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 1287 IPR2025-01034

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.

IPR2025-01034
Tesla EX1002 Page 1288

[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 1289 IPR2025-01034

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 1290 IPR2025-01034

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 1291 IPR2025-01034

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 1292 IPR2025-01034

[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

IPR2025-01034
Tesla EX1002 Page 1293

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.

IPR2025-01034
Tesla EX1002 Page 1294

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

IPR2025-01034
Tesla EX1002 Page 1295

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 1296 IPR2025-01034

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)

IPR2025-01034
Tesla EX1002 Page 1297

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 1298 IPR2025-01034

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 1299 IPR2025-01034

[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

IPR2025-01034
Tesla EX1002 Page 1300

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.

IPR2025-01034
Tesla EX1002 Page 1301

[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 1302 IPR2025-01034

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 1303 IPR2025-01034

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 1304 IPR2025-01034

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

IPR2025-01034
Tesla EX1002 Page 1305

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

IPR2025-01034
Tesla EX1002 Page 1307

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 1308 IPR2025-01034

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

IPR2025-01034
Tesla EX1002 Page 1311

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.

IPR2025-01034
Tesla EX1002 Page 1312

[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

IPR2025-01034
Tesla EX1002 Page 1313

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 1314 IPR2025-01034

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.

IPR2025-01034
Tesla EX1002 Page 1315

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

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

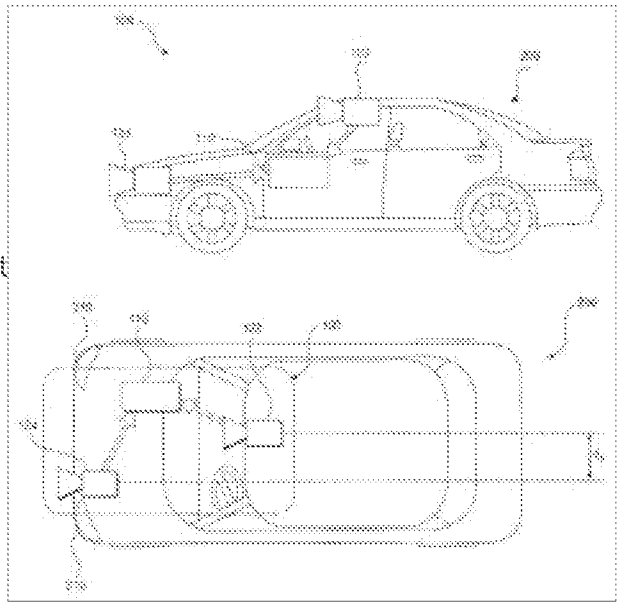
Also published as: CN106461774 (A) CN106461774 (B) CN109884618 (A) US10274598 (B2) US10690770 (B2) more

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

IPR2025-01034
Tesla EX1002 Page 1320

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 1323 IPR2025-01034

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 1324 IPR2025-01034

[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 1325 IPR2025-01034

[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 1326 IPR2025-01034

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 1327 IPR2025-01034

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 1328 IPR2025-01034

[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

IPR2025-01034
Tesla EX1002 Page 1329

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 1330 IPR2025-01034

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 1331 IPR2025-01034

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

IPR2025-01034
Tesla EX1002 Page 1332

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 1333 IPR2025-01034

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 1334 IPR2025-01034

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 1335 IPR2025-01034

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.

IPR2025-01034
Tesla EX1002 Page 1336

[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

IPR2025-01034
Tesla EX1002 Page 1339

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,

IPR2025-01034
Tesla EX1002 Page 1340

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.

IPR2025-01034
Tesla EX1002 Page 1341

[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 1342 IPR2025-01034

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

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

IPR2025-01034
Tesla EX1002 Page 1344

IPR2025-01034
Tesla EX1002 Page 1345

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 1346 IPR2025-01034

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.

IPR2025-01034
Tesla EX1002 Page 1347

[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.

IPR2025-01034
Tesla EX1002 Page 1348

[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.

Tesla EX1002 Page 1350 IPR2025-01034

[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 1351 IPR2025-01034

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 1352 IPR2025-01034

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

IPR2025-01034
Tesla EX1002 Page 1354

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 1355 IPR2025-01034

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 1356 IPR2025-01034

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.

IPR2025-01034
Tesla EX1002 Page 1357

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

IPR2025-01034
Tesla EX1002 Page 1358

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)

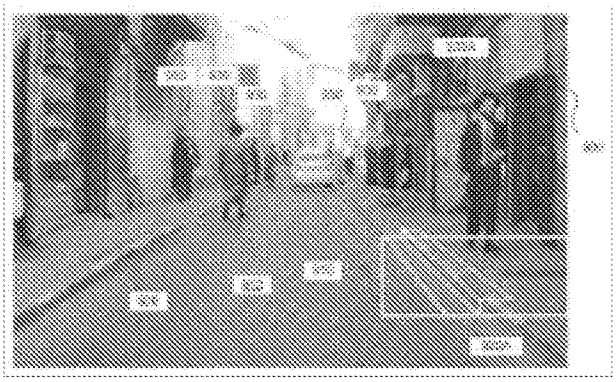
Classification: - **international:** ***G06K9/00; G08G1/16***
- **cooperative:** **B60W30/09 (US)**; **B60W50/14 (US)**; **G06K9/00798 (EP, CN, US)**; **G06K9/00805 (EP, CN, US)**; **G06K9/4604 (US)**; **G06K9/6267 (US)**; **G06T3/0093 (US)**; **G06T7/13 (EP, US)**; **G06T7/97 (EP, US)**; **G08G1/166 (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.

IPR2025-01034
Tesla EX1002 Page 1364

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.

IPR2025-01034
Tesla EX1002 Page 1366

[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 1367 IPR2025-01034

[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 1369 IPR2025-01034

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.

IPR2025-01034
Tesla EX1002 Page 1370

IPR2025-01034
Tesla EX1002 Page 1371

[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

IPR2025-01034
Tesla EX1002 Page 1372

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 1374 IPR2025-01034

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 1375 IPR2025-01034

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 1376 IPR2025-01034

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

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

IPR2025-01034
Tesla EX1002 Page 1379

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

IPR2025-01034
Tesla EX1002 Page 1381

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

IPR2025-01034
Tesla EX1002 Page 1382

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

IPR2025-01034
Tesla EX1002 Page 1383

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

IPR2025-01034
Tesla EX1002 Page 1384

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

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

IPR2025-01034 Tesla EX1002 Page 1386

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 1387 IPR2025-01034

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

IPR2025-01034
Tesla EX1002 Page 1388

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 1389 IPR2025-01034

[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

IPR2025-01034
Tesla EX1002 Page 1391

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.

IPR2025-01034
Tesla EX1002 Page 1392

[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

IPR2025-01034
Tesla EX1002 Page 1393

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

(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 1395 IPR2025-01034

[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 1396 IPR2025-01034

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 1397 IPR2025-01034

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.

IPR2025-01034
Tesla EX1002 Page 1398

IPR2025-01034
Tesla EX1002 Page 1399

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 free

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

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

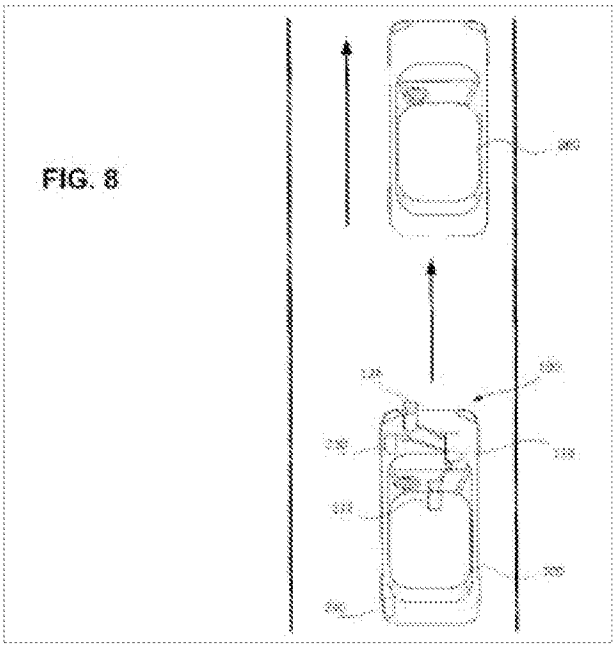
Also published as: EP3146506 (B1) CN106536299 (A) CN106536299 (B) CN110171405 (A) CN110171405 (B) more

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.

IPR2025-01034
Tesla EX1002 Page 1404

FIG. 8



IPR2025-01034
Tesla EX1002 Page 1405

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

IPR2025-01034
Tesla EX1002 Page 1406

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 1407 IPR2025-01034

[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

IPR2025-01034
Tesla EX1002 Page 1408

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.

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 1410 IPR2025-01034

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

IPR2025-01034
Tesla EX1002 Page 1411

IPR2025-01034
Tesla EX1002 Page 1412

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

IPR2025-01034
Tesla EX1002 Page 1413

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 1414 IPR2025-01034

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

IPR2025-01034
Tesla EX1002 Page 1415

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.

IPR2025-01034
Tesla EX1002 Page 1416

[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 1417 IPR2025-01034

[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.

IPR2025-01034
Tesla EX1002 Page 1419

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.

IPR2025-01034
Tesla EX1002 Page 1420

[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 1421 IPR2025-01034

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

IPR2025-01034
Tesla EX1002 Page 1422

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 1423 IPR2025-01034

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

IPR2025-01034
Tesla EX1002 Page 1424

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

IPR2025-01034
Tesla EX1002 Page 1425

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 1426 IPR2025-01034

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.

IPR2025-01034
Tesla EX1002 Page 1428

[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 1429 IPR2025-01034

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

IPR2025-01034
Tesla EX1002 Page 1430

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

IPR2025-01034
Tesla EX1002 Page 1431

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

IPR2025-01034
Tesla EX1002 Page 1432

IPR2025-01034
Tesla EX1002 Page 1433

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 1434 IPR2025-01034

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

IPR2025-01034
Tesla EX1002 Page 1435

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

IPR2025-01034
Tesla EX1002 Page 1436

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

IPR2025-01034
Tesla EX1002 Page 1437

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 1438 IPR2025-01034

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 1439 IPR2025-01034

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 1440 IPR2025-01034

[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 1441 IPR2025-01034

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 1442 IPR2025-01034

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.

IPR2025-01034
Tesla EX1002 Page 1443

[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

IPR2025-01034
Tesla EX1002 Page 1444

IPR2025-01034
Tesla EX1002 Page 1445

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 1448 IPR2025-01034

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).

Printed by Jouve, 75001 PARIS (FR)

IPR2025-01034
Tesla EX1002 Page 1452

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

IPR2025-01034
Tesla EX1002 Page 1453

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

IPR2025-01034
Tesla EX1002 Page 1454

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-

IPR2025-01034
Tesla EX1002 Page 1455

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

IPR2025-01034
Tesla EX1002 Page 1456

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

IPR2025-01034
Tesla EX1002 Page 1457

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

IPR2025-01034
Tesla EX1002 Page 1458

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-

IPR2025-01034
Tesla EX1002 Page 1459

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

IPR2025-01034
Tesla EX1002 Page 1460

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.

IPR2025-01034
Tesla EX1002 Page 1461

[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

IPR2025-01034
Tesla EX1002 Page 1462

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

IPR2025-01034
Tesla EX1002 Page 1463

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,

IPR2025-01034
Tesla EX1002 Page 1464

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-

IPR2025-01034
Tesla EX1002 Page 1465

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

IPR2025-01034
Tesla EX1002 Page 1466

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

IPR2025-01034
Tesla EX1002 Page 1467

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

IPR2025-01034
Tesla EX1002 Page 1468

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

IPR2025-01034
Tesla EX1002 Page 1469

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

IPR2025-01034
Tesla EX1002 Page 1470

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-

IPR2025-01034
Tesla EX1002 Page 1471

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

IPR2025-01034
Tesla EX1002 Page 1472

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-

IPR2025-01034
Tesla EX1002 Page 1473

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

IPR2025-01034
Tesla EX1002 Page 1474

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

IPR2025-01034
Tesla EX1002 Page 1475

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.

IPR2025-01034
Tesla EX1002 Page 1476

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

IPR2025-01034
Tesla EX1002 Page 1477

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;

IPR2025-01034
Tesla EX1002 Page 1478

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

IPR2025-01034
Tesla EX1002 Page 1479

(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

IPR2025-01034
Tesla EX1002 Page 1480

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

IPR2025-01034
Tesla EX1002 Page 1481

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 ;

IPR2025-01034
Tesla EX1002 Page 1482

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 :

IPR2025-01034
Tesla EX1002 Page 1483

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

IPR2025-01034
Tesla EX1002 Page 1484

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.

