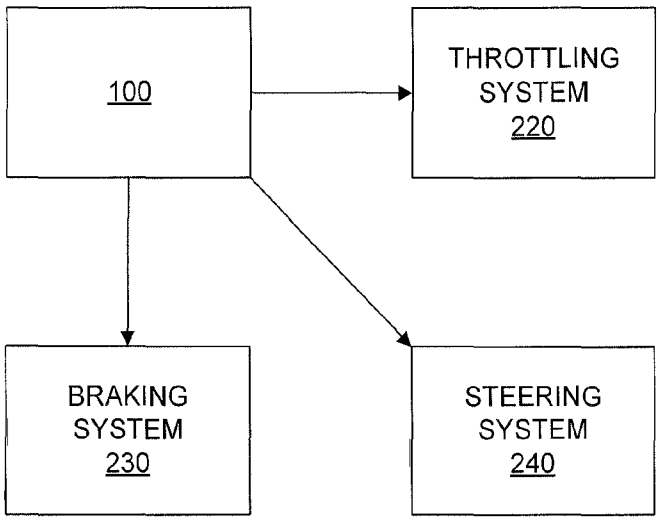


FIG. 2F

IPR2025-00944
Tesla EX1002 Page 1512

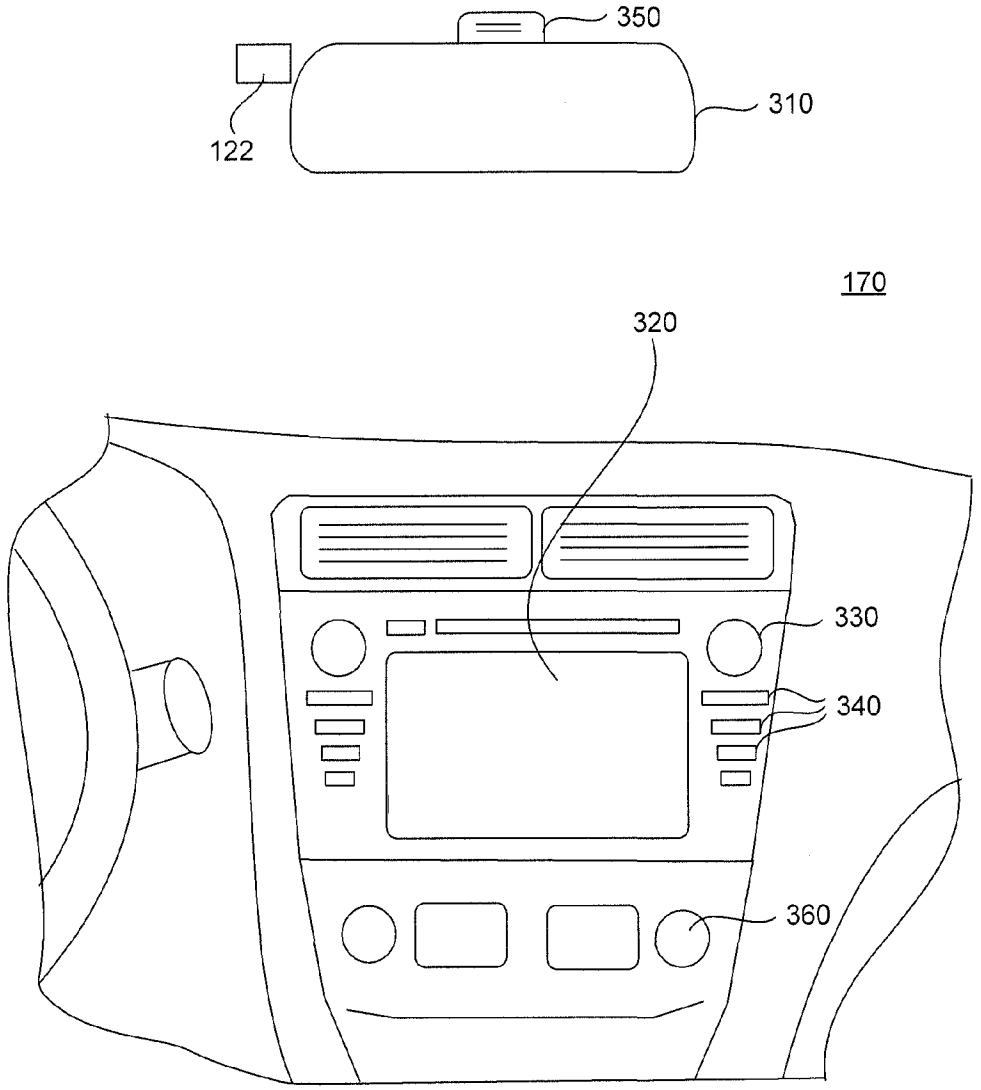


FIG. 3A

IPR2025-00944
Tesla EX1002 Page 1513

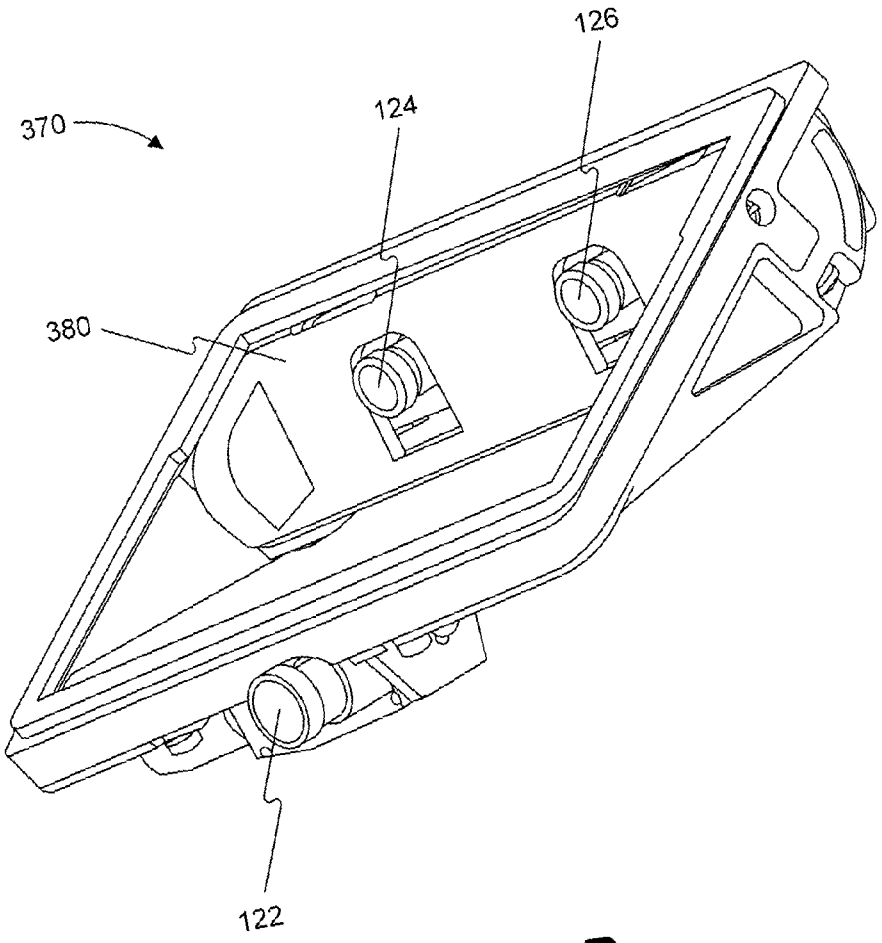


FIG. 3B

IPR2025-00944
Tesla EX1002 Page 1514

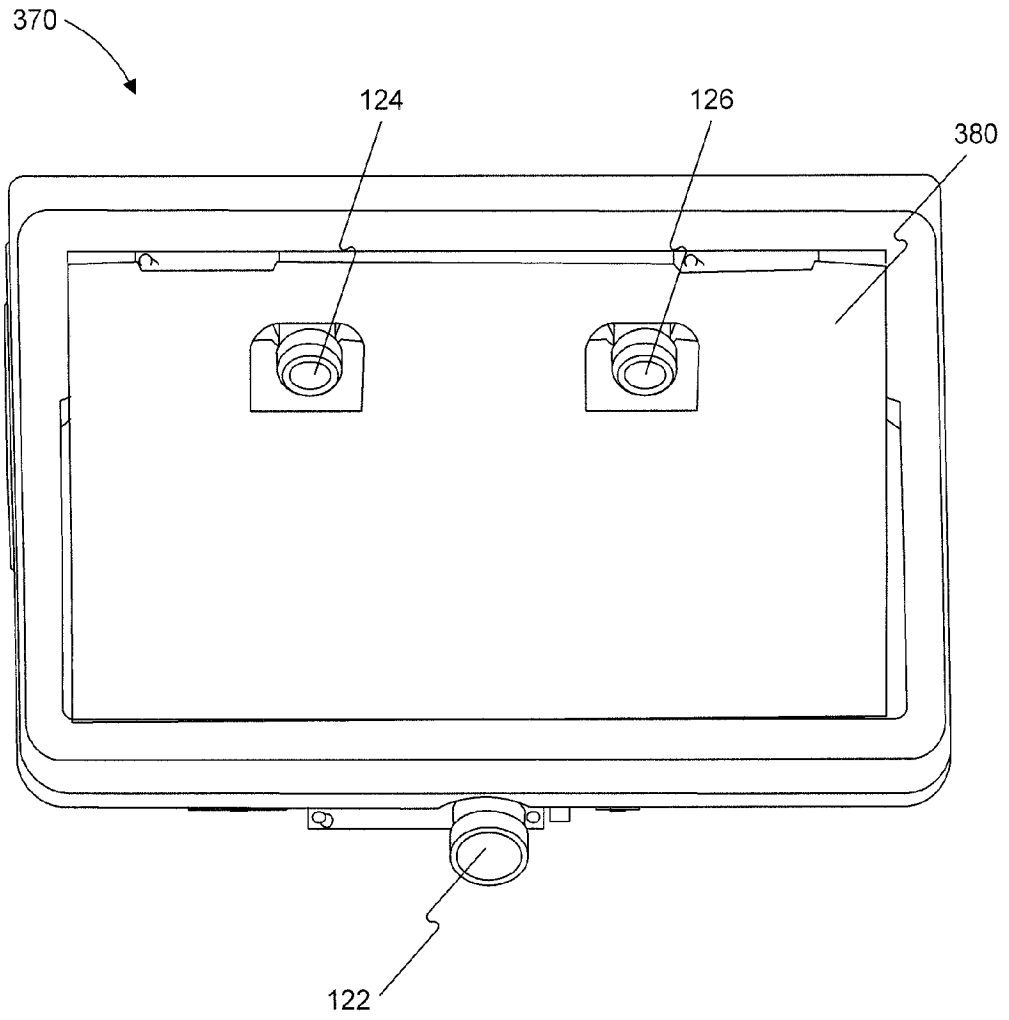


FIG. 3C

IPR2025-00944
Tesla EX1002 Page 1515

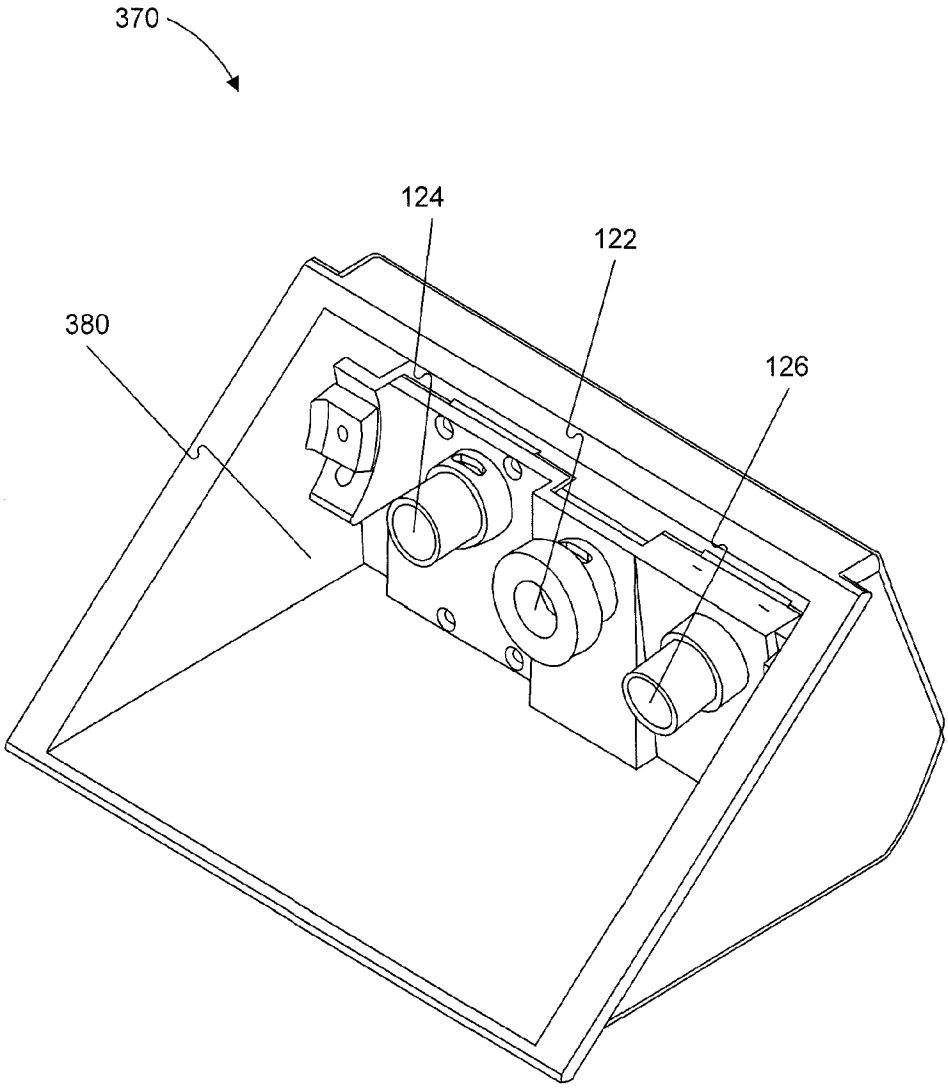


FIG. 3D

IPR2025-00944
Tesla EX1002 Page 1516

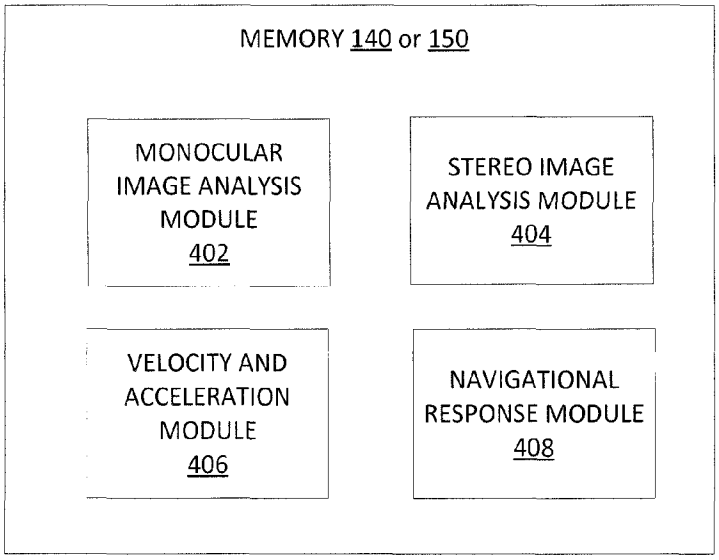


FIG. 4

IPR2025-00944
Tesla EX1002 Page 1517

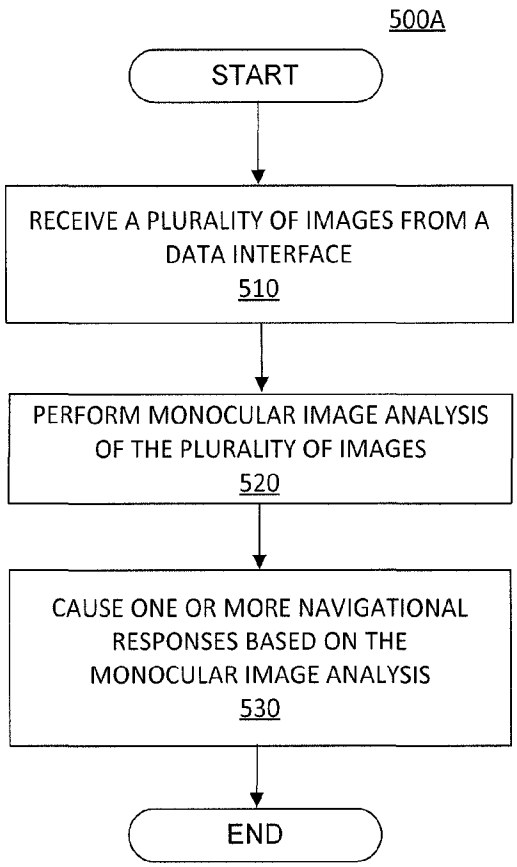


FIG. 5A

IPR2025-00944
Tesla EX1002 Page 1518

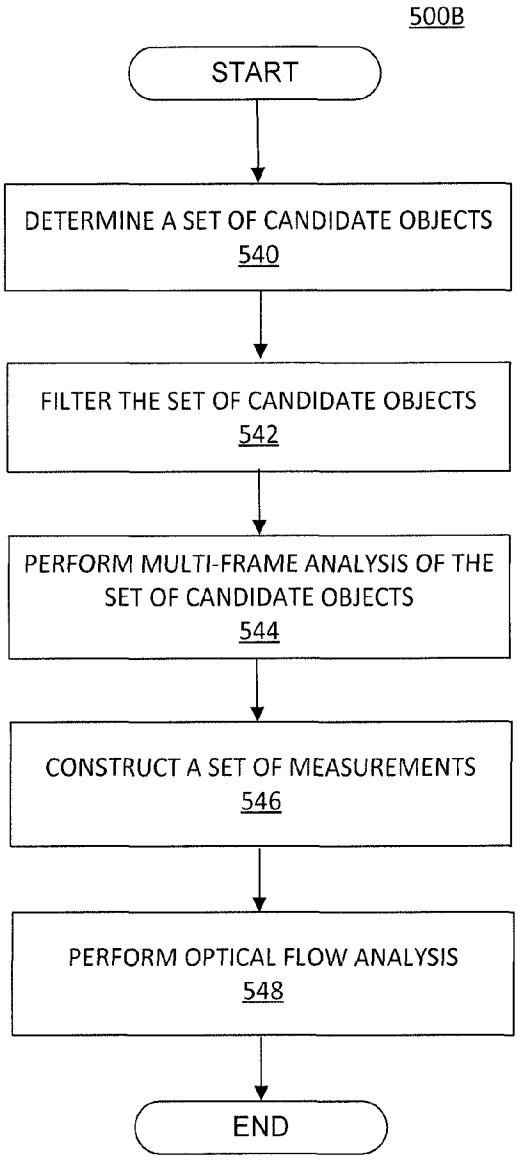


FIG. 5B

IPR2025-00944
Tesla EX1002 Page 1519

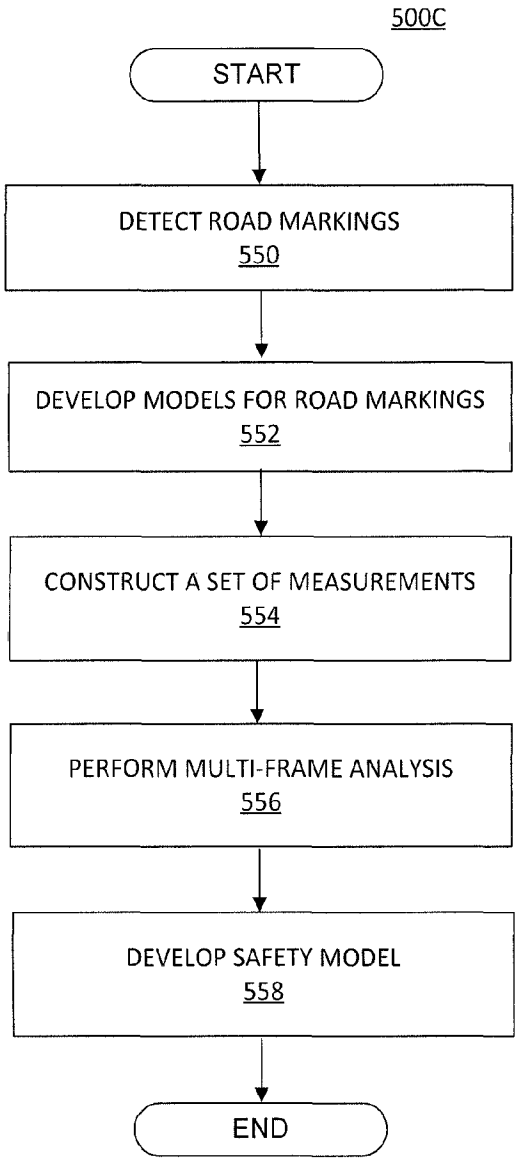


FIG. 5C

IPR2025-00944
Tesla EX1002 Page 1520

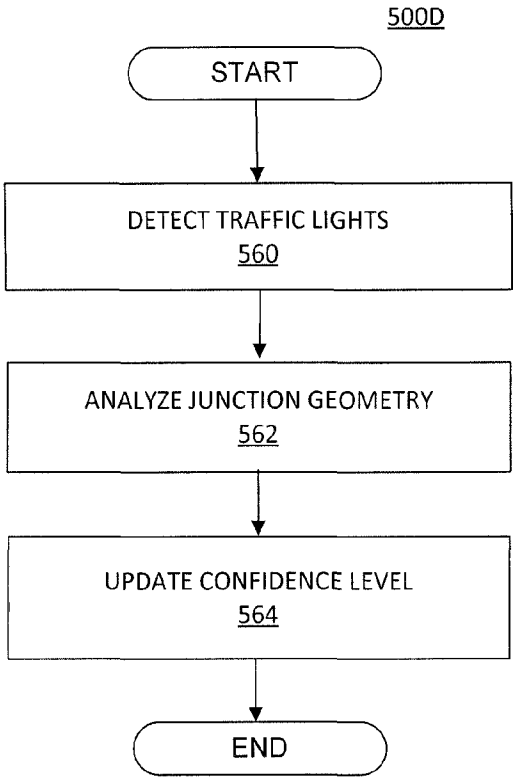


FIG. 5D

IPR2025-00944
Tesla EX1002 Page 1521

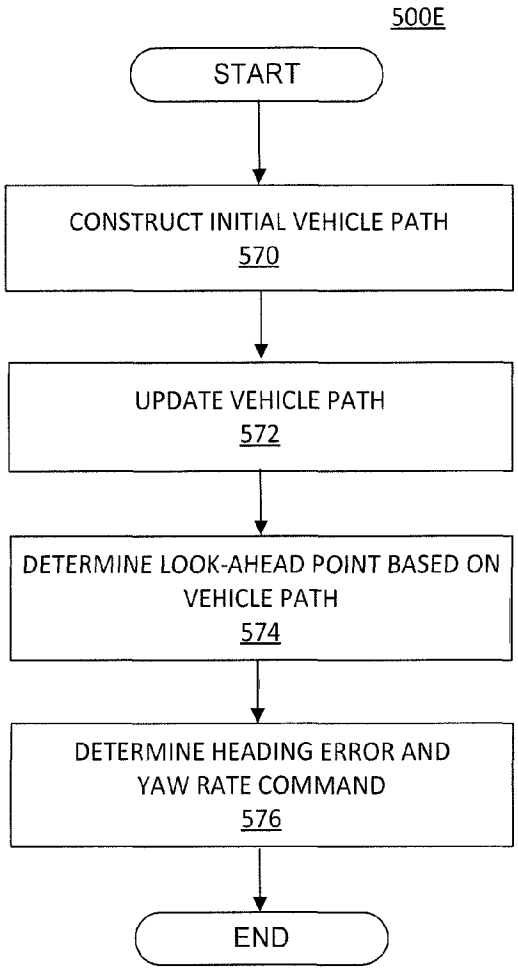


FIG. 5E

IPR2025-00944
Tesla EX1002 Page 1522

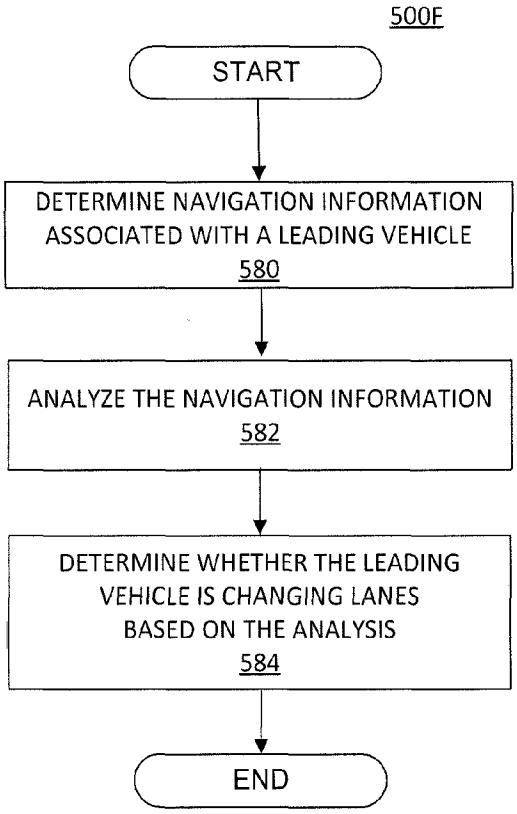


FIG. 5F

IPR2025-00944
Tesla EX1002 Page 1523

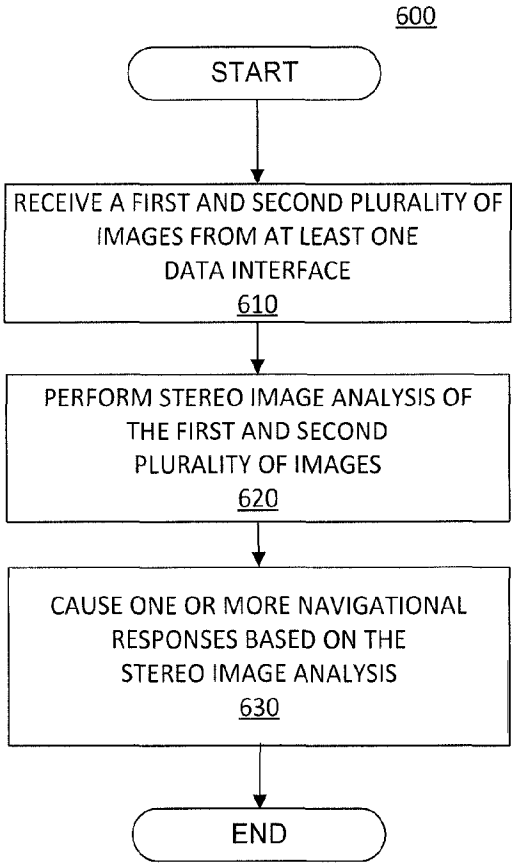


FIG. 6

IPR2025-00944
Tesla EX1002 Page 1524

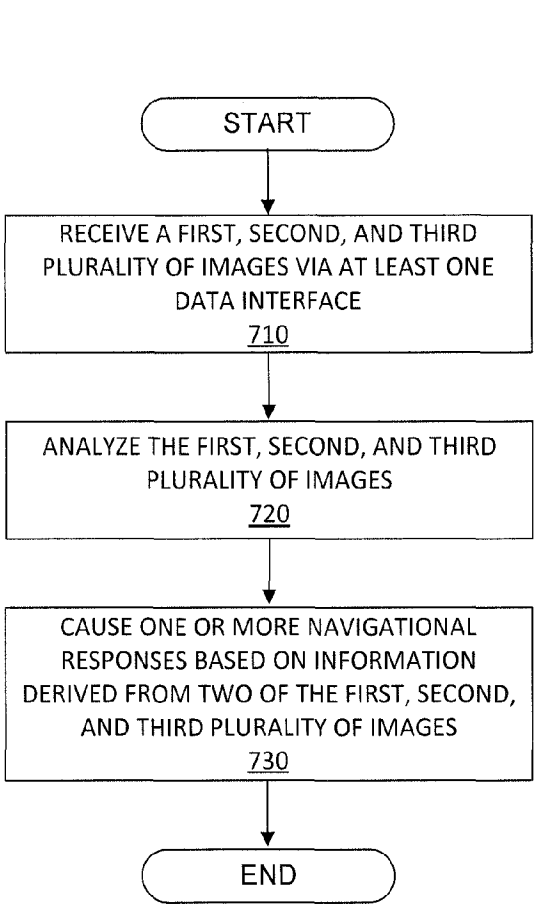


FIG. 7

IPR2025-00944
Tesla EX1002 Page 1525

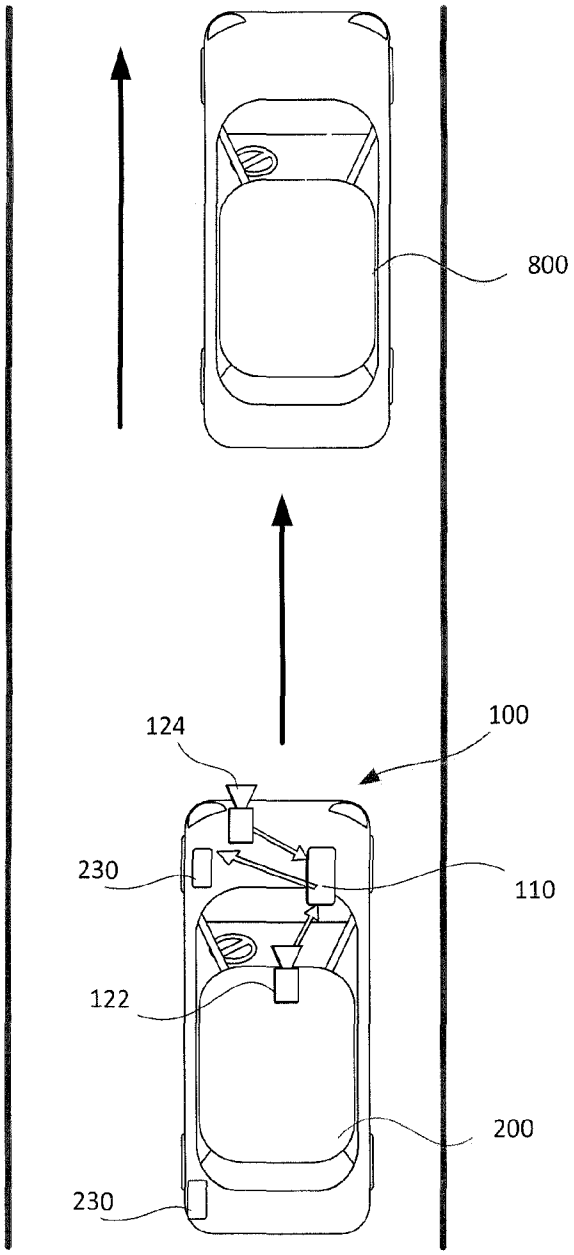


FIG. 8

IPR2025-00944
Tesla EX1002 Page 1526

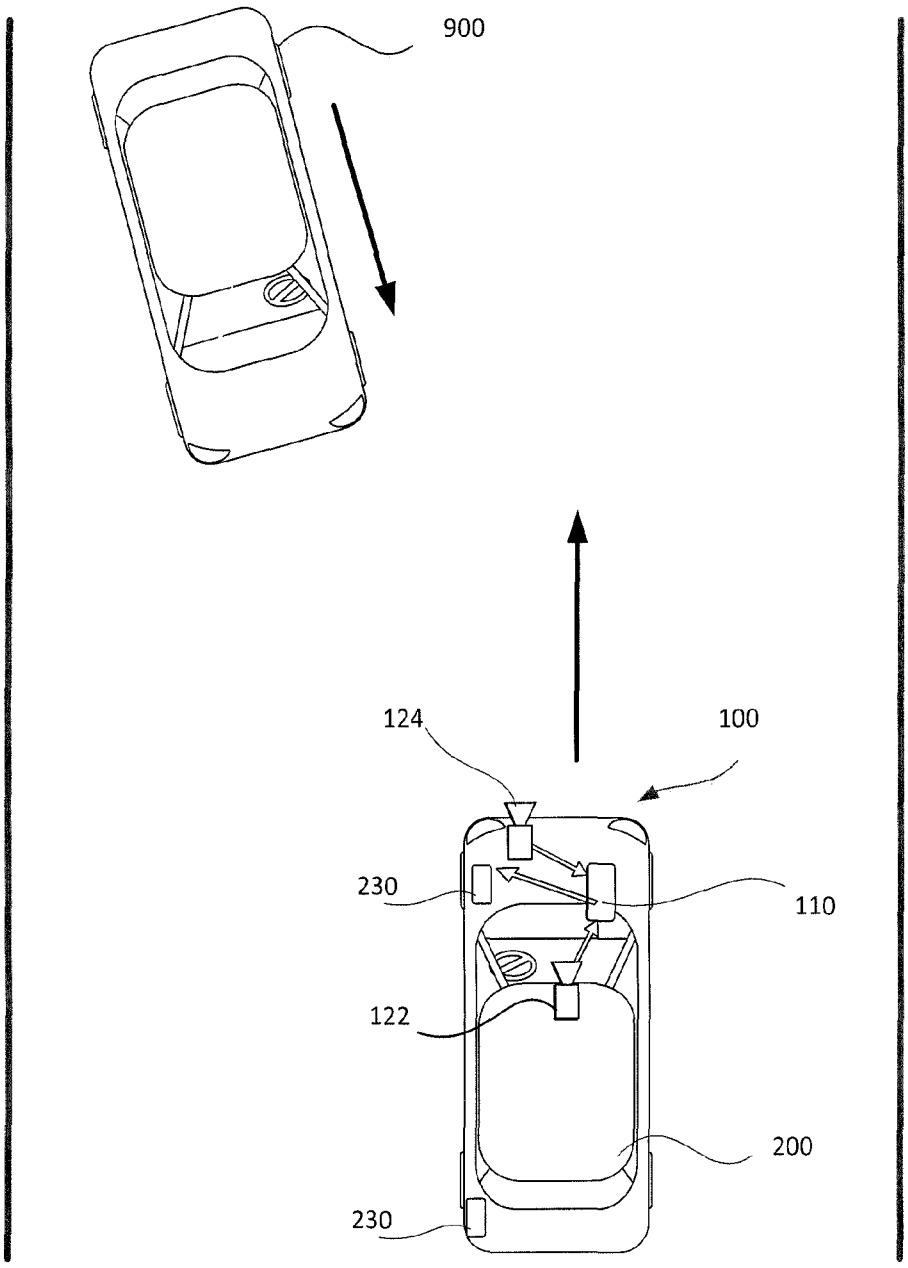


FIG. 9

IPR2025-00944
Tesla EX1002 Page 1527

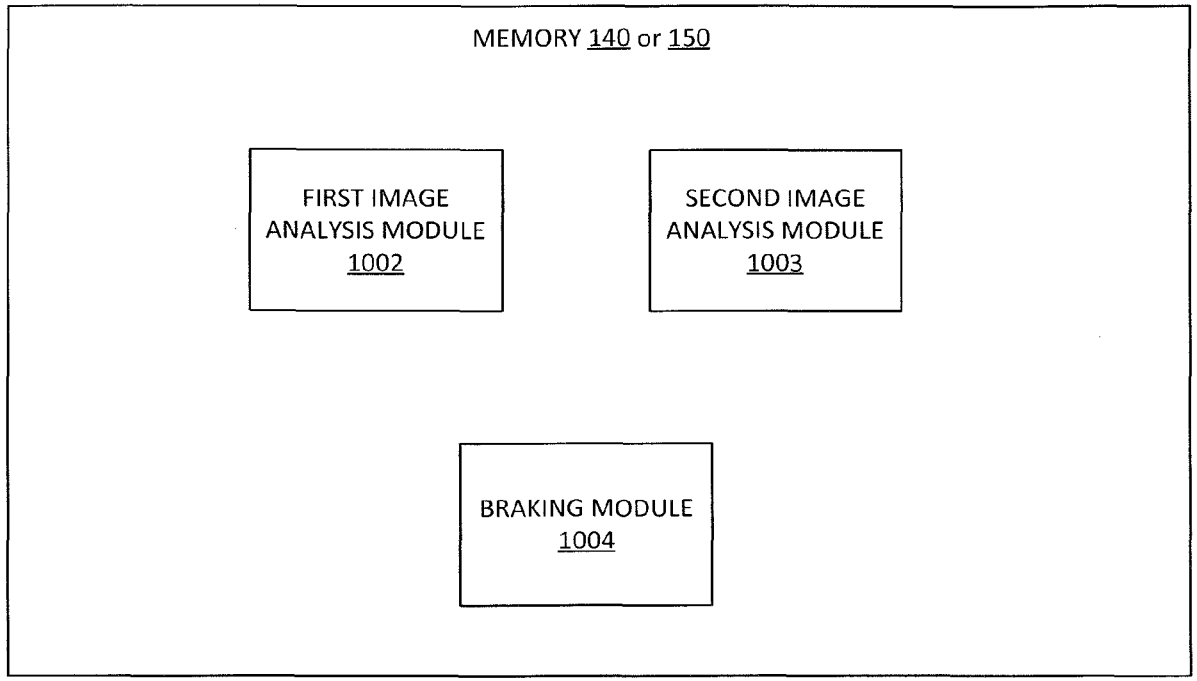

MEMORY 140 or 150
FIRST IMAGE ANALYSIS MODULE 1002
SECOND IMAGE ANALYSIS MODULE 1003
BRAKING MODULE 1004

FIG. 10

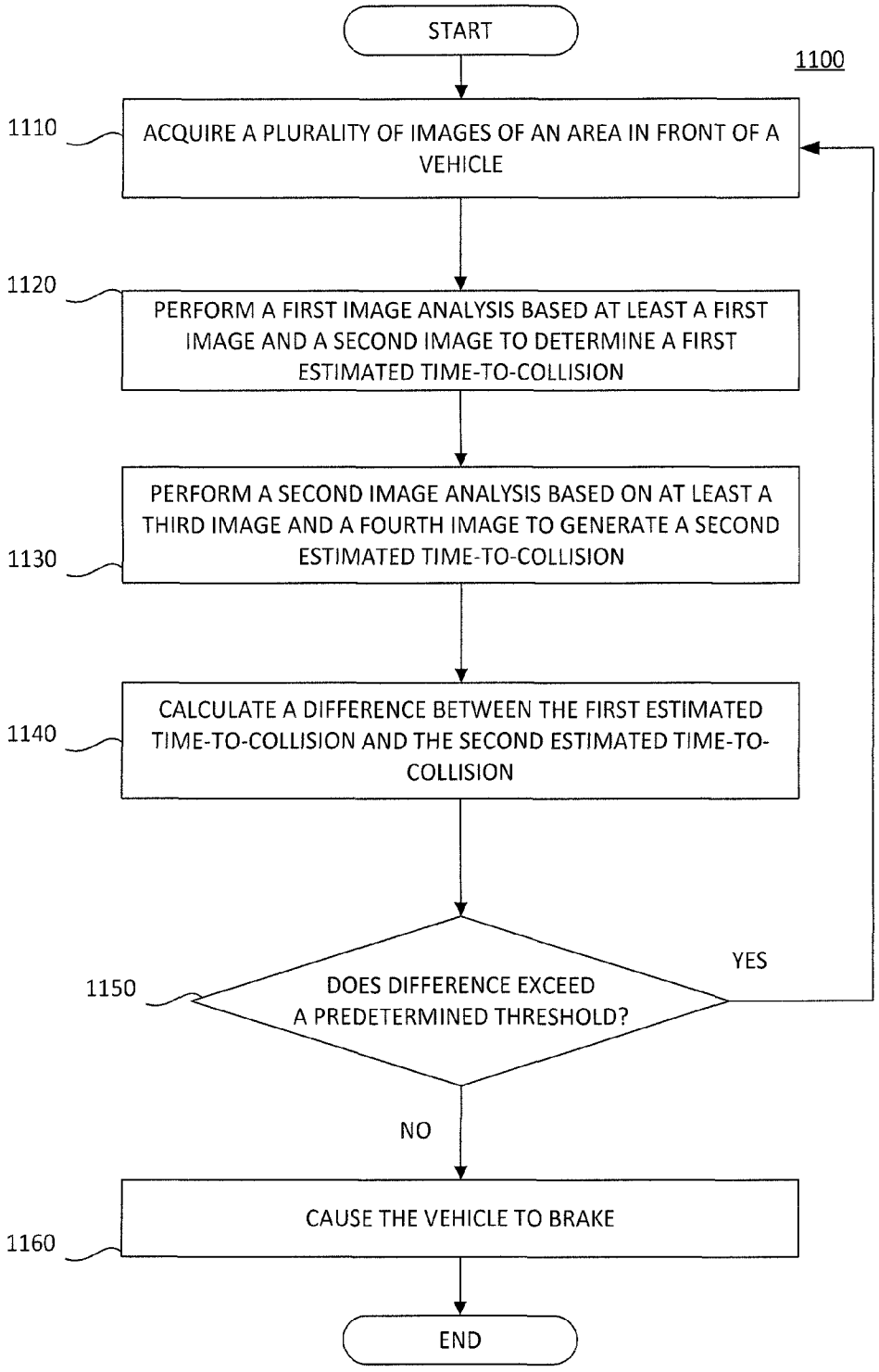


FIG. 11

IPR2025-00944
Tesla EX1002 Page 1529

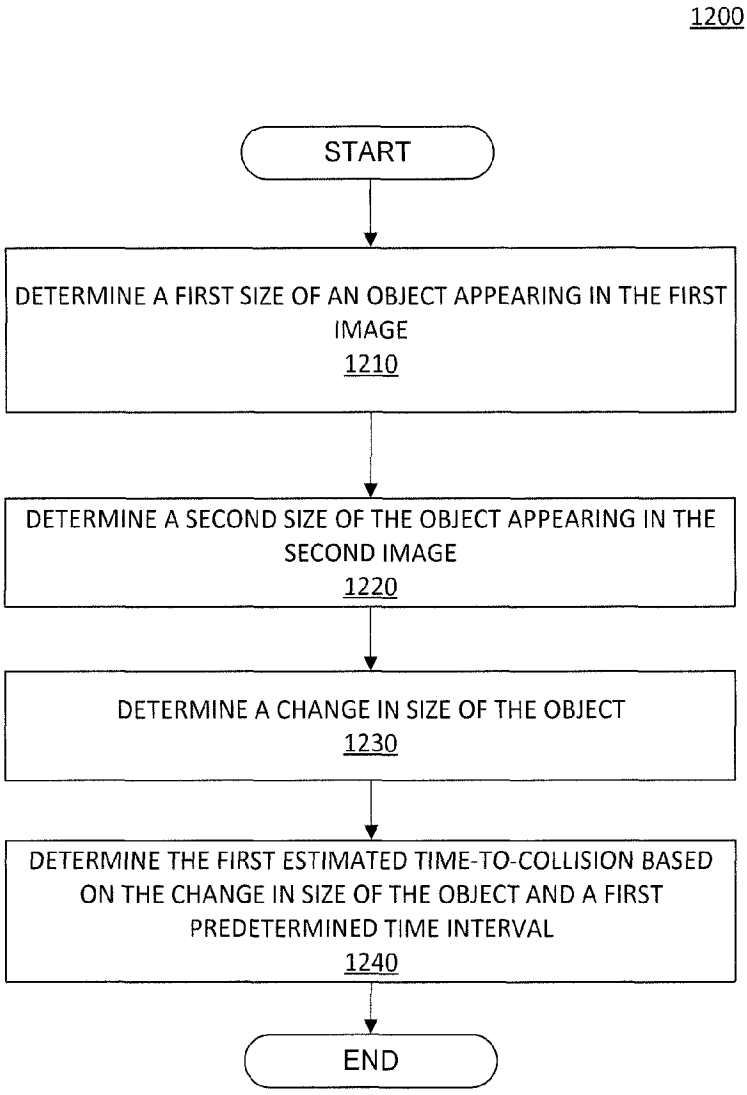


FIG. 12

IPR2025-00944
Tesla EX1002 Page 1530

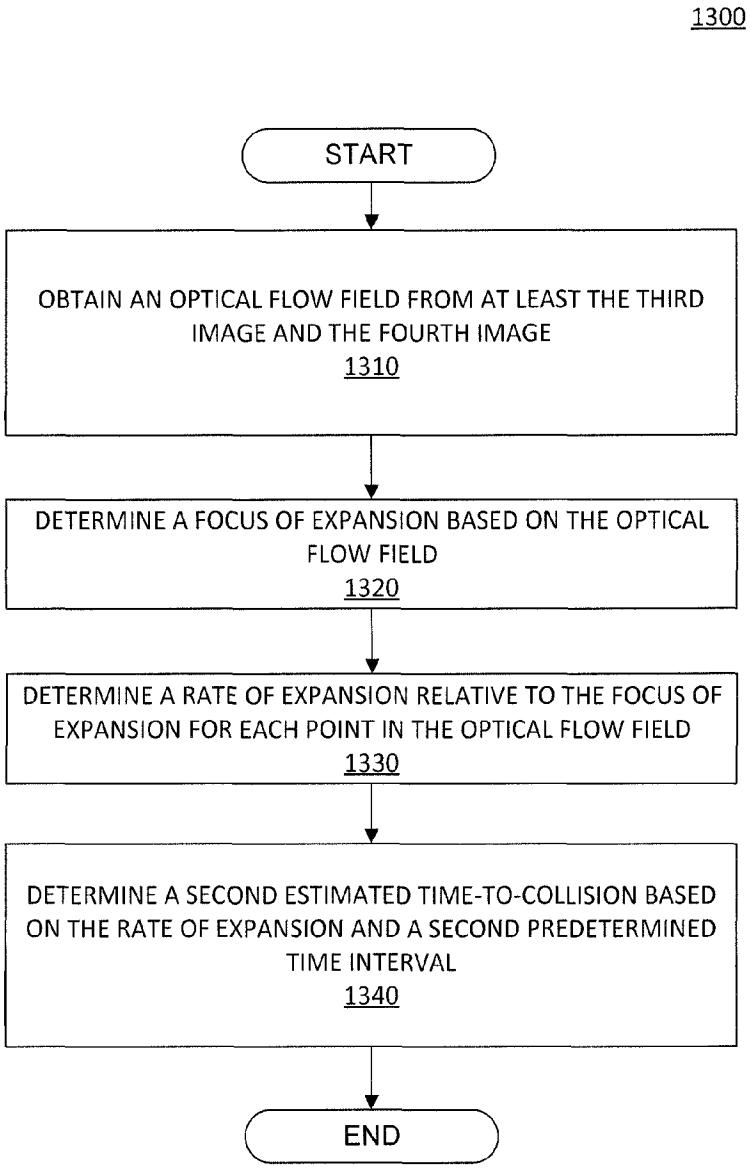


FIG. 13

IPR2025-00944
Tesla EX1002 Page 1531

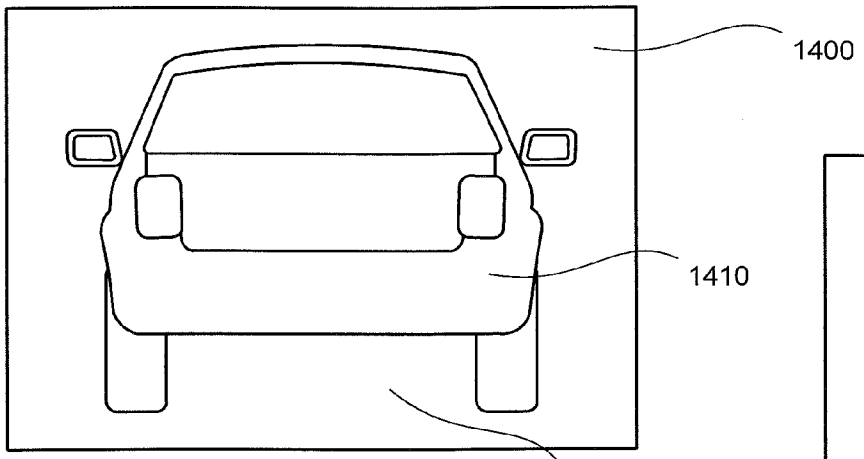


FIG. 14A

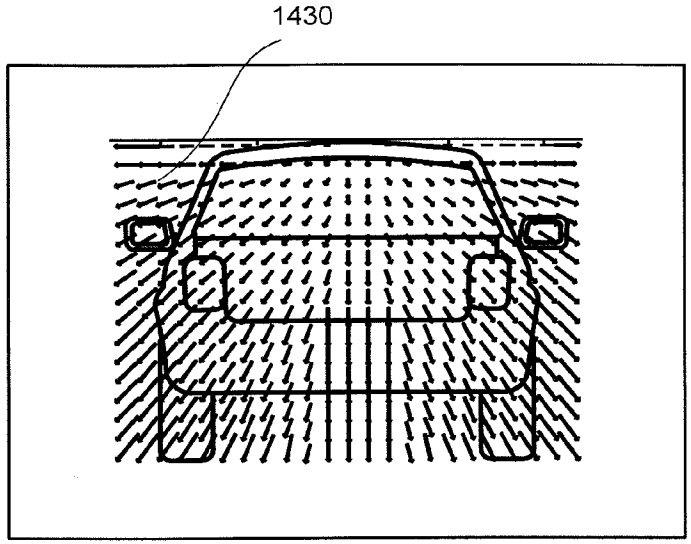


FIG. 14B

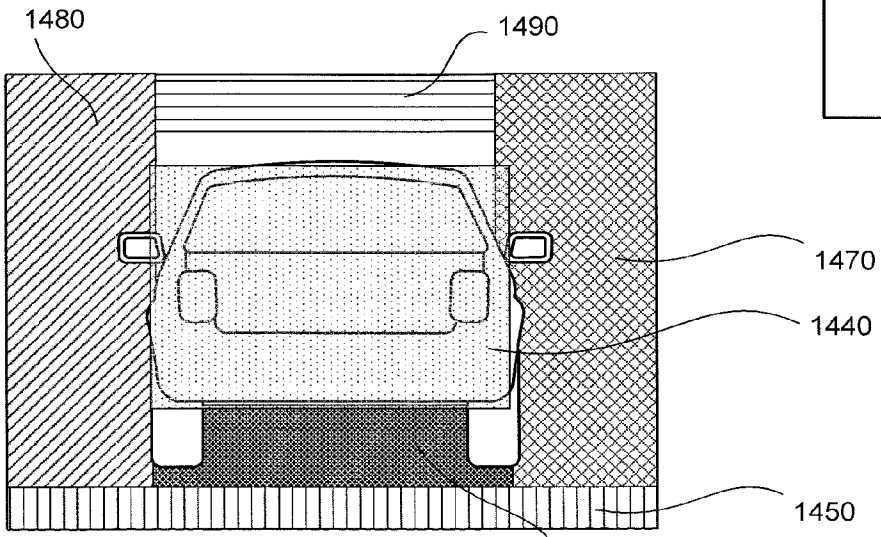


FIG. 14C

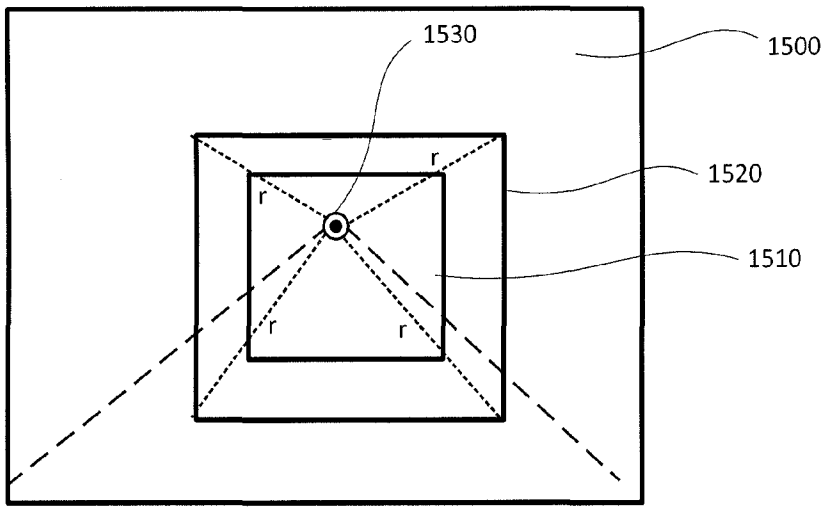


FIG. 15A

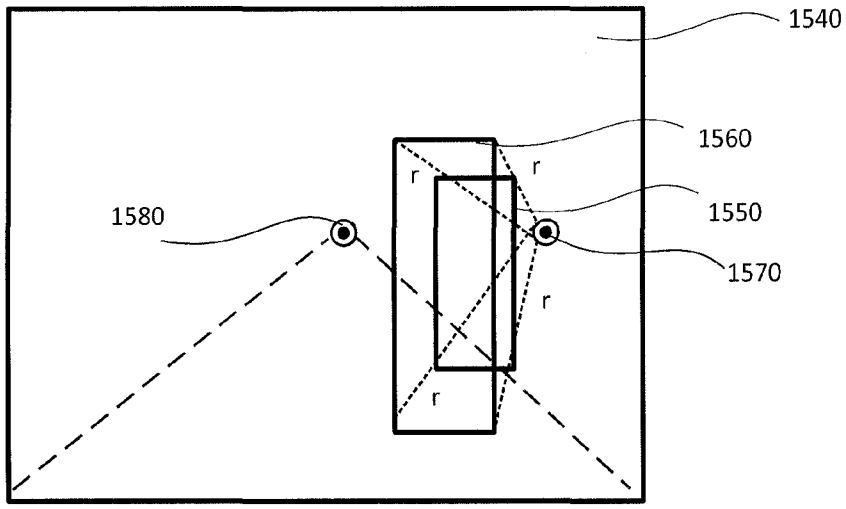


FIG. 15B

IPR2025-00944
Tesla EX1002 Page 1533

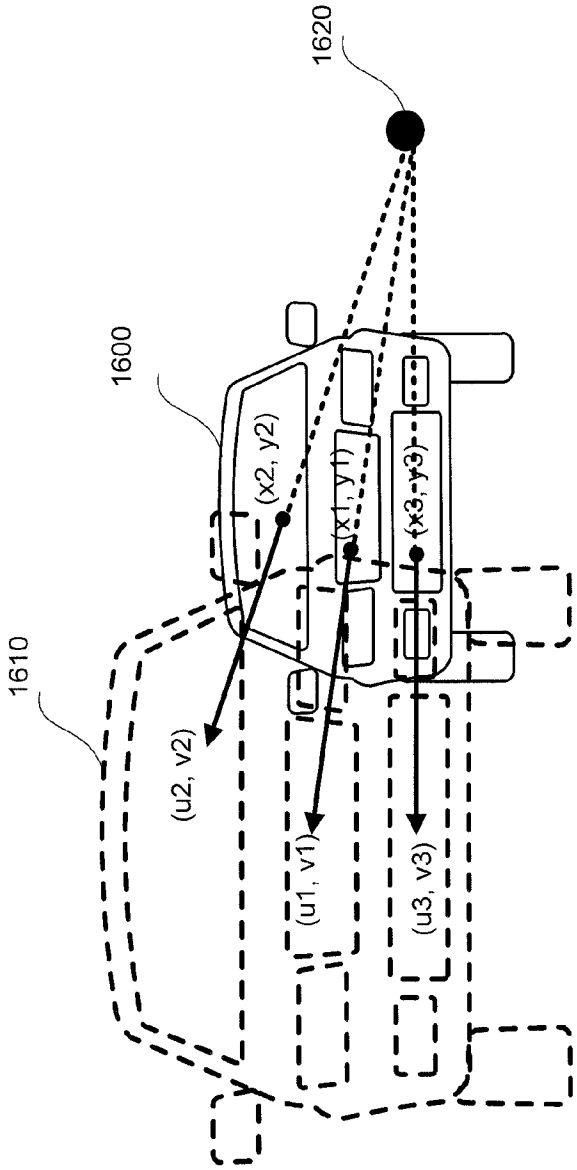


FIG. 16

IPR2025-00944
Tesla EX1002 Page 1534

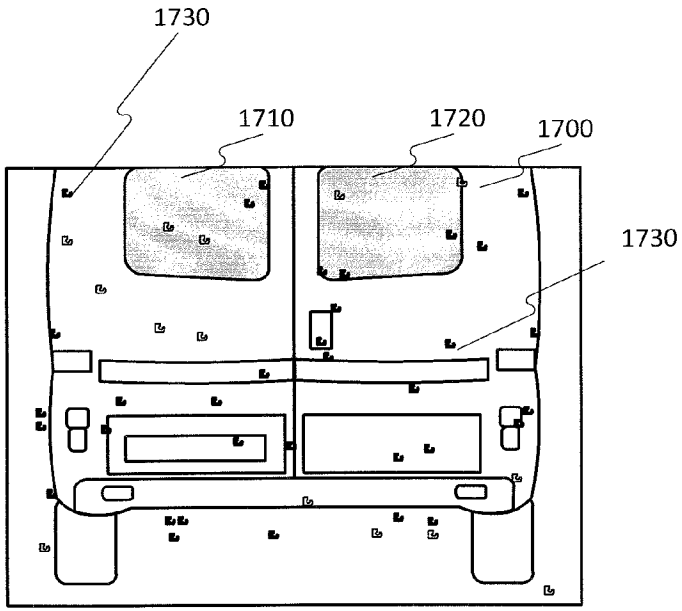


FIG. 17A

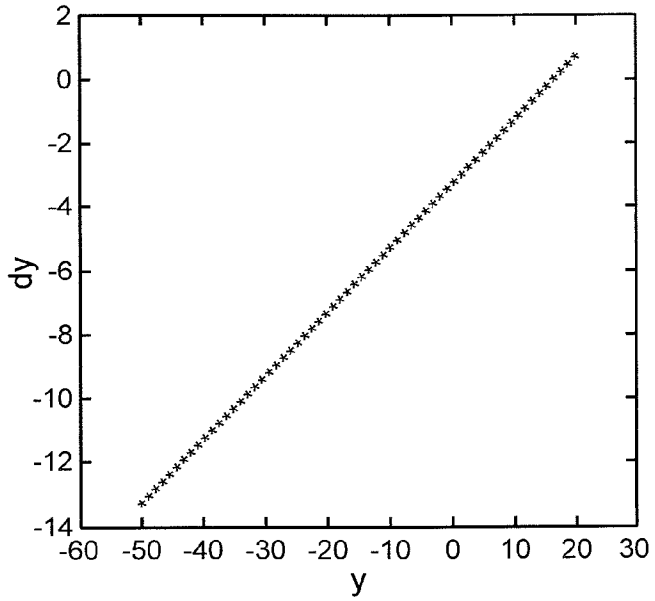


FIG. 17B

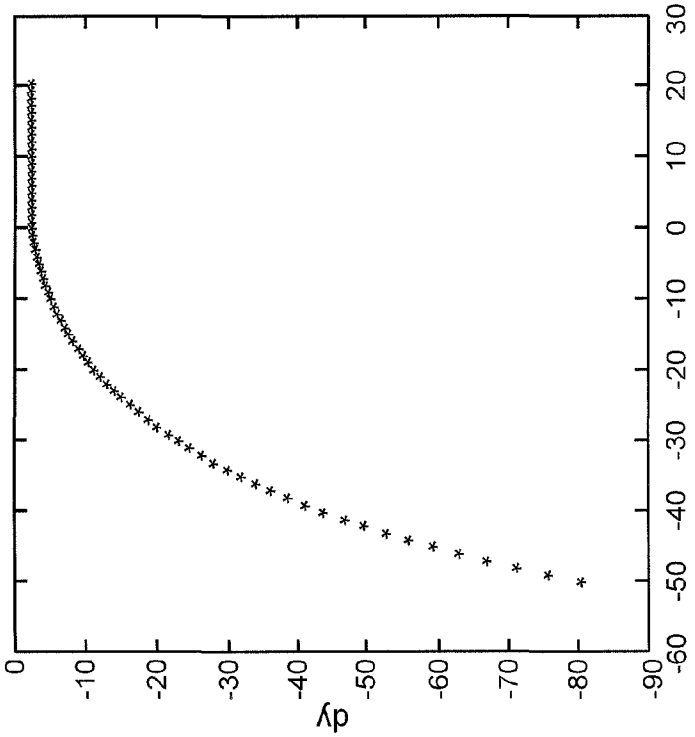


FIG. 18B

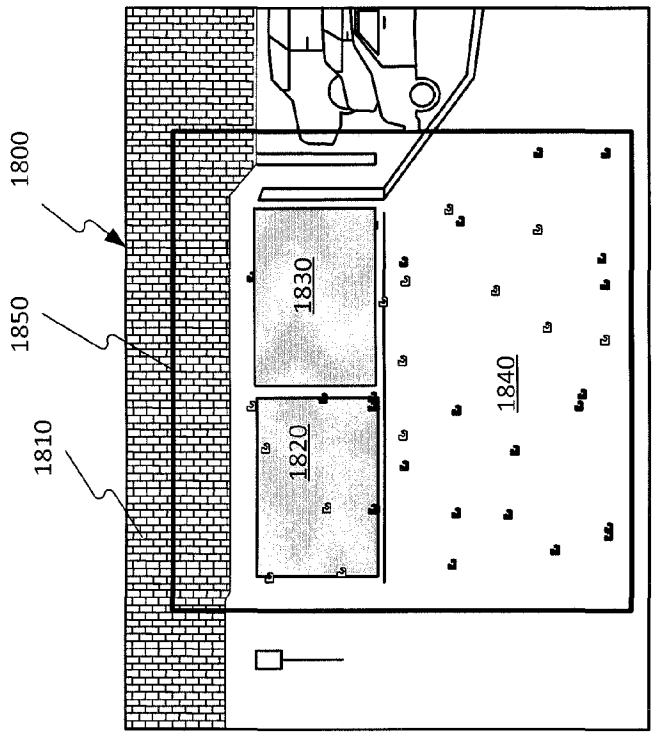


FIG. 18A

IPR2025-00944
Tesla EX1002 Page 1536

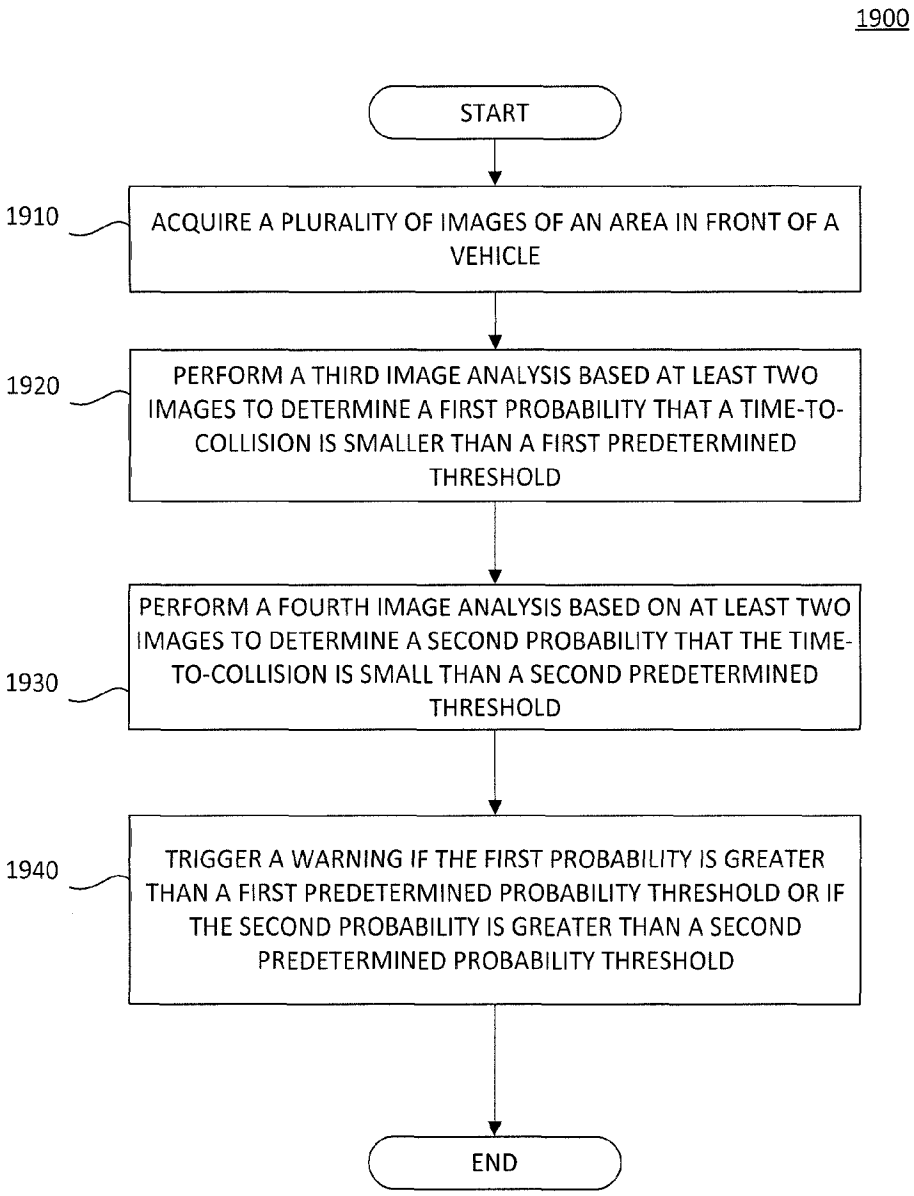


FIG. 19

IPR2025-00944
Tesla EX1002 Page 1537

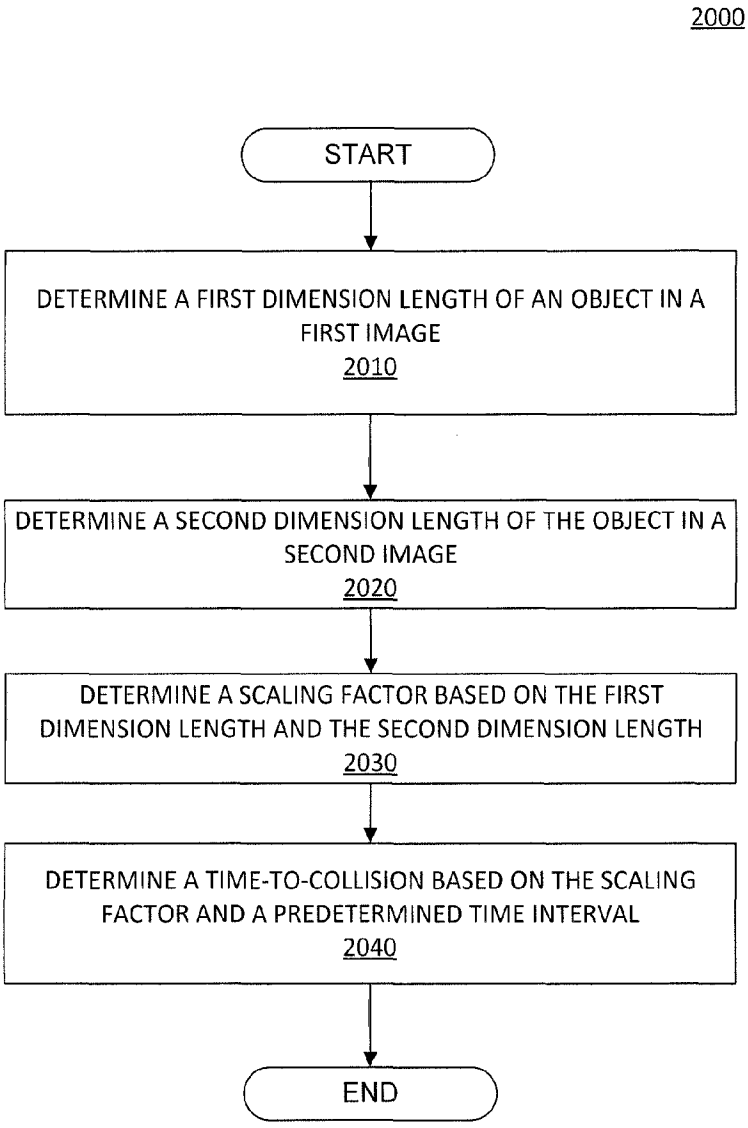


FIG. 20

IPR2025-00944
Tesla EX1002 Page 1538

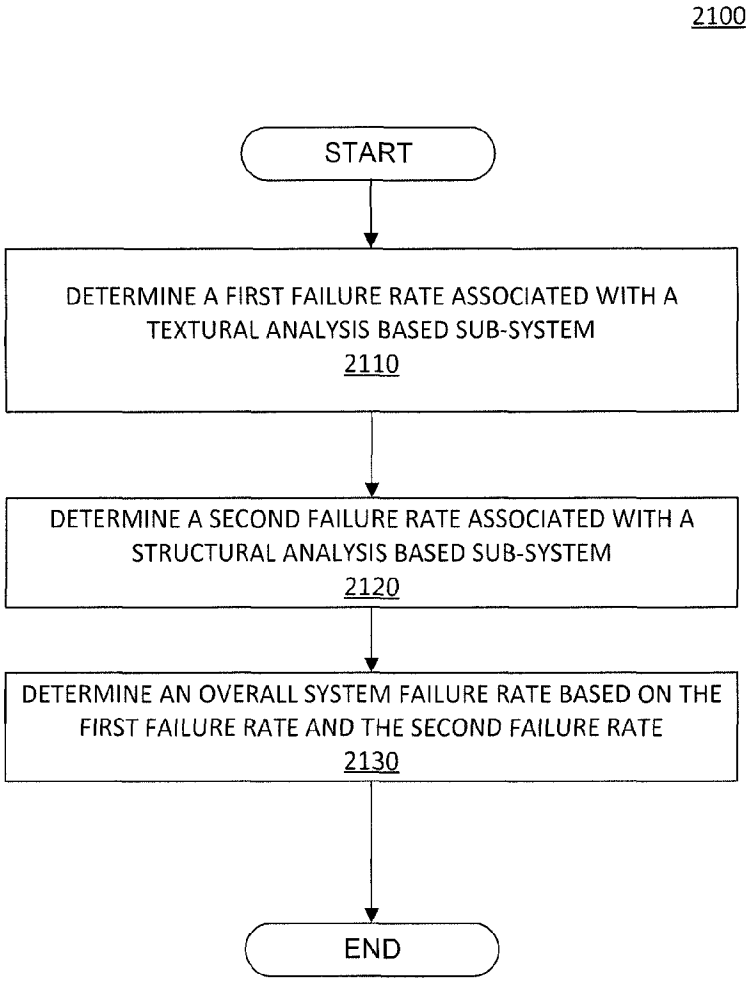


FIG. 21

IPR2025-00944
Tesla EX1002 Page 1539

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- US 20120314071 A1 **[0003]**
- US 20060178830 A1 **[0004]**
- US 5559695 A **[0005]**

Non-patent literature cited in the description

- **KYUNG-HO SHIN.** *A study on the reliability improvement for applying GNSS to railroad systems* **[0006]**

IPR2025-00944
Tesla EX1002 Page 1540

Espacenet

Bibliographic data: EP3152704 (A2) — 2017-04-12

SYSTEMS AND METHODS FOR DETECTING AN OBJECT

Inventor(s): SHASHUA AMNON [IL]; DAGAN EREZ [IL]; BABA TOMER [IL]; MYERS YONI [IL]; HOLLANDER YOSSI [IL] ± (SHASHUA, Amnon, ; DAGAN, Erez, ; BABA, Tomer, ; MYERS, Yoni, ; HOLLANDER, Yossi)

Applicant(s): MOBILEYE VISION TECHNOLOGIES LTD [IL] ± (Mobileye Vision Technologies Ltd)

Classification: **- international:** ***G06K9/00; G06T7/20***
- cooperative: **B60K37/02 (US)**; **B60R1/00 (US)**; **B60R11/04 (US)**; **G06K9/00805 (EP, US)**; **G06T7/215 (EP, US)**; B60K2370/1438 (EP, US); B60K2370/1446 (EP, US); B60K2370/148 (EP, US); B60K2370/21 (EP, US); B60R2300/105 (US); B60R2300/202 (US); B60R2300/303 (US); B60R2300/8026 (US); B60R2300/8033 (US); B60R2300/804 (US); B60R2300/8066 (US); B60R2300/8093 (US); B60Y2400/92 (US); G06T2207/10021 (EP, US); G06T2207/30261 (EP, US)

Application number: EP20150778386 20150603 Global Dossier

Priority number(s): US201462006912P 20140603 ; WO2015IB01593 20150603

Also published as: CN106796648 (A) CN106796648 (B) CN112528737 (A) US10572744 (B2) US10832063 (B2) more

Abstract not available for EP3152704 (A2)
Abstract of corresponding document: WO2015186002 (A2)

Systems and methods are provided for detecting an object in front of a vehicle. In one implementation, an object detecting system includes an image capture device configured to acquire a plurality of images of an area, a data interface, and a processing device programmed to compare a first image to a second image to determine displacement vectors between pixels, to search for a region of coherent expansion that is a set of pixels in at least one of the first image and the second image, for which there exists a common focus of expansion and a common scale magnitude

IPR2025-00944
Tesla EX1002 Page 1541

such that the set of pixels satisfy a relationship between pixel positions, displacement vectors, the common focus of expansion, and the common scale magnitude, and to identify presence of a substantially upright object based on the set of pixels.

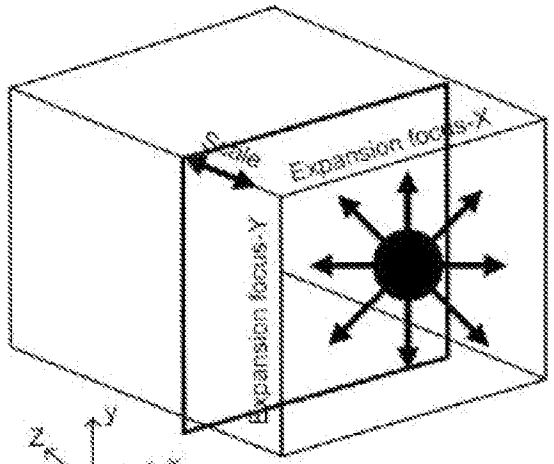


FIG. 10B

Espacenet

Description: EP3152704 (A2) — 2017-04-12

SYSTEMS AND METHODS FOR DETECTING AN OBJECT

Description not available for EP3152704 (A2)
Description of corresponding document: WO2015186002 (A2)

A high quality text as facsimile in your desired language may be available amongst the following family members:

CN106796648 (A) US2017098131 (A1) WO2015186002 (A2) CN112528737 (A) US2020234065 (A1) US2021042539 (A1)

The EPO does not accept any responsibility for the accuracy of data and information originating from other authorities than the EPO; in particular, the EPO does not guarantee that they are complete, up-to-date or fit for specific purposes.

SYSTEMS AND METHODS FOR DETECTING AN OBJECT

Cross Reference to Related Application

[001] This application claims the benefit of priority of United States Provisional Patent Application No. 62/006,912, filed on June 3, 2014. The foregoing application is incorporated herein by reference in its entirety.

BACKGROUND

I. Technical Field

[002] The present disclosure relates generally to autonomous vehicle navigation and, more specifically, to systems and methods that use cameras to detect an object in front of the vehicle.

II. Background Information

[003] As technology continues to advance, the goal of a fully autonomous vehicle that is capable of navigating on roadways is on the horizon. Primarily, an autonomous vehicle may be able to identify its environment and navigate without input from a human operator. Autonomous vehicles may also take into account a variety of factors and make appropriate decisions based on those factors to safely and accurately reach an intended destination. For example, various objects— such as other vehicles and

pedestrians---- are encountered when a vehicle typically travels a roadway. Autonomous driving systems may recognize these objects in a vehicle's environment and take appropriate and timely action to avoid collisions. Additionally, autonomous driving systems may identify other indicators---- such as traffic signals, traffic signs, and lane markings---- that regulate vehicle movement (e.g., when the vehicle must stop and may go, a speed at which the vehicle must not exceed, where the vehicle must be positioned on the roadway, etc.). Autonomous driving systems may need to determine when a vehicle should change lanes, turn at intersections, change roadways, etc. As is evident from these examples, many factors may need to be addressed in order to provide an autonomous vehicle that is capable of navigating safely and accurately.

SUMMARY

[004] Embodiments consistent with the present disclosure provide systems and methods for autonomous vehicle navigation. The disclosed embodiments may use cameras to provide autonomous vehicle navigation features. For example, consistent with the disclosed embodiments, the disclosed systems may include one, two, or more cameras that monitor the environment of a vehicle and cause a navigational response based on an analysis of images captured by one or more of the cameras.

[005] Consistent with a disclosed embodiment, an object detecting system for a vehicle is provided. The system may include at least one image capture device configured to acquire a plurality of images of an area in front of the vehicle. The system may include a data interface, and at least one processing device programmed to receive the plurality of images via the data interface. The at least one processing device may also be programmed to compare a first image to a second image to determine displacement vectors between pixels of the first image and pixels of the second image. The at least one processing device may further be programmed to search for a region of coherent expansion that is a set of pixels in at least one of the first image and the second image, for which there exists a common focus of expansion and a common scale magnitude such that the set of pixels satisfy a relationship between pixel positions, displacement vectors, the common focus of expansion, and the common scale magnitude. The at least one processing device may further be programmed to identify presence of a substantially upright object in at least one of the first image and the second image based on the set of pixels.

[006] Consistent with another disclosed embodiment, a vehicle is provided. The vehicle may include a body. The vehicle may include at least one image capture device configured to acquire a plurality of images of an area in front of the vehicle. The system may include a data interface and at least one processing device programmed to receive the plurality of images via the data interface. The at least one processing device may also be programmed to compare a first image to a second image to determine displacement vectors between pixels of the first image and pixels of the second image. The at least one processing device may further be programmed to search for a region of coherent expansion that is a set of pixels in at least one of the first image and the second image, for which there exists a common focus of expansion and a common scale magnitude such that the set of pixels satisfy a relationship between pixel positions, displacement vectors, the common focus of expansion, and the common scale magnitude. The at least one processing device may further be programmed to identify presence of a substantially upright object in at least one of the first image and the second image based on the set of pixels.

Tesla EX1002 Page 1544 IPR2025-00944

[007] Consistent with yet another disclosed embodiment, a method for detecting an object in front of a vehicle is provided. The method may include acquiring, via at least one image capture device, a plurality of images of an area in front of the vehicle. The method may also include receiving, via a processing device, the plurality of images. The method may also include comparing, via the processing device, a first image to a second image to determine displacement vectors between pixels of the first image and pixels of the second image. The method may further include searching for, via the processing device, a region of coherent expansion that is a set of pixels in at least one of the first image and the second image, for which there exists a common focus of expansion and a common scale magnitude such that the set of pixels satisfy a relationship between pixel positions, displacement vectors, the common focus of expansion, and the common scale magnitude. The method may further include identifying, via the processing device, presence of a substantially upright object in at least one of the first image and the second image based on the set of pixels.

[008] Consistent with another disclosed embodiment, an object detecting system for a vehicle is provided. The object detecting system may include at least one image capture device configured to acquire a plurality of images of an area in front of the vehicle. The system may also include a data interface and at least one processing device programmed to receive the plurality of images via the data interface. The at least one processing device may also be programmed to compare a first image to a second image, and based on the comparison, determine a focus of expansion of a crossing object moving in a direction that crosses a travel direction of the vehicle. The at least one processing device may also be programmed to determine whether a horizontal coordinate of the focus of expansion stays around a center in the horizontal dimension of the crossing object during a period of time. The at least one processing device may further be programmed to cause the vehicle to brake based on the determination that the horizontal coordinate of the focus of expansion stays around the center in the horizontal dimension of the crossing object during the period of time.

[009] Consistent with another disclosed embodiment, a vehicle is provided. The vehicle may include a body, and at least one image capture device configured to acquire a plurality of images of an area in front of the vehicle. The system may also include a data interface and at least one processing device programmed to receive the plurality of images via the data interface. The at least one processing device may also be programmed to compare a first image to a second image, and based on the comparison, determine a focus of expansion of a crossing object moving in a direction that crosses a travel direction of the vehicle. The at least one processing device may also be programmed to determine whether a horizontal coordinate of the focus of expansion stays around a center in the horizontal dimension of the crossing object during a period of time. The at least one processing device may further be programmed to cause the vehicle to brake based on the determination that the horizontal coordinate of the focus of expansion stays around the center in the horizontal dimension of the crossing object during the period of time.

[010] Consistent with another disclosed embodiment, a method for detecting an object in front of a vehicle is provided. The method may include acquiring, via at least one image capture device, a plurality of images of an area in front of the vehicle. The method may include receiving, via a processing device, the plurality of images. The method may also include comparing, via the processing device, a first image to a second image, and based on the comparison, determining a focus of expansion of a crossing object moving in a direction that crosses a travel direction of the vehicle. The

method may also include determining, via the processing device, whether a horizontal coordinate of the focus of expansion stays around a center in the horizontal dimension of the crossing object during a period of time. The method may also include causing, via the processing device, the vehicle to brake based on the determination that the horizontal coordinate of the focus of expansion stays around a center in the horizontal dimension of the crossing object during the period of time.