IPR2025-01034
Tesla EX1002 Page 1485

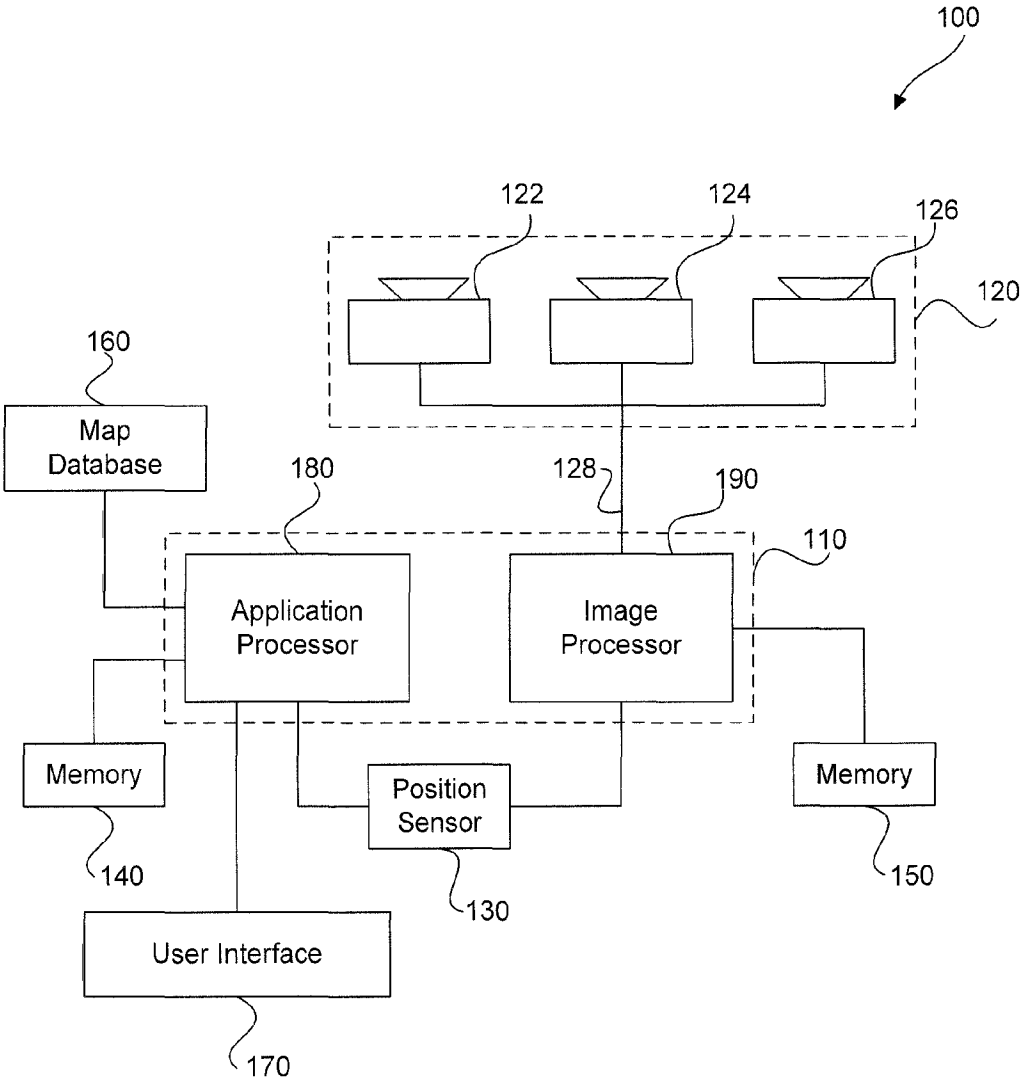


FIG. 1

IPR2025-01034
Tesla EX1002 Page 1486

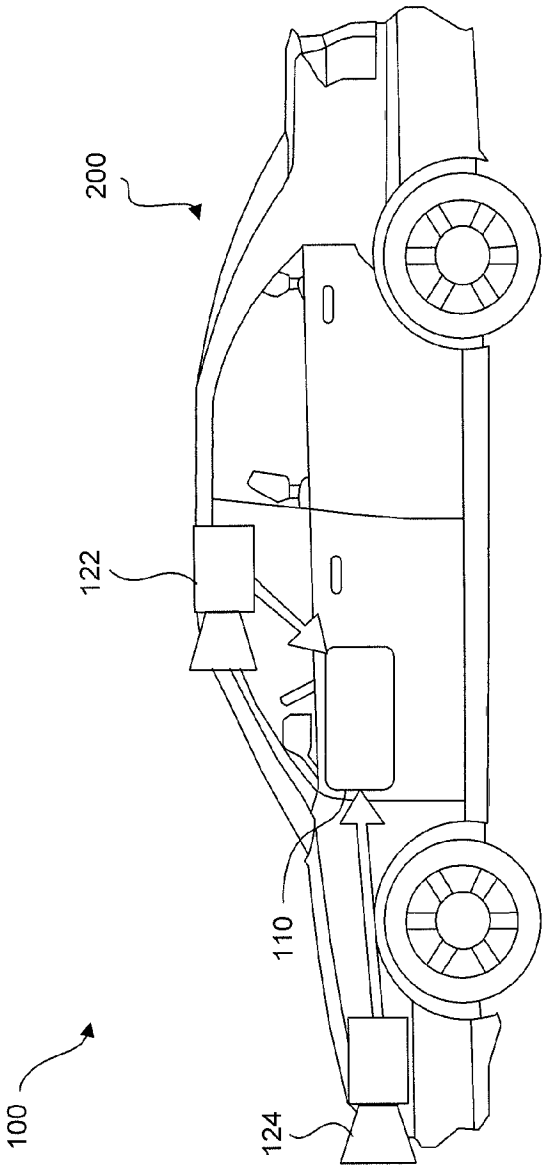


FIG. 2A

IPR2025-01034
Tesla EX1002 Page 1487

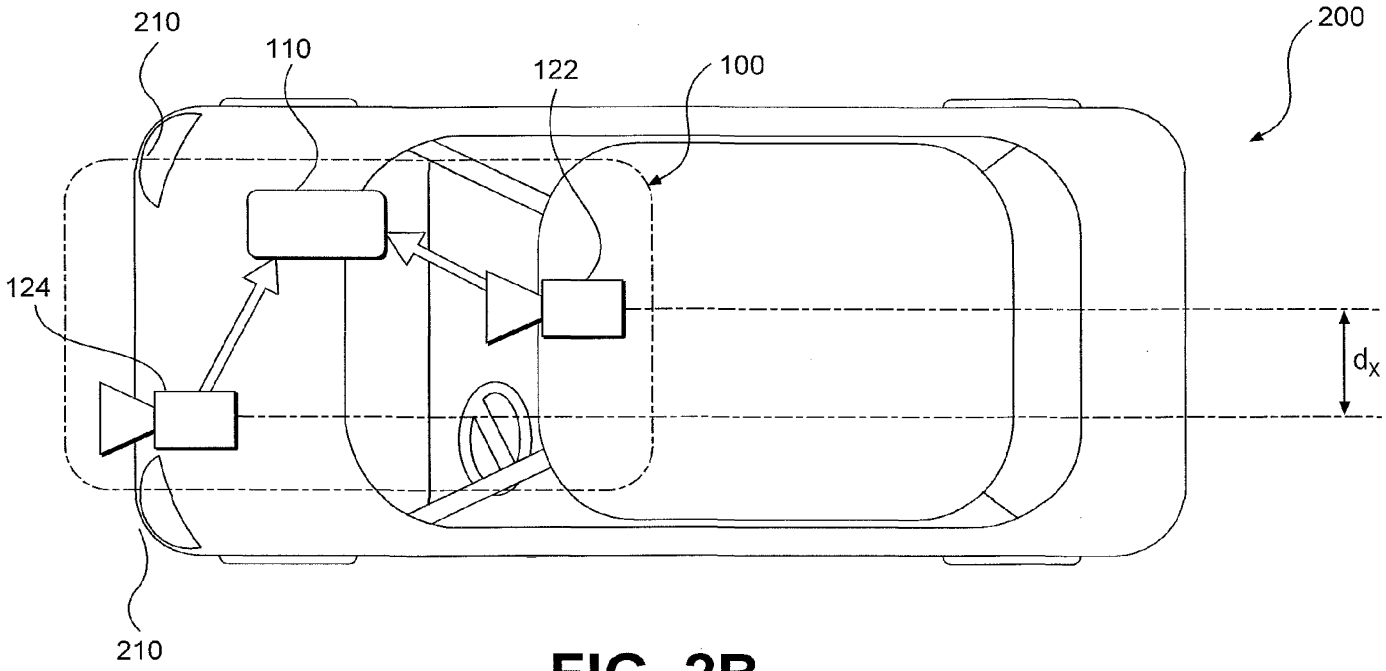


FIG. 2B

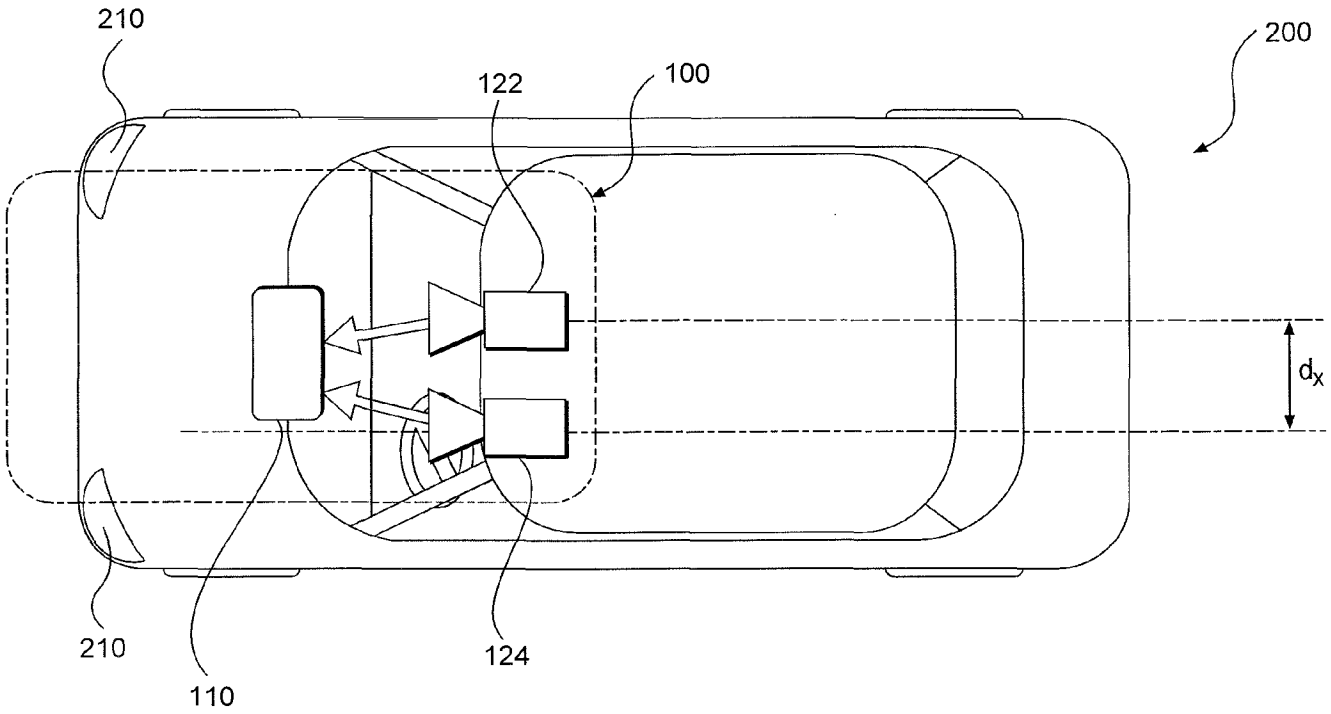


FIG. 2C

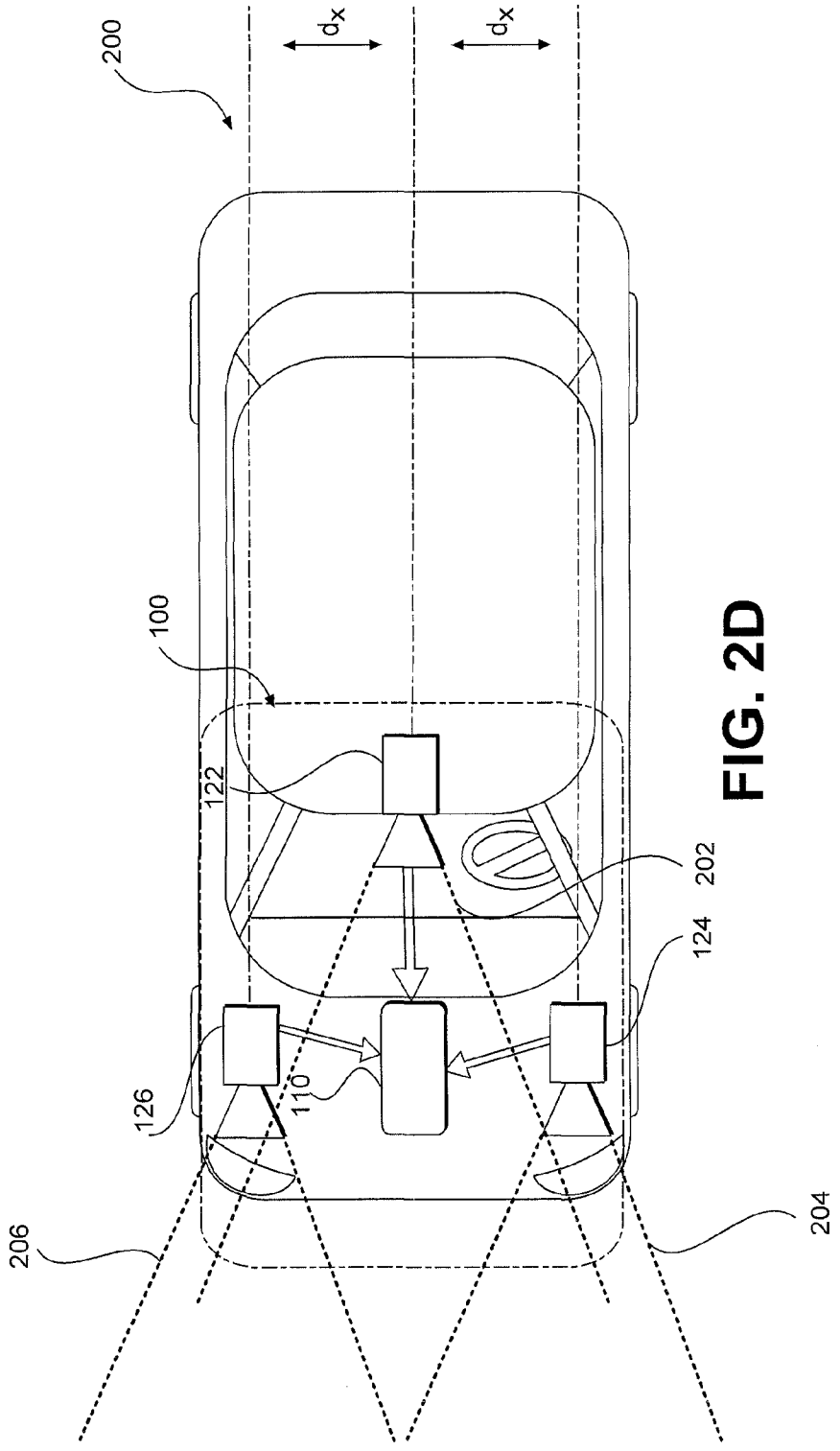


FIG. 2D

IPR2025-01034
Tesla EX1002 Page 1490

IPR2025-01034
Tesla EX1002 Page 1491

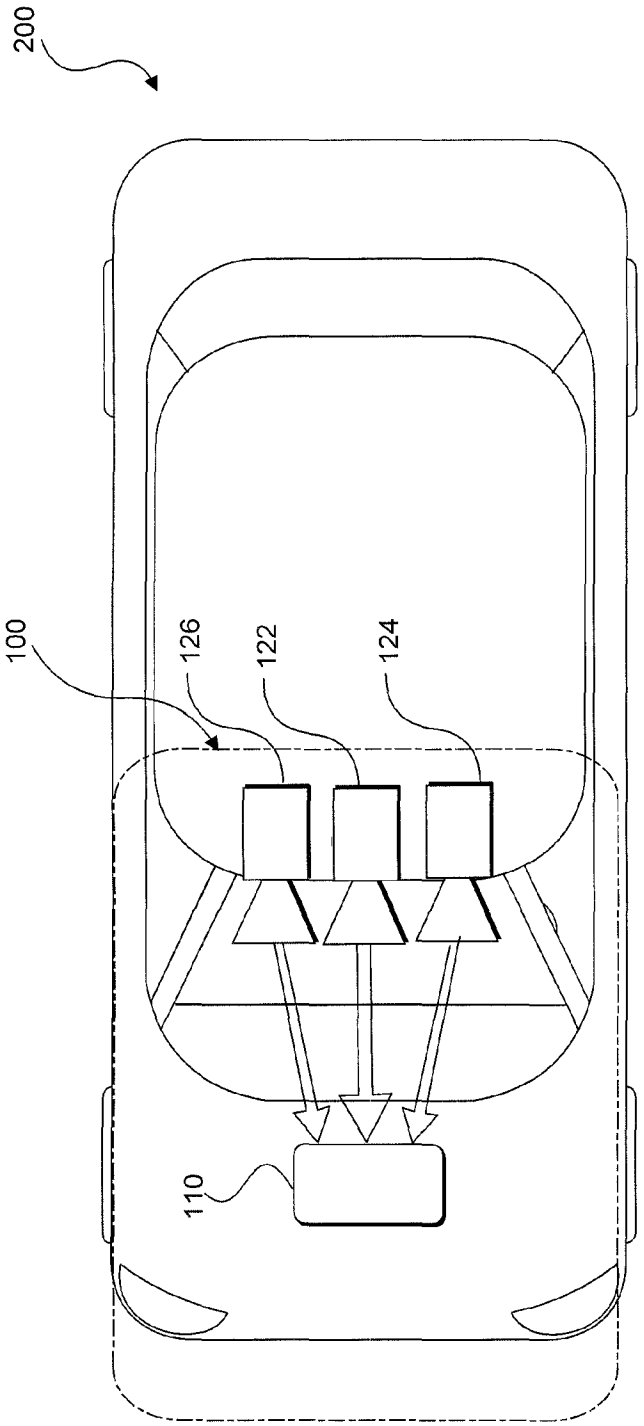


FIG. 2E

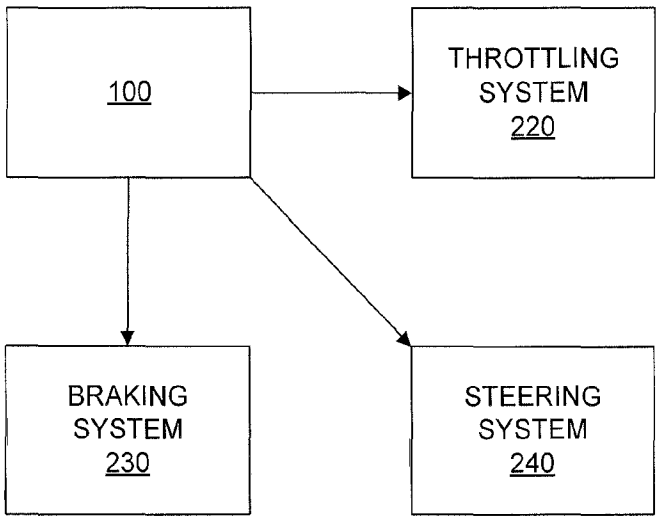


FIG. 2F

IPR2025-01034
Tesla EX1002 Page 1492

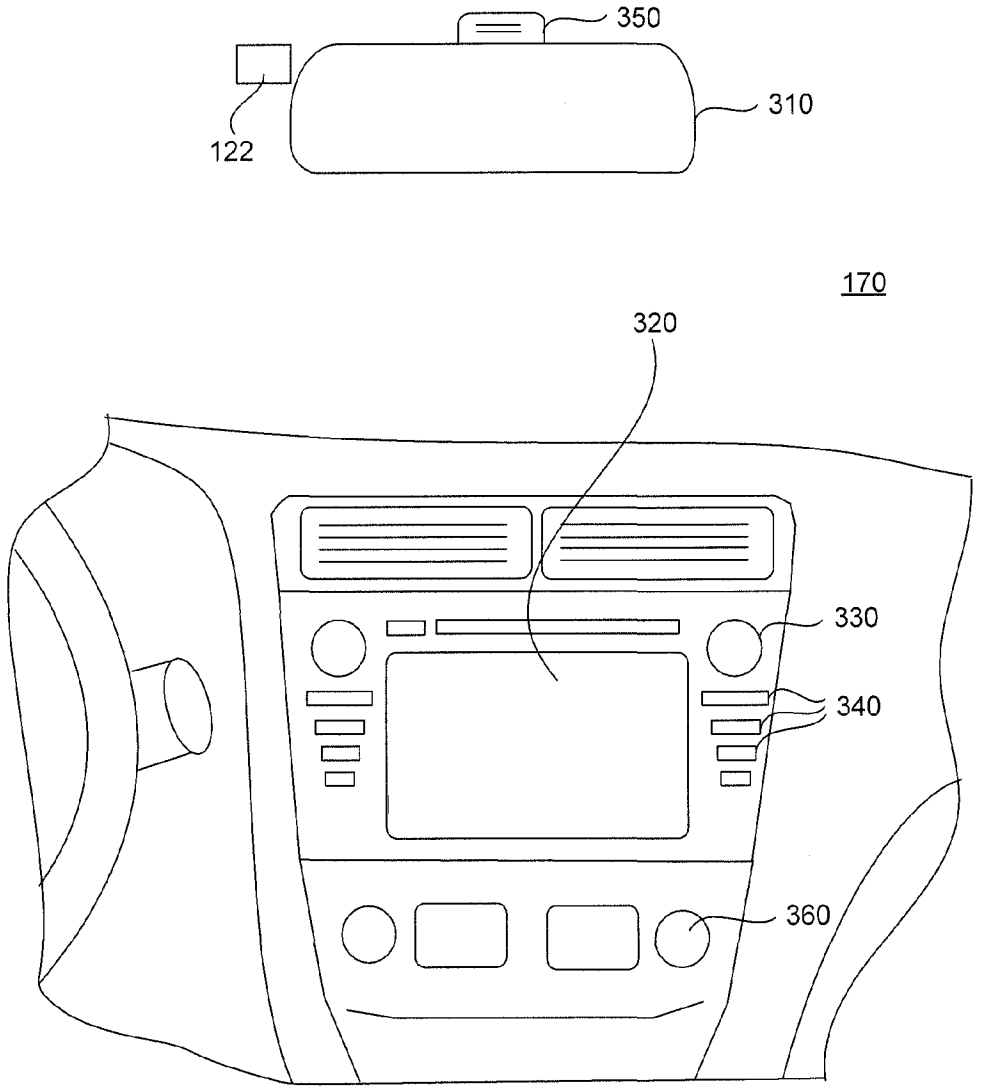


FIG. 3A

IPR2025-01034
Tesla EX1002 Page 1493

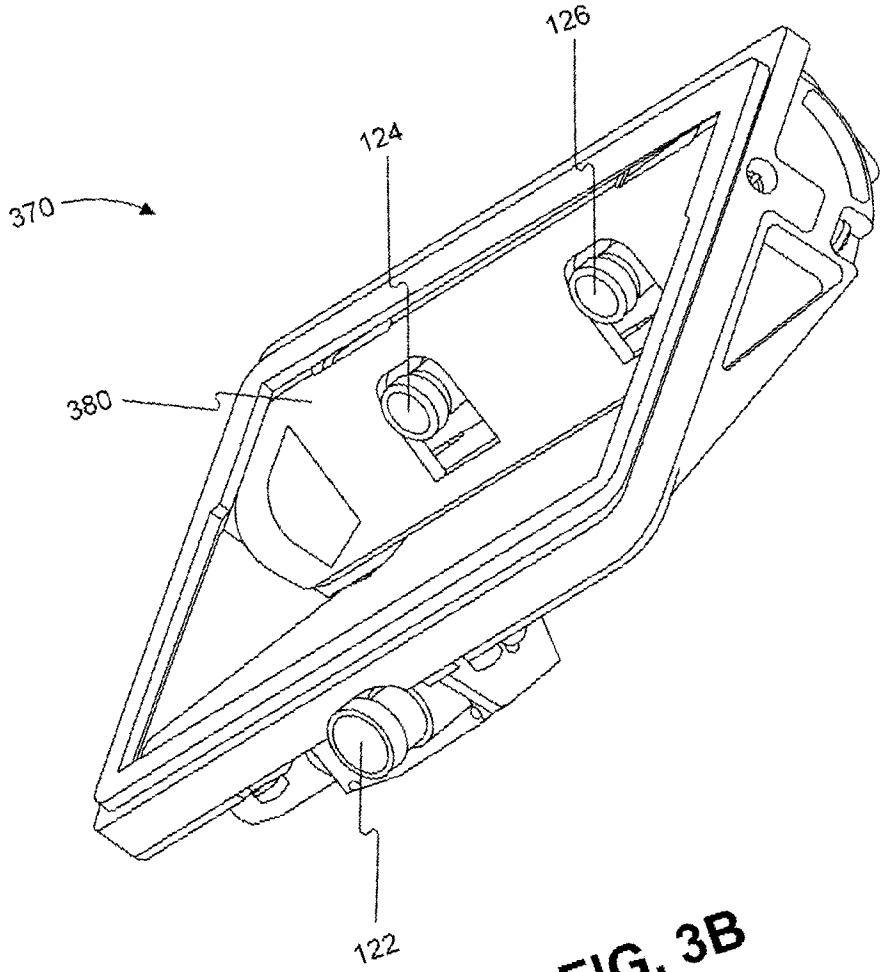


FIG. 3B

IPR2025-01034
Tesla EX1002 Page 1494

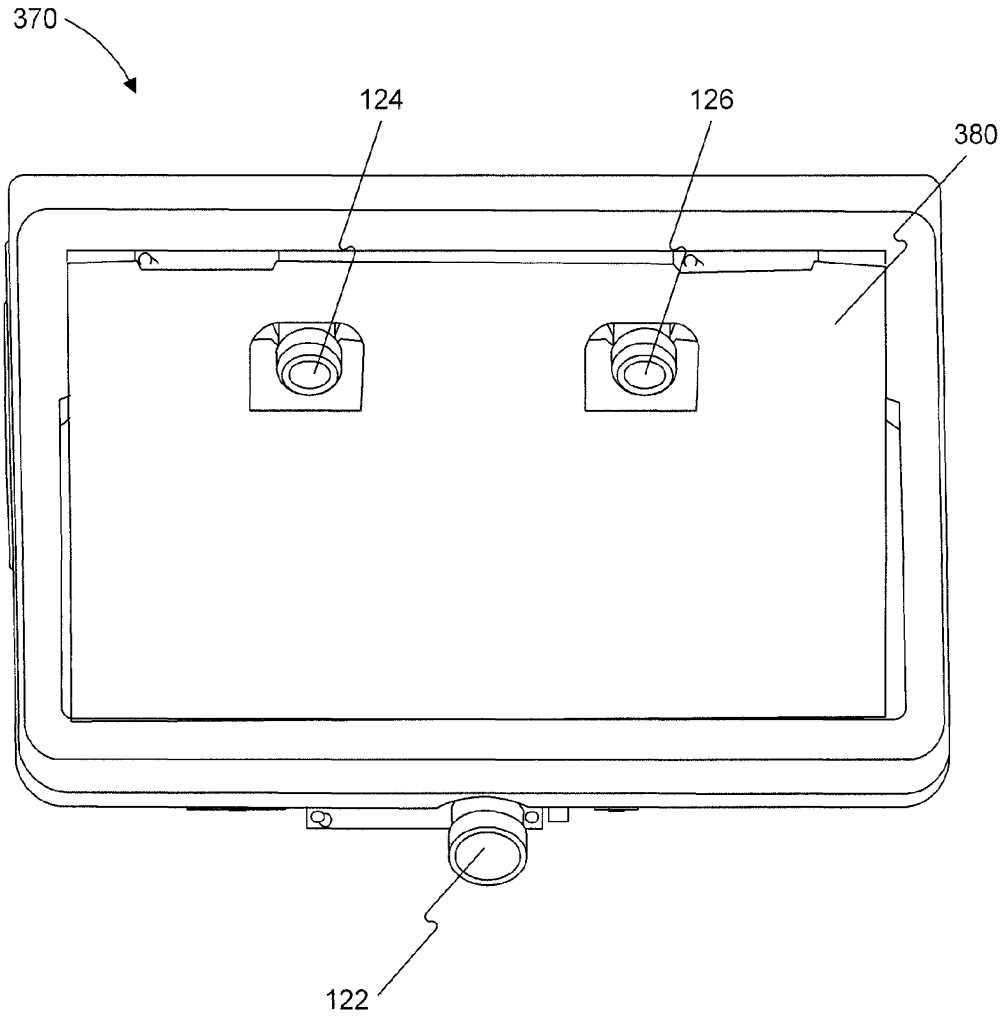


FIG. 3C

IPR2025-01034
Tesla EX1002 Page 1495

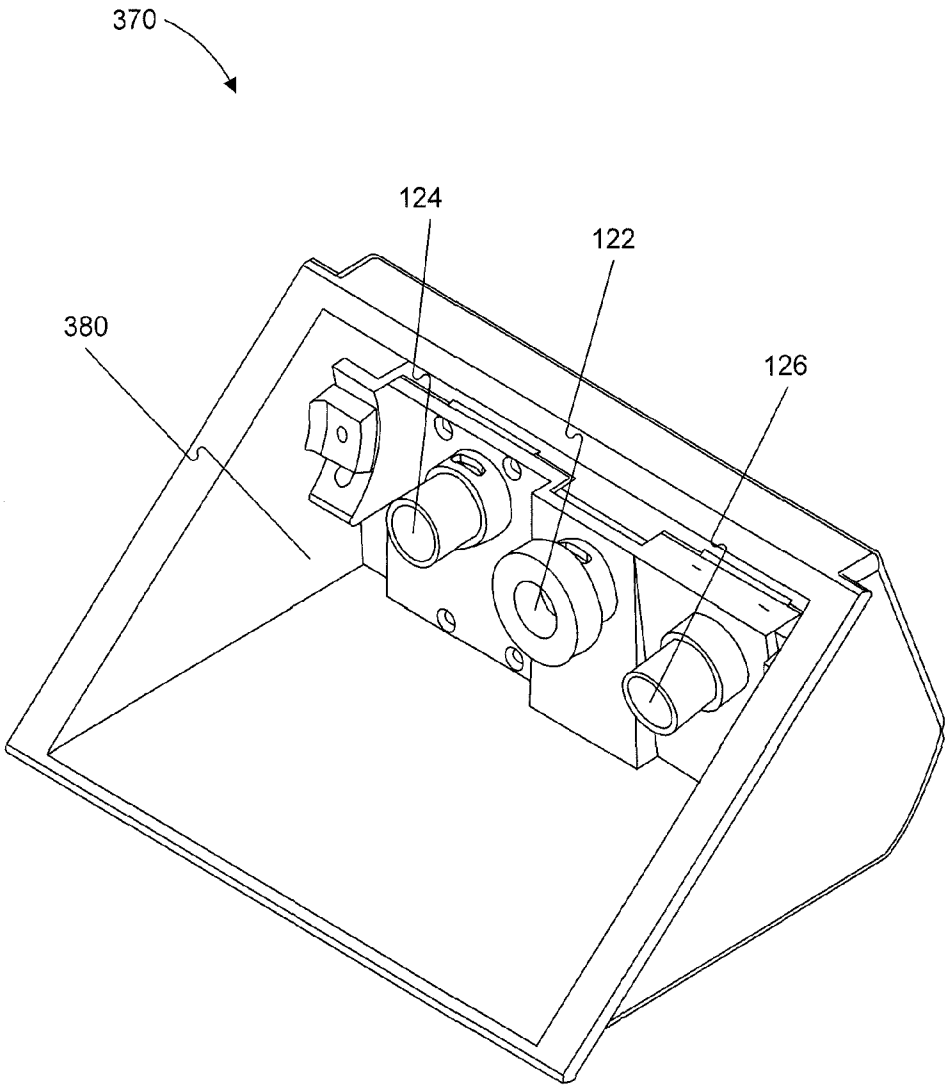


FIG. 3D

IPR2025-01034
Tesla EX1002 Page 1496

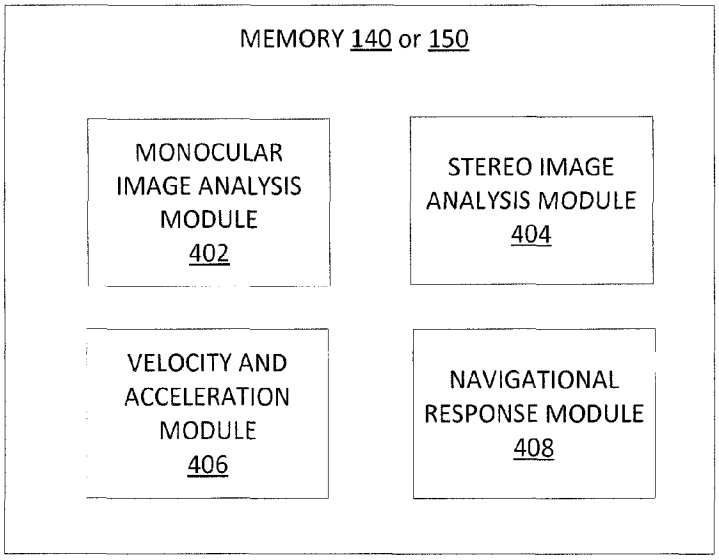


FIG. 4

IPR2025-01034
Tesla EX1002 Page 1497

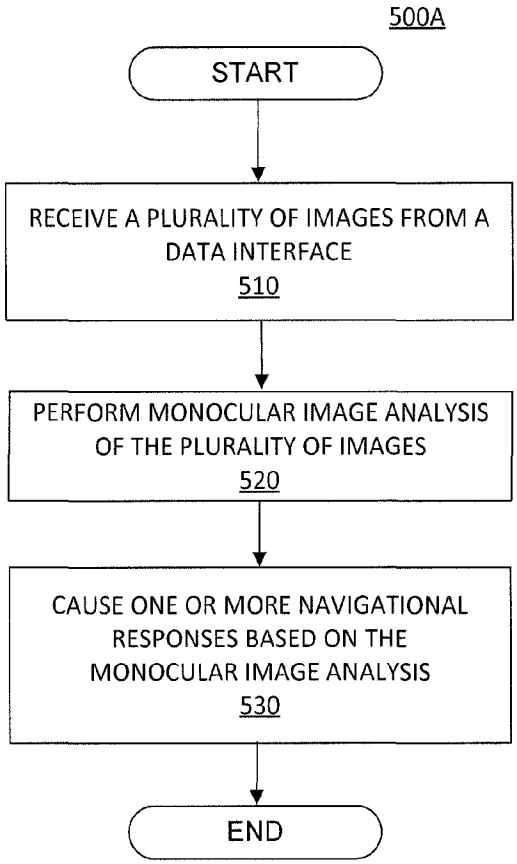


FIG. 5A

IPR2025-01034
Tesla EX1002 Page 1498

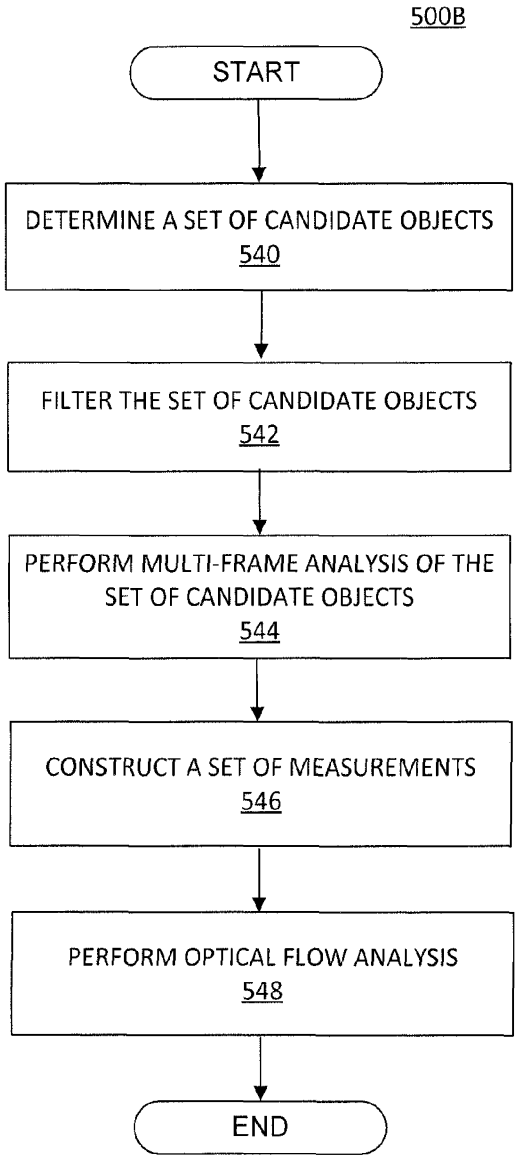


FIG. 5B

IPR2025-01034
Tesla EX1002 Page 1499

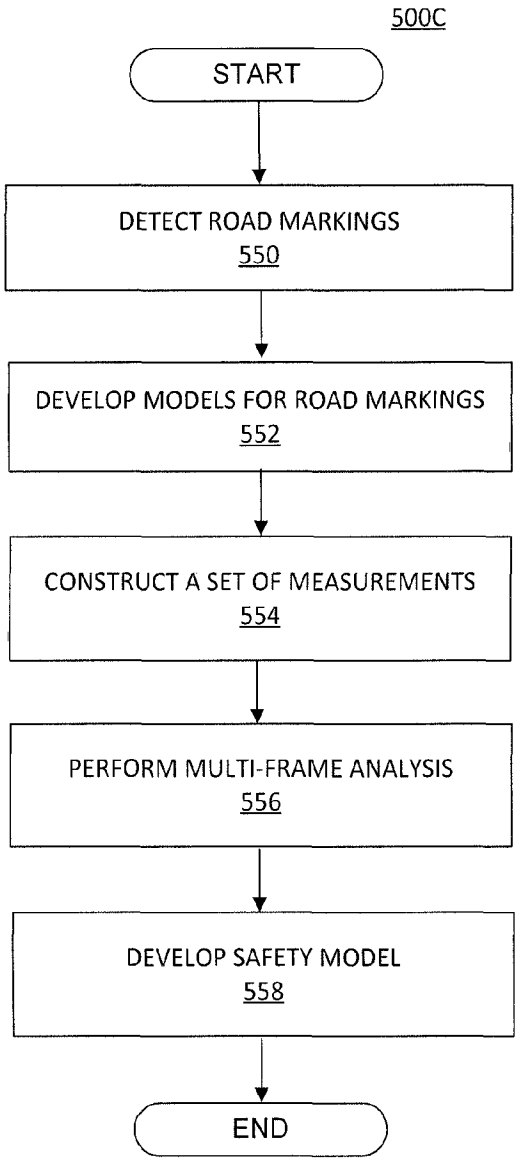


FIG. 5C

IPR2025-01034
Tesla EX1002 Page 1500

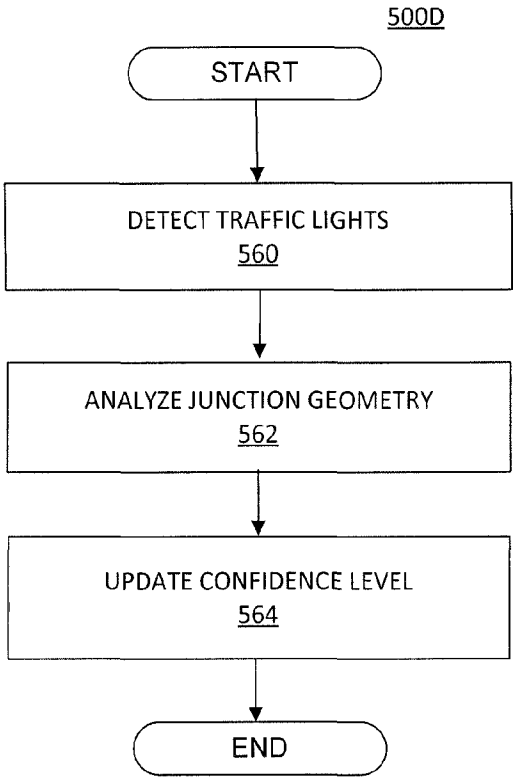


FIG. 5D

IPR2025-01034
Tesla EX1002 Page 1501

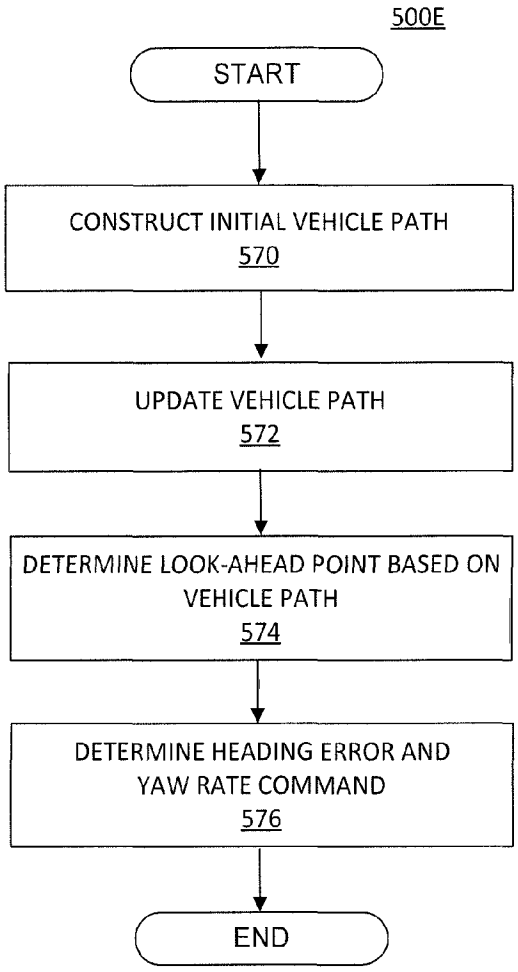


FIG. 5E

IPR2025-01034
Tesla EX1002 Page 1502

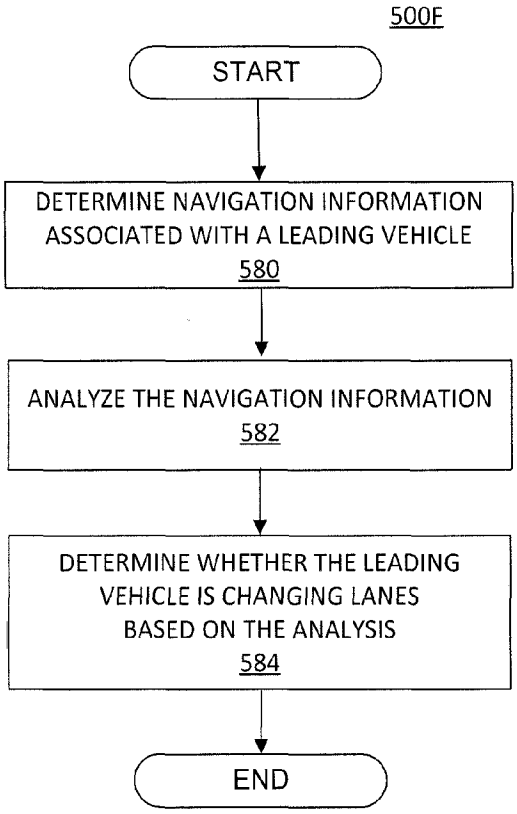


FIG. 5F

IPR2025-01034
Tesla EX1002 Page 1503

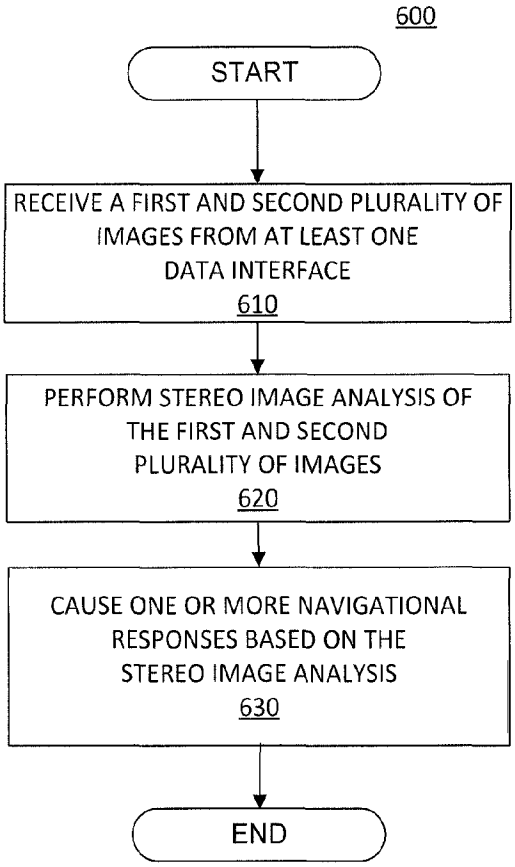


FIG. 6

IPR2025-01034
Tesla EX1002 Page 1504

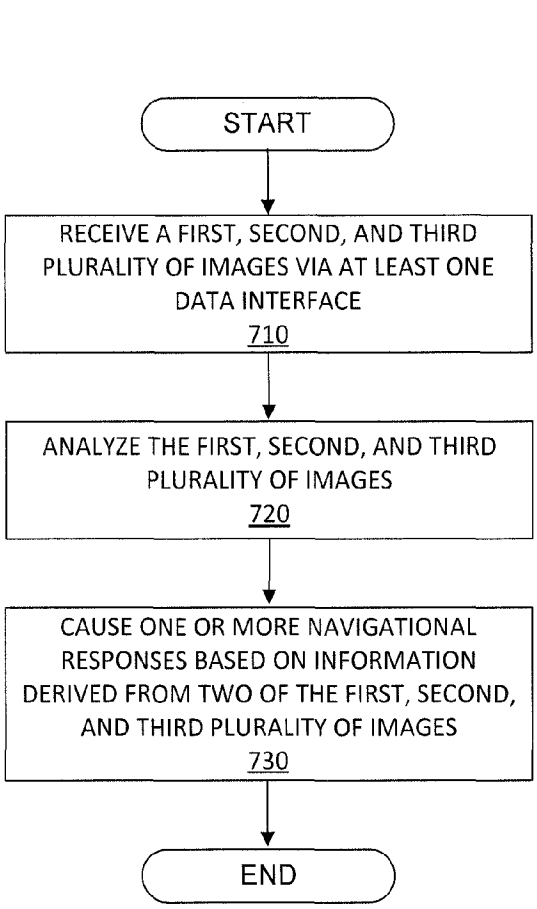


FIG. 7

IPR2025-01034
Tesla EX1002 Page 1505

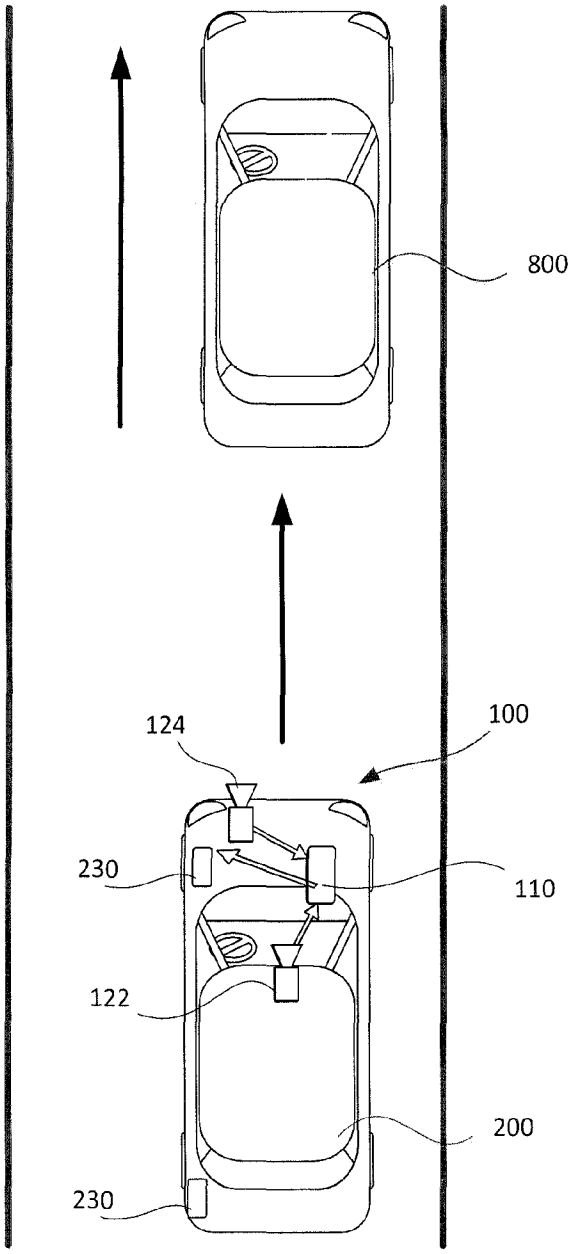


FIG. 8

IPR2025-01034
Tesla EX1002 Page 1506

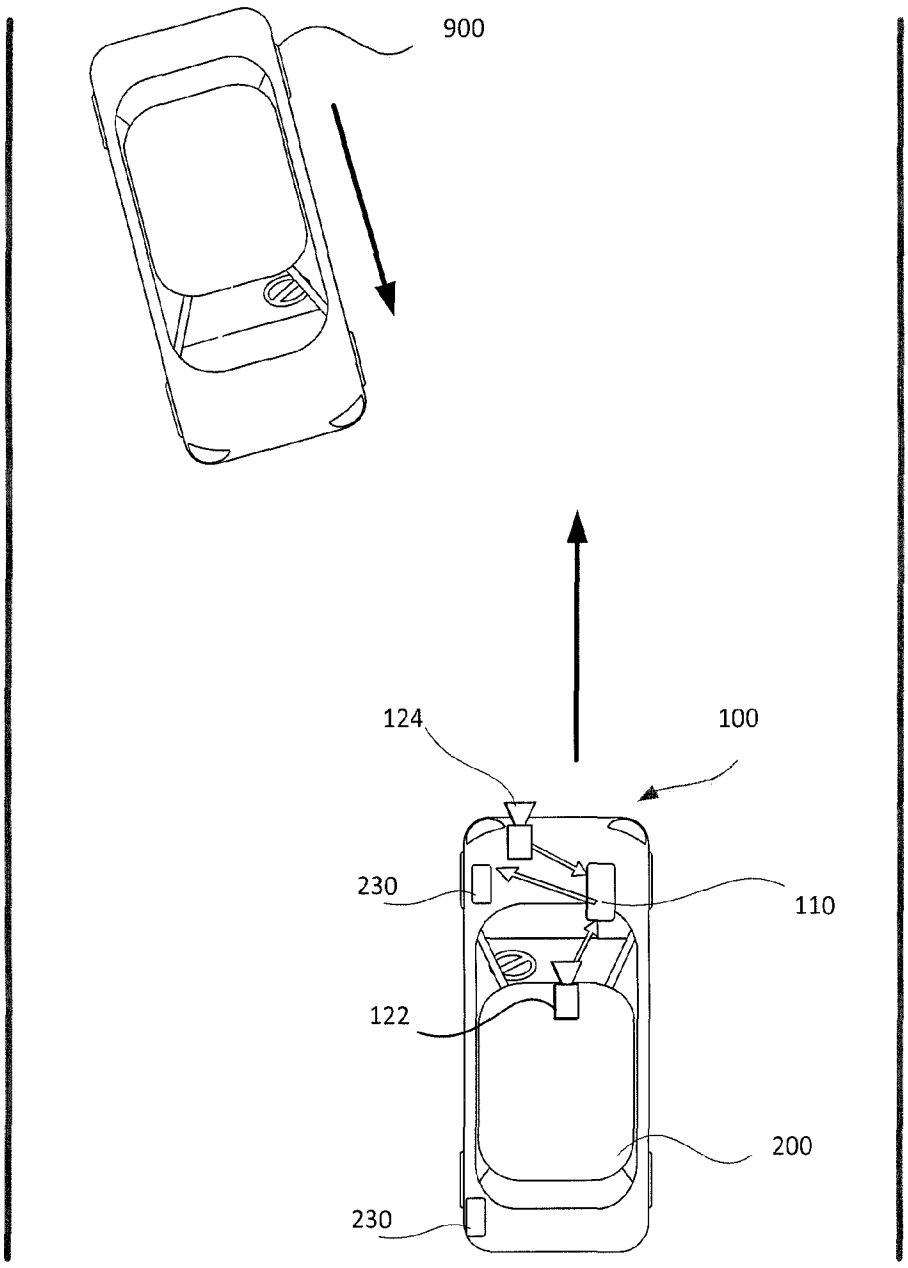


FIG. 9

IPR2025-01034
Tesla EX1002 Page 1507

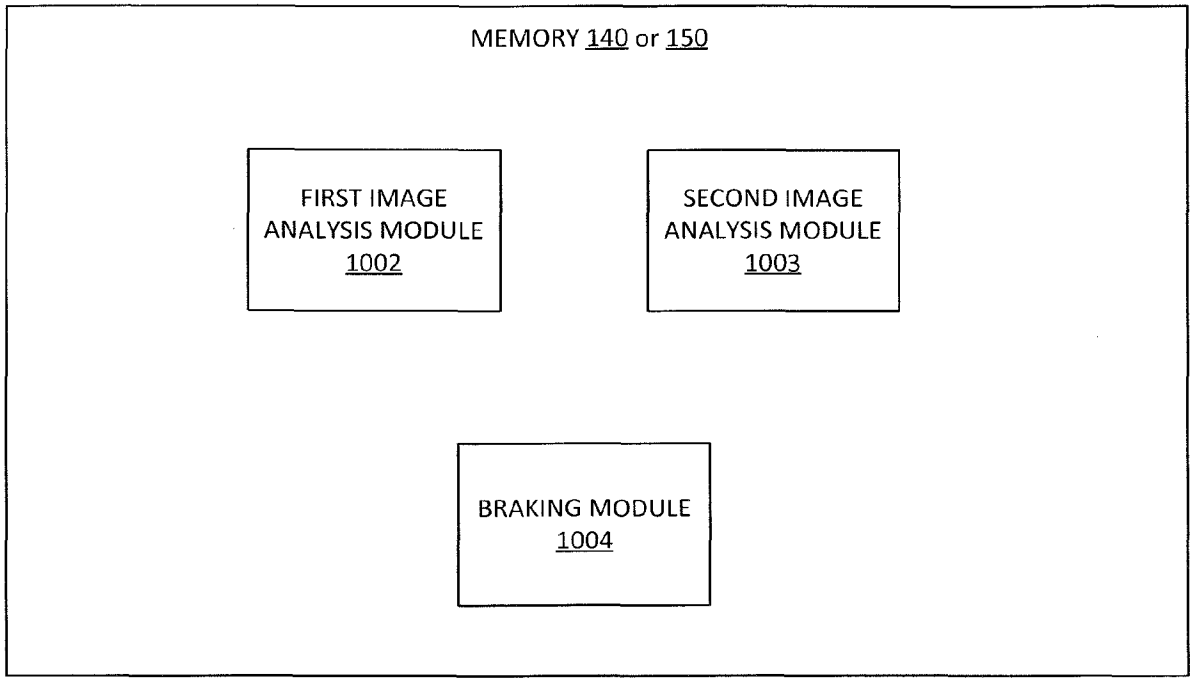

MEMORY 140 or 150
FIRST IMAGE ANALYSIS MODULE 1002
SECOND IMAGE ANALYSIS MODULE 1003
BRAKING MODULE 1004

FIG. 10

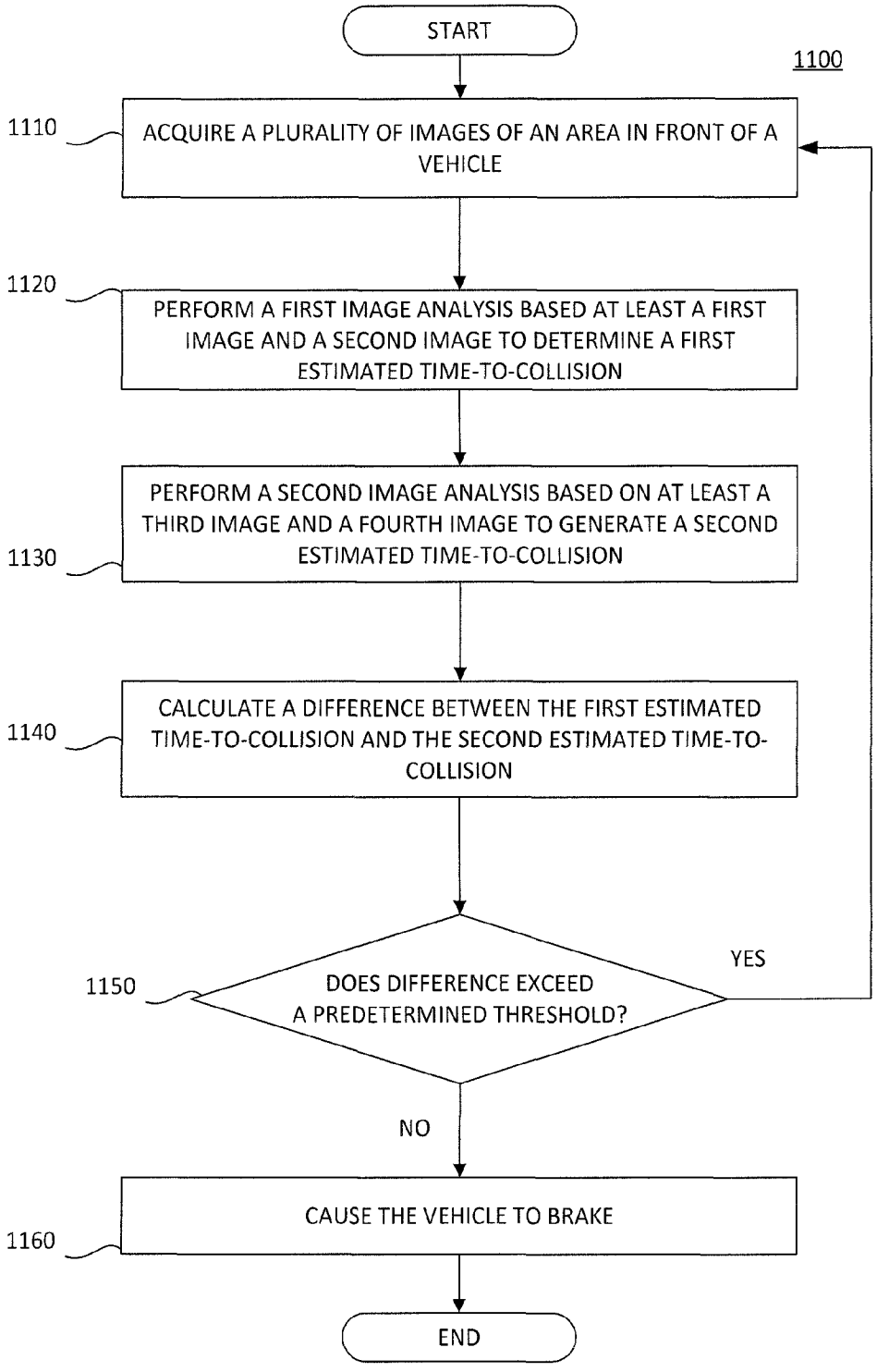


FIG. 11

IPR2025-01034
Tesla EX1002 Page 1509

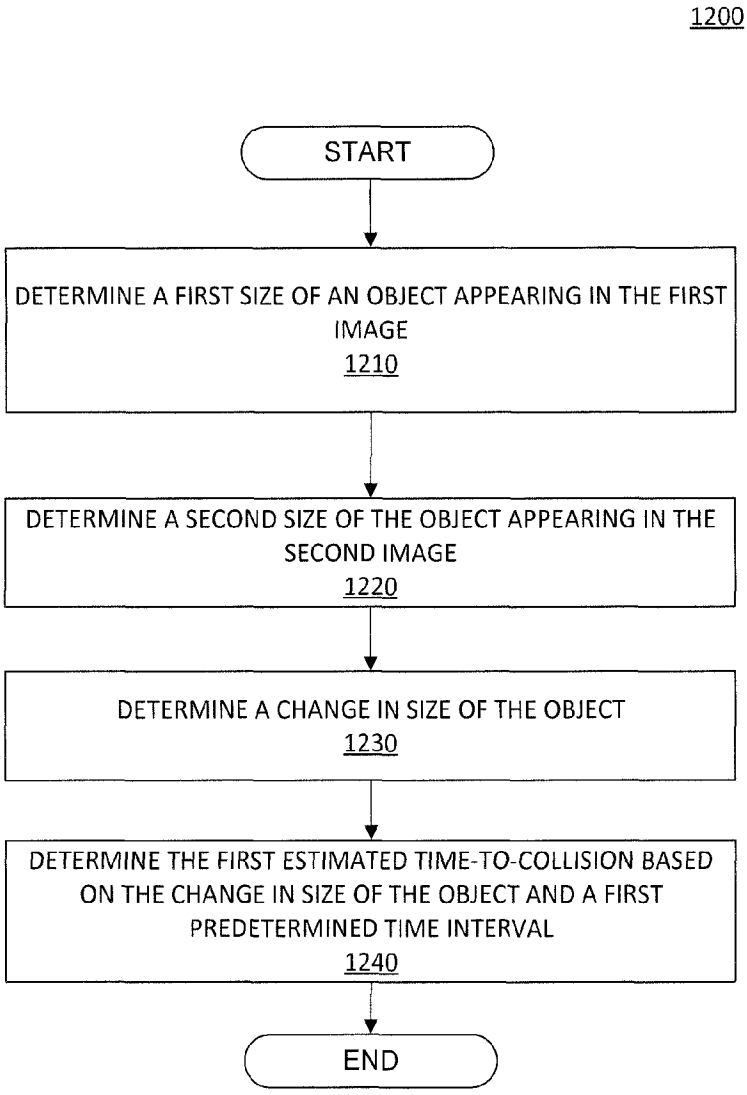


FIG. 12

IPR2025-01034
Tesla EX1002 Page 1510

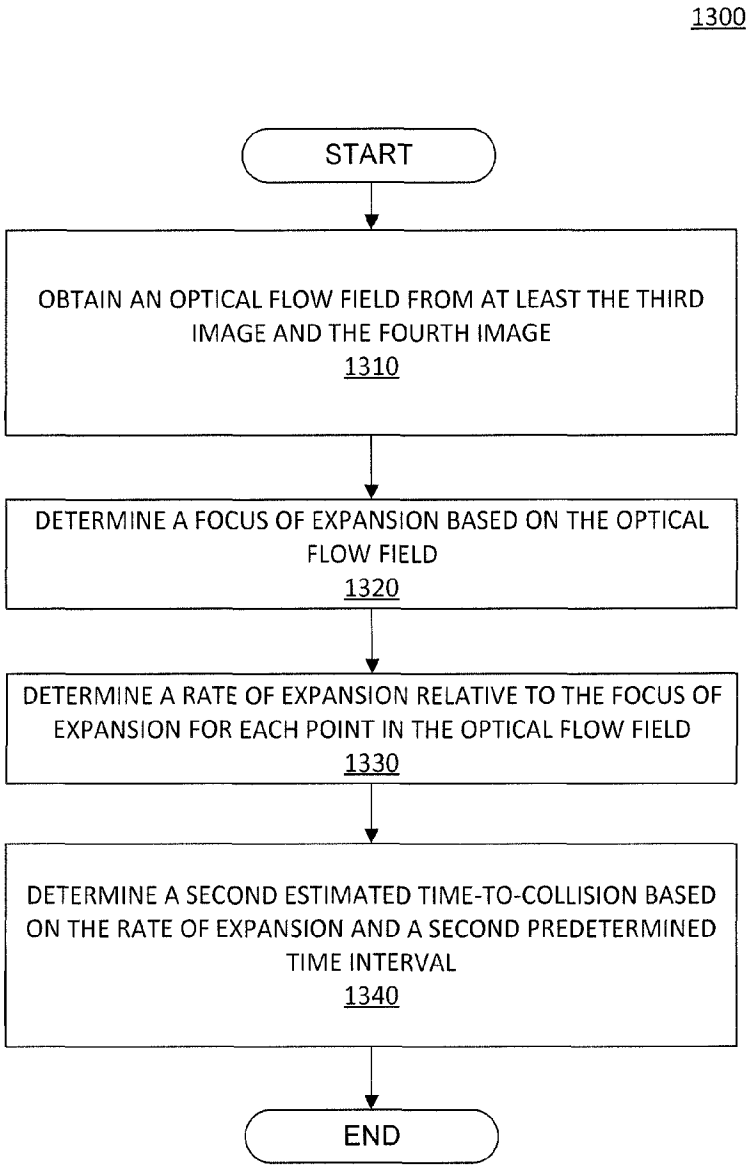


FIG. 13

IPR2025-01034
Tesla EX1002 Page 1511

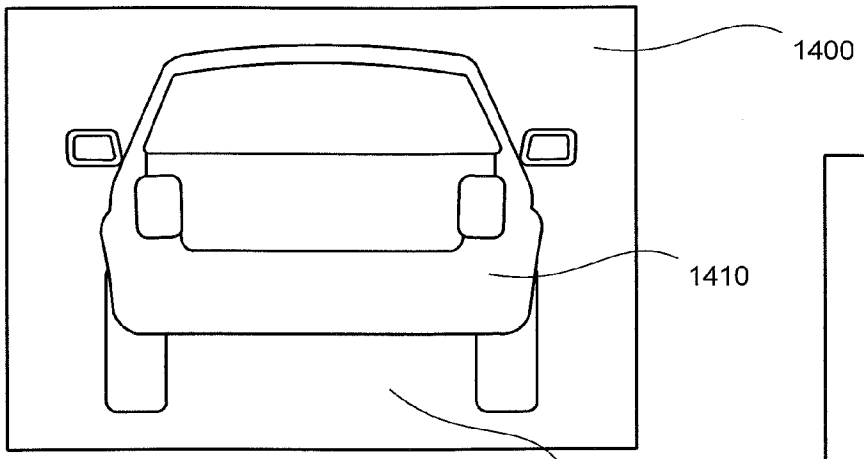


FIG. 14A

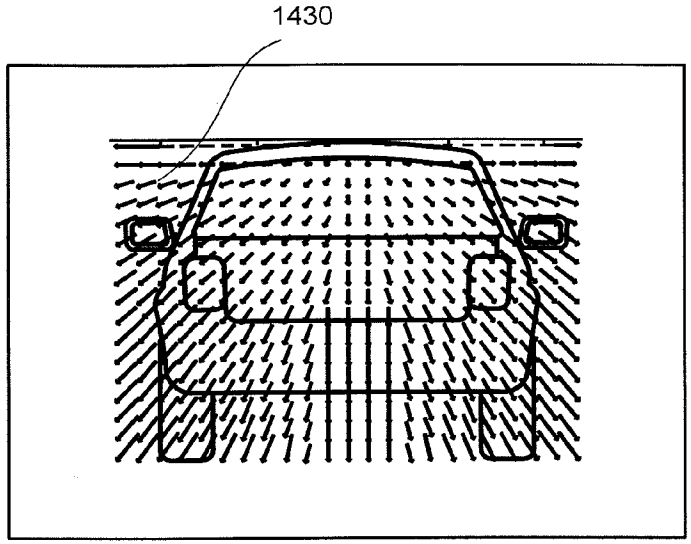


FIG. 14B

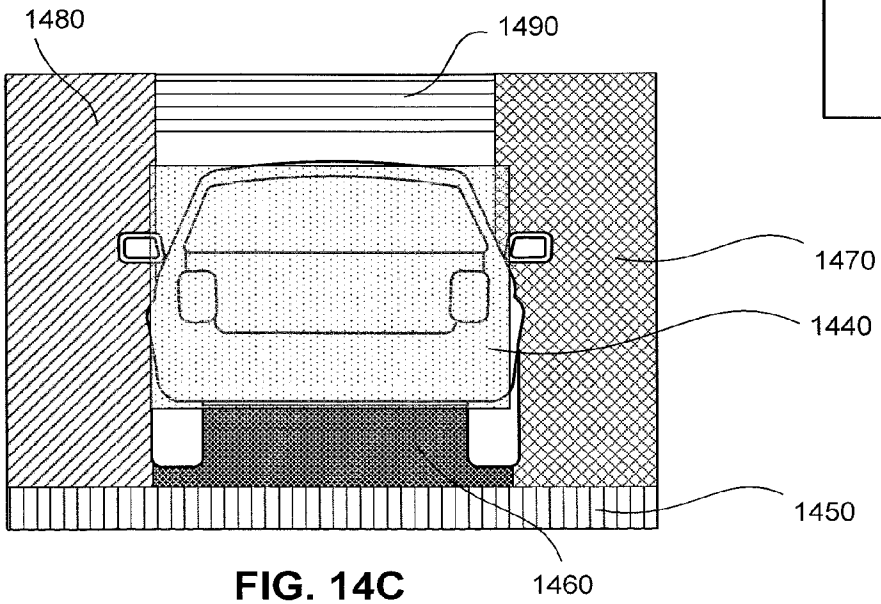


FIG. 14C

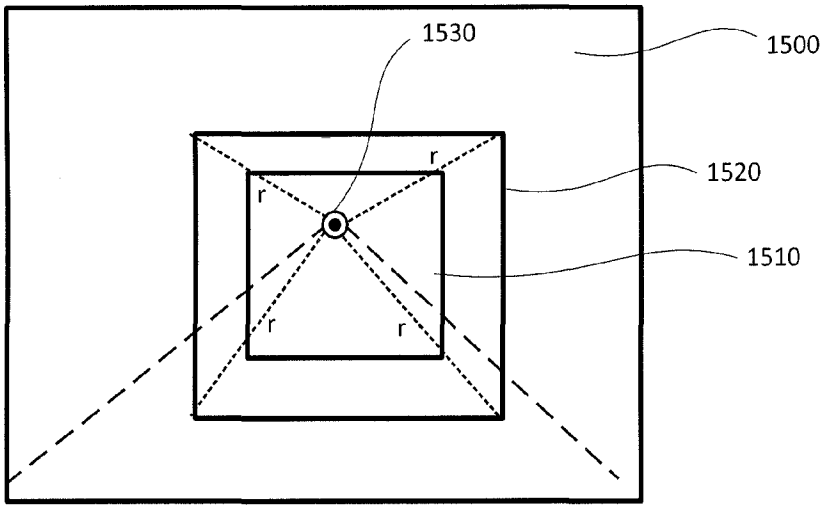


FIG. 15A

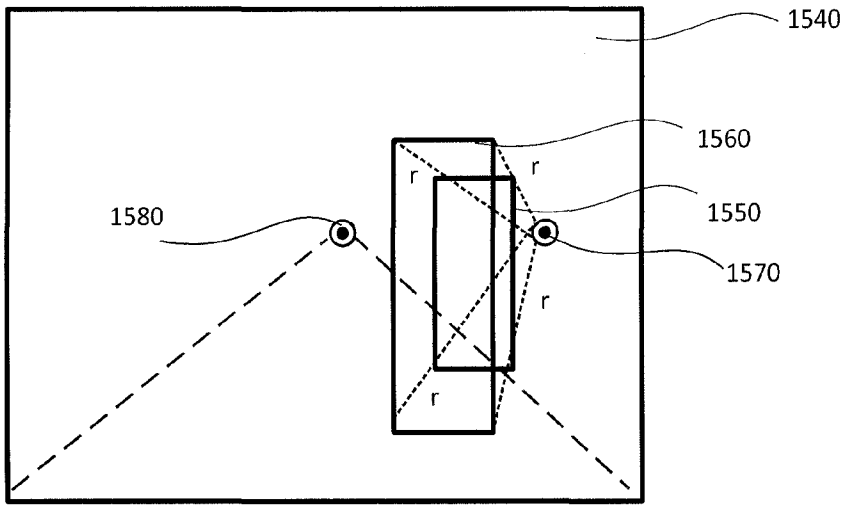


FIG. 15B

IPR2025-01034
Tesla EX1002 Page 1513

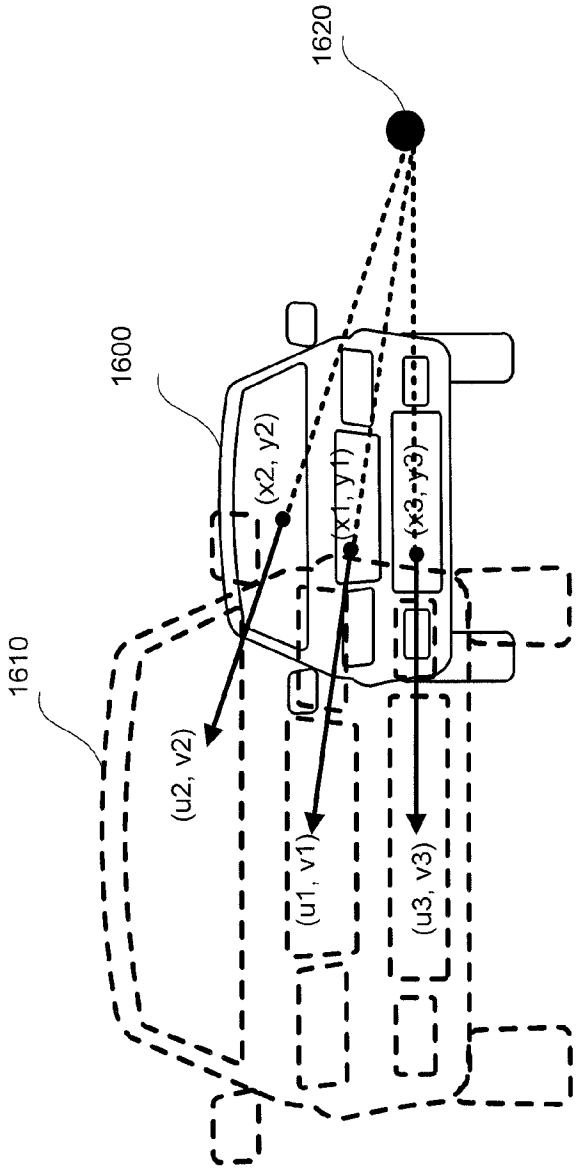


FIG. 16

IPR2025-01034
Tesla EX1002 Page 1514

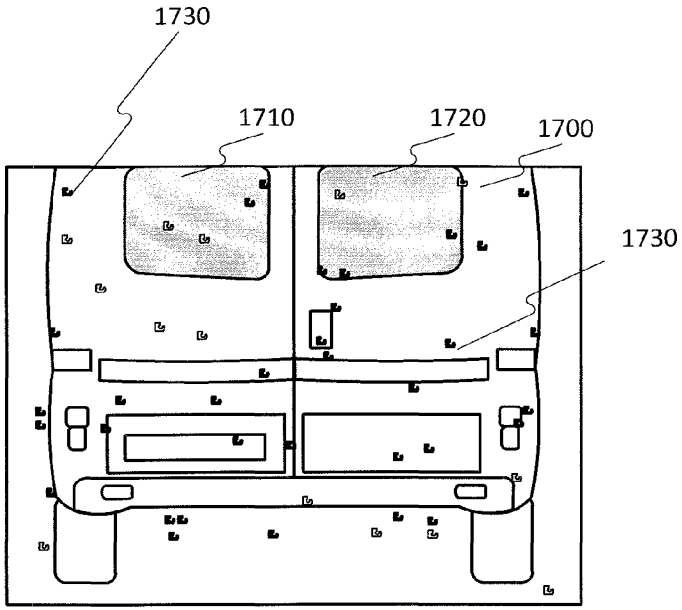


FIG. 17A

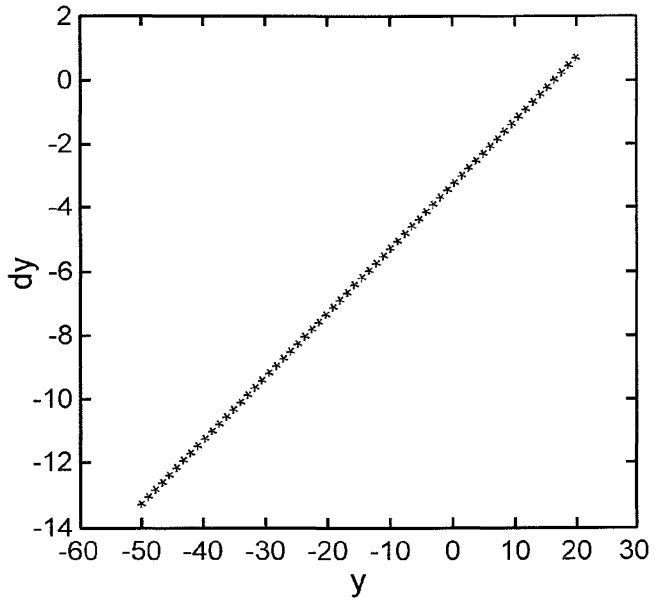


FIG. 17B