[01 1] Consistent with other disclosed embodiments, non-transitory computer-readable storage media may store program instructions, which are executed by at least one processing device and perform any of the methods described herein.

[012] The foregoing general description and the following detailed description are exemplary and explanatory only and are not restrictive of the claims.

BRIEF DESCRIPTION OF THE DRAWINGS

[013] The accompanying drawings, which are incorporated in and constitute a part of this disclosure, illustrate various disclosed embodiments. In the drawings:

[014] FIG. 1 is a diagrammatic representation of an exemplary system consistent with the disclosed embodiments.

[015] FIG. 2A is a diagrammatic side view representation of an exemplary vehicle including a system consistent with the disclosed embodiments. [016] FIG. 2B is a diagrammatic top view representation of the vehicle and system shown in FIG. 2A consistent with the disclosed embodiments.

[017] FIG. 2C is a diagrammatic top view representation of another embodiment of a vehicle including a system consistent with the disclosed embodiments.

[018] FIG. 2D is a diagrammatic top view representation of yet another embodiment of a vehicle including a system consistent with the disclosed embodiments.

[019] FIG. 2E is a diagrammatic top view representation of yet another embodiment of a vehicle including a system consistent with the disclosed embodiments.

[020] FIG. 2F is a diagrammatic representation of exemplary vehicle control systems consistent with the disclosed embodiments.

[021] FIG. 3 A is a diagrammatic representation of an interior of a vehicle including a rearview mirror and a user interface for a vehicle imaging system consistent with the disclosed embodiments.

[022] FIG. 3B is an illustration of an example of a camera mount that is configured to be positioned behind a rearview mirror and against a vehicle windshield consistent with the disclosed embodiments.

[023] FIG. 3C is an illustration of the camera mount shown in FIG. 3B from a different perspective consistent with the disclosed embodiments.

[024] FIG. 3D is an illustration of an example of a camera mount that is configured to be positioned behind a rearview mirror and against a vehicle windshield consistent with

IPR2025-00944
Tesla EX1002 Page 1546

the disclosed embodiments.

[025] FIG. 4 is an exemplary block diagram of a memory configured to store instructions for performing one or more operations consistent with the disclosed embodiments.

[026] FIG. 5A is a flowchart showing an exemplary process for causing one or more navigational responses based on monocular image analysis consistent with disclosed embodiments.

[027] FIG. 5B is a flowchart showing an exemplary process for detecting one or more vehicles and/or pedestrians in a set of images consistent with the disclosed embodiments.

[028] FIG. 5C is a flowchart showing an exemplary process for detecting road marks and/or lane geometry information in a set of images consistent with the disclosed embodiments.

[029] FIG. 5D is a flowchart showing an exemplary process for detecting traffic lights in a set of images consistent with the disclosed embodiments.

[030] FIG. 5E is a flowchart showing an exemplary process for causing one or more navigational responses based on a vehicle path consistent with the disclosed embodiments.

[031] FIG. 5F is a flowchart showing an exemplary process for determining whether a leading vehicle is changing lanes consistent with the disclosed embodiments.

[032] FIG. 6 is a flowchart showing an exemplary process for causing one or more navigational responses based on stereo image analysis consistent with the disclosed embodiments.

[033] FIG. 7 is a flowchart showing an exemplary process for causing one or more navigational responses based on an analysis of three sets of images consistent with the disclosed embodiments. [034] FIG. 8 is a diagrammatic representation of an exemplary vehicle including an object detecting system traveling on a road behind another vehicle consistent with the disclosed embodiments.

[035] FIG. 9 is an exemplary block diagram of a memory that may store instructions for performing one or more operations for detecting an object in front of the vehicle consistent with the disclosed embodiments.

[036] FIG. 10A is an exemplary image of an area showing vectors associated with pixels consistent with the disclosed embodiments.

[037] FIG. 10B is a schematic illustration of an exemplary relationship between the focus of expansion and the scale magnitude consistent with the disclosed embodiments.

[038] FIG. IOC shows an exemplary method of identifying a substantially upright object based on selected pixels consistent with the disclosed embodiments.

[039] FIG. 1 1 is a flowchart showing an exemplary process for detecting presence of an object in front of the vehicle consistent with the disclosed embodiments.

[040] FIG. 12A is an image of a static environment consistent with the disclosed embodiments.

[041] FIG. 12B is a three-dimensional cube space illustrating an exemplary relationship between the focus of expansion and the scale magnitude corresponding to the static environment of FIG. 12A consistent with the disclosed embodiments.

[042] FIG. 13A is an image of a environment having another vehicle static in the environment or traveling in a direction parallel to the traveling direction of the vehicle where object detecting system is installed consistent with the disclosed embodiments.

[043] FIG. 13B is a three-dimensional cube space illustrating an exemplary relationship between the focus of expansion and the scale magnitude corresponding to the image of FIG. 13A consistent with the disclosed embodiments.

[044] FIG. 13C is an image of a environment having another vehicle traveling longitudinally in front of and in alignment with the vehicle where object detecting system is installed consistent with the disclosed embodiments.

[045] FIG. 13D is a three-dimensional cube space illustrating an exemplary relationship between the focus of expansion and the scale magnitude corresponding to the image of FIG. 13C consistent with the disclosed embodiments.

[046] FIG. 14A is an image of an environment superimposed with an exemplary optical flow field consistent with the disclosed embodiments.

[047] FIG. 14B is a three-dimensional cube space illustrating two exemplary focuses of expansion corresponding to the vehicle object and the pedestrian object shown in FIG. 14A consistent with the disclosed embodiments.

[048] FIG. 15A shows a diagram illustrating an exemplary method of detecting a crossing-to- impact object consistent with the disclosed embodiment.

[049] FIG. 15B shows an image having a pedestrian crossing the road in front of the vehicle consistent with the disclosed embodiment. [050] FIG. 15C shows an image having a pedestrian crossing the road in front of the vehicle acquired at a time after the image shown in FIG. 15B is acquired, consistent with the disclosed embodiments.

[051] FIG. 16 is a flowchart showing an exemplary process for detecting an object in front of the vehicle consistent with the disclosed embodiments.

[052] FIG. 17 illustrates an exemplary method of detecting an object consistent with the disclosed embodiments.