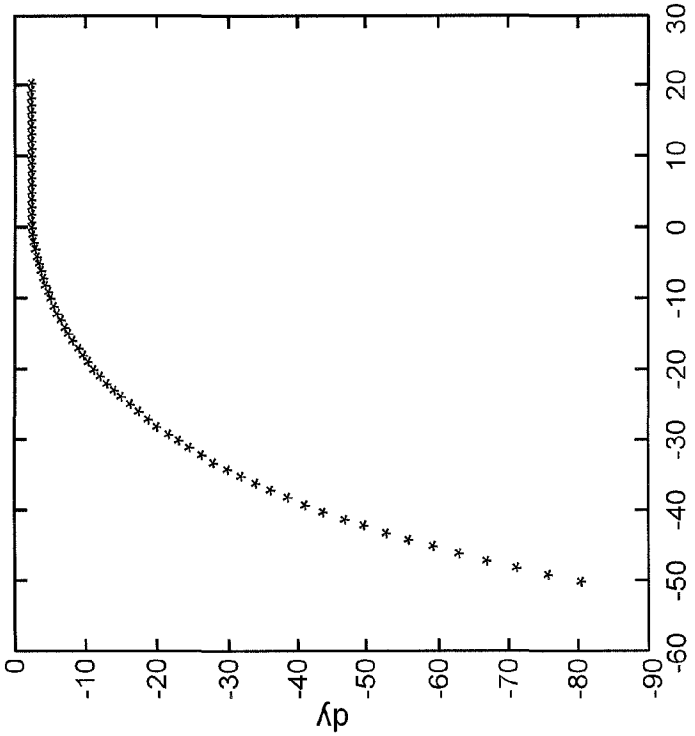


FIG. 18B

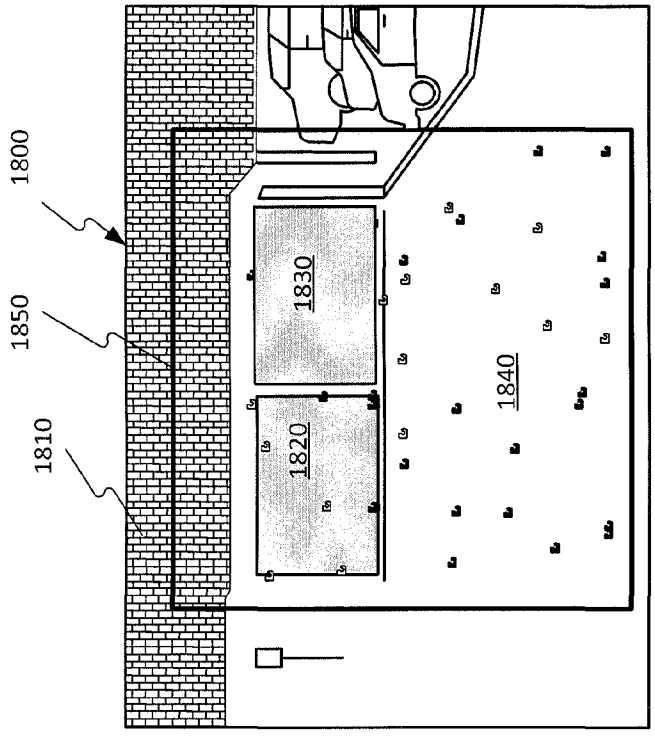


FIG. 18A

IPR2025-01034
Tesla EX1002 Page 1516

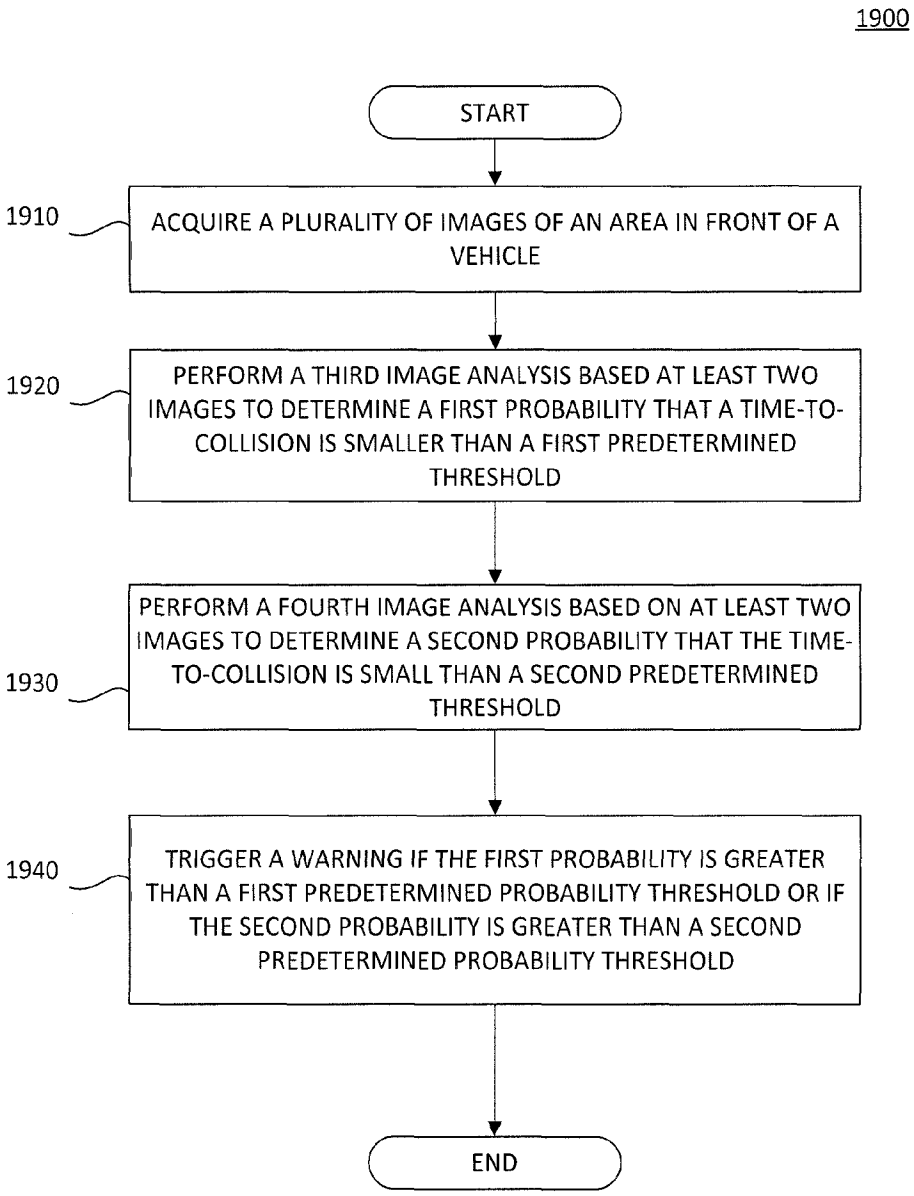


FIG. 19

IPR2025-01034
Tesla EX1002 Page 1517

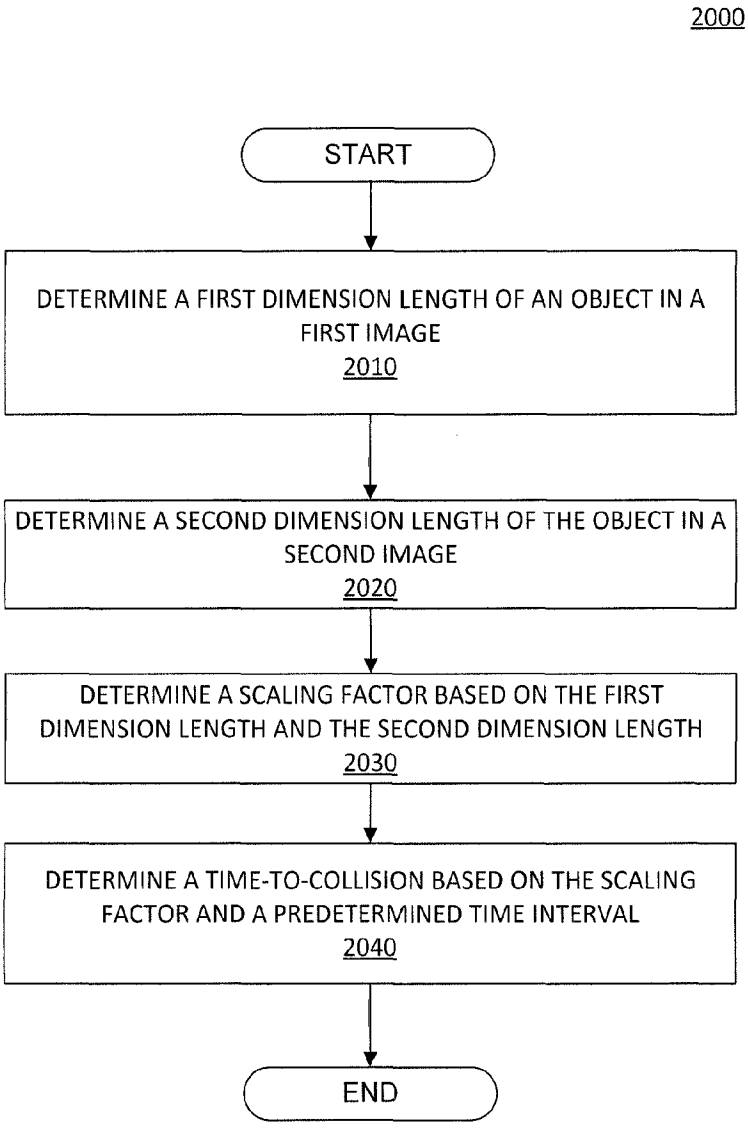


FIG. 20

IPR2025-01034
Tesla EX1002 Page 1518

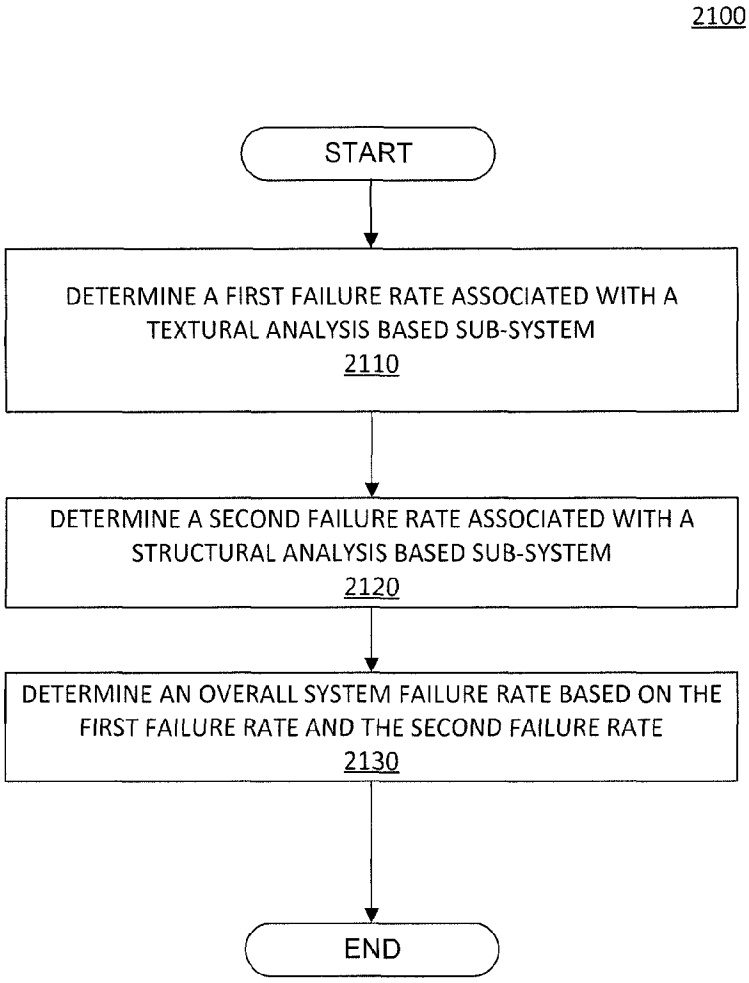


FIG. 21

IPR2025-01034
Tesla EX1002 Page 1519

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]**

IPR2025-01034
Tesla EX1002 Page 1520

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

IPR2025-01034
Tesla EX1002 Page 1521

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.

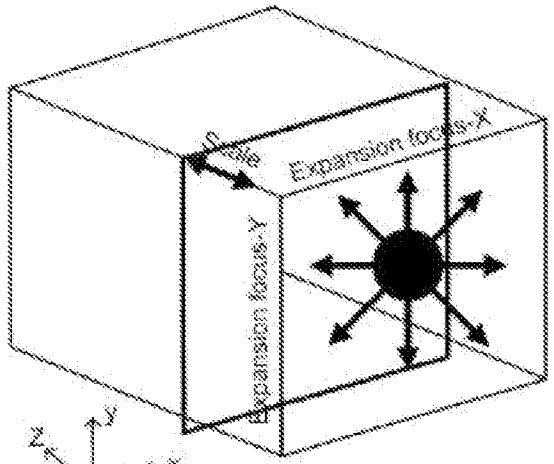


FIG. 10B

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 1524 IPR2025-01034

[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

IPR2025-01034
Tesla EX1002 Page 1526

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

IPR2025-01034 Tesla EX1002 Page 1528

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.

IPR2025-01034
Tesla EX1002 Page 1530

[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 1531 IPR2025-01034

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,

IPR2025-01034
Tesla EX1002 Page 1532

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).

IPR2025-01034
Tesla EX1002 Page 1534

IPR2025-01034
Tesla EX1002 Page 1535

[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,

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 1536 IPR2025-01034

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

IPR2025-01034
Tesla EX1002 Page 1537

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

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

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 1540 IPR2025-01034

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

IPR2025-01034
Tesla EX1002 Page 1541

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.

IPR2025-01034
Tesla EX1002 Page 1542

[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

IPR2025-01034
Tesla EX1002 Page 1544

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

IPR2025-01034
Tesla EX1002 Page 1547

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.

[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

IPR2025-01034
Tesla EX1002 Page 1549

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 1550 IPR2025-01034

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 1551 IPR2025-01034

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

IPR2025-01034
Tesla EX1002 Page 1553

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.

IPR2025-01034
Tesla EX1002 Page 1554

[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 1555 IPR2025-01034

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

IPR2025-01034
Tesla EX1002 Page 1556

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 1557 IPR2025-01034