DETAILED DESCRIPTION

[053] The following detailed description refers to the accompanying drawings. Wherever possible, the same reference numbers are used in the drawings and the following description to refer to the same or similar parts. While several illustrative

IPR2025-00944 Tesla EX1002 Page 1548

embodiments are described herein, modifications, adaptations and other implementations are possible. For example, substitutions, additions or modifications may be made to the components illustrated in the drawings, and the illustrative methods described herein may be modified by substituting, reordering, removing, or adding steps to the disclosed methods. Accordingly, the following detailed description is not limited to the disclosed embodiments and examples. Instead, the proper scope is defined by the appended claims.

[054] FIG. 1 is a block diagram representation of a system 100 consistent with the exemplary disclosed embodiments. System 100 may include various components depending on the requirements of a particular implementation. In some embodiments, system 100 may include a processing unit 1 10, an image acquisition unit 120, a position sensor 130, one or more memory units 140, 150, a map database 160, and a user interface 170. Processing unit 1 10 may include one or more processing devices. In some embodiments, processing unit 1 10 may include an applications processor 180, an image processor 190, or any other suitable processing device. Similarly, image acquisition unit 120 may include any number of image acquisition devices and components depending on the requirements of a particular application. In some embodiments, image acquisition unit 120 may include one or more image capture devices (e.g., cameras), such as image capture device 122, image capture device 124, and image capture device 126. System 100 may also include a data interface 128 communicatively connecting processing device 1 10 to image acquisition device 120. For example, data interface 128 may include any wired and/or wireless link or links for transmitting image data acquired by image accusation device 120 to processing unit 1 10.

[055] Both applications processor 180 and image processor 190 may include various types of processing devices. For example, either or both of applications processor 180 and image processor 190 may include a microprocessor, preprocessors (such as an image preprocessor), graphics processors, a central processing unit (CPU), support circuits, digital signal processors, integrated circuits, memory, or any other types of devices suitable for running applications and for image processing and analysis. In some embodiments, applications processor 180 and/or image processor 190 may include any type of single or multi-core processor, mobile device microcontroller, central processing unit, etc. Various processing devices may be used, including, for example, processors available from manufacturers such as Intel®, AMD®, etc. and may include various architectures (e.g., x86 processor, ARM®, etc.). [056] In some embodiments, applications processor 180 and/or image processor 190 may include any of the EyeQ series of processor chips available from Mobileye®. These processor designs each include multiple processing units with local memory and instruction sets. Such processors may include video inputs for receiving image data from multiple image sensors and may also include video out capabilities. In one example, the EyeQ2® uses 90nm-micron technology operating at 332Mhz. The EyeQ2® architecture consists of two floating point, hyper-thread 32-bit RISC CPUs (MIPS32® 34K® cores), five Vision Computing Engines (VCE), three Vector Microcode Processors (VMP®), Denali 64- bit Mobile DDR Controller, 128-bit internal Sonics Interconnect, dual 16-bit Video input and 18-bit Video output controllers, 16 channels DMA and several peripherals. The MIPS34K CPU manages the five VCEs, three VMP™ and the DMA, the second MIPS34K CPU and the multi-channel DMA as well as the other peripherals. The five VCEs, three VMP® and the MIPS34 CPU can perform intensive vision computations required by multi-function bundle applications. In another example, the EyeQ3®, which is a third generation processor and is six times more powerful that the EyeQ2®, may be

used in the disclosed embodiments.

[057] Any of the processing devices disclosed herein may be configured to perform certain functions. Configuring a processing device, such as any of the described EyeQ processors or other controller or microprocessor, to perform certain functions may include programming of computer executable instructions and making those instructions available to the processing device for execution during operation of the processing device. In some embodiments, configuring a processing device may include programming the processing device directly with architectural instructions. In other embodiments, configuring a processing device may include storing executable instructions on a memory that is accessible to the processing device during operation. For example, the processing device may access the memory to obtain and execute the stored instructions during operation.

[058] While FIG. 1 depicts two separate processing devices included in processing unit 1 10, more or fewer processing devices may be used. For example, in some embodiments, a single processing device may be used to accomplish the tasks of applications processor 180 and image processor 190. In other embodiments, these tasks may be performed by more than two processing devices.

[059] Processing unit 1 10 may comprise various types of devices. For example, processing unit 1 10 may include various devices, such as a controller, an image preprocessor, a central processing unit (CPU), support circuits, digital signal processors, integrated circuits, memory, or any other types of devices for image processing and analysis. The image preprocessor may include a video processor for capturing, digitizing and processing the imagery from the image sensors. The CPU may comprise any number of microcontrollers or microprocessors. The support circuits may be any number of circuits generally well known in the art, including cache, power supply, clock and input-output circuits. The memory may store software that, when executed by the processor, controls the operation of the system. The memory may include databases and image processing software. The memory may comprise any number of random access memories, read only memories, flash memories, disk drives, optical storage, tape storage, removable storage and other types of storage. In one instance, the memory may be separate from the processing unit 1 10. In another instance, the memory may be integrated into the processing unit 1 10.

[060] Each memory 140, 150 may include software instructions that when executed by a processor (e.g., applications processor 180 and/or image processor 190), may control operation of various aspects of system 100. These memory units may include various databases and image processing software. The memory units may include random access memory, read only memory, flash memory, disk drives, optical storage, tape storage, removable storage and/or any other types of storage. In some embodiments, memory units 140, 150 may be separate from the applications processor 180 and/or image processor 190. In other embodiments, these memory units may be integrated into applications processor 180 and/or image processor 190.

[061] Position sensor 130 may include any type of device suitable for determining a location associated with at least one component of system 100. In some embodiments, position sensor 130 may include a GPS receiver. Such receivers can determine a user position and velocity by processing signals broadcasted by global positioning system satellites. Position information from position sensor 130 may be made available to applications processor 180 and/or image processor 190.

IPR2025-00944
Tesla EX1002 Page 1550

[062] User interface 170 may include any device suitable for providing information to or for receiving inputs from one or more users of system 100. In some embodiments, user interface 170 may include user input devices, including, for example, a touchscreen, microphone, keyboard, pointer devices, track wheels, cameras, knobs, buttons, etc. With such input devices, a user may be able to provide information inputs or commands to system 100 by typing instructions or information, providing voice commands, selecting menu options on a screen using buttons, pointers, or eye-tracking capabilities, or through any other suitable techniques for communicating information to system 100.

[063] User interface 170 may be equipped with one or more processing devices configured to provide and receive information to or from a user and process that information for use by, for example, applications processor 180. In some embodiments, such processing devices may execute instructions for recognizing and tracking eye movements, receiving and interpreting voice commands, recognizing and interpreting touches and/or gestures made on a touchscreen, responding to keyboard entries or menu selections, etc. In some embodiments, user interface 170 may include a display, speaker, tactile device, and/or any other devices for providing output information to a user.

[064] Map database 160 may include any type of database for storing map data useful to system

100. In some embodiments, map database 160 may include data relating to the position, in a reference coordinate system, of various items, including roads, water features, geographic features, businesses, points of interest, restaurants, gas stations, etc. Map database 160 may store not only the locations of such items, but also descriptors relating to those items, including, for example, names associated with any of the stored features. In some embodiments, map database 160 may be physically located with other components of system 100. Alternatively or additionally, map database 160 or a portion thereof may be located remotely with respect to other components of system 100 (e.g., processing unit 1 10). In such embodiments, information from map database 160 may be downloaded over a wired or wireless data connection to a network (e.g., over a cellular network and/or the Internet, etc.).

[065] Image capture devices 122, 124, and 126 may each include any type of device suitable for capturing at least one image from an environment. Moreover, any number of image capture devices may be used to acquire images for input to the image processor. Some embodiments may include only a single image capture device, while other embodiments may include two, three, or even four or more image capture devices. Image capture devices 122, 124, and 126 will be further described with reference to FIGS. 2B-2E, below.

[066] System 100, or various components thereof, may be incorporated into various different platforms. In some embodiments, system 100 may be included on a vehicle 200, as shown in FIG. 2A.

For example, vehicle 200 may be equipped with a processing unit 1 10 and any of the other components of system 100, as described above relative to FIG. 1. While in some embodiments vehicle 200 may be equipped with only a single image capture device (e.g., camera), in other embodiments, such as those discussed in connection with

Tesla EX1002 Page 1551 IPR2025-00944

FIGS. 2B-2E, multiple image capture devices may be used. For example, either of image capture devices 122 and 124 of vehicle 200, as shown in FIG. 2A, may be part of an ADAS (Advanced Driver Assistance Systems) imaging set.

[067] The image capture devices included on vehicle 200 as part of the image acquisition unit 120 may be positioned at any suitable location. In some embodiments, as shown in FIGS. 2A-2E and 3A- 3C, image capture device 122 may be located in the vicinity of the rearview mirror. This position may provide a line of sight similar to that of the driver of vehicle 200, which may aid in determining what is and is not visible to the driver. Image capture device 122 may be positioned at any location near the rearview mirror, but placing image capture device 122 on the driver side of the mirror may further aid in obtaining images representative of the driver's field of view and/or line of sight.

[068] Other locations for the image capture devices of image acquisition unit 120 may also be used. For example, image capture device 124 may be located on or in a bumper of vehicle 200. Such a location may be especially suitable for image capture devices having a wide field of view. The line of sight of bumper-located image capture devices can be different from that of the driver and, therefore, the bumper image capture device and driver may not always see the same objects. The image capture devices (e.g., image capture devices 122, 124, and 126) may also be located in other locations. For example, the image capture devices may be located on or in one or both of the side mirrors of vehicle 200, on the roof of vehicle 200, on the hood of vehicle 200, on the trunk of vehicle 200, on the sides of vehicle 200, mounted on, positioned behind, or positioned in front of any of the windows of vehicle 200, and mounted in or near light figures on the front and/or back of vehicle 200, etc.

[069] In addition to image capture devices, vehicle 200 may include various other components of system 100. For example, processing unit 1 10 may be included on vehicle 200 either integrated with or separate from an engine control unit (ECU) of the vehicle. Vehicle 200 may also be equipped with a position sensor 130, such as a GPS receiver and may also include a map database 160 and memory units 140 and 150.
[070] FIG. 2A is a diagrammatic side view representation of an exemplary vehicle imaging system consistent with the disclosed embodiments. FIG. 2B is a diagrammatic top view illustration of the embodiment shown in FIG. 2A. As illustrated in FIG. 2B, the disclosed embodiments may include a vehicle 200 including in its body a system 100 with a first image capture device 122 positioned in the vicinity of the rearview mirror and/or near the driver of vehicle 200, a second image capture device 124 positioned on or in a bumper region (e.g., one of bumper regions 210) of vehicle 200, and a processing unit 1 10.

[071] As illustrated in FIG. 2C, image capture devices 122 and 124 may both be positioned in the vicinity of the rearview mirror and/or near the driver of vehicle 200. Additionally, while two image capture devices 122 and 124 are shown in FIGS. 2B and 2C, it should be understood that other embodiments may include more than two image capture devices. For example, in the embodiments shown in FIGS. 2D and 2E, first, second, and third image capture devices 122, 124, and 126, are included in the system 100 of vehicle 200.

[072] As illustrated in FIG. 2D, image capture device 122 may be positioned in the vicinity of the rearview mirror and/or near the driver of vehicle 200, and image capture devices 124 and 126 may be positioned on or in a bumper region (e.g., one of bumper regions 210) of vehicle 200. And as shown in FIG. 2E, image capture devices 122, 124,

IPR2025-00944
Tesla EX1002 Page 1552

and 126 may be positioned in the vicinity of the rearview mirror and/or near the driver seat of vehicle 200. The disclosed embodiments are not limited to any particular number and configuration of the image capture devices, and the image capture devices may be positioned in any appropriate location within and/or on vehicle 200.

[073] It is to be understood that the disclosed embodiments are not limited to vehicles and could be applied in other contexts. It is also to be understood that disclosed embodiments are not limited to a particular type of vehicle 200 and may be applicable to all types of vehicles including automobiles, trucks, trailers, and other types of vehicles.

[074] The first image capture device 122 may include any suitable type of image capture device. Image capture device 122 may include an optical axis. In one instance, the image capture device 122 may include an Aptina M9V024 WVGA sensor with a global shutter. In other embodiments, image capture device 122 may provide a resolution of 1280x960 pixels and may include a rolling shutter. Image capture device 122 may include various optical elements. In some embodiments one or more lenses may be included, for example, to provide a desired focal length and field of view for the image capture device. In some embodiments, image capture device 122 may be associated with a 6mm lens or a 12mm lens. In some embodiments, image capture device 122 may be configured to capture images having a desired field-of-view (FOV) 202, as illustrated in FIG. 2D. For example, image capture device 122 may be configured to have a regular FOV, such as within a range of 40 degrees to 56 degrees, including a 46 degree FOV, 50 degree FOV, 52 degree FOV, or greater. Alternatively, image capture device 122 may be configured to have a narrow FOV in the range of 23 to 40 degrees, such as a 28 degree FOV or 36 degree FOV. In addition, image capture device 122 may be configured to have a wide FOV in the range of 100 to 180 degrees. In some embodiments, image capture device 122 may include a wide angle bumper camera or one with up to a 180 degree FOV.

[075] The first image capture device 122 may acquire a plurality of first images relative to a scene associated with the vehicle 200. Each of the plurality of first images may be acquired as a series of image scan lines, which may be captured using a rolling shutter. Each scan line may include a plurality of pixels.

[076] The first image capture device 122 may have a scan rate associated with acquisition of each of the first series of image scan lines. The scan rate may refer to a rate at which an image sensor can acquire image data associated with each pixel included in a particular scan line.

[077] Image capture devices 122, 124, and 126 may contain any suitable type and number of image sensors, including CCD sensors or CMOS sensors, for example. In one embodiment, a CMOS image sensor may be employed along with a rolling shutter, such that each pixel in a row is read one at a time, and scanning of the rows proceeds on a row-by-row basis until an entire image frame has been captured. In some embodiments, the rows may be captured sequentially from top to bottom relative to the frame.

[078] The use of a rolling shutter may result in pixels in different rows being exposed and captured at different times, which may cause skew and other image artifacts in the captured image frame. On the other hand, when the image capture device 122 is configured to operate with a global or synchronous shutter, all of the pixels may be

exposed for the same amount of time and during a common exposure period. As a result, the image data in a frame collected from a system employing a global shutter represents a snapshot of the entire FOV (such as FOV 202) at a particular time. In contrast, in a rolling shutter application, each row in a frame is exposed and data is capture at different times. Thus, moving objects may appear distorted in an image capture device having a rolling shutter. This phenomenon will be described in greater detail below.

[079] The second image capture device 124 and the third image capturing device 126 may be any type of image capture device. Like the first image capture device 122, each of image capture devices 124 and 126 may include an optical axis. In one embodiment, each of image capture devices 124 and 126 may include an Aptina M9V024 WVGA sensor with a global shutter. Alternatively, each of image capture devices 124 and 126 may include a rolling shutter. Like image capture device 122, image capture devices 124 and 126 may be configured to include various lenses and optical elements. In some embodiments, lenses associated with image capture devices 124 and 126 may provide FOVs (such as FOVs 204 and 206) that are the same as, or narrower than, a FOV (such as FOV 202) associated with image capture device 122. For example, image capture devices 124 and 126 may have FOVs of 40 degrees, 30 degrees, 26 degrees, 23 degrees, 20 degrees, or less.

[080] Image capture devices 124 and 126 may acquire a plurality of second and third images relative to a scene associated with the vehicle 200. Each of the plurality of second and third images may be acquired as a second and third series of image scan lines, which may be captured using a rolling shutter. Each scan line or row may have a plurality of pixels. Image capture devices 124 and 126 may have second and third scan rates associated with acquisition of each of image scan lines included in the second and third series.

[081] Each image capture device 122, 124, and 126 may be positioned at any suitable position and orientation relative to vehicle 200. The relative positioning of the image capture devices 122, 124, and 126 may be selected to aid in fusing together the information acquired from the image capture devices. For example, in some embodiments, a FOV (such as FOV 204) associated with image capture device 124 may overlap partially or fully with a FOV (such as FOV 202) associated with image capture device 122 and a FOV (such as FOV 206) associated with image capture device 126.

[082] Image capture devices 122, 124, and 126 may be located on vehicle 200 at any suitable relative heights. In one instance, there may be a height difference between the image capture devices 122, 124, and 126, which may provide sufficient parallax information to enable stereo analysis. For example, as shown in FIG. 2A, the two image capture devices 122 and 124 are at different heights. There may also be a lateral displacement difference between image capture devices 122, 124, and 126, giving additional parallax information for stereo analysis by processing unit 1 10, for example. The difference in the lateral displacement may be denoted by dx, as shown in FIGS. 2C and 2D. In some embodiments, fore or aft displacement (e.g., range displacement) may exist between image capture devices 122, 124, and 126. For example, image capture device 122 may be located 0.5 to 2 meters or more behind image capture device 124 and/or image capture device 126. This type of displacement may enable one of the image capture devices to cover potential blind spots of the other image capture device(s).

IPR2025-00944
Tesla EX1002 Page 1554

[083] Image capture devices 122 may have any suitable resolution capability (e.g., number of pixels associated with the image sensor), and the resolution of the image sensor(s) associated with the image capture device 122 may be higher, lower, or the same as the resolution of the image sensor(s) associated with image capture devices 124 and 126. In some embodiments, the image sensor(s) associated with image capture device 122 and/or image capture devices 124 and 126 may have a resolution of 640 x 480, 1024 x 768, 1280 x 960, or any other suitable resolution.

[084] The frame rate (e.g., the rate at which an image capture device acquires a set of pixel data of one image frame before moving on to capture pixel data associated with the next image frame) may be controllable. The frame rate associated with image capture device 122 may be higher, lower, or the same as the frame rate associated with image capture devices 124 and 126. The frame rate associated with image capture devices 122, 124, and 126 may depend on a variety of factors that may affect the timing of the frame rate. For example, one or more of image capture devices 122, 124, and 126 may include a selectable pixel delay period imposed before or after acquisition of image data associated with one or more pixels of an image sensor in image capture device 122, 124, and/or 126. Generally, image data corresponding to each pixel may be acquired according to a clock rate for the device (e.g., one pixel per clock cycle). Additionally, in embodiments including a rolling shutter, one or more of image capture devices 122, 124, and 126 may include a selectable horizontal blanking period imposed before or after acquisition of image data associated with a row of pixels of an image sensor in image capture device 122, 124, and/or 126. Further, one or more of image capture devices 122, 124, and/or 126 may include a selectable vertical blanking period imposed before or after acquisition of image data associated with an image frame of image capture device 122, 124, and 126.

[085] These timing controls may enable synchronization of frame rates associated with image capture devices 122, 124, and 126, even where the line scan rates of each are different. Additionally, as will be discussed in greater detail below, these selectable timing controls, among other factors (e.g., image sensor resolution, maximum line scan rates, etc.) may enable synchronization of image capture from an area where the FOV of image capture device 122 overlaps with one or more FOVs of image capture devices 124 and 126, even where the field of view of image capture device 122 is different from the FOVs of image capture devices 124 and 126.

[086] Frame rate timing in image capture device 122, 124, and 126 may depend on the resolution of the associated image sensors. For example, assuming similar line scan rates for both devices, if one device includes an image sensor having a resolution of 640 x 480 and another device includes an image sensor with a resolution of 1280 x 960, then more time will be required to acquire a frame of image data from the sensor having the higher resolution.

[087] Another factor that may affect the timing of image data acquisition in image capture devices 122, 124, and 126 is the maximum line scan rate. For example, acquisition of a row of image data from an image sensor included in image capture device 122, 124, and 126 will require some minimum amount of time. Assuming no pixel delay periods are added, this minimum amount of time for acquisition of a row of image data will be related to the maximum line scan rate for a particular device. Devices that offer higher maximum line scan rates have the potential to provide higher frame rates than devices with lower maximum line scan rates. In some embodiments,

Tesla EX1002 Page 1555 IPR2025-00944

one or more of image capture devices 124 and 126 may have a maximum line scan rate that is higher than a maximum line scan rate associated with image capture device 122. In some embodiments, the maximum line scan rate of image capture device 124 and/or 126 may be 1.25, 1.5, 1.75, or 2 times or more than a maximum line scan rate of image capture device 122.

[088] In another embodiment, image capture devices 122, 124, and 126 may have the same maximum line scan rate, but image capture device 122 may be operated at a scan rate less than or equal to its maximum scan rate. The system may be configured such that one or more of image capture devices 124 and 126 operate at a line scan rate that is equal to the line scan rate of image capture device 122. In other instances, the system may be configured such that the line scan rate of image capture device 124 and/or image capture device 126 may be 1.25, 1.5, 1.75, or 2 times or more than the line scan rate of image capture device 122.

[089] In some embodiments, image capture devices 122, 124, and 126 may be asymmetric. That is, they may include cameras having different fields of view (FOV) and focal lengths. The fields of view of image capture devices 122, 124, and 126 may include any desired area relative to an environment of vehicle 200, for example. In some embodiments, one or more of image capture devices 122, 124, and 126 may be configured to acquire image data from an environment in front of vehicle 200, behind vehicle 200, to the sides of vehicle 200, or combinations thereof. [090] Further, the focal length associated with each image capture device 122, 124, and/or 126 may be selectable (e.g., by inclusion of appropriate lenses etc.) such that each device acquires images of objects at a desired distance range relative to vehicle 200. For example, in some embodiments image capture devices 122, 124, and 126 may acquire images of close-up objects within a few meters from the vehicle. Image capture devices 122, 124, and 126 may also be configured to acquire images of objects at ranges more distant from the vehicle (e.g., 25 m, 50 m, 100 m, 150 m, or more). Further, the focal lengths of image capture devices 122, 124, and 126 may be selected such that one image capture device (e.g., image capture device 122) can acquire images of objects relatively close to the vehicle (e.g., within 10 m or within 20 m) while the other image capture devices (e.g., image capture devices 124 and 126) can acquire images of more distant objects (e.g., greater than 20 m, 50 m, 100 m, 150 m, etc.) from vehicle 200.

[091] According to some embodiments, the FOV of one or more image capture devices 122, 124, and 126 may have a wide angle. For example, it may be advantageous to have a FOV of 140 degrees, especially for image capture devices 122, 124, and 126 that may be used to capture images of the area in the vicinity of vehicle 200. For example, image capture device 122 may be used to capture images of the area to the right or left of vehicle 200 and, in such embodiments, it may be desirable for image capture device 122 to have a wide FOV (e.g., at least 140 degrees).

[092] The field of view associated with each of image capture devices 122, 124, and 126 may depend on the respective focal lengths. For example, as the focal length increases, the corresponding field of view decreases.

[093] Image capture devices 122, 124, and 126 may be configured to have any suitable fields of view. In one particular example, image capture device 122 may have a horizontal FOV of 46 degrees, image capture device 124 may have a horizontal FOV of 23 degrees, and image capture device 126 may have a horizontal FOV in between 23 and 46 degrees. In another instance, image capture device 122 may have a horizontal

Tesla EX1002 Page 1556 IPR2025-00944

FOV of 52 degrees, image capture device 124 may have a horizontal FOV of 26 degrees, and image capture device 126 may have a horizontal FOV in between 26 and 52 degrees. In some embodiments, a ratio of the FOV of image capture device 122 to the FOVs of image capture device 124 and/or image capture device 126 may vary from 1.5 to 2.0. In other embodiments, this ratio may vary between 1.25 and 2.25.

[094] System 100 may be configured so that a field of view of image capture device 122 overlaps, at least partially or fully, with a field of view of image capture device 124 and/or image capture device 126. In some embodiments, system 100 may be configured such that the fields of view of image capture devices 124 and 126, for example, fall within (e.g., are narrower than) and share a common center with the field of view of image capture device 122. In other embodiments, the image capture devices 122, 124, and 126 may capture adjacent FOVs or may have partial overlap in their FOVs. In some embodiments, the fields of view of image capture devices 122, 124, and 126 may be aligned such that a center of the narrower FOV image capture devices 124 and/or 126 may be located in a lower half of the field of view of the wider FOV device 122. [095] FIG. 2F is a diagrammatic representation of exemplary vehicle control systems, consistent with the disclosed embodiments. As indicated in FIG. 2F, vehicle 200 may include throttling system 220, braking system 230, and steering system 240. System 100 may provide inputs (e.g., control signals) to one or more of throttling system 220, braking system 230, and steering system 240 over one or more data links (e.g., any wired and/or wireless link or links for transmitting data). For example, based on analysis of images acquired by image capture devices 122, 124, and/or 126, system 100 may provide control signals to one or more of throttling system 220, braking system 230, and steering system 240 to navigate vehicle 200 (e.g., by causing an acceleration, a turn, a lane shift, etc.). Further, system 100 may receive inputs from one or more of throttling system 220, braking system 230, and steering system 24 indicating operating conditions of vehicle 200 (e.g., speed, whether vehicle 200 is braking and/or turning, etc.). Further details are provided in connection with FIGS. 4-7, below.

[096] As shown in FIG. 3 A, vehicle 200 may also include a user interface 170 for interacting with a driver or a passenger of vehicle 200. For example, user interface 170 in a vehicle application may include a touch screen 320, knobs 330, buttons 340, and a microphone 350. A driver or passenger of vehicle 200 may also use handles (e.g., located on or near the steering column of vehicle 200 including, for example, turn signal handles), buttons (e.g., located on the steering wheel of vehicle 200), and the like, to interact with system 100. In some embodiments, microphone 350 may be positioned adjacent to a rearview mirror 310. Similarly, in some embodiments, image capture device 122 may be located near rearview mirror 310. In some embodiments, user interface 170 may also include one or more speakers 360 (e.g., speakers of a vehicle audio system). For example, system 100 may provide various notifications (e.g., alerts) via speakers 360.

[097] FIGS. 3B-3D are illustrations of an exemplary camera mount 370 configured to be positioned behind a rearview mirror (e.g., rearview mirror 310) and against a vehicle windshield, consistent with disclosed embodiments. As shown in FIG. 3B, camera mount 370 may include image capture devices 122, 124, and 126. Image capture devices 124 and 126 may be positioned behind a glare shield 380, which may be flush against the vehicle windshield and include a composition of film and/or anti-reflective materials. For example, glare shield 380 may be positioned such that it aligns against a vehicle windshield having a matching slope. In some embodiments, each of image capture devices 122, 124, and 126 may be positioned behind glare shield 380, as

IPR2025-00944
Tesla EX1002 Page 1557

depicted, for example, in FIG. 3D. The disclosed embodiments are not limited to any particular configuration of image capture devices 122, 124, and 126, camera mount 370, and glare shield 380. FIG. 3C is an illustration of camera mount 370 shown in FIG. 3B from a front perspective.

[098] As will be appreciated by a person skilled in the art having the benefit of this disclosure, numerous variations and/or modifications may be made to the foregoing disclosed embodiments. For example, not all components are essential for the operation of system 100. Further, any component may be located in any appropriate part of system 100 and the components may be rearranged into a variety of configurations while providing the functionality of the disclosed embodiments. Therefore, the foregoing configurations are examples and, regardless of the configurations discussed above, system 100 can provide a wide range of functionality to analyze the surroundings of vehicle 200 and navigate vehicle 200 in response to the analysis.

[099] As discussed below in further detail and consistent with various disclosed embodiments, system 100 may provide a variety of features related to autonomous driving and/or driver assist technology. For example, system 100 may analyze image data, position data (e.g., GPS location information), map data, speed data, and/or data from sensors included in vehicle 200. System 100 may collect the data for analysis from, for example, image acquisition unit 120, position sensor 130, and other sensors. Further, system 100 may analyze the collected data to determine whether or not vehicle 200 should take a certain action, and then automatically take the determined action without human intervention. For example, when vehicle 200 navigates without human intervention, system 100 may automatically control the braking, acceleration, and/or steering of vehicle 200 (e.g., by sending control signals to one or more of throttling system 220, braking system 230, and steering system 240). Further, system 100 may analyze the collected data and issue warnings and/or alerts to vehicle occupants based on the analysis of the collected data. Additional details regarding the various embodiments that are provided by system 100 are provided below.

[0100] Forward-Facing Multi-Imaging System

[0101] As discussed above, system 100 may provide drive assist functionality that uses a multi- camera system. The multi-camera system may use one or more cameras facing in the forward direction of a vehicle. In other embodiments, the multi-camera system may include one or more cameras facing to the side of a vehicle or to the rear of the vehicle. In one embodiment, for example, system 100 may use a two-camera imaging system, where a first camera and a second camera (e.g., image capture devices 122 and 124) may be positioned at the front and/or the sides of a vehicle (e.g., vehicle 200). The first camera may have a field of view that is greater than, less than, or partially overlapping with, the field of view of the second camera. In addition, the first camera may be connected to a first image processor to perform monocular image analysis of images provided by the first camera, and the second camera may be connected to a second image processor to perform monocular image analysis of images provided by the second camera. The outputs (e.g., processed information) of the first and second image processors may be combined. In some embodiments, the second image processor may receive images from both the first camera and second camera to perform stereo analysis. In another embodiment, system 100 may use a three-camera imaging system where each of the cameras has a different field of view. Such a system may, therefore, make decisions based on information derived from

IPR2025-00944
Tesla EX1002 Page 1558

objects located at varying distances both forward and to the sides of the vehicle. References to monocular image analysis may refer to instances where image analysis is performed based on images captured from a single point of view (e.g., from a single camera). Stereo image analysis may refer to instances where image analysis is performed based on two or more images captured with one or more variations of an image capture parameter. For example, captured images suitable for performing stereo image analysis may include images captured: from two or more different positions, from different fields of view, using different focal lengths, along with parallax information, etc.
[0102] For example, in one embodiment, system 100 may implement a three camera configuration using image capture devices 122- 126. In such a configuration, image capture device 122 may provide a narrow field of view (e.g., 34 degrees, or other values selected from a range of about 20 to 45 degrees, etc.), image capture device 124 may provide a wide field of view (e.g., 150 degrees or other values selected from a range of about 100 to about 180 degrees), and image capture device 126 may provide an intermediate field of view (e.g., 46 degrees or other values selected from a range of about 35 to about 60 degrees). In some embodiments, image capture device 126 may act as a main or primary camera. Image capture devices 122-126 may be positioned behind rearview mirror 310 and positioned substantially side-by-side (e.g., 6 cm apart). Further, in some embodiments, as discussed above, one or more of image capture devices 122- 126 may be mounted behind glare shield 380 that is flush with the windshield of vehicle 200. Such shielding may act to minimize the impact of any reflections from inside the car on image capture devices 122- 126.