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.

IPR2025-01034
Tesla EX1002 Page 1558

IPR2025-01034
Tesla EX1002 Page 1559

[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,

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.

IPR2025-01034
Tesla EX1002 Page 1560

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 1562 IPR2025-01034

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

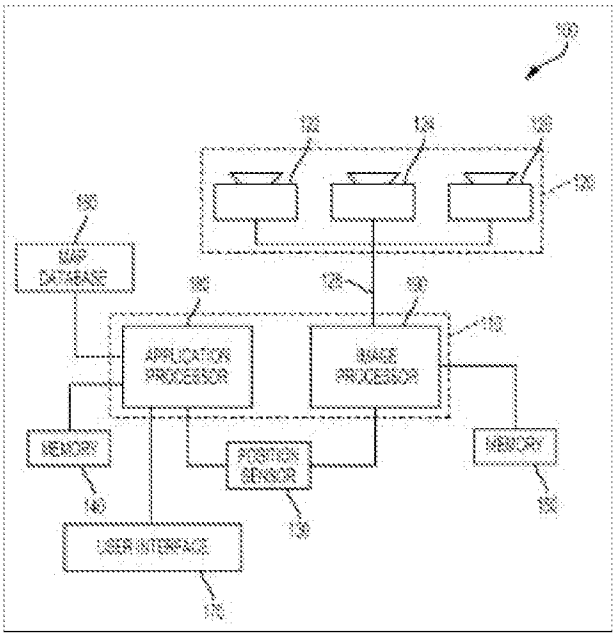
Priority number (s): US201462010003P 20140610 ; US201562173216P 20150609 ; WO2015IL50590 20150610

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.

IPR2025-01034
Tesla EX1002 Page 1566



IPR2025-01034
Tesla EX1002 Page 1567

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

IPR2025-01034
Tesla EX1002 Page 1569

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.

Tesla EX1002 Page 1570 IPR2025-01034

[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

IPR2025-01034
Tesla EX1002 Page 1572

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

IPR2025-01034
Tesla EX1002 Page 1573

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

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 1574 IPR2025-01034

[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 1575 IPR2025-01034

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

IPR2025-01034
Tesla EX1002 Page 1576

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.

IPR2025-01034
Tesla EX1002 Page 1577

[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 1578 IPR2025-01034

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 1579 IPR2025-01034

[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.

IPR2025-01034
Tesla EX1002 Page 1580

[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

IPR2025-01034
Tesla EX1002 Page 1581

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

IPR2025-01034
Tesla EX1002 Page 1582

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

IPR2025-01034
Tesla EX1002 Page 1583

memory 140, one of skill in the art will recognize that instructions may be stored in memory 140 and/or 150.

[0094] 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.

[0095] 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.

[0096] 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.

[0097] 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

Tesla EX1002 Page 1584 IPR2025-01034

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.

[0098] 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.

[0099] 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.

[0100] 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

Tesla EX1002 Page 1585 IPR2025-01034

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.

[0101] 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.

[0102] 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.

[0103] 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.

[0104] 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.

IPR2025-01034
Tesla EX1002 Page 1586

[0105] 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.

[0106] 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.

[0107] 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.

[0108] 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

Tesla EX1002 Page 1587 IPR2025-01034

vehicle 200 by tracking a set of feature points in the one or more images.

[0109] 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.

[0110] 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.

[0111] 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.

[0112] 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.

IPR2025-01034
Tesla EX1002 Page 1588

[0113] 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.

[0114] 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).

[0115] 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)

[0116] At step 574, processing unit 110 may determine a look-ahead point (expressed in coordinates as (x1, z1)) 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.

[0117] At step 576, processing unit 110 may determine a heading error and yaw rate

IPR2025-01034
Tesla EX1002 Page 1589

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 (x1/z1). 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).

[0118] 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.

[0119] 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.

[0120] 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

IPR2025-01034
Tesla EX1002 Page 1590

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.

[0121] 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.

[0122] 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.

[0123] 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,

IPR2025-01034
Tesla EX1002 Page 1591

IPR2025-01034
Tesla EX1002 Page 1592

movement characteristics, etc. of features associated with an object appearing in one or both of the image streams.

[0124] 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.

[0125] 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.

[0126] 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.

[0127] 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."

[0128] 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.

[0129] 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.

[0130] 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.

[0131] Top-Down Refinement

[0132] System 100 may provide vehicle 200 with top-down refinement of a lane marking model to improve lane marking detection. Top-down refinement of an updated model, before it is matched with bottom-up detection, can lead to more accurate results than matching an updated model, without top-down refinement, with bottom-up detection.