[0103] In another embodiment, as discussed above in connection with FIGS. 3B and 3C, the wide field of view camera (e.g., image capture device 124 in the above example) may be mounted lower than the narrow and main field of view cameras (e.g., image devices 122 and 126 in the above example). This configuration may provide a free line of sight from the wide field of view camera. To reduce reflections, the cameras may be mounted close to the windshield of vehicle 200, and may include polarizers on the cameras to damp reflected light.

[0104] A three camera system may provide certain performance characteristics. For example, some embodiments may include an ability to validate the detection of objects by one camera based on detection results from another camera. In the three camera configuration discussed above, processing unit 1 10 may include, for example, three processing devices (e.g., three EyeQ series of processor chips, as discussed above), with each processing device dedicated to processing images captured by one or more of image capture devices 122-126.

[0105] In a three camera system, a first processing device may receive images from both the main camera and the narrow field of view camera, and perform vision processing of the narrow FOV camera to, for example, detect other vehicles, pedestrians, lane marks, traffic signs, traffic lights, and other road objects. Further, the first processing device may calculate a disparity of pixels between the images from the main camera and the narrow camera and create a 3D reconstruction of the environment of vehicle 200. The first processing device may then combine the 3D reconstruction with 3D map data or with 3D information calculated based on information from another camera.

[0106] The second processing device may receive images from main camera and perform vision processing to detect other vehicles, pedestrians, lane marks, traffic signs, traffic lights, and other road objects. Additionally, the second processing device

IPR2025-00944
Tesla EX1002 Page 1559

may calculate a camera displacement and, based on the displacement, calculate a disparity of pixels between successive images and create a 3D

reconstruction of the scene (e.g., a structure from motion). The second processing device may send the structure from motion based 3D reconstruction to the first processing device to be combined with the stereo 3D images. [0107] The third processing device may receive images from the wide FOV camera and process the images to detect vehicles, pedestrians, lane marks, traffic signs, traffic lights, and other road objects. The third processing device may further execute additional processing instructions to analyze images to identify objects moving in the image, such as vehicles changing lanes, pedestrians, etc.

[0108] In some embodiments, having streams of image-based information captured and processed independently may provide an opportunity for providing redundancy in the system. Such redundancy may include, for example, using a first image capture device and the images processed from that device to validate and/or supplement information obtained by capturing and processing image information from at least a second image capture device.

[0109] In some embodiments, system 100 may use two image capture devices (e.g., image capture devices 122 and 124) in providing navigation assistance for vehicle 200 and use a third image capture device (e.g., image capture device 126) to provide redundancy and validate the analysis of data received from the other two image capture devices. For example, in such a configuration, image capture devices 122 and 124 may provide images for stereo analysis by system 100 for navigating vehicle 200, while image capture device 126 may provide images for monocular analysis by system 100 to provide redundancy and validation of information obtained based on images captured from image capture device 122 and/or image capture device 124. That is, image capture device 126 (and a corresponding processing device) may be considered to provide a redundant sub-system for providing a check on the analysis derived from image capture devices 122 and 124 (e.g., to provide an automatic emergency braking (AEB) system).

[01 10] One of skill in the art will recognize that the above camera configurations, camera placements, number of cameras, camera locations, etc., are examples only. These components and others described relative to the overall system may be assembled and used in a variety of different

configurations without departing from the scope of the disclosed embodiments. Further details regarding usage of a multi-camera system to provide driver assist and/or autonomous vehicle functionality follow below.

[01 1 1] FIG. 4 is an exemplary functional block diagram of memory 140 and/or 150, which may be stored/programmed with instructions for performing one or more operations consistent with the disclosed embodiments. Although the following refers to memory 140, one of skill in the art will recognize that instructions may be stored in memory 140 and/or 150.

[01 12] As shown in FIG. 4, memory 140 may store a monocular image analysis module 402, a stereo image analysis module 404, a velocity and acceleration module 406, and a navigational response module 408. The disclosed embodiments are not limited to any particular configuration of memory 140. Further, application processor

Tesla EX1002 Page 1560 IPR2025-00944

180 and/or image processor 190 may execute the instructions stored in any of modules 402-408 included in memory 140. One of skill in the art will understand that references in the following discussions to processing unit 1 10 may refer to application processor 180 and image processor 190 individually or collectively. Accordingly, steps of any of the following processes may be performed by one or more processing devices. [01 13] In one embodiment, monocular image analysis module 402 may store instructions (such as computer vision software) which, when executed by processing unit 1 10, performs monocular image analysis of a set of images acquired by one of image capture devices 122, 124, and 126. In some embodiments, processing unit 1 10 may combine infonnation from a set of images with additional sensory information (e.g., information from radar) to perform the monocular image analysis. As described in connection with FIGS. 5A-5D below, monocular image analysis module 402 may include instructions for detecting a set of features within the set of images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, hazardous objects, and any other feature associated with an environment of a vehicle. Based on the analysis, system 100 (e.g., via processing unit 1 10) may cause one or more navigational responses in vehicle 200, such as a turn, a lane shift, a change in acceleration, and the like, as discussed below in connection with navigational response module 408.

[01 14] In one embodiment, stereo image analysis module 404 may store instructions (such as computer vision software) which, when executed by processing unit 1 10, performs stereo image analysis of first and second sets of images acquired by a combination of image capture devices selected from any of image capture devices 122, 124, and 126. In some embodiments, processing unit 1 10 may combine information from the first and second sets of images with additional sensory information (e.g., information from radar) to perform the stereo image analysis. For example, stereo image analysis module 404 may include instructions for performing stereo image analysis based on a first set of images acquired by image capture device 124 and a second set of images acquired by image capture device 126. As described in connection with FIG. 6 below, stereo image analysis module 404 may include instructions for detecting a set of features within the first and second sets of images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, hazardous objects, and the like. Based on the analysis, processing unit 1 10 may cause one or more navigational responses in vehicle 200, such as a turn, a lane shift, a change in acceleration, and the like, as discussed below in connection with navigational response module 408.

[01 15] In one embodiment, velocity and acceleration module 406 may store software configured to analyze data received from one or more computing and electromechanical devices in vehicle 200 that are configured to cause a change in velocity and/or acceleration of vehicle 200. For example, processing unit 1 10 may execute instructions associated with velocity and acceleration module 406 to calculate a target speed for vehicle 200 based on data derived from execution of monocular image analysis module 402 and/or stereo image analysis module 404. Such data may include, for example, a target position, velocity, and/or acceleration, the position and/or speed of vehicle 200 relative to a nearby vehicle, pedestrian, or road object, position information for vehicle 200 relative to lane markings of the road, and the like. In addition, processing unit 1 10 may calculate a target speed for vehicle 200 based on sensory input (e.g., information from radar) and input from other systems of vehicle 200, such as throttling system 220, braking system 230, and/or steering system 240 of vehicle 200. Based on the calculated target speed, processing unit 1 10 may transmit

IPR2025-00944
Tesla EX1002 Page 1561

electronic signals to throttling system 220, braking system 230, and/or steering system 240 of vehicle 200 to trigger a change in velocity and/or acceleration by, for example, physically depressing the brake or easing up off the accelerator of vehicle 200.

[01 16] In one embodiment, navigational response module 408 may store software executable by processing unit 1 10 to determine a desired navigational response based on data derived from execution of monocular image analysis module 402 and/or stereo image analysis module 404. Such data may include position and speed information associated with nearby vehicles, pedestrians, and road objects, target position information for vehicle 200, and the like. Additionally, in some embodiments, the navigational response may be based (partially or fully) on map data, a predetermined position of vehicle 200, and/or a relative velocity or a relative acceleration between vehicle 200 and one or more objects detected from execution of monocular image analysis module 402 and/or stereo image analysis module 404.

Navigational response module 408 may also determine a desired navigational response based on sensory input (e.g., information from radar) and inputs from other systems of vehicle 200, such as throttling system 220, braking system 230, and steering system 240 of vehicle 200. Based on the desired navigational response, processing unit 1 10 may transmit electronic signals to throttling system 220, braking system 230, and steering system 240 of vehicle 200 to trigger a desired navigational response by, for example, turning the steering wheel of vehicle 200 to achieve a rotation of a predetermined angle. In some embodiments, processing unit 1 10 may use the output of navigational response module 408 (e.g., the desired navigational response) as an input to execution of velocity and acceleration module 406 for calculating a change in speed of vehicle 200.

[01 17] FIG. 5 A is a flowchart showing an exemplary process 500A for causing one or more navigational responses based on monocular image analysis, consistent with disclosed embodiments. At step 510, processing unit 1 10 may receive a plurality of images via data interface 128 between processing unit 1 10 and image acquisition unit 120. For instance, a camera included in image acquisition unit 120 (such as image capture device 122 having field of view 202) may capture a plurality of images of an area forward of vehicle 200 (or to the sides or rear of a vehicle, for example) and transmit them over a data connection (e.g., digital, wired, USB, wireless, Bluetooth, etc.) to processing unit 1 10. Processing unit 1 10 may execute monocular image analysis module 402 to analyze the plurality of images at step 520, as described in further detail in connection with FIGS. 5B-5D below. By performing the analysis, processing unit 1 10 may detect a set of features within the set of images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, and the like.

[01 18] Processing unit 1 10 may also execute monocular image analysis module 402 to detect various road hazards at step 520, such as, for example, parts of a truck tire, fallen road signs, loose cargo, small animals, and the like. Road hazards may vary in structure, shape, size, and color, which may make detection of such hazards more challenging. In some embodiments, processing unit 1 10 may execute monocular image analysis module 402 to perform multi-frame analysis on the plurality of images to detect road hazards. For example, processing unit 1 10 may estimate camera motion between consecutive image frames and calculate the disparities in pixels between the frames to construct a 3D-map of the road. Processing unit 1 10 may then use the 3D-map to detect the road surface, as well as hazards existing above the road surface.

[01 19] At step 530, processing unit 1 10 may execute navigational response module 408 to cause one or more navigational responses in vehicle 200 based on the analysis performed at step 520 and the techniques as described above in connection with FIG. 4. Navigational responses may include, for example, a turn, a lane shift, a change in acceleration, and the like. In some embodiments, processing unit 1 10 may use data derived from execution of velocity and acceleration module 406 to cause the one or more navigational responses. Additionally, multiple navigational responses may occur simultaneously, in sequence, or any combination thereof. For instance, processing unit 1 10 may cause vehicle 200 to shift one lane over and then accelerate by, for example, sequentially transmitting control signals to steering system 240 and throttling system 220 of vehicle 200. Alternatively, processing unit 1 10 may cause vehicle 200 to brake while at the same time shifting lanes by, for example, simultaneously transmitting control signals to braking system 230 and steering system 240 of vehicle 200.

[0120] FIG. 5B is a flowchart showing an exemplary process 500B for detecting one or more vehicles and/or pedestrians in a set of images, consistent with disclosed embodiments. Processing unit 1 10 may execute monocular image analysis module 402 to implement process 500B. At step 540, processing unit 1 10 may determine a set of candidate objects representing possible vehicles and/or pedestrians. For example, processing unit 1 10 may scan one or more images, compare the images to one or more predetermined patterns, and identify within each image possible locations that may contain objects of interest (e.g., vehicles, pedestrians, or portions thereof). The predetermined patterns may be designed in such a way to achieve a high rate of "false hits" and a low rate of "misses." For example, processing unit 1 10 may use a low threshold of similarity to predetermined patterns for identifying candidate objects as possible vehicles or pedestrians. Doing so may allow processing unit 1 10 to reduce the probability of missing (e.g., not identifying) a candidate object representing a vehicle or pedestrian.

[0121] At step 542, processing unit 1 10 may filter the set of candidate objects to exclude certain candidates (e.g., irrelevant or less relevant objects) based on classification criteria. Such criteria may be derived from various properties associated with object types stored in a database (e.g., a database stored in memory 140). Properties may include object shape, dimensions, texture, position (e.g., relative to vehicle 200), and the like. Thus, processing unit 1 10 may use one or more sets of criteria to reject false candidates from the set of candidate objects.

[0122] At step 544. processing unit 1 10 may analyze multiple frames of images to determine whether objects in the set of candidate objects represent vehicles and/or pedestrians. For example, processing unit 1 10 may track a detected candidate object across consecutive frames and accumulate frame-by-frame data associated with the detected object (e.g., size, position relative to vehicle 200, etc.). Additionally, processing unit 1 10 may estimate parameters for the detected object and compare the object's frame-by-frame position data to a predicted position.

[0123] At step 546, processing unit 1 10 may construct a set of measurements for the detected objects. Such measurements may include, for example, position, velocity, and acceleration values (relative to vehicle 200) associated with the detected objects. In some embodiments, processing unit 1 10 may construct the measurements based on estimation techniques using a series of time-based observations such as Kalman filters or linear quadratic estimation (LQE), and/or based on available modeling data for

different object types (e.g., cars, trucks, pedestrians, bicycles, road signs, etc.). The Kalman filters may be based on a measurement of an object's scale, where the scale measurement is proportional to a time to collision (e.g., the amount of time for vehicle 200 to reach the object). Thus, by performing steps 540- 546, processing unit 1 10 may identify vehicles and pedestrians appearing within the set of captured images and derive information (e.g., position, speed, size) associated with the vehicles and pedestrians. Based on the identification and the derived information, processing unit 1 10 may cause one or more navigational responses in vehicle 200, as described in connection with FIG. 5A, above.

[0124] At step 548, processing unit 1 10 may perform an optical flow analysis of one or more images to reduce the probabilities of detecting a "false hit" and missing a candidate object that represents a vehicle or pedestrian. The optical flow analysis may refer to, for example, analyzing motion patterns relative to vehicle 200 in the one or more images associated with other vehicles and pedestrians, and that are distinct from road surface motion. Processing unit 1 10 may calculate the motion of candidate objects by observing the different positions of the objects across multiple image frames, which are captured at different times. Processing unit 1 10 may use the position and time values as inputs into mathematical models for calculating the motion of the candidate objects. Thus, optical flow analysis may provide another method of detecting vehicles and pedestrians that are nearby vehicle 200. Processing unit 1 10 may perform optical flow analysis in combination with steps 540-546 to provide redundancy for detecting vehicles and pedestrians and increase the reliability of system 100.

[0125] FIG. 5C is a flowchart showing an exemplary process 500C for detecting road marks and/or lane geometry information in a set of images, consistent with disclosed embodiments. Processing unit 1 10 may execute monocular image analysis module 402 to implement process 500C. At step 550, processing unit 1 10 may detect a set of objects by scanning one or more images. To detect segments of lane markings, lane geometry information, and other pertinent road marks, processing unit 1 10 may filter the set of objects to exclude those determined to be irrelevant (e.g., minor potholes, small rocks, etc.). At step 552, processing unit 1 10 may group together the segments detected in step 550 belonging to the same road mark or lane mark. Based on the grouping, processing unit 1 10 may develop a model to represent the detected segments, such as a mathematical model.

[0126] At step 554, processing unit 1 10 may construct a set of measurements associated with the detected segments. In some embodiments, processing unit 1 10 may create a projection of the detected segments from the image plane onto the real-world plane. The projection may be characterized using a 3rd-degree polynomial having coefficients corresponding to physical properties such as the position, slope, curvature, and curvature derivative of the detected road. In generating the projection, processing unit 1 10 may take into account changes in the road surface, as well as pitch and roll rates associated with vehicle 200. In addition, processing unit 1 10 may model the road elevation by analyzing position and motion cues present on the road surface. Further, processing unit 1 10 may estimate the pitch and roll rates associated with vehicle 200 by tracking a set of feature points in the one or more images.

[0127] At step 556, processing unit 1 10 may perform multi-frame analysis by, for example, tracking the detected segments across consecutive image frames and accumulating frame-by-frame data associated with detected segments. As processing unit 1 10 performs multi-frame analysis, the set of measurements constructed at step

IPR2025-00944
Tesla EX1002 Page 1564

554 may become more reliable and associated with an increasingly higher confidence level. Thus, by performing steps 550-556, processing unit 1 10 may identify road marks appearing within the set of captured images and derive lane geometry information. Based on the identification and the derived information, processing unit 1 10 may cause one or more navigational responses in vehicle 200, as described in connection with FIG. 5A, above.

[0128] At step 558, processing unit 1 10 may consider additional sources of information to further develop a safety model for vehicle 200 in the context of its surroundings. Processing unit 1 10 may use the safety model to define a context in which system 100 may execute autonomous control of vehicle 200 in a safe manner. To develop the safety model, in some embodiments, processing unit 1 10 may consider the position and motion of other vehicles, the detected road edges and barriers, and/or general road shape descriptions extracted from map data (such as data from map database 160). By considering additional sources of information, processing unit 1 10 may provide redundancy for detecting road marks and lane geometry and increase the reliability of system 100.

[0129] FIG. 5D is a flowchart showing an exemplary process 500D for detecting traffic lights in a set of images, consistent with disclosed embodiments. Processing unit 1 10 may execute monocular image analysis module 402 to implement process 500D. At step 560, processing unit 1 10 may scan the set of images and identify objects appearing at locations in the images likely to contain traffic lights. For example, processing unit 1 10 may filter the identified objects to construct a set of candidate objects, excluding those objects unlikely to correspond to traffic lights. The filtering may be done based on various properties associated with traffic lights, such as shape, dimensions, texture, position (e.g., relative to vehicle 200), and the like. Such properties may be based on multiple examples of traffic lights and traffic control signals and stored in a database. In some embodiments, processing unit 1 10 may perform multi-frame analysis on the set of candidate objects reflecting possible traffic lights. For example, processing unit 1 10 may track the candidate objects across consecutive image frames, estimate the real- world position of the candidate objects, and filter out those objects that are moving (which are unlikely to be traffic lights). In some embodiments, processing unit 1 10 may perform color analysis on the candidate objects and identify the relative position of the detected colors appearing inside possible traffic lights.

[0130] At step 562, processing unit 1 10 may analyze the geometry of a junction. The analysis may be based on any combination of: (i) the number of lanes detected on either side of vehicle 200, (ii) markings (such as arrow marks) detected on the road, and (iii) descriptions of the junction extracted from map data (such as data from map database 160). Processing unit 1 10 may conduct the analysis using information derived from execution of monocular analysis module 402. In addition, Processing unit 1 10 may determine a correspondence between the traffic lights detected at step 560 and the lanes appearing near vehicle 200.

[0131] As vehicle 200 approaches the junction, at step 564, processing unit 1 10 may update the confidence level associated with the analyzed junction geometry and the detected traffic lights. For instance, the number of traffic lights estimated to appear at the junction as compared with the number actually appearing at the junction may impact the confidence level. Thus, based on the confidence level, processing unit 1 10 may delegate control to the driver of vehicle 200 in order to improve safety conditions.

By performing steps 560-564, processing unit 1 10 may identify traffic lights appearing within the set of captured images and analyze junction geometry information. Based on the identification and the analysis, processing unit 1 10 may cause one or more navigational responses in vehicle 200, as described in connection with FIG. 5A, above.

[0132] FIG. 5E is a flowchart showing an exemplary process 500E for causing one or more navigational responses in vehicle 200 based on a vehicle path, consistent with the disclosed embodiments. At step 570, processing unit 1 10 may construct an initial vehicle path associated with vehicle 200. The vehicle path may be represented using a set of points expressed in coordinates (x, z), and the distance dtbetween two points in the set of points may fall in the range of 1 to 5 meters. In one embodiment, processing unit 1 10 may construct the initial vehicle path using two polynomials, such as left and right road polynomials. Processing unit 1 10 may calculate the geometric midpoint between the two polynomials and offset each point included in the resultant vehicle path by a predetermined offset (e.g., a smart lane offset), if any (an offset of zero may correspond to travel in the middle of a lane). The offset may be in a direction perpendicular to a segment between any two points in the vehicle path. In another embodiment, processing unit 1 10 may use one polynomial and an estimated lane width to offset each point of the vehicle path by half the estimated lane width plus a predetermined offset (e.g., a smart lane offset).

[0133] At step 572, processing unit 1 10 may update the vehicle path constructed at step 570.

Processing unit 1 10 may reconstruct the vehicle path constructed at step 570 using a higher resolution, such that the distance dkbetween two points in the set of points representing the vehicle path is less than the distance d, described above. For example, the distance dkmay fall in the range of 0.1 to 0.3 meters. Processing unit 1 10 may reconstruct the vehicle path using a parabolic spline algorithm, which may yield a cumulative distance vector S corresponding to the total length of the vehicle path (i.e., based on the set of points representing the vehicle path).

[0134] At step 574, processing unit 1 10 may determine a look-ahead point (expressed in coordinates as (xt, z/)) based on the updated vehicle path constructed at step 572. Processing unit 1 10 may extract the look-ahead point from the cumulative distance vector S, and the look-ahead point may be associated with a look-ahead distance and look-ahead time. The look-ahead distance, which may have a lower bound ranging from 10 to 20 meters, may be calculated as the product of the speed of vehicle 200 and the look-ahead time. For example, as the speed of vehicle 200 decreases, the look-ahead distance may also decrease (e.g., until it reaches the lower bound). The look-ahead time, which may range from 0.5 to 1 .5 seconds, may be inversely proportional to the gain of one or more control loops associated with causing a navigational response in vehicle 200, such as the heading error tracking control loop. For example, the gain of the heading error tracking control loop may depend on the bandwidth of a yaw rate loop, a steering actuator loop, car lateral dynamics, and the like. Thus, the higher the gain of the heading error tracking control loop, the lower the look-ahead time.

[0135] At step 576, processing unit 1 10 may determine a heading error and yaw rate command based on the look-ahead point determined at step 574. Processing unit 1 10 may determine the heading error by calculating the arctangent of the look-ahead point, e.g., arctan (xtl zj). Processing unit 1 10 may determine the yaw rate command as the

product of the heading error and a high-level control gain. The high-level control gain may be equal to: (2 / look-ahead time), if the look-ahead distance is not at the lower bound. Otherwise, the high-level control gain may be equal to: (2 * speed of vehicle 200 / look- ahead distance).

[0136] FIG. 5F is a flowchart showing an exemplary process 500F for determining whether a leading vehicle is changing lanes, consistent with the disclosed embodiments. At step 580, processing unit 1 10 may determine navigation information associated with a leading vehicle (e.g., a vehicle traveling ahead of vehicle 200). For example, processing unit 1 10 may determine the position, velocity (e.g., direction and speed), and/or acceleration of the leading vehicle, using the techniques described in connection with FIGS. 5A and 5B, above. Processing unit 1 10 may also determine one or more road polynomials, a look-ahead point (associated with vehicle 200), and/or a snail trail (e.g., a set of points describing a path taken by the leading vehicle), using the techniques described in connection with FIG. 5E, above.

[0137] At step 582, processing unit 1 10 may analyze the navigation information determined at step 580. In one embodiment, processing unit 1 10 may calculate the distance between a snail trail and a road polynomial (e.g., along the trail). If the variance of this distance along the trail exceeds a predetermined threshold (for example, 0.1 to 0.2 meters on a straight road, 0.3 to 0.4 meters on a moderately curvy road, and 0.5 to 0.6 meters on a road with sharp curves), processing unit 1 10 may determine that the leading vehicle is likely changing lanes. In the case where multiple vehicles are detected traveling ahead of vehicle 200, processing unit 1 10 may compare the snail trails associated with each vehicle. Based on the comparison, processing unit 1 10 may determine that a vehicle whose snail trail does not match with the snail trails of the other vehicles is likely changing lanes. Processing unit 1 10 may additionally compare the curvature of the snail trail (associated with the leading vehicle) with the expected curvature of the road segment in which the leading vehicle is traveling. The expected curvature may be extracted from map data (e.g., data from map database 160), from road polynomials, from other vehicles' snail trails, from prior knowledge about the road, and the like. If the difference in curvature of the snail trail and the expected curvature of the road segment exceeds a predetermined threshold, processing unit 1 10 may determine that the leading vehicle is likely changing lanes.

[0138] In another embodiment, processing unit 1 10 may compare the leading vehicle's instantaneous position with the look-ahead point (associated with vehicle 200) over a specific period of time (e.g., 0.5 to 1.5 seconds). If the distance between the leading vehicle's instantaneous position and the look-ahead point varies during the specific period of time, and the cumulative sum of variation exceeds a predetermined threshold (for example, 0.3 to 0.4 meters on a straight road, 0.7 to 0.8 meters on a moderately curvy road, and 1.3 to 1.7 meters on a road with sharp curves), processing unit 1 10 may determine that the leading vehicle is likely changing lanes. In another embodiment, processing unit 1 10 may analyze the geometry of the snail trail by comparing the lateral distance traveled along the trail with the expected curvature of the snail trail. The expected radius of curvature may be determined according to the calculation: (δζ<2>+ δχ<2>) 1 2 1 (δχ), where δχrepresents the lateral distance traveled and δζrepresents the longitudinal distance traveled. If the difference between the lateral distance traveled and the expected curvature exceeds a predetermined threshold (e.g., 500 to 700 meters), processing unit 1 10 may determine that the leading vehicle is likely changing lanes. In another embodiment, processing unit 1 10 may analyze the position of the leading vehicle. If the position of the leading vehicle

IPR2025-00944
Tesla EX1002 Page 1567

obscures a road polynomial (e.g., the leading vehicle is overlaid on top of the road polynomial), then processing unit 1 10 may determine that the leading vehicle is likely changing lanes. In the case where the position of the leading vehicle is such that, another vehicle is detected ahead of the leading vehicle and the snail trails of the two vehicles are not parallel, processing unit 1 10 may determine that the (closer) leading vehicle is likely changing lanes.

[0139] At step 584, processing unit 1 10 may determine whether or not leading vehicle 200 is changing lanes based on the analysis performed at step 582. For example, processing unit 1 10 may make the determination based on a weighted average of the individual analyses performed at step 582. Under such a scheme, for example, a decision by processing unit 1 10 that the leading vehicle is likely changing lanes based on a particular type of analysis may be assigned a value of "1 " (and "0" to represent a determination that the leading vehicle is not likely changing lanes). Different analyses performed at step 582 may be assigned different weights, and the disclosed embodiments are not limited to any particular combination of analyses and weights.

[0140] FIG. 6 is a flowchart showing an exemplary process 600 for causing one or more navigational responses based on stereo image analysis, consistent with disclosed embodiments. At step 610, processing unit 1 10 may receive a first and second plurality of images via data interface 128. For example, cameras included in image acquisition unit 120 (such as image capture devices 122 and 124 having fields of view 202 and 204) may capture a first and second plurality of images of an area forward of vehicle 200 and transmit them over a digital connection (e.g., USB, wireless, Bluetooth, etc.) to processing unit 1 10. In some embodiments, processing unit 1 10 may receive the first and second plurality of images via two or more data interfaces. The disclosed embodiments are not limited to any particular data interface configurations or protocols.

[0141] At step 620, processing unit 1 10 may execute stereo image analysis module 404 to perform stereo image analysis of the first and second plurality of images to create a 3D map of the road in front of the vehicle and detect features within the images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, road hazards, and the like. Stereo image analysis may be performed in a manner similar to the steps described in connection with FIGS. 5A-5D, above. For example, processing unit 1 10 may execute stereo image analysis module 404 to detect candidate objects (e.g., vehicles, pedestrians, road marks, traffic lights, road hazards, etc.) within the first and second plurality of images, filter out a subset of the candidate objects based on various criteria, and perform multi-frame analysis, construct measurements, and determine a confidence level for the remaining candidate objects. In performing the steps above, processing unit 1 10 may consider information from both the first and second plurality of images, rather than information from one set of images alone. For example, processing unit 1 10 may analyze the differences in pixel-level data (or other data subsets from among the two streams of captured images) for a candidate object appearing in both the first and second plurality of images. As another example, processing unit 1 10 may estimate a position and/or velocity of a candidate object (e.g., relative to vehicle 200) by observing that the object appears in one of the plurality of images but not the other or relative to other differences that may exist relative to objects appearing if the two image streams. For example, position, velocity, and/or acceleration relative to vehicle 200 may be determined based on trajectories, positions, movement characteristics, etc. of features associated with an object appearing in one or both of the image streams.