[0133] In road processing an updated multi-frame analysis may be combined with

IPR2025-01034
Tesla EX1002 Page 1593

bottom-up detection to build a model (e.g., parametric) of lane markings. The multi-frame analysis may update itself by relying on camera odometry to estimate, based on the model data, a predicted position of the lane markings at a subsequent time. The camera odometry, such as ego-motion or sensor data such as speed and yaw rate, may indicate how far the camera has moved relative to the roadway since a prior frame. The model of lane markings may comprise mathematical constructs, such as polynomials, to represent the lane markings on a roadway. Multi-frame analysis can be viewed as a model layer analysis that need not process raw image data (e.g., pixels), but rather prior model information. The multi-frame analysis may additionally or alternatively utilize any suitable

[0134] A bottom-up detection system may comprise a single frame analysis that generates a model of lane markings based on image data. The model generated from the single frame analysis can then be matched to the lane marking model to redetect, eliminate of make new lane markings in the lane marking model. Single frame analysis can be viewed as an image layer analysis that processes raw image data. For example, it may include identifying edgelets (e.g., using the Sobel method and non-maximal suppression), creating segments (e.g., using local Hough-peaks and iteratively reweighted least squares method), growing a graph (e.g., using a statistically tuned semi-greedy scan line algorithm—approximating best subset choosing).

[0135] It should be understood that for the purposes of the present disclosure each frame analyzed in either the multi-frame or single frame analysis need not be limited to image data from only one camera at a particular time but can also comprise image data derived from multiple cameras at a particular time.

[0136] FIGS. 8A-8D are diagrammatic representations of exemplary lane detection analyses without top-down refinement, such as those that may be implemented by a bottom-up detection system. FIGS. 8A and 8B illustrates how well lane markings can be detected by a bottom-up detection system when an image of a roadway in a vicinity of a vehicle is in good condition (e.g., clear weather conditions, clear windshield in front of camera(s), new road with no tar marks/seams, bright and clear lane markings). FIG. 8A illustrates an image of a roadway in good condition, with lane markings comprising solid lines 802, 804 and 806 and dashed line 800. FIG. 8B shows sections 812, 814 and 816 as identified by the bottom-up detection system, accurately representing solid line 802, 804 and 806, and section 810 as identified by the system, accurately representing dashed line 800. The accurate identification of sections 810, 812, 814 and 816 result in solid lines 802, 804 and 806 and dashed line 800 being correctly detected by the system.

[0137] However, lane detection can become more difficult when an image of a roadway in a vicinity of a vehicle comprises a tough scene, such as those having low SNR (signal to noise ration), low sun, shadows, washed-out markings, curves—especially with dashed lines, and only a single or partial dash of a dashed line in view. FIG. 8C illustrates an image of a roadway in a tough scene, with lane markings comprising curved dashed lines 820 and 822, light pole shadows 826 and 828, and sun glare region 824 on the vehicle windshield in front of the camera(s). FIG. 8D shows sections 830 and 832 as identified by the bottom-up detection system, accurately representing dashed lines 820 and 822, but sections 834, 836 and 838 are also identified by the system, which represent sun glare region 824 and light pole shadows 826 and 828. The accurate identification of sections 830 and 832, but the inaccurate identification of sections 834, 836 and 838, result in dashed lines 820 and 822 being only partially

correctly detected by the system. For example, the system may inaccurately detect dashed lines 820 and 822 as being longer than they actually are.

[0138] This example can reflect problems with bottom-up detection, such as solving a big combinatorial search, needing to be cheap and greedy (e.g., by setting limits on edgelets, segments, and graph growing), causing missing and unclean paths, making it difficult to classify objects (especially dashes), and sometimes costing too much. Bottom-up detection is designed to do bottom up processing that assumes very little about the scene (e.g., mostly just calibration and horizon), and tries to build a graph of paths representing the road structures for each frame (e.g., painted lanes, road edge, barriers) independently. These structures are then linked up (e.g., using a more constrained 3<rd >degree polynomial model) over time to form the multi-frame candidates which may then be approved and become part of the lane marking model.

[0139] This type of bottom-up detection system does a lot of work (since it assumes very little) and has to take shortcuts, mostly in the form of making early greedy decisions. However, once the system has a set of good candidates (from the previous frame, for example), the system can do better by going back into the image and verifying, extending, and refining those candidates. This is a goal of refinement.

[0140] FIG. 9 is a diagrammatic representation of an exemplary top-down refinement architecture consistent with the disclosed embodiments. As shown in FIG. 9, multi-frame modeling 900 may undergo a top-down refinement (902) of its update of the lane marking model in the image layer, prior to matching (906) in the model layer against the single frame modeling generated from bottom-up detection (904). By introducing refinement to the updated lane marking model, system 100 can search the image at a location that the multi-frame analysis expects lane marks to be, and it can search the image for a lane mark appearance known by system 100. The appearance model can comprise any suitable feature or combination of features, such as dimensions (e.g., width, cross-section, distance along its length, profile, shape), texture, sharpness of edges, brightness, etc. In other words, top-down refinement allows system 100 to use the history of what a lane marking looked like in the past in order to help identify the candidate lane marks in the current image. In the context of dashed lines, system 100 may use an appearance model in combination any other suitable feature, such as one or more of a spacing between the dashes, a lateral position of the dashes relative to the roadway, and the existence of a parallel lane mark on one side of a dash, which together forms a double lane mark (e.g., like a double yellow line in the U.S.). Top-down refinement can provide robustness by being semi-independent and provide cleaner detections, extended view range, more accurate and richer modeling and load balancing.

[0141] Based on the top-down refinement of the lane marking model, system 100 may 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 collected data based on the top-down refinement of the lane marking model and issue warnings and/or alerts to vehicle occupants based on the analysis of the collected data.

[0142] Any lane mark candidate proposal system which specifies an approximate

IPR2025-01034
Tesla EX1002 Page 1595

position (either parametrically or non-parametrically) for a given frame may constitute the input hypothesis that is being refined. Such a proposal may comprise additional information besides the approximate position of the lane mark that can then be used by the refinement algorithm (such as width, brightness, dash mark spacing, edge sharpness, etc.) but these are not necessary to run the refinement algorithm. Thus, such hypotheses can come from:
if no lane marks were found, then assumed lane mark location and orientation relative to vehicle 200 (e.g., parallel to the vehicle on either side)
if a lane mark was detected by some method somewhere in the image, the location of other lane marks that are parallel and are at some lane-width distance away can be hypothesized
any other suitable lane detection systems, and/or
any such algorithms (including refinement) that found lane marks on the previous frame and, using ego-motion, estimated their position on the current frame.

[0147] FIG. 10A is an exemplary block diagram of memory 140 or 150, which may store instructions for performing one or more operations consistent with disclosed embodiments. As illustrated in FIG. 10A, memory 140 or 150 may store one or more modules for performing the top-down refinement of a lane marking model described herein. For example, memory 140 or 150 may store a lane detection module 1000 and an action module 1002.

[0148] Lane detection module 1000 may store instructions which, when executed by processing unit 110, may perform top-down refinement of a lane marking model. For example, processing unit 110 may receive from one or more cameras 122, 124 and 126 one or more images of a roadway in a vicinity of a vehicle. The roadway may comprise a lane marking comprising a dashed line. Processing unit 110 may update a model of the lane marking based on odometry of the one or more cameras relative to the roadway, and 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.

[0149] The camera odometry may comprise any suitable mechanism to determine the distance traveled by the camera relative to the roadway, such as ego-motion, sensor data such as speed and yaw rate, and/or visual analysis of multiple successive image frames.

[0150] The refinement of the updated model may be further based on a lateral position, relative to the roadway, of dashes derived from the received one or more images. The refinement of the updated model may be further based on existence of a parallel lane marking on one side of dashes derived from the received one or more images. The appearance of the dashes can comprise any suitable identifying feature, such as one or more of a width of the dashes, a brightness of the dashes and an edge sharpness of the dashes. The refinement of the updated model may comprise a dynamic programming technique.

[0151] Processing unit 110 may also generate a model of the lane marking based on the received one or more images, and match the generated model to the refined model to approve the lane marking.

[0152] Action module 1002 may store instructions which, when executed by processing unit 110, may cause one or more navigational responses in the vehicle based on the

Tesla EX1002 Page 1596 IPR2025-01034

refinement of the updated model. The one or more one navigational responses may comprise a steering of the vehicle and/or an audible alert.

[0153] FIG. 10B illustrates a process 1050 for performing the top-down refinement of a lane marking model consistent with disclosed embodiments. According to some embodiments, process 1050 may be implemented by one or more components of navigation system 100, such as at least one processing unit 110. At step 1052, process 1050 may receive, 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 may comprise a lane marking comprising a dashed line. At step 1054, process 1050 may update, by one or more processors, a model of the lane marking based on odometry of the one or more cameras relative to the roadway. At step 1056, process 1050 may refine, 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. At step 1058, process 1050 may cause, by one or more processors, one or more navigational responses in the vehicle based on the refinement of the updated model.

[0154] In process 1050 the camera odometry may comprise any suitable mechanism to determine the distance traveled by the camera relative to the roadway, such as ego-motion, sensor data such as speed and yaw rate, and/or visual analysis of multiple successive image frames. The refinement of the updated model may be further based on a lateral position, relative to the roadway, of dashes derived from the received one or more images. The refinement of the updated model may be further based on existence of a parallel lane marking on one side of dashes derived from the received one or more images. The appearance of the dashes can comprise any suitable identifying feature, such as one or more of a width of the dashes, a brightness of the dashes and an edge sharpness of the dashes. The refinement of the updated model may comprise a dynamic programming technique. A model of the lane marking may also be generated based on the received one or more images, and the generated model may be matched to the refined model to approve the lane marking.

[0155] FIG. 11A illustrates a process 1100 for implementing top-down refinement consistent with the disclosed embodiments. FIGS. 11B-11N are diagrammatic representations of the top-down refinement implementation consistent with the disclosed embodiments.

[0156] According to some embodiments, process 1100 may be implemented by one or more components of navigation system 100, such as at least one processing unit 110. At step 1102, process 1100 may find refined pair points (RPPs). For example, system 100 may analyze received image data as shown in FIG. 11B and identify RPPs along the edges of the left and right lane markings, such as RPPs 1122, as shown in FIG. 11C. The vertical dots, such as boundary 1120, mark the search area around the predicted position of the lane marking where system 100 looks for RPPs based on the multi-frame analysis. The roadway in the image data may be straightened to provide a bird's eye view that does not converge at the horizon.

[0157] At step 1104, the detected RPPs may subsequently be clustered into sections, one of which is section 1130, as shown in FIG. 11D, and as indicated by the curly braces in FIG. 11E. The idea to use sections as an intermediate granularity object between local edge detections (e.g., edgelets) and the full lane mark "path" that is made up of a collection of segments. A key point is that the sections are locally linear

Tesla EX1002 Page 1597 IPR2025-01034

one-sided objects, that can be made by clustering local edge detections and can be described compactly using a position and orientation (together with optional appearance attributes such as edge strength). This choice may make both forming the sections from local evidence possible and may allow the growing algorithm that connects the patch sections (described below) to treat them as atomic units with a simple parametric representation.

[0158] At step 1106, process 1100 may match the created sections to previous frame sections. As shown in FIG. 11F, the arrows indicate that a newly created section (on the right side of the image) matches to a previously modeled frame section (on the left side of the image). The circled section on the right side of the image indicates a section that was not matched to a previously modeled frame section.

[0159] At step 1108, process 1100 may select the path sections to be connected into the best path for the lane markings. The image data in FIG. 11G is provided to illustrate how system 100 can start from the left image which comprises 7 candidate path sections (ordered from 1 to 7 from the bottom up) and end by selecting the best path (e.g., a combination of sections or candidate lane markings) as shown by the 3 section dashed line path in the right image.

[0160] To generalize the problem, the input comprises n sections (and a set of features on them) and the output comprises a path comprising a subset of k sections that best describe the lane marking. The optimal solution is too expensive because there are 2<n >possible solutions. A few features may be used to make it run fast and yet allow paths that can split and merge. For example, the algorithm may run from bottom to top of image. Each current path end may try to find the next best section that continues it. If a section looks like a very good connection, it makes a greedy decision to connect, otherwise it continues to look for the best one. Additionally, a reverse process may find segments that were not chosen by paths from below and try to connect backwards to a good path that they connect well to, but happened not have been the best connection for that path when growing forward. This may allow for the finding of lane marking that split (e.g., around a traffic island).

[0161] Thus, the problem of the high cost of the optimal solution can be solved by defining a simpler problem and solving it by dynamic programming, which has n<2 >in the worst case). An example dynamic programming algorithm follows:

Define the connection ends at section #1. (0~>1)

For each section i (=2:n),

Go over j (=0:i−1) previous sections,

Define the path-section connection ends at j~>i

Find the best connection ends at section i.

IPR2025-01034 Tesla EX1002 Page 1598

[0162] In other words, find the best path that ends at section k given the best paths that end at sections (1, . . . , k−1). The best path is then chosen by choosing the end section.

[0163] FIGS. 11H, 11I and 11J each show the same seven path sections that reflect the path section candidates for the dashed line of the roadway shown in the image of FIG. 11F. FIG. 11H shows the best paths (determined from the running of the dynamic programming algorithm) at each section, whereas FIG. 11I shows the best path for section 7 only. FIG. 11J shows part of the process of the algorithm when selecting section 6. This example shows connections being attempted between path section 6 and path sections 5, 4 and 3, respectively.

[0164] These tests may utilize any suitable number of classifiers, such as a step classifier, first classifier, terminal classifier, optimal classifier and path classifier. Features of the step classifier may include sec, sec-sec, and sec-paty, appearance features (e.g., brightness, width, edge sharpness, etc.), lateral-geometry (orientation, position, etc.) and longitudinal model (dash/solid-painted/gap lengths. The tests may also take into account one or more of appearance, spacing between sections and lateral position (spline). FIG. 11K shows an example where the attempted connecting of sections 19 and 20 fail because the lateral position test veers the path sharply to the left, whereas FIG. 11L shows the attempted connecting of sections 19 and 21 which passes, since it provides a better fit with the sections below it.

[0165] At step 1110, process 1100 may estimate line intensity model of the connected path sections. System 100 can sample several scan lines along the path, estimate the cross section model parameters from each sample, and average the results to get a single model per path.

[0166] FIG. 11M shows the results of lane detection of the image data of FIG. 8C using a bottom-up detection system without top-down refinement. FIG. 11N shows the results of lane detection of the image data of FIG. 8C using system 100 with top-down refinement. As shown in FIG. 11M, the bottom-up detection system misses far dash marks such as dash 1142 and falsely detects glare on the windshield as lane mark 1140. However, FIG. 11N shows that with refinement these problems are resolved due to the use of the cross section model which matches the far dash marks and doesn't match the glare.

[0167] 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.

[0168] 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

IPR2025-01034
Tesla EX1002 Page 1599

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. Further, it is to be understood that the functionality of the program modules described herein is not limited to the particular type of program module named herein, but rather can be programmed in any suitable manner across among one or more modules or other software constructs.

[0169] 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.

IPR2025-01034
Tesla EX1002 Page 1600

Espacenet

Claims: EP3154835 (A1) — 2017-04-19

TOP-DOWN REFINEMENT IN LANE MARKING NAVIGATION

Claims not available for EP3154835 (A1)
Claims of corresponding document: US2015354976 (A1)

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

US10317231 (B2) WO2015189847 (A1) US2020003573 (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 computer system comprising:
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.

2. The computer system of claim 1, wherein the refinement of the updated model is further based on a lateral position, relative to the roadway, of dashes derived from the received one or more images.

3. The computer system of claim 1, wherein the refinement of the updated model is further based on existence of a parallel lane marking on one side of dashes derived from the received one or more images.

4. The computer system of claim 1, wherein the appearance of the dashes comprises a width of the dashes.

5. The computer system of claim 1, wherein the appearance of the dashes comprises a brightness of the dashes.

6. The computer system of claim 1, wherein the appearance of the dashes comprises an edge sharpness of the dashes.

7. The computer system of claim 1, wherein the refinement of the updated model comprises a dynamic programming technique.

8. The computer system of claim 1, wherein the one or more processors are configured to execute the instructions to cause the system to:
generate a model of the lane marking based on the received one or more images, and
match the generated model to the refined model to approve the lane marking.

9. The computer system of claim 1, wherein the one or more one navigational responses comprises a steering of the vehicle.

10. The computer system of claim 1, wherein the one or more one navigational responses comprises an audible alert.

11. A vehicle, comprising:
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 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.

12. The vehicle of claim 11, wherein the refinement of the updated model is further based on a lateral position, relative to the roadway, of dashes derived from the received one or more images.

13. The vehicle of claim 11, wherein the refinement of the updated model is further based on existence of a parallel lane marking on one side of dashes derived from the received one or more images.

14. The vehicle of claim 11, wherein the appearance of the dashes comprises a width of the dashes.

15. The vehicle of claim 11, wherein the appearance of the dashes comprises a brightness of the dashes.

16. The vehicle of claim 11, wherein the appearance of the dashes comprises an edge sharpness of the dashes.

Tesla EX1002 Page 1602 IPR2025-01034

17. The vehicle of claim 11, wherein the refinement of the updated model comprises a dynamic programming technique.

18. The vehicle of claim 11, wherein the one or more processors are configured to execute the instructions to cause the vehicle to:
generate a model of the lane marking based on the received one or more images, and
match the generated model to the refined model to approve the lane marking.

19. The vehicle of claim 11, wherein the one or more one navigational responses comprises a steering of the vehicle.

20. The vehicle of claim 11, wherein the one or more one navigational responses comprises an audible alert.

21. A method comprising:
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 processors, one or more navigational responses in the vehicle based on the refinement of the updated model.

22. The method of claim 21, wherein the refinement of the updated model is further based on a lateral position, relative to the roadway, of dashes derived from the received one or more images.

IPR2025-01034
Tesla EX1002 Page 1603

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

(11) Veröffentlichungsnummer:
(11) Publication number: **EP 3 154 835 A0**
(11) Numéro de publication:

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

WO 2015/189847 (Art. 153(3) EPÜ).

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

WO 2015/189847 (Art. 153(3) EPC).

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

WO 2015/189847 (art. 153(3) CBE).

Espacenet

Bibliographic data: EP3183688 (A1) — 2017-06-28

RECOGNITION AND PREDICTION OF LANE CONSTRAINTS AND CONSTRUCTION AREAS IN NAVIGATION

Inventor(s): AHARONY HILA [IL]; DYSHEL MARIA [IL]; HAREL AMIR [IL]; RABANI MEITAL [IL]; SHAMAH SHAKED [IL]; HORN KOBI [IL] ± (AHARONY, Hila, ; DYSHEL, Maria, ; HAREL, Amir, ; RABANI, Meital, ; SHAMAH, Shaked, ; HORN, Kobi)

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

Classification: **- international:** ***G06K9/00***
- cooperative: **B60W10/04 (US); B60W10/18 (US); B60W10/20 (US); B60W30/0953 (US); B60W30/0956 (US); G06K9/00798 (EP, CN, US); G08G1/09623 (EP, US);** B60W2420/403 (US); B60W2710/18 (US); B60W2710/20 (US); B60W2720/10 (US)

Application number: EP20150793913 20150817 Global Dossier

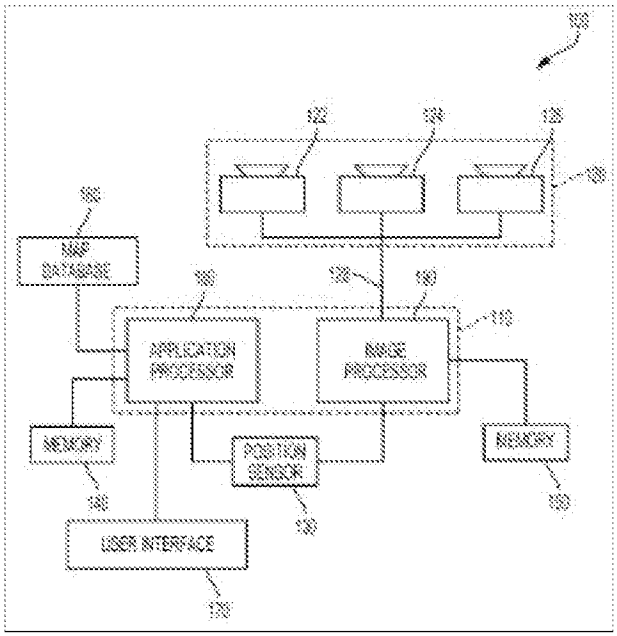
Priority number(s): US201462038426P 20140818 ; WO2015IL50833 20150817

Also published as: CN107111742 (A) CN107111742 (B) CN113158820 (A) US10150473 (B2) US10926763 (B2) more

Abstract not available for EP3183688 (A1)
Abstract of corresponding document: US2016046290 (A1)

Systems and methods use cameras to provide autonomous and/or driver-assist navigation features. In some implementations, techniques for predicting the location of first roadway lane constraints are provided. The system may receive multiple images of a roadway in a vicinity of a vehicle, recognize a first roadway lane constraint, and, when lane prediction conditions are determined to be satisfied, predict a location of a second roadway lane constraint. In some implementations, techniques for detecting and responding to construction zones are provided. The system may receive multiple images of a roadway in a vicinity of a vehicle, recognize indicators of a construction zone in the images, determine that the vehicle is proximate to a construction zone, and output a signal indicating that the vehicle is proximate to a construction zone.

IPR2025-01034
Tesla EX1002 Page 1605


MAP DATABASE
APPLICATION PROCESSOR
BASE PROCESSOR
MEMORY
POSITION SENSOR
MEMORY
USER INTERFACE

Espacenet

Description: EP3183688 (A1) — 2017-06-28

RECOGNITION AND PREDICTION OF LANE CONSTRAINTS AND CONSTRUCTION AREAS IN NAVIGATION

Description not available for EP3183688 (A1)
Description of corresponding document: US2016046290 (A1)

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

CN107111742 (A) US2016046290 (A1) WO2016027270 (A1) CN113158820 (A) US2019077400 (A1) US2021291827 (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. 62/038,426, the entirety of which is incorporated herein by reference.

FIELD OF THE DISCLOSURE

[0002] This disclosure 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 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] Current autonomous or driver-assist (e.g., partially autonomous) vehicle systems may be unable to adequately identify a lane when lane lines (e.g., the marked lines that divide lanes) or other markings do not clearly indicate the locations of lane constraints on a roadway. For example, current systems may not function properly when a lane constraint cannot be identified from lines or markings detected in received images of a roadway; or, current systems may inaccurately identify a lane constraint where a lane constraint does not in reality exist. Current autonomous or driver-assist vehicle systems may also be unable to adequately identify or respond to construction zones on roadways, where modified vehicle control may be required, and where responding to certain roadway elements in a different manner may be advantageous. The inventors have recognized these and other problems which would be apparent from the present disclosure, which are addressed by the techniques disclosed herein.

[0005] Some embodiments consistent with the present disclosure provide systems and methods for autonomous or driver-assist vehicle navigation. Some disclosed embodiments may use cameras to provide autonomous vehicle navigation features. For example, consistent with some 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.

[0006] In accordance with some embodiments, a computer system comprises a memory that stores instructions, and a processor that executes the instructions to cause the system to: receive, from a camera, multiple images of a roadway in a vicinity of a moving vehicle, recognize, based on the multiple images, a first roadway lane constraint, determine, based on the multiple images, that one or more defined lane prediction conditions are satisfied, in accordance with a determination that one or more defined lane prediction conditions are satisfied, predict a location of a second roadway lane constraint, and enable the moving vehicle to avoid the first roadway lane constraint and the second roadway lane constraint.

[0007] In accordance with some embodiments, a computer system comprises a memory that stores instructions and a processor that executes the instructions to cause the system to: receive, from a camera, multiple images of a roadway in a vicinity of a moving vehicle, recognize, in the multiple images, one or more indicators of a construction zone, determine, based on the one or more indicators, that the vehicle is proximate to a construction zone, and, in accordance with a determination that the vehicle is proximate to a construction zone, output a signal indicating that the vehicle is proximate to a construction zone.

BRIEF DESCRIPTION OF THE DRAWINGS

[0008] FIG. 1 is a diagrammatic representation of an exemplary system consistent with

Tesla EX1002 Page 1608 IPR2025-01034

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.

[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

Tesla EX1002 Page 1609 IPR2025-01034

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-8C are depictions of a vehicle traveling along a roadway in accordance with some embodiments.

[0029] FIGS. 9A-9C are depictions of a vehicle traveling along a roadway in accordance with some embodiments.

[0030] FIG. 10A is a depiction of a perspective view from a camera associated with a system in accordance with some embodiments.

[0031] FIG. 10B is a depiction of a vehicle traveling along a roadway in accordance with some embodiments.

[0032] FIG. 11 is a diagrammatic representation of the memory of an exemplary system consistent with some disclosed embodiments.

[0033] FIG. 12 is a flowchart of an exemplary process in accordance with some embodiments.

[0034] FIG. 13 is a depiction of a vehicle traveling along a roadway in accordance with some embodiments.

[0035] FIG. 14 is a diagrammatic representation of the memory of an exemplary system consistent with some disclosed embodiments.

[0036] FIG. 15 is a flowchart of an exemplary process in accordance with some embodiments.

DETAILED DESCRIPTION

[0037] 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

IPR2025-01034
Tesla EX1002 Page 1610

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.

[0038] Disclosed embodiments provide systems and methods that use cameras to provide autonomous navigation, vehicle control, and/or driver assist technology features. Driver assist technology, as opposed to fully autonomous driving, refers to any suitable technology to assist drivers in the navigation and control of their vehicles, such as LKA (lane keeping assist), LDW (lane departure warning), acceleration of the vehicle, deceleration of the vehicle, steering of the vehicle, controlling vehicle braking, preparing the vehicle for an unavoidable collision, and/or any other suitable manner of controlling any aspect of the operation of the vehicle or assisting the driver in controlling any such aspect. In various embodiments, the system may include one, two or more cameras that monitor the environment of a vehicle. In some embodiments, the system may provide techniques for predicting the location of roadway lane constraints and/or techniques for detecting and responding to construction zones.

[0039] 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.

[0040] 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.).

[0041] 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

IPR2025-01034
Tesla EX1002 Page 1611

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 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.

[0042] 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.

[0043] 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.

[0044] 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.

[0045] 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.

[0046] User interface 170 may include any device suitable for providing information to

IPR2025-01034
Tesla EX1002 Page 1612

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.

[0047] 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.

[0048] 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.).

[0049] 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.

[0050] 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.

[0051] The image capture devices included on vehicle 200 as part of the image

Tesla EX1002 Page 1613 IPR2025-01034

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.

[0052] 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.

[0053] 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.

[0054] 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.

[0055] 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.

[0056] 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.

Tesla EX1002 Page 1614 IPR2025-01034

[0057] 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, motorcycles, bicycles, self-balancing transport devices and other types of vehicles.

[0058] 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.

[0059] 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.

[0060] 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.

[0061] 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.

[0062] 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

IPR2025-01034
Tesla EX1002 Page 1615

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.

[0063] 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 different than (e.g., 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.

[0064] 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.

[0065] 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.

[0066] 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).

[0067] 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

IPR2025-01034
Tesla EX1002 Page 1616

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.

[0068] 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.

[0069] 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.

[0070] 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.

[0071] 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

Tesla EX1002 Page 1617 IPR2025-01034

image capture device 122.

[0072] 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.

[0073] 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.

[0074] 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.

[0075] 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).

[0076] In some embodiments, a system may define the “vicinity” of the vehicle according to predefined constraints or dynamically determined constraints. The vicinity of the vehicle may, for example, be considered any area within a certain distance of the vehicle in any one or more directions, such as 1 meter, 5 meters, 10 meters, 25 meters, 50 meters, 100 meters, 250 meters, 500 meters, or 1000 meters. In some embodiments, the area defined as the vicinity of the vehicle may be irregularly shaped, such that it extends for different distances in different directions; the vicinity in the forward direction as defined by the vehicle's travel may, for example, extend for several hundred meters, while the vicinity in the backward or sideways directions may only extend, for example, for several dozen meters or less. The defined vicinity of the

IPR2025-01034
Tesla EX1002 Page 1618

vehicle may, in some embodiments, be varied dynamically according to the information captured by the one or more image capture devices, such that the vicinity may vary according to, for example, what parts of a roadway are visible in the FOV of the image capture devices. The vicinity of the vehicle may, in some embodiments, be dynamically varied according to, for example, any perceptible condition or state of the roadway or vehicle, including vehicle speed, posted speed limit, roadway conditions, roadway congestion, roadway size, visibility distance, environment brightness, time of day, presence of other vehicles, presence of pedestrians, or any other suitable factor.

[0077] 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.

[0078] 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.

[0079] 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.

[0080] 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 240 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.

[0081] As shown in FIG. 3A, vehicle 200 may also include a user interface 170 for

Tesla EX1002 Page 1619 IPR2025-01034

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.

[0082] 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.

[0083] 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, in response to this analysis, navigate and/or otherwise control and/or operate vehicle 200. Navigation, control, and/or operation of vehicle 200 may include enabling and/or disabling (directly or via intermediary controllers, such as the controllers mentioned above) various features, components, devices, modes, systems, and/or subsystems associated with vehicle 200. Navigation, control, and/or operation may alternately or additionally include interaction with a user, driver, passenger, passerby, and/or other vehicle or user, which may be located inside or outside vehicle 200, for example by providing visual, audio, haptic, and/or other sensory alerts and/or indications.

[0084] 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

IPR2025-01034
Tesla EX1002 Page 1620

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, indications, recommendations, alerts, or instructions to a driver, passenger, user, or other person inside or outside of the vehicle (or to other vehicles) based on the analysis of the collected data.

[0085] 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 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.

[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 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 (e.g., a depth map) 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 or with the depth information obtained by stereo processing.

[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 execute additional

Tesla EX1002 Page 1622 IPR2025-01034

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.,

Tesla EX1002 Page 1623 IPR2025-01034

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. In some embodiments, an environment of a vehicle may include, for any given frame, the area or space covered by any one or by any combination of the imaging devices (the FOV of the imaging devices). 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. Furthermore, based on the analysis, system 100 (e.g., via processing unit 110) may cause one or more navigational and/or vehicle control responses, including but not limited to the various manners of controlling vehicle 200 and any associated systems and/or subsystems discussed above. In some embodiments, any or all of the responses or actions mentioned above may be referred to, for simplicity, as a navigational action or a navigational response.

[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 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,

IPR2025-01034
Tesla EX1002 Page 1624

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 or of certain areas of the road (e.g., within a region of interest defined around image data which is suspected to show a road hazard). Processing unit 110 may then use the 3D-map to detect road hazards. For example, segmentation techniques and classification techniques can be used to detect the road surface, as well as hazards existing above the road surface. For example, as explained in additional detail below, a system may detect features of a road surface such as

IPR2025-01034
Tesla EX1002 Page 1625

painted lane lines or lane markings, and may additionally detect features of a roadway existing above the road surface, such as cones/barrels/poles/barriers in a construction zone, guard rails, other vehicles, road signs, and the like.

[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, size, position, velocity, and acceleration values (relative to vehicle 200) associated with the detected

IPR2025-01034
Tesla EX1002 Page 1626

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

Tesla EX1002 Page 1627 IPR2025-01034

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.

[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

IPR2025-01034
Tesla EX1002 Page 1628

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.

[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, in some embodiments, the distance di between two points in the set of points may fall in the range of 1 to 5 meters, or in further examples, a timing measure can be used to provide the set of points, whereby the points in the set of points are 0.5 second to 5 seconds apart from one another (e.g., 1.5 seconds apart). 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 (x1, z1)) 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

Tesla EX1002 Page 1629 IPR2025-01034

may determine the heading error by calculating the arctangent of the look-ahead point, e.g., arctan (x1/z1). 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).

[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

Tesla EX1002 Page 1630 IPR2025-01034

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.

[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

Tesla EX1002 Page 1631 IPR2025-01034

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 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

Tesla EX1002 Page 1632 IPR2025-01034

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."

[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] Lane Constraint Prediction

[0136] System 100 may provide vehicle 200 with lane constraint prediction techniques, such that the system predicts the location of one or more lane constraints.

[0137] In some embodiments, lane constraint prediction may be understood as one or more techniques for predicting the location of a lane constraint. A lane constraint, in some embodiments, may be defined as the boundary at the edge of a lane, roadway, parking space, or other path or space in which a vehicle should be caused to travel. In some situations, a lane constraint may be marked on a roadway, such as by painted

IPR2025-01034
Tesla EX1002 Page 1633

lines, a physical barrier, Botts' dots, or one or more additional indicators. In some embodiments, a lane constraint may be defined by the edge of a roadway, a curb, a guard-rail, or the like. In some embodiments, a lane constraint may be defined by one or more other vehicles, such as a row of parallel-parked vehicles adjacent to the lane, or a row of parked vehicles in a parking lot. In some embodiments, a lane constraint may be unmarked, such as on an unmarked roadway or in a parking lot without lane lines.

[0138] In some embodiments, a self-driving or driver-assist system may use image capture devices and image analysis to determine the location of one or more lane constraints based on objects detected in the images, such as lane markings, edges of roadways, and/or other vehicles. In most driving situations, the system may be able to detect at least two lane constraints—one on each side of the lane in which the vehicle is traveling—based on the visible characteristics (e.g., lane lines, etc.) of the captured images.

[0139] In certain situations, however, including those in which only one lane constraint on one side of a lane or on one side of vehicle 200 can be sufficiently detected based on visible characteristics in the captured images (e.g., lane lines, etc.), system 100 may predict the position of a lane constraint on the other side of the lane or on the other side of the vehicle, allowing the vehicle to continue to travel within a lane defined by the detected constraint and the predicted constraint opposite it.

[0140] In some embodiments, lane constraint prediction may be performed by system 100 subject to one or more conditions being satisfied. For example, in some embodiments, system 100 predicts the location of a lane constraint when a previously-detected lane constraint ceases to be detected as vehicle 200 progresses along a roadway. Thus, in some embodiments if a lane marking or other detected lane constraint suddenly ends or becomes undetectable, then lane constraint prediction may be performed. In other examples, system 100 predicts the location of a lane constraint when the system determines that guard-rail shadow conditions are present, such that a guard rail is present in the vicinity of the roadway or a guardrail is casting a shadow onto the roadway.

[0141] In some embodiments, system 100 may only perform lane constraint prediction when system 100 has access to valid and sufficient historical information about the width of a lane in which vehicle 200 is traveling; in situations in which a lane constraint on one side of vehicle 200 is newly detected as vehicle 200 progresses along a roadway, system 100 may in some embodiments refrain from performing lane constraint prediction.

[0142] In some embodiments, lane constraint prediction may be performed for a maximum travel distance, such that, if vehicle 200 travels greater than a predefined maximum travel distance while system 100 is continuously performing lane constraint prediction, then lane constraint prediction may then be ceased. In some embodiments, the predefined maximum travel distance may be set or modified in accordance with one or more various factors detected by system 100. In some embodiments, the predefined maximum travel distance may be turned off (e.g., it may be functionally set to "infinite") in accordance with one or more various factors detected by system 100.

[0143] In some embodiments, lane constraint prediction may be ceased in accordance with various cancellation conditions detected by system 100. For example, in some

embodiments, if a driver of vehicle 200 changes lanes, then lane constraint prediction may be ceased. In some embodiments, if vehicle 200 is traversing a curve in the roadway that exceeds a predefined curve sharpness threshold, then lane constraint prediction may be ceased.

[0144] FIGS. 8A-8C illustrate an embodiment in which vehicle 200 travels on a roadway 800. In some embodiments, vehicle 200 may have one or more fields of vision provided by one or more image capture devices or cameras, such as image capture devices 122, 124, and/or 126. In the depicted example, vehicle 200 may have a field of vision facing frontward from the vehicle, as depicted by the angled lines forming an upward-opening triangle in front of vehicle 200; in some embodiments, vehicle 200 may have one or more fields of vision facing in multiple different directions from vehicle 200. In the depicted example, the disclosed systems and methods for lane constraint prediction (e.g., predicting the location of a lane constraint when one is not detected) and operating vehicle 200 within a lane defined by a predicted lane constraint may be used. As shown, roadway 800 may have lane 802 in which vehicle 200 is traveling. Lane 802 may be defined on a first side by first lane constraint 804 and on a second side, opposite the first side, by lane constraint 806. These lane constraints 804 and 806 may be detected by system 100, as indicated by the bolded lines showing lane constraints 804 and 806. Vehicle 200 may have a first vehicle side 808 on the same side of vehicle 200 as first lane constraint 804, and a second vehicle side 810 on the same side of vehicle 200 as second lane constraint 806.

[0145] Processing unit 110 may be configured to determine first lane constraint 804 and second lane constraint 806 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 804 and/or second lane constraint 806 may be identified by visible lane boundaries, such as dashed or solid lines marked on a road surface, such as lane lines 803a (solid line) and 803b (dashed line). Additionally or alternatively, first lane constraint 804 and/or second lane constraint 806 may be identified by an edge of a road surface or a barrier. Additionally or alternatively, first lane constraint 804 and/or second lane constraint 806 may be identified by markers (e.g., Botts' dots). According to some embodiments, processing unit 110 may determine first lane constraint 804 and/or second lane constraint 806 by identifying a midpoint of a road surface width. According to some embodiments, if processing unit 110 identifies only first lane constraint 804, processing unit 110 may estimate second lane constraint 806, such as based on an estimated lane width or road width. Processing unit 110 may predict lane constraints in this manner when, for example, lines designating road lanes are not painted or otherwise labeled.

[0146] In FIGS. 8A-8C, lane constraints 804 and 806 are marked on roadway 800 by solid and dashed lines 803a and 803b, respectively. In a portion of roadway 800 toward the center of FIGS. 8A-8C, dashed line 803b corresponding to lane constraint 806 is not painted on roadway 800. In FIGS. 8A-8C, the detection of lane constraints 804 and 806 by system 100 is depicted by the bolded lines that overlay the respective lane lines in front of vehicle 200. As depicted by the bolded lines in FIG. 8A, system 100 detects lane constraint 804 as a solid-line lane constraint on the left side of vehicle 200 for a certain visible distance in front of vehicle 200. Similarly, system 100 detects lane constraint 806 as a dashed-line lane constraint on the right side of vehicle 200 for a certain visible distance in front of vehicle 200. In accordance with the detection of lane constraints 804 and 806, system 100 causes vehicle 200 to travel inside lane 802 as defined by lane constraints 804 and 806.

IPR2025-01034
Tesla EX1002 Page 1635

[0147] FIG. 8B depicts vehicle 200 continuing to progress along roadway 800 in lane 802 following the point in time depicted in FIG. 8A. As vehicle 200 reaches the portion of roadway 800 at which dashed line 803b is not painted on roadway 800, lane constraint prediction may be performed by system 100. As shown, system 100 predicts the location of lane constraint 806 as being located a distance 812 from opposite lane constraint 804. This distance 812 may be determined, in some embodiments, based on the historical information detected and stored by system 100 about the previous widths of lane 802 and/or the average historical width of lanes in general. As shown, system 100 predicts that lane constraint 806 continues to be a dashed lane constraint, based on the fact that lane constraint 806 was marked by dashed lane line 803b at the last time that lane constraint 806 was detected by system 100 before lane constraint prediction was activated. In accordance with the continued detection of lane constraint 804 and the prediction of lane constraint 806, system 100 causes vehicle 200 to continue to travel inside lane 802 as defined by lane constraints 804 and 806. In some embodiments, when lane constraint prediction is activated, and possibly under certain circumstances (e.g., there are other vehicles nearby), a certain default type of lane constraint is used to control the motion of the vehicle which is not necessarily the same as the last lane constraint that was detected previous to the activation of the lane constraint prediction. For example, in some embodiments, a system may default to predicting solid-line lane constraints rather than dashed-line lane constraints.

[0148] FIG. 8C depicts vehicle 200 continuing to progress along roadway 800 in lane 802 following the point in time depicted in FIG. 8B. As vehicle 200 reaches the portion of roadway 800 at which dashed line 803b is once again painted on roadway 800, lane constraint prediction may cease to be performed by system 100, as lane constraint 806 is once again detected (rather than predicted) as a dashed lane constraint by system 100. As depicted in FIG. 8C, vehicle 200 has traveled distance 814 while performing lane constraint prediction. In some embodiments, lane constraint prediction may only be performed continuously or substantially continuously for a predefined maximum distance (and/or predefined maximum time). In the depicted example of FIG. 8C, distance 814 is less than the predefined maximum distance for performing lane constraint prediction, such that lane constraint prediction was able to be performed continuously as vehicle 200 traversed the distance between its position in FIG. 8B and its position in FIG. 8C. In some embodiments, the predefined maximum distance for performing lane constraint prediction may be set to a default of 18 meters. In some embodiments, the predefined maximum distance may be varied by vehicle manufacturers, drivers, and/or any other party.

[0149] FIGS. 9A-9C illustrate an embodiment in which vehicle 200 travels on a roadway 900. In the depicted example, the disclosed systems and methods for lane constraint prediction (e.g., predicting the location of a lane constraint when one is not detected) and operating vehicle 200 within a lane defined by a predicted lane constraint may be used. The example depicted in FIGS. 9A-9C differs from the example depicted in FIGS. 8A-8C at least in that, on roadway 900, two lanes are merging into one lane. As explained in greater detail below, in some embodiments, lane constraint prediction may be performed when two lanes are merging into one lane and lane lines are not detected on the side of the lane into which a second lane is merging.

[0150] As shown, roadway 900 may have lane 902 in which vehicle 200 is traveling. Lane 902 may be defined on a first side by first lane constraint 904 and on a second side, opposite the first side, by lane constraint 906. These lane constraints 904 and 906

Tesla EX1002 Page 1636 IPR2025-01034

may be detected by system 100, as indicated by the bolded lines showing lane constraints 904 and 906. Vehicle 200 may have a first vehicle side 908 on the same side of vehicle 200 as first lane constraint 904, and a second vehicle side 910 on the same side of vehicle 200 as second lane constraint 906.

[0151] Processing unit 110 may be configured to determine first lane constraint 904 and second lane constraint 906 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 904 and/or second lane constraint 906 may be identified, detected, estimated, or predicted in any or all of the manners discussed above with respect to FIGS. 8A-8C.

[0152] As shown, roadway 900 may have lane 916. Lane 916 may be bounded by lane constraints. In the example shown, lane 916 is marked by solid lane line 918 on its left side and solid lane line 120 on its right side. In the example shown, lane 916 merges into lane 902 as roadway 900 continues in the direction of travel of vehicle 200.

[0153] In FIGS. 9A-9C, lane constraint 904 is continuously marked by solid line 903 on roadway 900. Lane constraint 906, on the other hand, is inconsistently marked; at the bottom-most portion of the figures, lane constraint 906 is marked by solid line 905a on roadway 900; at the portion above that, as lane 916 begins to merge into lane 902, lane constraint 906 is marked by dashed line 905b on roadway 900; at the portion above that, as lane 916 finishes merging into lane 902, lane constraint 906 is not marked; finally, at the top-most portion of the figures, lane constraint 906 is again marked by a solid line, this time by solid line 905c on roadway 900.