IPR2025-00944
Tesla EX1002 Page 1568

[0142] At step 630, processing unit 1 10 may execute navigational response module 408 to cause one or more navigational responses in vehicle 200 based on the analysis performed at step 620 and the techniques as described above in connection with FIG. 4. Navigational responses may include, for example, a turn, a lane shift, a change in acceleration, a change in velocity, braking, and the like. In some embodiments, processing unit 1 10 may use data derived from execution of velocity and acceleration module 406 to cause the one or more navigational responses. Additionally, multiple navigational responses may occur simultaneously, in sequence, or any combination thereof.

[0143] FIG. 7 is a flowchart showing an exemplary process 700 for causing one or more navigational responses based on an analysis of three sets of images, consistent with disclosed embodiments. At step 710, processing unit 1 10 may receive a first, second, and third plurality of images via data interface 128. For instance, cameras included in image acquisition unit 120 (such as image capture devices 122, 124, and 126 having fields of view 202, 204, and 206) may capture a first, second, and third plurality of images of an area forward and/or to the side of vehicle 200 and transmit them over a digital connection (e.g., USB, wireless, Bluetooth, etc.) to processing unit 1 10. In some embodiments, processing unit 1 10 may receive the first, second, and third plurality of images via three or more data interfaces. For example, each of image capture devices 122, 124, 126 may have an associated data interface for communicating data to processing unit 1 10. The disclosed embodiments are not limited to any particular data interface configurations or protocols.

[0144] At step 720, processing unit 1 10 may analyze the first, second, and third plurality of images to detect features within the images, such as lane markings, vehicles, pedestrians, road signs, highway exit ramps, traffic lights, road hazards, and the like. The analysis may be performed in a manner similar to the steps described in connection with FIGS. 5A-5D and 6, above. For instance, processing unit 1 10 may perform monocular image analysis (e.g., via execution of monocular image analysis module 402 and based on the steps described in connection with FIGS. 5A-5D, above) on each of the first, second, and third plurality of images. Alternatively, processing unit 1 10 may perform stereo image analysis (e.g., via execution of stereo image analysis module 404 and based on the steps described in connection with FIG. 6, above) on the first and second plurality of images, the second and third plurality of images, and/or the first and third plurality of images. The processed information corresponding to the analysis of the first, second, and/or third plurality of images may be combined. In some embodiments, processing unit 1 10 may perform a combination of monocular and stereo image analyses. For example, processing unit 1 10 may perform monocular image analysis (e.g., via execution of monocular image analysis module 402) on the first plurality of images and stereo image analysis (e.g., via execution of stereo image analysis module 404) on the second and third plurality of images. The configuration of image capture devices 122, 124, and 126— including their respective locations and fields of view 202, 204, and 206— may influence the types of analyses conducted on the first, second, and third plurality of images. The disclosed embodiments are not limited to a particular configuration of image capture devices 122, 124, and 126, or the types of analyses conducted on the first, second, and third plurality of images.

[0145] In some embodiments, processing unit 1 10 may perform testing on system 100 based on the images acquired and analyzed at steps 710 and 720. Such testing may provide an indicator of the overall performance of system 100 for certain configurations of image capture devices 122, 124, and 126. For example, processing unit 1 10 may

IPR2025-00944
Tesla EX1002 Page 1569

determine the proportion of "false hits" (e.g., cases where system 100 incorrectly determined the presence of a vehicle or pedestrian) and "misses."

[0146] At step 730, processing unit 1 10 may cause one or more navigational responses in vehicle

200 based on information derived from two of the first, second, and third plurality of images. Selection of two of the first, second, and third plurality of images may depend on various factors, such as, for example, the number, types, and sizes of objects detected in each of the plurality of images. Processing unit 1 10 may also make the selection based on image quality and resolution, the effective field of view reflected in the images, the number of captured frames, the extent to which one or more objects of interest actually appear in the frames (e.g., the percentage of frames in which an object appears, the proportion of the object that appears in each such frame, etc.), and the like.

[0147] In some embodiments, processing unit 1 10 may select information derived from two of the first, second, and third plurality of images by determining the extent to which information derived from one image source is consistent with information derived from other image sources. For example, processing unit 1 10 may combine the processed information derived from each of image capture devices 122, 124, and 126 (whether by monocular analysis, stereo analysis, or any combination of the two) and determine visual indicators (e.g., lane markings, a detected vehicle and its location and/or path, a detected traffic light, etc.) that are consistent across the images captured from each of image capture devices 122, 124, and 126. Processing unit 1 10 may also exclude information that is inconsistent across the captured images (e.g., a vehicle changing lanes, a lane model indicating a vehicle that is too close to vehicle 200, etc.). Thus, processing unit 1 10 may select information derived from two of the first, second, and third plurality of images based on the determinations of consistent and inconsistent information. [0148] Navigational responses may include, for example, a turn, a lane shift, a change in acceleration, and the like. Processing unit 1 10 may cause the one or more navigational responses based on the analysis performed at step 720 and the techniques as described above in connection with FIG. 4. Processing unit 1 10 may also use data derived from execution of velocity and acceleration module 406 to cause the one or more navigational responses. In some embodiments, processing unit 1 10 may cause the one or more navigational responses based on a relative position, relative velocity, and/or relative acceleration between vehicle 200 and an object detected within any of the first, second, and third plurality of images. Multiple navigational responses may occur simultaneously, in sequence, or any combination thereof.

[0149] Object Detecting System

[0150] In some embodiments, vehicle 200 may include an object detecting system configured to detect an object in front of vehicle 200 and in some circumstances, cause vehicle 200 to brake, e.g., using braking system 230, based on the detection of the object. For example, the object detecting system may cause vehicle 200 to brake automatically, or autonomously, without or with limited driver intervention. The automatic or autonomous object detecting and braking system may or may not allow a driver to intervene to take over control of braking system 230.

[01 51] The object detecting system may include a plurality of other systems discussed

Tesla EX1002 Page 1570 IPR2025-00944

above. For example, the object detecting and braking system may include system 100, braking system 230, and one or more of image capture devices 122, 124, and 126. The one or more of image capture devices 122, 124, and 126 may acquire one or more images of an environment or an area in front of vehicle 200, which may include an object, such as another vehicle, a pedestrian, etc. Processing unit 1 10 may compare at least two images to obtain optical flow information, which may include a vector field formed by a plurality of vectors (e.g., displacement vectors) each representing a displacement between a first pixel in the first image and a corresponding second pixel in the second image. Based on the optical flow information, processing unit 1 10 may detect an object, such as an upright object in the images, and may verify the structure of the object. The upright object may have a low (e.g., short) time to collision with vehicle 200. Thus, the upright object may also be referred to as a low time-to-collision upright object. In some embodiments, processing unit may search for a region of coherent expansion that is a set of pixels in, e.g., the first image, for which there exists a common focus of expansion and a common scale magnitude such that the set of pixels satisfy a relationship between pixel positions, displacement vectors, the common focus of expansion, and the common scale magnitude. The region of coherent expansion may represent a low time-to-collision upright object. The scale magnitude may be defined as a magnitude of a displacement vector associated with the pixel (if denoted as d) divided by a distance (if denoted as r) from the pixel to the focus of expansion, i.e., scale magnitude = d/r. The scale magnitude may also be referred to as the expansion rate.

[0152] In some embodiments, processing unit 1 10 may calculate a plurality of combinations of possible scale magnitudes and possible focus of expansions for a plurality of pixels from at least one of the first image and the second image. The plurality of pixels may include the set of pixels to be searched for. In some embodiments, processing unit 1 10 may determine the common focus of expansion and the common scale magnitude based on a distribution of the plurality of combinations of possible scale magnitudes and possible focus of expansions. The plurality of combinations of possible scale magnitudes and possible focus of expansions may be distributed in a three-dimensional space. In some embodiments, processing unit 1 10 may determine the common focus of expansion and common scale magnitude using a suitable clustering technique, such as one that may determine the center of mass (e.g., center of density) about the distribution. Processing unit 1 10 may determine from the distribution that the density of distribution around a particular pair of scale magnitude and focus of expansion is higher than a predetermined threshold, and may select that particular pair of scale magnitude and focus of expansion as the common scale magnitude and common focus of expansion.

[0153] Processing unit 1 10 may select the region of coherent expansion (i.e., the set of pixels) based on the common focus of expansion and common scale magnitude, wherein the set of pixels satisfy the relationship between pixel positions, displacement vectors, the common focus of expansion, and the common scale magnitude. The set of pixels may be selected from the first image or the second image. Based on the selected set of pixels (or the region of coherent expansion), processing unit 1 10 may identify the presence of a substantially upright low time-to-collision object in the first image (or the second image). For example, the set of pixels may show a portion of the substantially upright low time-to- collision object. Processing unit 1 10 may identify such a low time-to-collision upright object regardless of whether the object is static relative to the road, and/or regardless of the moving direction of the object, and/or regardless of whether the object has a low (e.g., short) time to collision with vehicle 200, as long as

Tesla EX1002 Page 1571 IPR2025-00944

the object has a relative speed as compared to vehicle 200 such that there exists a focus of expansion for the object. The focus of expansion may not exist for the object when the object appears static in an image plane relative to vehicle 200 because there would be no expansion or shrinking in the optical flow field.

[0154] Processing unit 1 10 may further determine that the identified object is a crossing object

(which may also be referred to as a crossing-to-impact object hereinafter) that is moving at a constant speed in a direction crossing the travel direction of vehicle 200, and which would likely collide with vehicle 200 should vehicle 200 and the crossing object continue to move at their respective constant speeds. The crossing object may be a low time-to-collision upright object. Processing unit 1 10 may determine whether a horizontal coordinate of the focus of expansion of the crossing object stays around a center in the horizontal dimension of the crossing object during a period of time. This may indicate that a viewing angle from a camera installed on vehicle 200 is constant. When a viewing angle is constant, if the crossing object moves linearly at a constant speed, and vehicle 200 moves linearly at a constant speed, they may collide at certain point. Thus, processing unit 1 10 may determine whether crossing object and vehicle 200 may collide based on a determination of whether the horizontal coordinate of the focus of expansion of the crossing object stays around a center in the horizontal dimension of the crossing object during the period of time. When processing unit 1 10 determines that the horizontal coordinate of the focus of expansion of the crossing object stays around a center in the horizontal dimension of the crossing object during the period of time, processing unit may cause vehicle 200 to brake using, e.g., braking system 230. In some embodiments, the braking may be autonomous without driver intervention.

[0155] FIG. 8 is a diagrammatic representation of an exemplary vehicle 200 including an object detecting system (or referred to as an emergency braking decisioning system). The object detecting system may include system 100 and braking system 230. The object detecting system may also include at least one of image capture devices 122 and 124 (and 126, which is not shown in FIG. 8). The object detecting system may include processing unit 1 10, which may be part of system 100. At least one of image capture devices 122 and 124 may acquire one or a plurality of images of an area including an object in front of vehicle 200. The object may be another vehicle, such as a vehicle 800 that may be traveling in front of vehicle 200, or a pedestrian, an animal, etc. As shown in FIG. 8, in some embodiments, vehicle 800 may be traveling in a direction that is parallel to a traveling direction of vehicle 200. A parallel direction may include a situation in which vehicle 800 is traveling in the same direction as and in alignment or substantially in alignment with vehicle 200 (e.g., vehicle 800 traveling in front of vehicle 200), as shown in FIG. 8. A parallel direction may also include a situation in which vehicle 800 is traveling in an opposite direction relative to vehicle 200 and in alignment with vehicle 200 (e.g., vehicle 800 moving toward vehicle 200 in a direction that is 180 degrees, or substantially 180 degrees, opposite to the traveling direction of vehicle 200). A parallel direction may also include a situation in which vehicle 800 is not traveling in alignment with vehicle 200, but is traveling in a direction parallel to the traveling direction of vehicle 200 (e.g., vehicle 800 traveling on a front side of vehicle 200 at an offset distance).

[01 6] In some embodiments, the object may be a pedestrian who may be standing in the road, or walking or running across the road in front of vehicle 200. The area

included in the images may include other objects, such as a road surface, a house, a wall, a sign, etc., which may form a static environment. When an object is not traveling in a direction parallel to the traveling direction of vehicle 200, the traveling direction of the object may transverse or cross the traveling direction of vehicle 200 with an angle that is neither zero (or substantially zero) nor 180 degrees (or substantially 180 degrees). The images may include both moving and/or static objects.

[0157] Processing unit 1 10 may communicate with braking system 230 and at least one of image capture devices 122 and 124. For example, processing unit 1 10 may receive images from at least one of image capture devices 122 and 124, and may process the images using various methods to detect the object (e.g., another vehicle, pedestrian, or animal, etc.). Although two image capture devices 122 and 124 are shown in FIG. 8, the use of additional image capture devices (e.g., 3, 4, 5, etc., image capture devices) is consistent with the disclosed embodiments.

[0158] FIG. 9 is an exemplary block diagram of memory 140 or 150 that may store instructions for performing one or more operations for detecting an object in front of vehicle 200 and/or causing vehicle 200 to brake based on the detection. As shown in FIG. 9, memory 140 or 150 may store one or more modules for performing the operations, as described herein. For example, memory 140 or 150 may store an image analysis module 900 configured to analyze the images. Memory 140 or 150 may store a braking module 905 configured to cause vehicle 200 to brake by controlling, e.g., braking system 230. Application processor 180 and/or image processor 190 may execute the instructions stored in any of modules 900 and 905 included in memory 140 or 150. One of skill in the art will understand that references in the following discussions to processing unit 1 10 may refer to application processor 180 and image processor 190 individually or collectively. Accordingly, steps of any of the following processes may be performed by one or more processing devices.

[0159] Image analysis module 900 may store instructions which, when executed by processing unit 1 10, may perform an image analysis to detect an object traveling in front of vehicle 200 (e.g., vehicle 800 or a pedestrian crossing the road). Image analysis module 900 may analyze at least two images acquired by at least one of image capture devices 122 and 124. For example, image analysis module 900 may compare at least a first image to a second image to obtain optical flow information. The optical flow information may include an optical flow field formed by a plurality of displacement vectors between each first pixel in the first image and each corresponding second pixel in the second image. Image analysis module 900 may search for a region of coherent expansion that is a set of pixels in at least one of the first image and the second image, for which there exists a common focus of expansion and a common scale magnitude such that the set of pixels satisfy a relationship between pixel positions, displacement vectors, the common focus of expansion, and the common scale magnitude. Image analysis module 900 may further identify presence of a substantially upright object in at least one of the first image and the second image based on the set of pixels.

[0160] In some embodiments, image analysis module 900 may store instructions which, when executed by processing unit 1 10, may perform an image analysis to detect an object crossing the road in front of vehicle 200, which would likely collide with vehicle 200 if vehicle 200 and the object continue to move linearly (e.g., in a straight line) at their respective constant speeds in their respective directions that cross each other at an angle which is neither zero nor 180 degrees. Image analysis module 900 may

IPR2025-00944
Tesla EX1002 Page 1573

determine whether a horizontal coordinate of the focus of expansion stays around a center in the horizontal dimension of the crossing object during a period of time.

[0161] Braking module 905 may store instructions which, when executed by processing unit 1 10, may cause vehicle 200 to brake using braking system 230. For example, when image analysis module 900 determines that a horizontal coordinate of the focus of expansion of the crossing object stays around a center in the horizontal dimension of the crossing object during a period of time, braking module 905 may cause vehicle 200 to brake to avoid collision with the crossing object.

[0162] FIG. 10A is an exemplary image of an area showing vectors associated with pixels. Although only one image of a vehicle 1000 is shown for illustration purposes, it is understood that a plurality of image may be acquired by one of the image capture devices 122-126 at a time interval after the first image is acquired. Processing unit 1 10 may compare a first image to a second image to obtain optical flow information including a vector field formed by a plurality of vectors, each vector representing displacement (or movement) associated with each first pixel in the first image and each corresponding second pixel in the second image. The vectors may also be referred to as displacement vectors. In the example displacement vector field shown in FIG. 10A, coordinates of three pixels in the first image are shown, (xl , y 1), (x2, y2), and (x3, y3). Displacement vectors are shown as arrows.

Coordinates (ul , vl), (u2, v2). and (u3, v3) may represent horizontal and vertical displacements associated with each displacement vector, or may represent coordinates of corresponding second pixels on the second image in some embodiments.

[0163] Processing unit 1 10 may search for a region of coherent expansion that is a set of pixels in at least one of the first image and the second image, for which there exists a common focus of expansion and a common scale magnitude such that the set of pixels satisfy a relationship between pixel positions, displacement vectors, the common focus of expansion, and the common scale magnitude. The relationship may be expressed as equations (1) and (2) for an ith pixel in the first image (or in the second image) having coordinates (xi, yi), where i is a integer number:

[0164] SS * (^ - FOEJ = u, (1)

[0165] SS * (y, - FOEy) = v, (2)

[0166] Where SS is the scale magnitude, FOExand FOEyare coordinates of the focus of expansion. Because SS, FOEx, and FOEyare unknown variables, there may be an infinite number of possible solutions for equations (1) and (2). The possible solutions of equations (1) and (2) may be distributed in a three-dimensional cube space shown in FIG. 10B (which may represent a portion of the infinite space), where the x axis represents FOEx, and the y-axis represents FOEy, and the z-axis represents scale magnitude SS. There may exist a set of pixels among the pixels of the first image (or the second image) for which there exists a pair of common scale magnitude SSO and common focus of expansion (FOEx0, FOEy0) such that all or most of the set of pixels satisfy the relationship represented by equations (1) and (2) with SS being replaced by common scale magnitude SSO and FOExand FOEybeing replaced by FOEx0, FOEy0of the common focus of expansion. This set of pixels may be referred to a region of coherent expansion. The pixels in the region of coherent expansion may share the same common scale magnitude and the same focus of expansion.

IPR2025-00944
Tesla EX1002 Page 1574

[0167] A plurality of methods may be used to determine a common focus of expansion from the plurality of focuses of expansion. For example, a clustering technique may be used to determine the standard deviation of the distribution, the center of mass, or the center of the density of distribution. Such information obtained using the clustering technique may be used to determine the common scale magnitudes, SSO, and the common focus of expansion (FOEx0, FOEy0).

[0168] In one embodiment, the three-dimensional cube space shown in FIG. 10B may be divided into a plurality of cells, each cell corresponding to a small range corresponding to a pair of (SS;, FOExi, FOEyi). Processing unit 1 10 may place the possible solutions into the cells and count the number of solutions within the cells. Processing unit 1 10 may determine which cell has the largest number of solutions. In one embodiment, the cell has the largest number of solutions may be used to determine the common focus of expansion and the common scale magnitude.

[0169] In another embodiment, processing unit 1 10 may one variable, such as one of the scale magnitude SS, FOEx, or FOEy, and solve the equations (1) and (2) for possible solutions for each pixel. Processing unit 1 10 may repeat this for all or a plurality of pixels in the first image to generate a plurality of possible solutions. The plurality of possible solutions (SS, FOEx, FOEy) may be arranged into a list or an array. Processing unit 1 10 may determine a distribution of the possible solutions using, e.g., a clustering technique, to find a dense clustering. For example, if the density around a particular solution (SS, FOEx, FOEy) is higher than a predetermined threshold, processing unit 1 10 may use that particular solution as the common focus of expansion and common scale magnitude for the plurality of pixels in the first image.

[0170] In some embodiments, processing unit 1 10 may determine whether the possible solutions "agree" with each other, e.g., whether all or most of the solutions or more than a predetermined number of solutions are located within a predetermined small space in the three-dimensional cube space shown in FIG. 10B. If the possible solutions "agree" with each other, processing unit 1 10 may determine the common focus of expansion and the commons scale magnitude based on the possible solutions within the predetermined space (e.g., taking the average of the all solutions within the predetermined space as the common scale magnitude and the common focus of expansion).

[0171] Based on the common focus of expansion and the common scale magnitude, processing unit 1 10 may identify the set of pixels that satisfy the relationship represented by equations (1) and (2) with the common scale magnitude and common focus of expansion. In other words, processing unit 1 10 may identify the set of pixels that share the common scale magnitude and the common focus of expansion. Processing unit 1 10 may detect the presence of an object, such as an upright object low time- to-collision object (e.g., a back door of a van shown in FIG. I OC) based on the identified set of pixels.

For example, the identified set of pixels may show an image of a substantial portion of the upright object (e.g., the back door of the van), as shown in FIG. I OC.

[0172] FIG. 1 1 is a flowchart showing an exemplary process for detecting presence of an object in front of vehicle 200. Method 1 100 may include acquiring a plurality of images of an area in front of vehicle 200 (step 1 1 10). For example, at least one of

Tesla EX1002 Page 1575 IPR2025-00944

image capture devices 122 and 124 may acquire one or more images of an area in front of vehicle 200. The area may include an object, such as another vehicle or a pedestrian. The area may include other static objects, such as a sign, a house, a mountain, a road surface, etc., which may form a static environment.

[0173] Method 1 100 may include receiving the plurality of images (step 1 120). For example, processing unit 1 10 may receive the plurality of images from at least one of the image capture devices 122 and 124. Method 1 100 may also include comparing a first image to a second image to determine displacement vectors between pixels of the first image and pixels of the second image (step 1 130). The first and second images are from the plurality of images, and may be acquired at a predetermined time interval, such as 0.01 second, 0.02 second, etc. Processing unit 1 10 may compare the first image and second image to obtain optical flow information, which may include displacement vectors between pixels of the first image and pixels of the second image. A displacement vector indicates a direction and a magnitude of displacement between a first pixel in the first image and a second pixel in the second image. The displacement vectors may converge into one or more common converging points, which may also be referred to as common focus of expansions. Processing unit 1 10 may determine the common focus of expansion based on any of the methods disclosed herein.

[0174] Method 1 100 may further include searching for a region of coherent expansion that is a set of pixels in at least one of the first image and the second image, for which there exists a common focus of expansion and a common scale magnitude such that the set of pixels satisfy a relationship between pixel positions, displacement vectors, the common focus of expansion, and the common scale magnitude (step 1 140). For example, processing unit 1 10 may search for the set of pixels that satisfy the relationship represented by equations (1) and (2) with the common focus of expansion and the common scale magnitude. Based on the set of pixels, processing unit 1 10 may identify the presence of a substantially upright object in at least one of the first image and the second image (step 1 150). The upright object may be low time-to-collision object, such as a moving vehicle or pedestrian with which vehicle 200 may collision at a future time.

[0175] FIG. 12A is an image of a static environment captured by at least one of image capture devices 122 and 124 installed on vehicle 200. The static environment may include a road, one or more trees along the road, and a portion of the sky. FIG. 12B shows a three-dimensional cube space illustrating the relationship between the focus of expansion and the scale magnitude. The cube space shows a space where possible solutions of (SS, FOEx, FOEy) may be distributed. As shown in FIG. 12B, distribution of the scale magnitudes and focus of expansions of different objects in the static environment may be represented by a cylindrical shape 1200 (or a straight line; the cylindrical shape is only for illustration purposes). Different objects in the image of the static environment may have different distances to the camera on vehicle 200. Thus, the scale magnitudes may be different for different objects in the image of the static environment. As indicated by the cylindrical shape 1200 in FIG. 12B, objects in the static environment share the same focus of expansion.

[0176] FIG. 13A is an image of a environment having another vehicle 1300 in front of vehicle 200. The object vehicle 1300 may be static relative to the road or may be moving in a direction parallel to the traveling direction of vehicle 200. As shown in FIG. 13B, the distribution of the scale magnitude and focus of expansion of different objects in the static environment may be represented by a cylindrical shape 1310. The focus of

IPR2025-00944
Tesla EX1002 Page 1576

expansion of object vehicle 1300 may be represented by the circle 1320 centered on the cylindrical shape 1310 at a certain scale magnitude. In other words, the focus of expansion of object vehicle 1300 is the same as the focus of expansion of the static environment.

[0177] FIG. 13C is an image of a environment having another object vehicle 1350 traveling longitudinally in front of and in alignment with vehicle 200. As shown in FIG. 13D, the distribution of scale magnitudes and focus of expansions for different objects in the static environment may be represented by a cylindrical shape 1360. The focus of expansion of object vehicle 1350 may be represented by a circle 1370 centered on the cylindrical shape 1360 at a certain scale magnitude. In other words, the focus of expansion of object vehicle 1350 is the same as the focus of expansion of the static environment. [0178] FIG. 14A is an image of an environment superimposed with an exemplary optical flow field (e.g., the displacement vector field). The image of the environment may include an object vehicle 1410 and a pedestrian 1420, who may be crossing the road or standing still on the road. An optical flow field having a plurality of vectors is superimposed to the figure to indicate that an optical flow field may be obtained from at least two images of the environment with object vehicle 1410 and pedestrian 1420 acquired at two different times.

[0179] FIG. 14B is a three-dimensional cube space illustrating two exemplary focus of expansions corresponding to the vehicle object and the pedestrian object shown in FIG. 14A. As shown in FIG. 14B, the distribution of scale magnitude and focus of expansion of different objects in the static environment may be represented by a cylindrical shape 1430. The focus of expansion of object vehicle 1410 may be represented by a circle 1440 centering around the cylindrical shape 1430 at a certain scale magnitude, indicating that the focus of expansion of object vehicle 1410 is the same as the focus of expansion of the static environment. FIG. 14B also shows a focus of expansion of pedestrian 1420, as indicated by the circle 1450. The focus of expansion of pedestrian 1420 may be located offset from the focus of expansion of the static environment and the focus of expansion of object vehicle 1410, which means at least one of the horizontal and the vertical coordinate of focus of expansion of pedestrian 1420 may be different from that of the focus of expansion of the static environment and/or object vehicle 1410. The focus of expansion of pedestrian 1420 may be located on a plane corresponding to a first common scale magnitude, which may be different from a second common scale magnitude corresponding to the focus of expansion of object vehicle 1410.

[0180] In some embodiments, system 100 may detect an object crossing the road in front of vehicle 200, and may cause vehicle 200 to brake to avoid collision. Such an object may be referred to as a crossing object or a crossing-to-impact object, which may have a low time to collision with vehicle 200. The crossing object may likely collide with vehicle 200 if the crossing object and vehicle 200 continue to move in their current linear directions at their current substantially constant speed. FIG. 15A shows a diagram illustrating an exemplary method of detecting a crossing-to-impact object. A crossing object 1500 may travel in a direction that crosses the travel direction of vehicle 200. If crossing object 1500 and vehicle continue to travel at their respective substantially constant speeds, the viewing angle from the camera installed on vehicle 200, a, will remain constant as crossing object 1500 and vehicle 200 travel to point 1510. They may likely collide in a future time at point 1510. System 100 may detect such a cross- to-impact object and may take actions to avoid collision by, for example, braking vehicle 200 before vehicle 200 and crossing object 1500 reach the collision

Tesla EX1002 Page 1577 IPR2025-00944

point 1510. Although FIG. 15A shows that crossing object 1500 and vehicle 200 travel in directions that are perpendicular to each other, their directions do not have to be perpendicular. For example, the angle between the direction of travel of crossing object 1500 and vehicle 200 may be any suitable value, such as 60 degrees, 45 degrees, etc.

[0181] In some embodiments, a constant viewing angle may be detected based on the optical flow information. For example, when object 1500 moves with a constant viewing angle, the object may appear fixed horizontally in an image plane and expands in a direction perpendicular to the image plane. When object 1500 moves with a constant viewing angle, the object 1500 may have a focus of expansion, which may stay around a center in the horizontal dimension of object 1500. Processing unit 1 10 may use such information to determine whether object 1500 may collide with camera installed on vehicle 200.

[0182] FIG. 15B shows an image having a pedestrian 1520 (an exemplary crossing object) crossing the road in front of vehicle 200. FIG. 15C shows an image having the pedestrian 1520 crossing the road in front of vehicle 200 acquired at a time after the image shown in FIG. 1 5B is acquired (thus, pedestrian 1520 appears larger in FIG. 15C than in FIG. 15B). A vertical dotted line 1530 in FIG. 15B and FIG. 15C indicates the horizontal coordinate of the focus of expansion of pedestrian 1520. If processing unit 1 10 detects that the vertical line 1530 stays around a center in the horizontal dimension of pedestrian 1520 and does not move away from the center in the horizontal direction during the period of time from the first image (e.g., FIG. 1 5B) to the second image (e.g., FIG. 15C), processing unit 1 10 may determine that a viewing angle from the camera (e.g., one in the image capture device 122, 124, or 126) installed in vehicle 200 is constant. A constant viewing angle may indicate that there is a high likelihood that vehicle 200 and pedestrian 1520 will collide if they continue to move linearly at substantially constant speeds in their respective directions that cross each other. When processing unit 1 10 determines that the horizontal coordinate of the focus of expansion of pedestrian 1520. as indicated by the vertical line 1530, stays around the center in the horizontal dimension of pedestrian 1520, processing unit 1 10 may cause vehicle 200 to brake to avoid the collision as vehicle 200 approaches pedestrian 1520. If processing unit 1 10 detects that the line 1530 does not stay around the center in the horizontal dimension of pedestrian 1520 (e.g., moves away from the center), which may indicate that the viewing angle is not constant, processing unit 1 10 may not cause vehicle 200 to brake because the likelihood to have a collision may not be high.

[0183] FIG. 16 is a flowchart showing an exemplary process for detecting an object in front of vehicle 200. Method 1600 may include acquiring a plurality of images of an area in front of vehicle 200 (step 1610). For example, at least one of image capture devices 122-126 may capture the plurality of images of the area in front of vehicle 200 at a predetermined interval, such as, for example, 0.01 second, 0.02 second, etc. Method 1600 may also include receiving the plurality of images (step 1620). For example, processing unit 1 10 may receive the plurality of images from at least one of image capture devices 122- 126. Method 1600 may also include comparing a first image to a second image to determine a focus of expansion of a crossing object moving in a direction that crosses a travel direction of the vehicle (step 1630). For example, processing unit 1 10 may compare the first and second images to obtain optical flow information, which includes displacement vectors relating to each first pixel in the first image and a corresponding second pixel in the second image. Processing unit 1 10 may determine the focus of expansion based on the optical flow information.

IPR2025-00944
Tesla EX1002 Page 1578

[0184] Method 1600 may also include determine whether a horizontal coordinate of the focus of expansion stays around a center in the horizontal dimension of the crossing object during a period of time (step 1640). Processing unit 1 10 may compare the horizontal coordinate of the focus of expansion of the crossing object with the horizontal center (i.e., horizontal coordinate) of the crossing object in the image, and determine whether they are the same or substantially the same (e.g., within a small error, such as one or two pixels). When processing unit 1 10 determines that the horizontal coordinate stays around the center in the horizontal dimension of the crossing object during the period of time (Yes, step 1640), processing unit 1 10 may cause vehicle 200 to brake using, e.g., braking system 230, to avoid collision with the crossing object (step 1650). When processing unit 1 10 determines that the horizontal coordinate does not stay within the crossing object during the period of time (No, step 1640), processing unit 1 10 may executes steps 1610- 1630.

[0185] The situation shown in FIG. 15A assumes that vehicle 200 and crossing object 1500 may be represented by points. In reality, when camera is installed on vehicle 200, which may not be treated as a point, the method discussed above for detecting a crossing-to-impact object may be generalized to take into account the sizes of vehicle 200 and the crossing object 1500. The generalized method may provide a visual indication of potential collision with a target object. For example, an image space of a target object may be extrapolated using any suitable extrapolation method to predict a future location of that target object in an image space, in a future time, relative to vehicle 200, thereby showing a boundary for avoiding collision with the target object.

[0186] As shown in FIG. 17, as vehicle 200 drives along the road where object vehicles 1710, 1720, 1730 are parked, images of vehicles 1710, 1720, 1730 may be extrapolated in time to predict the future position of vehicles 1710, 1720, and 1730 relative to vehicle 200 in the image space. For example, an image of a portion of vehicle 1710 within a box 1750 may be extrapolated in time to produce a predicted relative position of vehicle 1710, as indicated by box 1760. Similar extrapolation may be performed based on images of vehicles 1720 and 1730. Lines 1770 and 1780 may be superimposed on the real-time image captured by one of the image capture devices 122-126 to show the predicted boundary for avoiding potential collision with parked vehicles 1710-1730. To further assist the driver to prevent collision, in some embodiments, one or more lines 1740 representing a position of the front bumper of vehicle 200 may be superimposed in the live, real-time image. Thus, when lines 1740 representing the front bumper do not touch the boundary lines 1770 and/or 1780, it provides an indication to the driver that vehicle 200 is at a safe distance from vehicles 1710-1730 to avoid collision. If lines 1740 touch the boundary lines 1770 and/or 1780, it may indicate that vehicle 200 has a potential to collide with at least one of parked vehicle 1710, 1720, or 1730, and system 100 may provide a visual and/or audio warning to the driver.

[0187] The foregoing description has been presented for purposes of illustration. It is not exhaustive and is not limited to the precise forms or embodiments disclosed. Modifications and adaptations will be apparent to those skilled in the art from consideration of the specification and practice of the disclosed embodiments. Additionally, although aspects of the disclosed embodiments are described as being stored in memory, one skilled in the art will appreciate that these aspects can also be stored on other types of computer readable media, such as secondary storage devices, for example, hard disks or CD ROM, or other forms of RAM or ROM, USB media, DVD,

Tesla EX1002 Page 1579 IPR2025-00944

Blu-ray, or other optical drive media. [0188] Computer programs based on the written description and disclosed methods are within the skill of an experienced developer. The various programs or program modules can be created using any of the techniques known to one skilled in the art or can be designed in connection with existing software. For example, program sections or program modules can be designed in or by means of .Net Framework, .Net Compact Framework (and related languages, such as Visual Basic, C, etc.), Java, C++, Objective-C, HTML, HTML/ AJ AX combinations, XML, or HTML with included Java applets.

[0189] Moreover, while illustrative embodiments have been described herein, the scope of any and all embodiments having equivalent elements, modifications, omissions, combinations (e.g., of aspects across various embodiments), adaptations and/or alterations as would be appreciated by those skilled in the art based on the present disclosure. The limitations in the claims are to be interpreted broadly based on the language employed in the claims and not limited to examples described in the present specification or during the prosecution of the application. The examples are to be construed as non-exclusive.

Furthermore, the steps of the disclosed methods may be modified in any manner, including by reordering steps and/or inserting or deleting steps. It is intended, therefore, that the specification and examples be considered as illustrative only, with a true scope and spirit being indicated by the following claims and their full scope of equivalents.

Espacenet

Claims: EP3152704 (A2) — 2017-04-12

SYSTEMS AND METHODS FOR DETECTING AN OBJECT

Claims not available for EP3152704 (A2)
Claims of corresponding document: WO2015186002 (A2)

A high quality text as facsimile in your desired language may be available amongst the following family members:

CN106796648 (A) US2017098131 (A1) WO2015186002 (A2) CN112528737 (A) US2020234065 (A1) US2021042539 (A1)

- Original claims
- Claims tree

The EPO does not accept any responsibility for the accuracy of data and information originating from other authorities than the EPO; in particular, the EPO does not guarantee that they are complete, up-to-date or fit for specific purposes.

WHAT IS CLAIMED IS:

1. An object detecting system for a vehicle, comprising:
at least one image capture device configured to acquire a plurality of images of an area in front of the vehicle;
a data interface; and
at least one processing device programmed to:
receive the plurality of images via the data interface;
compare a first image to a second image to determine displacement vectors between pixels of the first image and pixels of the second image;
search for a region of coherent expansion that is a set of pixels in at least one of the first image and the second image, for which there exists a common focus of expansion and a common scale magnitude such that the set of pixels satisfy a relationship between pixel positions, displacement vectors, the common focus of expansion, and the common scale magnitude; and
identify presence of a substantially upright object in at least one of the first image and the second image based on the set of pixels.

2. The object detecting system of claim 1 , wherein the region of coherent expansion represents a substantially upright object that has a low time to collision with the vehicle.

3. The object detecting system of claim 1 , wherein the common focus of expansion represents a converging point for a plurality of displacement vectors associated with the set of pixels.

4. The object detecting system of claim 3, wherein the scale magnitude associated with a pixel in the first image is defined as a magnitude of a displacement vector associated with the pixel divided by a distance from the pixel to the focus of expansion.

5. The object detecting system of claim 4, wherein the at least one processing device is further programmed to:
calculate a plurality of combinations of scale magnitudes and focus of expansions for a plurality of pixels from at least one of the first image and the second image, the plurality of pixels include the set of pixels; and
determine the common focus of expansion and the common scale magnitude based on a distribution of the plurality of combinations of scale magnitudes and focus of expansions.
6. The object detecting system of claim 5, wherein determining the common focus of expansion and the common scale magnitude based on the distribution of the plurality of focuses of expansion includes: determining the common focus of expansion and the common scale magnitude by determining that a density of distribution around the pair of common focus of expansion and the common scale magnitude is higher than a predetermined threshold.
7. A vehicle, comprising:
a body;
at least one image capture device configured to acquire a plurality of images of an area in front of the vehicle; a data interface; and
at least one processing device programmed to:
receive the plurality of images via the data interface;
compare a first image to a second image to determine displacement vectors between pixels of the first image and pixels of the second image;
search for a region of coherent expansion that is a set of pixels in at least one of the first image and the second image, for which there exists a common focus of expansion and a common scale magnitude such that the set of pixels satisfy a relationship between pixel positions, displacement vectors, the common focus of expansion, and the common scale magnitude; and
identify presence of a substantially upright object in at least one of the first image and the second image based on the set of pixels.
8. A method for detecting an object in front of a vehicle, comprising:
acquiring, via at least one image capture device, a plurality of images of an area in front of the vehicle;
receiving, via a processing device, the plurality of images;
comparing, via the processing device, a first image to a second image to determine displacement vectors between pixels of the first image and pixels of the second image;
searching for, via the processing device, a region of coherent expansion that is a set of pixels in at least one of the first image and the second image, for which there exists a common focus of expansion and a common scale magnitude such that the set of pixels satisfy a relationship between pixel positions, displacement vectors, the common focus of expansion, and the common scale magnitude; and
identifying, via the processing device, presence of a substantially upright object in at least one of the first image and the second image based on the set of pixels.
9. The method of claim 8, further comprising:
calculating a plurality of combinations of scale magnitudes and focus of expansions for a plurality of pixels from at least one of the first image and the second image, the plurality of pixels include the set of pixels; and
determining the common focus of expansion and the common scale magnitude based on a distribution of the plurality of combinations of scale magnitudes and focus of expansions.
10. The method of claim 8, wherein determining the common focus of expansion and the common scale magnitude based on the distribution of the plurality of focuses of expansion includes:

Tesla EX1002 Page 1582 IPR2025-00944

determining the common focus of expansion and the common scale magnitude by determining that a density of distribution around the pair of common focus of expansion and the common scale magnitude is higher than a predetermined threshold.

1 1. An object detecting system for a vehicle, comprising:

at least one image capture device configured to acquire a plurality of images of an area in front of the vehicle; a data interface; and

at least one processing device programmed to:

receive the plurality of images via the data interface;

compare a first image to a second image, and based on the comparison, determine a focus of expansion of a crossing object moving in a direction that crosses a travel direction of the vehicle;

determine whether a horizontal coordinate of the focus of expansion stays around a center in the horizontal dimension of the crossing object during a period of time; and

cause the vehicle to brake based on the determination that the horizontal coordinate of the focus of expansion stays around the center in the horizontal dimension of the crossing object during the period of time.

12. The object detecting system of claim 1 1 , wherein the crossing object is an object that has a low time to collision with the vehicle, and

wherein when the crossing object moves linearly at a substantially constant or piecewise constant speed, and the vehicle also moves linearly at a substantially constant or piecewise constant speed, a viewing angle from a camera of the vehicle remains constant during the period of time.

13. The object detecting system of claim 12, wherein the at least one processing device is further programmed to:

determine that the viewing angle is constant based on a detection that the crossing object appears fixed on a horizontal axis of an image plane, and expands in a direction perpendicular to the image plane. 14. The object detecting system of claim 13, wherein the at least one processing device is further programmed to:

determine whether the crossing object will impact the camera based on either a determination that the viewing angle is constant or a determination that the horizontal coordinate of the focus of expansion of the crossing object stays around the center in the horizontal dimension.

15. A vehicle, comprising:

a body;

at least one image capture device configured to acquire a plurality of images of an area in front of the vehicle;

a data interface; and

at least one processing device programmed to:

receive the plurality of images via the data interface;

compare a first image to a second image, and based on the comparison, determine a focus of expansion of a crossing object moving in a direction that crosses a travel direction of the vehicle;

determine whether a horizontal coordinate of the focus of expansion stays around a center in the horizontal dimension of the crossing object during a period of time; and

cause the vehicle to brake based on the determination that the horizontal coordinate of the focus of expansion stays around a center in the horizontal dimension of the crossing object during the period of time.

16. A method for detecting an object in front of a vehicle, comprising:

acquiring, via at least one image capture device, a plurality of images of an area in front of the vehicle;

receiving, via a processing device, the plurality of images;

comparing, via the processing device, a first image to a second image, and based on

the comparison, determining a focus of expansion of a crossing object moving in a direction that crosses a travel direction of the vehicle;
determining, via the processing device, whether a horizontal coordinate of the focus of expansion stays around a center in the horizontal dimension of the crossing object during a period of time; and causing, via the processing device, the vehicle to brake based on the determination that the horizontal coordinate of the focus of expansion stays around a center in the horizontal dimension of the crossing object during the period of time.

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Veröffentlichungsnummer:
(11) Publication number: **EP 3 152 704 A0**
(11) Numéro de publication:

Internationale Anmeldung veröffentlicht durch die Weltorganisation für geistiges Eigentum unter der Nummer:

WO 2015/186002 (Art. 153(3) EPÜ).

International application published by the World Intellectual Property Organization under number:

WO 2015/186002 (Art. 153(3) EPC).

Demande internationale publiée par l'Organisation Mondiale de la Propriété Intellectuelle sous le numéro:

WO 2015/186002 (art. 153(3) CBE).

Espacenet

Bibliographic data: EP3154835 (A1) — 2017-04-19

TOP-DOWN REFINEMENT IN LANE MARKING NAVIGATION

Inventor(s): FERENCZ ANDRAS [IL]; REISMAN PINCHAS [IL] ± (FERENCZ, Andras, ; REISMAN, Pinchas)

Applicant(s): MOBILEYE VISION TECHNOLOGIES LTD [IL] ± (Mobileye Vision Technologies Ltd)

Classification: **- international:** ***B60W30/12; G06K9/00; G06K9/32; G06K9/62; G06T7/20***

- cooperative: **B60W30/12 (EP, US)**; **G01C21/3602 (US)**; **G05D1/0251 (EP, US)**; **G05D1/0253 (EP, US)**; **G06K9/00798 (EP, US)**; **G06K9/3241 (EP, US)**; **G06K9/6203 (EP, US)**; G05D2201/0213 (EP, US)

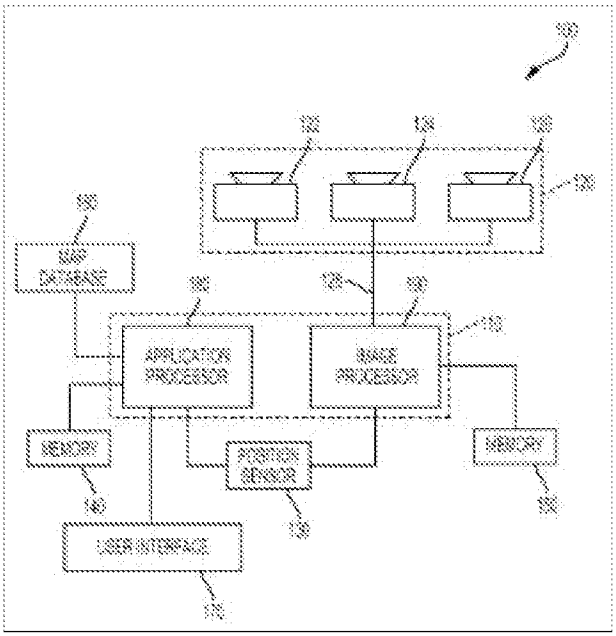
Application number: EP20150738138 20150610 Global Dossier

Priority number (s): US201462010003P 20140610 ; US201562173216P 20150609 ; WO2015IL50590 20150610

Also published as: US10317231 (B2) US10753758 (B2) US2015354976 (A1) US2020003573 (A1) WO2015189847 (A1)

Abstract not available for EP3154835 (A1)
Abstract of corresponding document: US2015354976 (A1)

Systems and methods use cameras to provide autonomous navigation features. In one implementation, top-down refinement in lane marking navigation is provided. The system may include one or more memories storing instructions and one or more processors configured to execute the instructions to cause the system to receive from one or more cameras one or more images of a roadway in a vicinity of a vehicle, the roadway comprising a lane marking comprising a dashed line, update a model of the lane marking based on odometry of the one or more cameras relative to the roadway, refine the updated model of the lane marking based on an appearance of dashes derived from the received one or more images and a spacing between dashes derived from the received one or more images, and cause one or more navigational responses in the vehicle based on the refinement of the updated model.



IPR2025-00944
Tesla EX1002 Page 1587

Espacenet

Description: EP3154835 (A1) — 2017-04-19

TOP-DOWN REFINEMENT IN LANE MARKING NAVIGATION

Description not available for EP3154835 (A1)
Description of corresponding document: US2015354976 (A1)

A high quality text as facsimile in your desired language may be available amongst the following family members:

US2015354976 (A1) WO2015189847 (A1) US2020003573 (A1)

The EPO does not accept any responsibility for the accuracy of data and information originating from other authorities than the EPO; in particular, the EPO does not guarantee that they are complete, up-to-date or fit for specific purposes.

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of U.S. Provisional Patent Application No. 62/010,003, filed Jun. 10, 2014, and U.S. Provisional Patent Application No. 62/173,216, filed Jun. 9, 2015, the entireties of which are incorporated herein by reference.

FIELD OF THE DISCLOSURE

[0002] This relates generally to autonomous driving and/or driver assist technology and, more specifically, to systems and methods that use cameras to provide autonomous driving and/or driver assist technology features.

BACKGROUND

[0003] As technology continues to advance, the goal of a fully autonomous vehicle that is capable of navigating on roadways is on the horizon. Primarily, an autonomous vehicle may be able to identify its environment and navigate without input from a human operator. Autonomous vehicles may also take into account a variety of factors and make appropriate decisions based on those factors to safely and accurately reach an intended destination. For example, various objects—such as other vehicles and pedestrians—are encountered when a vehicle typically travels a roadway. Autonomous driving systems may recognize these objects in a vehicle's environment and take appropriate and timely action to avoid collisions. Additionally, autonomous driving systems may identify other indicators—such as traffic signals, traffic signs, and lane

markings—that regulate vehicle movement (e.g., when the vehicle must stop and may go, a speed at which the vehicle must not exceed, where the vehicle must be positioned on the va-452177 roadway, etc.). Autonomous driving systems may need to determine when a vehicle should change lanes, turn at intersections, change roadways, etc. As is evident from these examples, many factors may need to be addressed in order to provide an autonomous vehicle that is capable of navigating safely and accurately.

SUMMARY

[0004] Embodiments consistent with the present disclosure provide systems and methods for autonomous vehicle navigation. The disclosed embodiments may use cameras to provide autonomous vehicle navigation features. For example, consistent with the disclosed embodiments, the disclosed systems may include one, two, or more cameras that monitor the environment of a vehicle and cause a navigational response based on an analysis of images captured by one or more of the cameras.

[0005] Consistent with a disclosed embodiment, top-down refinement in lane marking navigation is provided for a computer system. The system may include one or more memories storing instructions and one or more processors. The one or more processors may be configured to execute the instructions to cause the system to receive from one or more cameras one or more images of a roadway in a vicinity of a vehicle, the roadway comprising a lane marking comprising a dashed line, update a model of the lane marking based on odometry of the one or more cameras relative to the roadway, refine the updated model of the lane marking based on an appearance of dashes derived from the received one or more images and a spacing between dashes derived from the received one or more images, and cause one or more navigational responses in the vehicle based on the refinement of the updated model.

[0006] Consistent with another disclosed embodiment, top-down refinement in lane marking navigation is provided for a vehicle. The vehicle may include a body, one or more cameras configured to acquire one or more images of a roadway in a vicinity of a vehicle, the roadway comprising a lane marking comprising a dashed line, one or more memories storing instructions, and one or more processors. The one or more processors may be configured to execute the instructions to cause the vehicle to receive from one or more cameras the one or more acquired images of the roadway, update a model of the lane marking based on odometry of the one or more cameras relative to the roadway, refine the updated model of the lane marking based on an appearance of dashes derived from the received one or more images and a spacing between dashes derived from the received one or more images, and cause one or more navigational responses in the vehicle based on the refinement of the updated model.

[0007] Consistent with another disclosed embodiment, a method of top-down refinement in lane marking navigation is provided. The method may include receiving, by one or more processors, from one or more cameras one or more images of a roadway in a vicinity of a vehicle, the roadway comprising a lane marking comprising a dashed line, updating, by one or more processors, a model of the lane marking based on odometry of the one or more cameras relative to the roadway, refining, by one or more processors, the updated model of the lane marking based on an appearance of dashes derived from the received one or more images and a spacing between dashes derived from the received one or more images, and causing, by one or more

IPR2025-00944
Tesla EX1002 Page 1589

processors, one or more navigational responses in the vehicle based on the refinement of the updated model.

BRIEF DESCRIPTION OF THE DRAWINGS

[0008] FIG. 1 is a diagrammatic representation of an exemplary system consistent with the disclosed embodiments.

[0009] FIG. 2A is a diagrammatic side view representation of an exemplary vehicle including a system consistent with the disclosed embodiments.

[0010] FIG. 2B is a diagrammatic top view representation of the vehicle and system shown in FIG. 2A consistent with the disclosed embodiments.

[0011] FIG. 2C is a diagrammatic top view representation of another embodiment of a vehicle including a system consistent with the disclosed embodiments.

[0012] FIG. 2D is a diagrammatic top view representation of yet another embodiment of a vehicle including a system consistent with the disclosed embodiments.

[0013] FIG. 2E is a diagrammatic top view representation of yet another embodiment of a vehicle including a system consistent with the disclosed embodiments.

[0014] FIG. 2F is a diagrammatic representation of exemplary vehicle control systems consistent with the disclosed embodiments.

[0015] FIG. 3A is a diagrammatic representation of an interior of a vehicle including a rearview mirror and a user interface for a vehicle imaging system consistent with the disclosed embodiments.

[0016] FIG. 3B is an illustration of an example of a camera mount that is configured to be positioned behind a rearview mirror and against a vehicle windshield consistent with the disclosed embodiments.

[0017] FIG. 3C is an illustration of the camera mount shown in FIG. 3B from a different perspective consistent with the disclosed embodiments.

[0018] FIG. 3D is an illustration of an example of a camera mount that is configured to be positioned behind a rearview mirror and against a vehicle windshield consistent with the disclosed embodiments.

[0019] FIG. 4 is an exemplary block diagram of a memory configured to store instructions for performing one or more operations consistent with the disclosed embodiments.

[0020] FIG. 5A is a flowchart showing an exemplary process for causing one or more navigational responses based on monocular image analysis consistent with disclosed embodiments.

[0021] FIG. 5B is a flowchart showing an exemplary process for detecting one or more vehicles and/or pedestrians in a set of images consistent with the disclosed embodiments.

IPR2025-00944 Tesla EX1002 Page 1590

[0022] FIG. 5C is a flowchart showing an exemplary process for detecting road marks and/or lane geometry information in a set of images consistent with the disclosed embodiments.

[0023] FIG. 5D is a flowchart showing an exemplary process for detecting traffic lights in a set of images consistent with the disclosed embodiments.

[0024] FIG. 5E is a flowchart showing an exemplary process for causing one or more navigational responses based on a vehicle path consistent with the disclosed embodiments.

[0025] FIG. 5F is a flowchart showing an exemplary process for determining whether a leading vehicle is changing lanes consistent with the disclosed embodiments.

[0026] FIG. 6 is a flowchart showing an exemplary process for causing one or more navigational responses based on stereo image analysis consistent with the disclosed embodiments.

[0027] FIG. 7 is a flowchart showing an exemplary process for causing one or more navigational responses based on an analysis of three sets of images consistent with the disclosed embodiments.

[0028] FIGS. 8A-8D are diagrammatic representations of exemplary lane detection analyses without top-down refinement.

[0029] FIG. 9 is a diagrammatic representation of an exemplary top-down refinement architecture consistent with the disclosed embodiments.

[0030] FIG. 10A is a diagrammatic representation of the memory of an exemplary navigation system consistent with the disclosed embodiments.

[0031] FIG. 10B is a flowchart of an exemplary process for navigating a vehicle using top-down refinement consistent with the disclosed embodiments.

[0032] FIG. 11A is a flowchart of an exemplary process for implementing top-down refinement consistent with the disclosed embodiments.

[0033] FIGS. 11B-11N are diagrammatic representations of a top-down refinement implementation consistent with the disclosed embodiments.

DETAILED DESCRIPTION

[0034] The following detailed description refers to the accompanying drawings. Wherever possible, the same reference numbers are used in the drawings and the following description to refer to the same or similar parts. While several illustrative embodiments are described herein, modifications, adaptations and other implementations are possible. For example, substitutions, additions or modifications may be made to the components illustrated in the drawings, and the illustrative methods described herein may be modified by substituting, reordering, removing, or adding steps to the disclosed methods. Accordingly, the following detailed description is not limited to the disclosed embodiments and examples. Instead, the proper scope is

defined by the appended claims.

[0035] Disclosed embodiments provide systems and methods that use cameras to provide autonomous navigation and/or driver assist technology features. Driver assist technology refers to any suitable technology to assist drivers in the navigation of their vehicles, such as LKA (lane keeping assist) and LDW (lane departure warning), as opposed to fully autonomous driving. In various embodiments, the system may include one, two or more cameras that monitor the environment of a vehicle. In one embodiment, the system may provide top-down refinement in lane marking navigation.

[0036] FIG. 1 is a block diagram representation of a system 100 consistent with the exemplary disclosed embodiments. System 100 may include various components depending on the requirements of a particular implementation. In some embodiments, system 100 may include a processing unit 110, an image acquisition unit 120, a position sensor 130, one or more memory units 140, 150, a map database 160, and a user interface 170. Processing unit 110 may include one or more processing devices. In some embodiments, processing unit 110 may include an application processor 180, an image processor 190, or any other suitable processing device. Similarly, image acquisition unit 120 may include any number of image acquisition devices and components depending on the requirements of a particular application. In some embodiments, image acquisition unit 120 may include one or more image capture devices (e.g., cameras), such as image capture device 122, image capture device 124, and image capture device 126. System 100 may also include a data interface 128 communicatively connecting processing unit 110 to image acquisition device 120. For example, data interface 128 may include any wired and/or wireless link or links for transmitting image data acquired by image accusation device 120 to processing unit 110.

[0037] Both application processor 180 and image processor 190 may include various types of processing devices. For example, either or both of application processor 180 and image processor 190 may include one or more microprocessors, preprocessors (such as image preprocessors), graphics processors, central processing units (CPUs), support circuits, digital signal processors, integrated circuits, memory, or any other types of devices suitable for running applications and for image processing and analysis. In some embodiments, application processor 180 and/or image processor 190 may include any type of single or multi-core processor, mobile device microcontroller, central processing unit, etc. Various processing devices may be used, including, for example, processors available from manufacturers such as Intel®, AMD®, etc. and may include various architectures (e.g., x86 processor, ARM®, etc.).

[0038] In some embodiments, application processor 180 and/or image processor 190 may include any of the EyeQ series of processor chips available from Mobileye®. These processor designs each include multiple processing units with local memory and instruction sets. Such processors may include video inputs for receiving image data from multiple image sensors and may also include video out capabilities. In one example, the EyeQ2® uses 90 nm-micron technology operating at 332 Mhz. The EyeQ2® architecture has two floating point, hyper-thread 32-bit RISC CPUs (MIPS32® 34K® cores), five Vision Computing Engines (VCE), three Vector Microcode Processors (VMP®), Denali 64-bit Mobile DDR Controller, 128-bit internal Sonics Interconnect, dual 16-bit Video input and 18-bit Video output controllers, 16 channels DMA and several peripherals. The MIPS34K CPU manages the five VCEs, three VMP™ and the DMA, the second MIPS34K CPU and the multi-channel DMA as well as

Tesla EX1002 Page 1592 IPR2025-00944

the other peripherals. The five VCEs, three VMP® and the MIPS34K CPU can perform intensive vision computations required by multi-function bundle applications. In another example, the EyeQ3®, which is a third generation processor and is six times more powerful that the EyeQ2®, may be used in the disclosed embodiments.

[0039] While FIG. 1 depicts two separate processing devices included in processing unit 110, more or fewer processing devices may be used. For example, in some embodiments, a single processing device may be used to accomplish the tasks of application processor 180 and image processor 190. In other embodiments, these tasks may be performed by more than two processing devices.

[0040] Processing unit 110 may comprise various types of devices. For example, processing unit 110 may include various devices, such as a controller, an image preprocessor, a central processing unit (CPU), support circuits, digital signal processors, integrated circuits, memory, or any other types of devices for image processing and analysis. The image preprocessor may include a video processor for capturing, digitizing and processing the imagery from the image sensors. The CPU may comprise any number of microcontrollers or microprocessors. The support circuits may be any number of circuits generally well known in the art, including cache, power supply, clock and input-output circuits. The memory may store software that, when executed by the processor, controls the operation of the system. The memory may include databases and image processing software. The memory may comprise any number of random access memories, read only memories, flash memories, disk drives, optical storage, tape storage, removable storage and other types of storage. In one instance, the memory may be separate from the processing unit 110. In another instance, the memory may be integrated into the processing unit 110.

[0041] Each memory 140, 150 may include software instructions that when executed by a processor (e.g., application processor 180 and/or image processor 190), may control operation of various aspects of system 100. These memory units may include various databases and image processing software. The memory units may include random access memory, read only memory, flash memory, disk drives, optical storage, tape storage, removable storage and/or any other types of storage. In some embodiments, memory units 140, 150 may be separate from the application processor 180 and/or image processor 190. In other embodiments, these memory units may be integrated into application processor 180 and/or image processor 190.

[0042] Position sensor 130 may include any type of device suitable for determining a location associated with at least one component of system 100. In some embodiments, position sensor 130 may include a GPS receiver. Such receivers can determine a user position and velocity by processing signals broadcasted by global positioning system satellites. Position information from position sensor 130 may be made available to application processor 180 and/or image processor 190.

[0043] User interface 170 may include any device suitable for providing information to or for receiving inputs from one or more users of system 100. In some embodiments, user interface 170 may include user input devices, including, for example, a touchscreen, microphone, keyboard, pointer devices, track wheels, cameras, knobs, buttons, etc. With such input devices, a user may be able to provide information inputs or commands to system 100 by typing instructions or information, providing voice commands, selecting menu options on a screen using buttons, pointers, or eye-tracking capabilities, or through any other suitable techniques for communicating information to

Tesla EX1002 Page 1593 IPR2025-00944

system 100.

[0044] User interface 170 may be equipped with one or more processing devices configured to provide and receive information to or from a user and process that information for use by, for example, application processor 180. In some embodiments, such processing devices may execute instructions for recognizing and tracking eye movements, receiving and interpreting voice commands, recognizing and interpreting touches and/or gestures made on a touchscreen, responding to keyboard entries or menu selections, etc. In some embodiments, user interface 170 may include a display, speaker, tactile device, and/or any other devices for providing output information to a user.

[0045] Map database 160 may include any type of database for storing map data useful to system 100. In some embodiments, map database 160 may include data relating to the position, in a reference coordinate system, of various items, including roads, water features, geographic features, businesses, points of interest, restaurants, gas stations, etc. Map database 160 may store not only the locations of such items, but also descriptors relating to those items, including, for example, names associated with any of the stored features. In some embodiments, map database 160 may be physically located with other components of system 100. Alternatively or additionally, map database 160 or a portion thereof may be located remotely with respect to other components of system 100 (e.g., processing unit 110). In such embodiments, information from map database 160 may be downloaded over a wired or wireless data connection to a network (e.g., over a cellular network and/or the Internet, etc.).

[0046] Image capture devices 122, 124, and 126 may each include any type of device suitable for capturing at least one image from an environment. Moreover, any number of image capture devices may be used to acquire images for input to the image processor. Some embodiments may include only a single image capture device, while other embodiments may include two, three, or even four or more image capture devices. Image capture devices 122, 124, and 126 will be further described with reference to FIGS. 2B-2E, below.

[0047] System 100, or various components thereof, may be incorporated into various different platforms. In some embodiments, system 100 may be included on a vehicle 200, as shown in FIG. 2A. For example, vehicle 200 may be equipped with a processing unit 110 and any of the other components of system 100, as described above relative to FIG. 1. While in some embodiments vehicle 200 may be equipped with only a single image capture device (e.g., camera), in other embodiments, such as those discussed in connection with FIGS. 2B-2E, multiple image capture devices may be used. For example, either of image capture devices 122 and 124 of vehicle 200, as shown in FIG. 2A, may be part of an ADAS (Advanced Driver Assistance Systems) imaging set.

[0048] The image capture devices included on vehicle 200 as part of the image acquisition unit 120 may be positioned at any suitable location. In some embodiments, as shown in FIGS. 2A-2E and 3A-3C, image capture device 122 may be located in the vicinity of the rearview mirror. This position may provide a line of sight similar to that of the driver of vehicle 200, which may aid in determining what is and is not visible to the driver. Image capture device 122 may be positioned at any location near the rearview mirror, but placing image capture device 122 on the driver side of the mirror may further aid in obtaining images representative of the driver's field of view and/or line of sight.

Tesla EX1002 Page 1594 IPR2025-00944

[0049] Other locations for the image capture devices of image acquisition unit 120 may also be used. For example, image capture device 124 may be located on or in a bumper of vehicle 200. Such a location may be especially suitable for image capture devices having a wide field of view. The line of sight of bumper-located image capture devices can be different from that of the driver and, therefore, the bumper image capture device and driver may not always see the same objects. The image capture devices (e.g., image capture devices 122, 124, and 126) may also be located in other locations. For example, the image capture devices may be located on or in one or both of the side mirrors of vehicle 200, on the roof of vehicle 200, on the hood of vehicle 200, on the trunk of vehicle 200, on the sides of vehicle 200, mounted on, positioned behind, or positioned in front of any of the windows of vehicle 200, and mounted in or near light figures on the front and/or back of vehicle 200, etc.

[0050] In addition to image capture devices, vehicle 200 may include various other components of system 100. For example, processing unit 110 may be included on vehicle 200 either integrated with or separate from an engine control unit (ECU) of the vehicle. Vehicle 200 may also be equipped with a position sensor 130, such as a GPS receiver and may also include a map database 160 and memory units 140 and 150.

[0051] FIG. 2A is a diagrammatic side view representation of an exemplary vehicle imaging system consistent with the disclosed embodiments. FIG. 2B is a diagrammatic top view illustration of the embodiment shown in FIG. 2A. As illustrated in FIG. 2B, the disclosed embodiments may include a vehicle 200 including in its body a system 100 with a first image capture device 122 positioned in the vicinity of the rearview mirror and/or near the driver of vehicle 200, a second image capture device 124 positioned on or in a bumper region (e.g., one of bumper regions 210) of vehicle 200, and a processing unit 110.

[0052] As illustrated in FIG. 2C, image capture devices 122 and 124 may both be positioned in the vicinity of the rearview mirror and/or near the driver of vehicle 200. Additionally, while two image capture devices 122 and 124 are shown in FIGS. 2B and 2C, it should be understood that other embodiments may include more than two image capture devices. For example, in the embodiments shown in FIGS. 2D and 2E, first, second, and third image capture devices 122, 124, and 126, are included in the system 100 of vehicle 200.

[0053] As illustrated in FIG. 2D, image capture device 122 may be positioned in the vicinity of the rearview mirror and/or near the driver of vehicle 200, and image capture devices 124 and 126 may be positioned on or in a bumper region (e.g., one of bumper regions 210) of vehicle 200. And as shown in FIG. 2E, image capture devices 122, 124, and 126 may be positioned in the vicinity of the rearview mirror and/or near the driver seat of vehicle 200. The disclosed embodiments are not limited to any particular number and configuration of the image capture devices, and the image capture devices may be positioned in any appropriate location within and/or on vehicle 200.

[0054] It is to be understood that the disclosed embodiments are not limited to vehicles and could be applied in other contexts, including portable devices such as smartphones and wearable accessories such as helmets and self-contained mountable units. It is also to be understood that disclosed embodiments are not limited to a particular type of vehicle 200 and may be applicable to all types of vehicles including automobiles, trucks, trailers, motorcycles, bicycles, self-balancing transport devices and other types

Tesla EX1002 Page 1595 IPR2025-00944

of vehicles.

[0055] The first image capture device 122 may include any suitable type of image capture device. Image capture device 122 may include an optical axis. In one instance, the image capture device 122 may include an Aptina M9V024 WVGA sensor with a global shutter. In other embodiments, image capture device 122 may provide a resolution of 1280×960 pixels and may include a rolling shutter. Image capture device 122 may include various optical elements. In some embodiments one or more lenses may be included, for example, to provide a desired focal length and field of view for the image capture device. In some embodiments, image capture device 122 may be associated with a 6 mm lens or a 12 mm lens. In some embodiments, image capture device 122 may be configured to capture images having a desired field-of-view (FOV) 202, as illustrated in FIG. 2D. For example, image capture device 122 may be configured to have a regular FOV, such as within a range of 40 degrees to 56 degrees, including a 46 degree FOV, 50 degree FOV, 52 degree FOV, or greater. Alternatively, image capture device 122 may be configured to have a narrow FOV in the range of 23 to 40 degrees, such as a 28 degree FOV or 36 degree FOV. In addition, image capture device 122 may be configured to have a wide FOV in the range of 100 to 180 degrees. In some embodiments, image capture device 122 may include a wide angle bumper camera or one with up to a 180 degree FOV.

[0056] The first image capture device 122 may acquire a plurality of first images relative to a scene associated with the vehicle 200. Each of the plurality of first images may be acquired as a series of image scan lines, which may be captured using a rolling shutter. Each scan line may include a plurality of pixels.

[0057] The first image capture device 122 may have a scan rate associated with acquisition of each of the first series of image scan lines. The scan rate may refer to a rate at which an image sensor can acquire image data associated with each pixel included in a particular scan line.

[0058] Image capture devices 122, 124, and 126 may contain any suitable type and number of image sensors, including CCD sensors or CMOS sensors, for example. In one embodiment, a CMOS image sensor may be employed along with a rolling shutter, such that each pixel in a row is read one at a time, and scanning of the rows proceeds on a row-by-row basis until an entire image frame has been captured. In some embodiments, the rows may be captured sequentially from top to bottom relative to the frame.

[0059] The use of a rolling shutter may result in pixels in different rows being exposed and captured at different times, which may cause skew and other image artifacts in the captured image frame. On the other hand, when the image capture device 122 is configured to operate with a global or synchronous shutter, all of the pixels may be exposed for the same amount of time and during a common exposure period. As a result, the image data in a frame collected from a system employing a global shutter represents a snapshot of the entire FOV (such as FOV 202) at a particular time. In contrast, in a rolling shutter application, each row in a frame is exposed and data is capture at different times. Thus, moving objects may appear distorted in an image capture device having a rolling shutter. This phenomenon will be described in greater detail below.

[0060] The second image capture device 124 and the third image capturing device 126

IPR2025-00944
Tesla EX1002 Page 1596

may be any type of image capture device. Like the first image capture device 122, each of image capture devices 124 and 126 may include an optical axis. In one embodiment, each of image capture devices 124 and 126 may include an Aptina M9V024 WVGA sensor with a global shutter. Alternatively, each of image capture devices 124 and 126 may include a rolling shutter. Like image capture device 122, image capture devices 124 and 126 may be configured to include various lenses and optical elements. In some embodiments, lenses associated with image capture devices 124 and 126 may provide FOVs (such as FOVs 204 and 206) that are the same as, or narrower than, a FOV (such as FOV 202) associated with image capture device 122. For example, image capture devices 124 and 126 may have FOVs of 40 degrees, 30 degrees, 26 degrees, 23 degrees, 20 degrees, or less.

[0061] Image capture devices 124 and 126 may acquire a plurality of second and third images relative to a scene associated with the vehicle 200. Each of the plurality of second and third images may be acquired as a second and third series of image scan lines, which may be captured using a rolling shutter. Each scan line or row may have a plurality of pixels. Image capture devices 124 and 126 may have second and third scan rates associated with acquisition of each of image scan lines included in the second and third series.

[0062] Each image capture device 122, 124, and 126 may be positioned at any suitable position and orientation relative to vehicle 200. The relative positioning of the image capture devices 122, 124, and 126 may be selected to aid in fusing together the information acquired from the image capture devices. For example, in some embodiments, a FOV (such as FOV 204) associated with image capture device 124 may overlap partially or fully with a FOV (such as FOV 202) associated with image capture device 122 and a FOV (such as FOV 206) associated with image capture device 126.

[0063] Image capture devices 122, 124, and 126 may be located on vehicle 200 at any suitable relative heights. In one instance, there may be a height difference between the image capture devices 122, 124, and 126, which may provide sufficient parallax information to enable stereo analysis. For example, as shown in FIG. 2A, the two image capture devices 122 and 124 are at different heights. There may also be a lateral displacement difference between image capture devices 122, 124, and 126, giving additional parallax information for stereo analysis by processing unit 110, for example. The difference in the lateral displacement may be denoted by dx, as shown in FIGS. 2C and 2D. In some embodiments, fore or aft displacement (e.g., range displacement) may exist between image capture devices 122, 124, and 126. For example, image capture device 122 may be located 0.5 to 2 meters or more behind image capture device 124 and/or image capture device 126. This type of displacement may enable one of the image capture devices to cover potential blind spots of the other image capture device(s).

[0064] Image capture devices 122 may have any suitable resolution capability (e.g., number of pixels associated with the image sensor), and the resolution of the image sensor(s) associated with the image capture device 122 may be higher, lower, or the same as the resolution of the image sensor(s) associated with image capture devices 124 and 126. In some embodiments, the image sensor(s) associated with image capture device 122 and/or image capture devices 124 and 126 may have a resolution of 640×480, 1024×768, 1280×960, or any other suitable resolution.

IPR2025-00944
Tesla EX1002 Page 1597

[0065] The frame rate (e.g., the rate at which an image capture device acquires a set of pixel data of one image frame before moving on to capture pixel data associated with the next image frame) may be controllable. The frame rate associated with image capture device 122 may be higher, lower, or the same as the frame rate associated with image capture devices 124 and 126. The frame rate associated with image capture devices 122, 124, and 126 may depend on a variety of factors that may affect the timing of the frame rate. For example, one or more of image capture devices 122, 124, and 126 may include a selectable pixel delay period imposed before or after acquisition of image data associated with one or more pixels of an image sensor in image capture device 122, 124, and/or 126. Generally, image data corresponding to each pixel may be acquired according to a clock rate for the device (e.g., one pixel per clock cycle). Additionally, in embodiments including a rolling shutter, one or more of image capture devices 122, 124, and 126 may include a selectable horizontal blanking period imposed before or after acquisition of image data associated with a row of pixels of an image sensor in image capture device 122, 124, and/or 126. Further, one or more of image capture devices 122, 124, and/or 126 may include a selectable vertical blanking period imposed before or after acquisition of image data associated with an image frame of image capture device 122, 124, and 126.

[0066] These timing controls may enable synchronization of frame rates associated with image capture devices 122, 124, and 126, even where the line scan rates of each are different. Additionally, as will be discussed in greater detail below, these selectable timing controls, among other factors (e.g., image sensor resolution, maximum line scan rates, etc.) may enable synchronization of image capture from an area where the FOV of image capture device 122 overlaps with one or more FOVs of image capture devices 124 and 126, even where the field of view of image capture device 122 is different from the FOVs of image capture devices 124 and 126.

[0067] Frame rate timing in image capture device 122, 124, and 126 may depend on the resolution of the associated image sensors. For example, assuming similar line scan rates for both devices, if one device includes an image sensor having a resolution of 640×480 and another device includes an image sensor with a resolution of 1280×960, then more time will be required to acquire a frame of image data from the sensor having the higher resolution.

[0068] Another factor that may affect the timing of image data acquisition in image capture devices 122, 124, and 126 is the maximum line scan rate. For example, acquisition of a row of image data from an image sensor included in image capture device 122, 124, and 126 will require some minimum amount of time. Assuming no pixel delay periods are added, this minimum amount of time for acquisition of a row of image data will be related to the maximum line scan rate for a particular device. Devices that offer higher maximum line scan rates have the potential to provide higher frame rates than devices with lower maximum line scan rates. In some embodiments, one or more of image capture devices 124 and 126 may have a maximum line scan rate that is higher than a maximum line scan rate associated with image capture device 122. In some embodiments, the maximum line scan rate of image capture device 124 and/or 126 may be 1.25, 1.5, 1.75, or 2 times or more than a maximum line scan rate of image capture device 122.

[0069] In another embodiment, image capture devices 122, 124, and 126 may have the same maximum line scan rate, but image capture device 122 may be operated at a scan rate less than or equal to its maximum scan rate. The system may be configured

Tesla EX1002 Page 1598 IPR2025-00944

such that one or more of image capture devices 124 and 126 operate at a line scan rate that is equal to the line scan rate of image capture device 122. In other instances, the system may be configured such that the line scan rate of image capture device 124 and/or image capture device 126 may be 1.25, 1.5, 1.75, or 2 times or more than the line scan rate of image capture device 122.

[0070] In some embodiments, image capture devices 122, 124, and 126 may be asymmetric. That is, they may include cameras having different fields of view (FOV) and focal lengths. The fields of view of image capture devices 122, 124, and 126 may include any desired area relative to an environment of vehicle 200, for example. In some embodiments, one or more of image capture devices 122, 124, and 126 may be configured to acquire image data from an environment in front of vehicle 200, behind vehicle 200, to the sides of vehicle 200, or combinations thereof.

[0071] Further, the focal length associated with each image capture device 122, 124, and/or 126 may be selectable (e.g., by inclusion of appropriate lenses etc.) such that each device acquires images of objects at a desired distance range relative to vehicle 200. For example, in some embodiments image capture devices 122, 124, and 126 may acquire images of close-up objects within a few meters from the vehicle. Image capture devices 122, 124, and 126 may also be configured to acquire images of objects at ranges more distant from the vehicle (e.g., 25 m, 50 m, 100 m, 150 m, or more). Further, the focal lengths of image capture devices 122, 124, and 126 may be selected such that one image capture device (e.g., image capture device 122) can acquire images of objects relatively close to the vehicle (e.g., within 10 m or within 20 m) while the other image capture devices (e.g., image capture devices 124 and 126) can acquire images of more distant objects (e.g., greater than 20 m, 50 m, 100 m, 150 m, etc.) from vehicle 200.

[0072] According to some embodiments, the FOV of one or more image capture devices 122, 124, and 126 may have a wide angle. For example, it may be advantageous to have a FOV of 140 degrees, especially for image capture devices 122, 124, and 126 that may be used to capture images of the area in the vicinity of vehicle 200. For example, image capture device 122 may be used to capture images of the area to the right or left of vehicle 200 and, in such embodiments, it may be desirable for image capture device 122 to have a wide FOV (e.g., at least 140 degrees).

[0073] The field of view associated with each of image capture devices 122, 124, and 126 may depend on the respective focal lengths. For example, as the focal length increases, the corresponding field of view decreases.

[0074] Image capture devices 122, 124, and 126 may be configured to have any suitable fields of view. In one particular example, image capture device 122 may have a horizontal FOV of 46 degrees, image capture device 124 may have a horizontal FOV of 23 degrees, and image capture device 126 may have a horizontal FOV in between 23 and 46 degrees. In another instance, image capture device 122 may have a horizontal FOV of 52 degrees, image capture device 124 may have a horizontal FOV of 26 degrees, and image capture device 126 may have a horizontal FOV in between 26 and 52 degrees. In some embodiments, a ratio of the FOV of image capture device 122 to the FOVs of image capture device 124 and/or image capture device 126 may vary from 1.5 to 2.0. In other embodiments, this ratio may vary between 1.25 and 2.25.

Tesla EX1002 Page 1599 IPR2025-00944

[0075] System 100 may be configured so that a field of view of image capture device 122 overlaps, at least partially or fully, with a field of view of image capture device 124 and/or image capture device 126. In some embodiments, system 100 may be configured such that the fields of view of image capture devices 124 and 126, for example, fall within (e.g., are narrower than) and share a common center with the field of view of image capture device 122. In other embodiments, the image capture devices 122, 124, and 126 may capture adjacent FOVs or may have partial overlap in their FOVs. In some embodiments, the fields of view of image capture devices 122, 124, and 126 may be aligned such that a center of the narrower FOV image capture devices 124 and/or 126 may be located in a lower half of the field of view of the wider FOV device 122.

[0076] FIG. 2F is a diagrammatic representation of exemplary vehicle control systems, consistent with the disclosed embodiments. As indicated in FIG. 2F, vehicle 200 may include throttling system 220, braking system 230, and steering system 240. System 100 may provide inputs (e.g., control signals) to one or more of throttling system 220, braking system 230, and steering system 240 over one or more data links (e.g., any wired and/or wireless link or links for transmitting data). For example, based on analysis of images acquired by image capture devices 122, 124, and/or 126, system 100 may provide control signals to one or more of throttling system 220, braking system 230, and steering system 240 to navigate vehicle 200 (e.g., by causing an acceleration, a turn, a lane shift, etc.). Further, system 100 may receive inputs from one or more of throttling system 220, braking system 230, and steering system 24 indicating operating conditions of vehicle 200 (e.g., speed, whether vehicle 200 is braking and/or turning, etc.). Further details are provided in connection with FIGS. 4-7, below.

[0077] As shown in FIG. 3A, vehicle 200 may also include a user interface 170 for interacting with a driver or a passenger of vehicle 200. For example, user interface 170 in a vehicle application may include a touch screen 320, knobs 330, buttons 340, and a microphone 350. A driver or passenger of vehicle 200 may also use handles (e.g., located on or near the steering column of vehicle 200 including, for example, turn signal handles), buttons (e.g., located on the steering wheel of vehicle 200), and the like, to interact with system 100. In some embodiments, microphone 350 may be positioned adjacent to a rearview mirror 310. Similarly, in some embodiments, image capture device 122 may be located near rearview mirror 310. In some embodiments, user interface 170 may also include one or more speakers 360 (e.g., speakers of a vehicle audio system). For example, system 100 may provide various notifications (e.g., alerts) via speakers 360.

[0078] FIGS. 3B-3D are illustrations of an exemplary camera mount 370 configured to be positioned behind a rearview mirror (e.g., rearview mirror 310) and against a vehicle windshield, consistent with disclosed embodiments. As shown in FIG. 3B, camera mount 370 may include image capture devices 122, 124, and 126. Image capture devices 124 and 126 may be positioned behind a glare shield 380, which may be flush against the vehicle windshield and include a composition of film and/or anti-reflective materials. For example, glare shield 380 may be positioned such that it aligns against a vehicle windshield having a matching slope. In some embodiments, each of image capture devices 122, 124, and 126 may be positioned behind glare shield 380, as depicted, for example, in FIG. 3D. The disclosed embodiments are not limited to any particular configuration of image capture devices 122, 124, and 126, camera mount 370, and glare shield 380. FIG. 3C is an illustration of camera mount 370 shown in FIG. 3B from a front perspective.

IPR2025-00944
Tesla EX1002 Page 1600

[0079] As will be appreciated by a person skilled in the art having the benefit of this disclosure, numerous variations and/or modifications may be made to the foregoing disclosed embodiments. For example, not all components are essential for the operation of system 100. Further, any component may be located in any appropriate part of system 100 and the components may be rearranged into a variety of configurations while providing the functionality of the disclosed embodiments. Therefore, the foregoing configurations are examples and, regardless of the configurations discussed above, system 100 can provide a wide range of functionality to analyze the surroundings of vehicle 200 and navigate vehicle 200 in response to the analysis.

[0080] As discussed below in further detail and consistent with various disclosed embodiments, system 100 may provide a variety of features related to autonomous driving and/or driver assist technology. For example, system 100 may analyze image data, position data (e.g., GPS location information), map data, speed data, and/or data from sensors included in vehicle 200. System 100 may collect the data for analysis from, for example, image acquisition unit 120, position sensor 130, and other sensors. Further, system 100 may analyze the collected data to determine whether or not vehicle 200 should take a certain action, and then automatically take the determined action without human intervention. For example, when vehicle 200 navigates without human intervention, system 100 may automatically control the braking, acceleration, and/or steering of vehicle 200 (e.g., by sending control signals to one or more of throttling system 220, braking system 230, and steering system 240). Further, system 100 may analyze the collected data and issue warnings and/or alerts to vehicle occupants based on the analysis of the collected data.

[0081] Further, consistent with disclosed embodiments, the functionality provided by system 100 may cause vehicle 200 to take different actions to navigate vehicle 200 within a lane and/or relative to other vehicles and/or objects. For example, system 100 may adjust the positioning of vehicle 200 relative to a lane within which vehicle 200 is traveling and/or relative to objects positioned near vehicle 200, select a particular lane for vehicle 200 to use while traveling, and take action in response to an encroaching vehicle, such as a vehicle attempting to move into the lane within which vehicle 200 is traveling. Additionally, system 100 may control the speed of vehicle 200 in different scenarios, such as when vehicle 200 is making a turn. System 100 may further cause vehicle 200 to mimic the actions of a leading vehicle or monitor a target vehicle and navigate vehicle 200 so that it passes the target vehicle. Additional details regarding the various embodiments that are provided by system 100 are provided below.

[0082] Forward-Facing Multi-Imaging System

[0083] As discussed above, system 100 may provide drive assist functionality that uses a multi-camera system. The multi-camera system may use one or more cameras facing in the forward direction of a vehicle. In other embodiments, the multi-camera system may include one or more cameras facing to the side of a vehicle or to the rear of the vehicle. In one embodiment, for example, system 100 may use a two-camera imaging system, where a first camera and a second camera (e.g., image capture devices 122 and 124) may be positioned at the front and/or the sides of a vehicle (e.g., vehicle 200). The first camera may have a field of view that is greater than, less than, or partially overlapping with, the field of view of the second camera. In addition, the first camera may be connected to a first image processor to perform monocular image analysis of

IPR2025-00944
Tesla EX1002 Page 1601

images provided by the first camera, and the second camera may be connected to a second image processor to perform monocular image analysis of images provided by the second camera. The outputs (e.g., processed information) of the first and second image processors may be combined. In some embodiments, the second image processor may receive images from both the first camera and second camera to perform stereo analysis. In another embodiment, system 100 may use a three-camera imaging system where each of the cameras has a different field of view. Such a system may, therefore, make decisions based on information derived from objects located at varying distances both forward and to the sides of the vehicle. References to monocular image analysis may refer to instances where image analysis is performed based on images captured from a single point of view (e.g., from a single camera). Stereo image analysis may refer to instances where image analysis is performed based on two or more images captured with one or more variations of an image capture parameter. For example, captured images suitable for performing stereo image analysis may include images captured: from two or more different positions, from different fields of view, using different focal lengths, along with parallax information, etc.

[0084] For example, in one embodiment, system 100 may implement a three camera configuration using image capture devices 122-126. In such a configuration, image capture device 122 may provide a narrow field of view (e.g., 34 degrees, or other values selected from a range of about 20 to 45 degrees, etc.), image capture device 124 may provide a wide field of view (e.g., 150 degrees or other values selected from a range of about 100 to about 180 degrees), and image capture device 126 may provide an intermediate field of view (e.g., 46 degrees or other values selected from a range of about 35 to about 60 degrees). In some embodiments, image capture device 126 may act as a main or primary camera. Image capture devices 122-126 may be positioned behind rearview mirror 310 and positioned substantially side-by-side (e.g., 6 cm apart). Further, in some embodiments, as discussed above, one or more of image capture devices 122-126 may be mounted behind glare shield 380 that is flush with the windshield of vehicle 200. Such shielding may act to minimize the impact of any reflections from inside the car on image capture devices 122-126.

[0085] In another embodiment, as discussed above in connection with FIGS. 3B and 3C, the wide field of view camera (e.g., image capture device 124 in the above example) may be mounted lower than the narrow and main field of view cameras (e.g., image devices 122 and 126 in the above example). This configuration may provide a free line of sight from the wide field of view camera. To reduce reflections, the cameras may be mounted close to the windshield of vehicle 200, and may include polarizers on the cameras to damp reflected light.

[0086] A three camera system may provide certain performance characteristics. For example, some embodiments may include an ability to validate the detection of objects by one camera based on detection results from another camera. In the three camera configuration discussed above, processing unit 110 may include, for example, three processing devices (e.g., three EyeQ series of processor chips, as discussed above), with each processing device dedicated to processing images captured by one or more of image capture devices 122-126.

[0087] In a three camera system, a first processing device may receive images from both the main camera and the narrow field of view camera, and perform vision processing of the narrow FOV camera to, for example, detect other vehicles, pedestrians, lane marks, traffic signs, traffic lights, and other road objects. Further, the

IPR2025-00944
Tesla EX1002 Page 1602

first processing device may calculate a disparity of pixels between the images from the main camera and the narrow camera and create a 3D reconstruction of the environment of vehicle 200. The first processing device may then combine the 3D reconstruction with 3D map data or with 3D information calculated based on information from another camera.

[0088] The second processing device may receive images from main camera and perform vision processing to detect other vehicles, pedestrians, lane marks, traffic signs, traffic lights, and other road objects. Additionally, the second processing device may calculate a camera displacement and, based on the displacement, calculate a disparity of pixels between successive images and create a 3D reconstruction of the scene (e.g., a structure from motion). The second processing device may send the structure from motion based 3D reconstruction to the first processing device to be combined with the stereo 3D images.

[0089] The third processing device may receive images from the wide FOV camera and process the images to detect vehicles, pedestrians, lane marks, traffic signs, traffic lights, and other road objects. The third processing device may further execute additional processing instructions to analyze images to identify objects moving in the image, such as vehicles changing lanes, pedestrians, etc.

[0090] In some embodiments, having streams of image-based information captured and processed independently may provide an opportunity for providing redundancy in the system. Such redundancy may include, for example, using a first image capture device and the images processed from that device to validate and/or supplement information obtained by capturing and processing image information from at least a second image capture device.

[0091] In some embodiments, system 100 may use two image capture devices (e.g., image capture devices 122 and 124) in providing navigation assistance for vehicle 200 and use a third image capture device (e.g., image capture device 126) to provide redundancy and validate the analysis of data received from the other two image capture devices. For example, in such a configuration, image capture devices 122 and 124 may provide images for stereo analysis by system 100 for navigating vehicle 200, while image capture device 126 may provide images for monocular analysis by system 100 to provide redundancy and validation of information obtained based on images captured from image capture device 122 and/or image capture device 124. That is, image capture device 126 (and a corresponding processing device) may be considered to provide a redundant sub-system for providing a check on the analysis derived from image capture devices 122 and 124 (e.g., to provide an automatic emergency braking (AEB) system).

[0092] One of skill in the art will recognize that the above camera configurations, camera placements, number of cameras, camera locations, etc., are examples only. These components and others described relative to the overall system may be assembled and used in a variety of different configurations without departing from the scope of the disclosed embodiments. Further details regarding usage of a multi-camera system to provide driver assist and/or autonomous vehicle functionality follow below.

[0093] FIG. 4 is an exemplary functional block diagram of memory 140 and/or 150, which may be stored/programmed with instructions for performing one or more operations consistent with the disclosed embodiments. Although the following refers to

IPR2025-00944
Tesla EX1002 Page 1603