[0154] In FIGS. 9A-9C, the detection of lane constraints 904, 906, and 922 by system 100 is depicted by the bolded lines that overlay the respective lane lines in front of vehicle 200. As depicted by the bolded lines in FIG. 9A, system 100 detects lane constraint 904 as a solid-line lane constraint on the left side of vehicle 200 for a certain visible distance in front of vehicle 200. Similarly, system 100 detects lane constraint 906 as a solid lane constraint on the right side of vehicle 200 for a certain visible distance in front of vehicle 200. In accordance with the detection of lane constraints 904 and 906, system 100 causes vehicle 200 to travel inside lane 902 as defined by lane constraints 904 and 906.

[0155] FIG. 9B depicts vehicle 200 continuing to progress along roadway 900 in lane 902 following the point in time depicted in FIG. 9A. As vehicle 200 reaches the portion of roadway 900 at which lane 916 merges into lane 902 from the right side, where lane constraint 906 is marked briefly by dashed line 905b and then is not marked by any lines, lane constraint prediction may be performed by system 100. As shown, system 100 predicts the location of lane constraint 906 as being located a distance 912 from opposite lane constraint 904. This distance 912 may be determined, in some embodiments, based on the historical information detected and stored by system 100 about the previous widths of lane 902 and/or the average historical width of lanes in general. As shown, system 100 predicts that lane constraint 906 continues to be a dashed lane constraint, based on the fact that lane constraint 906 was marked by dashed line 905b at the last time that lane constraint 906 was detected by system 100 before lane constraint prediction was activated. In accordance with the continued detection of lane constraint 904 and the prediction of lane constraint 906, system 100 causes vehicle 200 to continue to travel inside lane 902 as defined by lane constraints 904 and 906.

Tesla EX1002 Page 1637 IPR2025-01034

[0156] Also shown in FIG. 9B is system 100 detecting lane constraint 922, which bounds the right side of lane 916, as lane 916 merges into lane 902. System 100 detects lane constraint 922 as a solid lane constraint based on the corresponding solid line 920 on roadway 900. In some embodiments, system 100 may determine that detected lane constraint 922 is not a constraint of lane 902 in which vehicle 200 is traveling, but rather is a lane constraint of a lane that is merging into lane 902. This determination may be made, in some embodiments, in accordance with a determination lane constraint 922 is detected as approaching lane 902 at above a threshold rate of approach. For example, if a detected lane constraint outside a constraint of a lane in which the vehicle is traveling makes an angle that is more than a predefined threshold angle above parallel with a detected or previously detected lane constraint of the lane in which the vehicle is travelling, then the outside lane constraint may be determined to correspond to an approaching merging lane.

[0157] In some embodiments, the determination that a lane is merging into another lane may be made in accordance with the detection and analysis of other aspects, characteristics, or objects in the roadway in or around a vehicle. For example, a second vehicle (e.g., different from the vehicle in which the system is provided) may be detected and tracked, and the system may determine that the other vehicle is traveling at an angle to a lane on the roadway that is indicative of a lane in which the other vehicle is traveling merging into another lane. In some embodiments, the system may detect that the other vehicle is approaching a lane, from an area outside the lane, at above a threshold rate of approach. In some embodiments, the system may make such a determination about one or more other vehicles, and may accordingly determine that lanes around the one or more other vehicles are merging. In some embodiments, information about the one or more other vehicles may be received via the multiple images captured by cameras associated with the system; in some embodiments, alternatively or additionally, information about the one or more other vehicles may be received via any suitable form of electronic communication, such as communication transmitted from one or more of the one or more other vehicles.

[0158] In yet some other embodiments, the determination that lanes are merging may be made in accordance with historical information about lane configuration and/or layout, such as historical information about the lanes at or around the current location of the vehicle.

[0159] FIG. 9C depicts vehicle 200 continuing to progress along roadway 900 in lane 902 following FIG. 9B. As vehicle 200 reaches the portion of roadway 900 at which the merge of lane 916 into lane 902 has been completed, lane constraint 906 is once again marked by a solid lines on roadway 900, this time by solid line 905c, and lane constraint prediction may cease to be performed by system 100, as lane constraint 906 is once again detected (rather than predicted) as a solid lane constraint by system 100. In some embodiments, system 100 may determine that it is detecting a lane constraint of the lane in which vehicle 200 is located (rather than a lane constraint of a merging lane) by determining that the lane constraint is no longer approaching the vehicle (or the lane in which the vehicle is traveling) at above the predefined threshold rate of approach. In some embodiments, system 100 may determine that it is detecting a lane constraint of the lane in which the vehicle is located (rather than a lane constraint of a merging lane) by determining that the lane constraint is within a predefined threshold distance of the predicted location of the lane constraint, as predicted by lane constraint prediction.

IPR2025-01034
Tesla EX1002 Page 1638

[0160] As depicted in FIG. 9C, vehicle 200 has traveled distance 914 while performing lane constraint prediction. In some embodiments, lane constraint prediction may only be performed continuously or substantially continuously for a predefined maximum distance (and/or predefined maximum time). According to some embodiments, the maximum distance/time that is allocated for lane constraint predication can be set dynamically. For example, the maximum distance/time that is allocated for lane constraint predication can be adjusted according to the vehicle's surrounding, the speed of the vehicle, the topology of the road, the brightness or the environment, the weather, the density and/or proximity of other vehicles on the roadway, and/or any other road conditions or characteristics. In some embodiments, when it is determined that an approaching merging lane is present in the vicinity of vehicle 200, system 100 may permit lane constraint prediction to be executed for a greater total distance than it would otherwise be permitted to be executed. For example, while a default maximum distance may be 18 meters, the maximum distance in situations when an ongoing lane merge is detected may be 1000 meters. This difference may account for the fact that lane merges may take several hundred meters to be completed. In the depicted example of FIG. 8C, distance 914 is less than the predefined maximum distance for performing lane constraint prediction in a lane merge situation, such that lane constraint prediction was able to be performed continuously as vehicle 200 traversed the distance between its position in FIG. 9B and its position in FIG. 9C.

[0161] FIGS. 10A and 10B illustrate an embodiment in which vehicle 200 travels on roadway 1000. FIG. 10A shows a three-dimensional view as seen from the perspective of vehicle 200, such as one of the one or more images received by system 100 as vehicle 200 travels on roadway 1000. (FIG. 10A depicts a lead vehicle driving on roadway 1000 in front of vehicle 200, not to be confused with vehicle 200 itself.) FIG. 10B shows an overhead diagram of vehicle 200 on the same portion of roadway 1000.

[0162] In the depicted example, the disclosed systems and methods for lane constraint prediction (e.g., predicting the location of a lane constraint when guard rail shadow conditions are present) and operating vehicle 200 within a lane defined by a predicted lane constraint may be used. As shown, roadway 1000 may have lane 1002 in which vehicle 200 is traveling. Lane 1002 may be defined on a first side by first lane constraint 1004 and on a second side, opposite the first side, by lane constraint 1006. These lane constraints 1004 and 1006 may, in some embodiments, be detected by system 100, as indicated by the bolded lines showing lane constraints 1004 and 1006. Vehicle 200 may have a first vehicle side 1008 on the same side of vehicle 200 as first lane constraint 1004, and a second vehicle side 1010 on the same side of vehicle 200 as second lane constraint 1006.

[0163] Roadway 1000 may include lane lines 1003a and 1003b, which physically mark the left and right sides, respectively, of lane 1002 on roadway 1000.

[0164] Roadway 1000 may include guardrail 1012, which is supported by posts 1014. Both guard rail 1012 and its posts 1014 are casting guard rail shadow 1016, which in the depicted example falls over lane line 1003a and into lane 1002.

[0165] As depicted by the bolded lines in FIGS. 10A and 10B, system 100 detects lane constraint 1006 as a solid-line constraint on the right side of vehicle 200 for a certain visible distance in front of vehicle 200. On the left side of roadway 1000, the bolded lines show how system 1000 detects several edges, one of which is attributable to lane

Tesla EX1002 Page 1639 IPR2025-01034

line 1003a and two of which (shadow edges 1005a and 1005b) are attributable to guard rail shadow 1016. In some embodiments, the detection of multiple edges on one side of a vehicle or lane may make it more difficult to identify a lane constraint from the images available to system 100, such that the lane constraint cannot be identified with a high level of certainty. Accordingly, in some embodiments, when system 100 determines that guard rail 1012 or guard rail posts 1014 are present, that guard rail shadow 1016 is being cast onto roadway 1000, or that multiple edges are detected on the same side of lane 1002, system 100 may determine that guard rail shadow conditions are present, and may accordingly engage a guard rail shadow mode. In some embodiment, a guard rail shadow mode may cause lane constraint prediction to be executed for the side of a lane that is proximate to the detected guard rail, shadow, or edges. By using lane constraint prediction, system 100 may predict the location of lane constraint 1004 in accordance with any of the techniques discussed above with respect to FIGS. 8A-8C and/or 9A-9C. In the depicted example, system 100 correctly predicts the location of lane constraint 1002 to coincide with the edge defined by painted lane line 1003a, rather than either of the edges defined by guard rail shadow 1016. In accordance with the detection of lane constraints 1004 and 1006, system 100 causes vehicle 200 to travel inside lane 1002 as defined by lane constraints 1004 and 1006.

[0166] In some embodiments, guard rail shadow prediction may cause lane constraint prediction to be engaged with no limitation as to how long lane constraint prediction can be carried out. Thus, for example, rather than the maximum distance for lane constraint prediction being set to 18 meters (as in the example of FIGS. 8A-8C) or 1000 meters (as in the example of FIGS. 9A-9C), system 100 may turn off the maximum distance, functionally configuring the system so that lane constraint prediction may continue for an indefinite distance (and/or an indefinite time) so long as guard-rail shadow conditions continue to be detected and so long as no lane constraint prediction cancellation conditions are detected.

[0167] FIG. 11 is an exemplary block diagram of memory 140 or 150, which may store instructions for performing one or more operations consistent with disclosed embodiments. As illustrated in FIG. 11, memory 140 or 150 may store one or more modules for performing the techniques described herein. For example, memory 140 or 150 may store image obtaining module 1100, roadway lane constraint recognizing module 1102, determining module 1104, roadway lane constraint predicting module 1106, vehicle enabling module 1108, and lane constraint prediction ceasing module 1110. As would be appreciated by a person of ordinary skill in the art, any of the previously recited modules may be combined with each other or with other modules, interfaced with one another or with other modules, and/or subdivided into further modules and/or sub-modules.

[0168] In some embodiments, processing unit 110 may execute instructions stored on the one or more modules in memory 140 or 150, and the execution by processing unit 110 may cause system 100 to perform the methods discussed herein.

[0169] In some embodiments, processing unit 110 may cause system 100 to obtain (e.g., with image obtaining module 1100) multiple images from a camera (e.g., image capture devices 122, 124, and/or 126), wherein the camera is adapted to capture, when mounted on a vehicle (e.g., vehicle 200) and while the vehicle is in motion, images of a roadway within a field of view of the camera, apply image processing to the multiple images to recognize (e.g., with roadway lane constraint recognizing module 1102) a first roadway lane constraint, determine (e.g., with determining module 1104),

IPR2025-01034
Tesla EX1002 Page 1640

based on the multiple images, that one or more defined lane prediction conditions are satisfied, in accordance with the a determination that one or more defined lane prediction conditions are satisfied, predict (e.g., with roadway lane constraint predicting module 1106) the location of a second roadway lane constraint, and issue (e.g., with issuing module 1108) a notification or a control signal to control a motion of the vehicle according to the predicted location of the second roadway lane constraint.

[0170] In some embodiments, processing unit 110 may cause system 100 to issue (e.g., with issuing module 1108) a notification or a control signal to direct a motion of the vehicle according to a location of the first roadway lane constraint.

[0171] In some embodiments, processing unit 110 may cause system 100 to issue (e.g., with issuing module 1108) a notification or a control signal to direct a motion of the vehicle along a lane or an area between the first roadway lane constraint and the second roadway lane constrain.

[0172] In some embodiments, processing unit 110 may cause system 100 to determine (e.g., with determining module 1104) whether one or more defined lane prediction cancellation conditions are satisfied, and in accordance with a determination that one or more defined lane prediction cancellation conditions are satisfied, cease to predict (e.g., with lane constraint prediction ceasing module 1110) the location of the second roadway lane constraint.

[0173] FIG. 12 illustrates a process 1200 for implementing lane constraint prediction consistent with the disclosed embodiments. In some embodiments, the steps in process 1200 may be illustrated by the situations illustrated in FIGS. 8A-8C, 9A-9C, and 10A-10B. According to some embodiments, process 1200 may be implemented by one or more components of navigation system 100, such as at least one processing unit 110. According to some embodiments, process 1200 may be implemented by system 100 associated with vehicle 200.

[0174] At step 1202, in some embodiments, a system receives multiple images of a roadway in the vicinity of a moving vehicle associated with the system. In some embodiments, the one or more images are acquired using at least one image capture device, such as image capture devices 122, 124, and/or 126. The plurality of images may be received by the system through a data interface such as data interface 128. In the example illustrated in FIGS. 8A-8C, system 100 may acquire a plurality of images of roadway 800 in the vicinity of vehicle 200.

[0175] At step 1204, in some embodiments, the system recognizes a first roadway lane constraint. In some embodiments, the system may use image analysis techniques to process the one or more images received to detect lane constraints by identifying objects, markings, or other characteristics in the images that are indicative of a lane constraint, such as lane markings, roadway edges, lane dividers, and/or other vehicles. In some embodiments, the system may attempt to identify at least one lane constraint on one side of the vehicle and another lane constraint on the other side of the vehicle (e.g., one lane constraint to each side of the center line of the vehicle in its traveling direction, front to back). In some embodiments, the system may attempt to identify at least one lane constraint on one side of a lane and another lane constraint on another side of the lane. In the example illustrated in FIGS. 8A-8C, system 100 may analyze the images received of roadway 800, and may detect lane constraint 804 on the left side of lane 802 as corresponding to and coinciding with solid lane line 803a. That is, in

Tesla EX1002 Page 1641 IPR2025-01034

some embodiments, lane line 803a may be recognized by system 100 in the multiple acquired images, and system 100 may determine that lane line 803a is a visual marker for lane constraint 804.

[0176] At step 1206, in some embodiments, the system determines, based on the multiple images, whether one or more defined lane prediction conditions are satisfied. In some embodiments, a lane prediction condition may be any predetermined condition by which system 100 may determine whether to predict the location of a lane constraint. In some embodiments, only one lane prediction condition need be satisfied, while in other embodiments, a minimum threshold number of lane prediction conditions (e.g., two or more) need to be satisfied. In some embodiments, the minimum threshold number of lane predication conditions can be determined in accordance with historical information regarding the roadway lane layout, the roadway topology, shadow conditions, visibility conditions, presence of guard rails and/or other barriers along the roadway, presence of other vehicles on the roadway, etc.

[0177] In some embodiments, the one or more defined lane prediction conditions comprise that the second roadway lane constraint ceases to be recognized, based on the multiple images, as the vehicle progresses. For example, in some embodiments, the system may continuously capture and consult multiple images corresponding to the vehicle's current location and vicinity; as the vehicle travels along a roadway and along a lane, a lane constraint that was recognizable at one instance in time based on one subset of the multiple images may cease to be recognizable at a subsequent instance in time based on a later subset of the multiple images. In the example of FIGS. 8A-8C, system 100 may analyze the images received of roadway 800 as vehicle 200 travels along roadway 800 from its position in FIG. 8A to its position in FIG. 8B. As vehicle 200 travels from its position in FIG. 8A to its position in FIG. 8B, it reaches the middle portion of roadway 800 at which dotted lane line 803b is not painted on roadway 800. Upon analyzing the received multiple images from this middle portion of roadway 800, system 100 may cease to recognize lane constraint 806 due to the absence of lane line 803b. This ceasing to recognize a lane constraint as vehicle 200 progresses along roadway 800 may constitute a lane prediction condition in some embodiments.

[0178] In some embodiments, the one or more defined lane prediction conditions comprise that the lane constraint on the first side of the lane has been recognized for at least a threshold time or a threshold distance. Optionally, a threshold time or a threshold distance associated with the one or more defined lane prediction conditions can be dynamic and can be adjusted according to historical information regarding the roadway lane layout, the roadway topology, shadow conditions, visibility conditions, presence of guard rails or other barriers along the roadway, presence of other vehicles on the roadway, etc. In some embodiments, such a lane prediction condition may serve to prevent the system from (or reduce the likelihood of the system) engaging in lane constraint prediction when the system detects a lane constraint on a single side of a lane (or a single side of a vehicle) for a first time. That is, in some embodiments, if a vehicle is traveling and comes upon a single lane line or other lane marking or indicator for the first time, the system may refrain from predicting the location of another lane constraint based on the single lane marking/indicator. In some embodiments, this lane prediction condition may serve to ensure (or increase the likelihood) that the system is making valid and accurate lane constraint predictions based on actual lanes that exist on a roadway, and not on isolated, stray, or erroneously or randomly placed lane markings/indicators. In some embodiments, a predefined threshold distance may be fixed in the system, or may be determined by the system based on a current speed of

the vehicle, the size or posted speed of the roadway on which the vehicle is traveling, or any other discernible variables regarding the vehicle or roadway.

[0179] In some embodiments, the one or more defined lane prediction conditions comprise that a guard rail is present in the vicinity of the vehicle. In some embodiments, the system may engage in lane constraint prediction when the system determines that the vehicle is in the vicinity of a guard rail. In some embodiments, this lane prediction condition may serve to prevent the system from erroneously determining that one or more edges detected in the one or more images correspond to a lane constraint, when in reality the one or more edges is associated with a guard rail or with a guard rail shadow. In some embodiments, a guard rail may cast a shadow onto a roadway, and an edge of the shadow on the roadway may be detected by the image analysis techniques used by the system. In some embodiments, the system may erroneously determine that the guard rail shadow is a marker of a lane constraint, and may cause the vehicle to travel in accordance with the guard rail shadow rather than with the actual lane constraint. In some embodiments, in order to prevent this error, the system may recognize that a guard rail is present in the vicinity of the vehicle, and may therefore engage in lane constraint prediction to predict the location of a lane constraint on the side of a lane with a guard rail, rather than attempting to recognize the lane constraint on that side of the lane and risking erroneously recognizing the guard rail shadow as a lane constraint.

[0180] In the example of FIGS. 10A and 10B, vehicle 200 (shown in FIG. 10B; not shown in FIG. 10A, as vehicle 200 constitutes the point of view in FIG. 10A) is traveling on roadway 1000 in lane 1002, which has guard rail 1012 to the left side of lane 1002. In some embodiments, system 100 may analyze the one or more images to detect that guard rail 1012 is present in the vicinity of vehicle 200 on the left side of lane 1002. In some embodiments, the presence of guard rail 1012 may satisfy a lane prediction condition.

[0181] In some embodiments, determining that a guard rail is present in the vicinity of the vehicle comprises detecting, in the one or more images, a pattern created by posts supporting the guard rail. In some embodiments, the system may be configured to detect one or more patterns in the one or more images that is known or suspected to be indicative of the presence of a guard rail. One such pattern may, in some embodiments, be referred to as a "rapid dash" pattern, and may be present when the system detects in the one or more images a pattern of rapid successive dashes or rapid successive lines caused by the posts supporting a guard rail. In some embodiments, the system may be configured to detect any other suitable pattern or visual indicia known to be characteristic of the posts that support guard rails. In some embodiments, the system may detect a pattern of the posts of a guard rail, while in some embodiments the system may alternatively or additionally detect a pattern of the shadows cast by the posts of a guard rail. In the example of FIGS. 10A and 10B, guard rail 1012 is supported by posts 1014 (shown in FIG. 10A; not visible from the overhead angle of FIG. 10B). In FIGS. 10A and 10B, posts 1014 are casting a shadow that falls onto roadway 1000 and into lane 1002. System 100 may detect the presence of a rapid dashed pattern or any other suitable pattern created in the multiple images by posts 1014 or their shadows, and this detection may be one manner in which system 100 determines that a guard rail is present in the vicinity of vehicle 200.

[0182] In some embodiments, system 100 may utilize depth information to identify a shape that is characteristic of a guard rail or of a portion of a guard rail. It would be

Tesla EX1002 Page 1643 IPR2025-01034

appreciated that a painted line would have substantially no volume, and its height is virtually zero relative to the roadway surface, and so, depth information can be useful to distinguish between a lane marking and a guard rail.

[0183] In some embodiments, the one or more defined lane prediction conditions comprise that a guard rail is casting a shadow onto the roadway. In some embodiments, a lane prediction condition may be based on the detection by the system of a shadow of a guard rail, in place of or in addition to detection by the system of a guard rail itself. In some embodiments, a lane prediction condition may be satisfied if a guard rail shadow is detected anywhere in the vicinity of the vehicle; in some embodiments, a lane prediction condition may be satisfied if a shadow is detected as being cast from a guard rail in the direction of a lane or a roadway in or on which the vehicle is traveling; in some embodiments, a lane prediction condition may be satisfied if a shadow is detected as being cast from a guard rail onto a roadway or lane in which the vehicle is traveling.

[0184] In some embodiments, determining that a guard rail is casting a shadow onto the roadway may involve determining whether a predicted location of the sun comports with the location of a detected edge that is possible a guard rail shadow. For example, in some embodiments, the system may determine a position of the sun and an orientation of the sun with respect to the multiple images available to the system. These determinations as to the location of the sun and the orientation of the vehicle may be made in accordance with various pieces of information including, for example, one or more of the multiple images, a location and/or orientation of the vehicle determined by a GPS, an orientation of the vehicle determined by a compass, and/or a time of day. The system may determine, based on the determined location of the sun and orientation of the vehicle that a shadow is predicted to be cast from a guard rail in a certain direction at a certain time.

[0185] In the example of FIGS. 10A and 10B, guard rail 1012 is casting shadow 1016 onto roadway 1000 and into lane 1002; shadow 1016 falling on roadway 1000 and in lane 1002 may be detected by system 100, and said detection may satisfy a lane prediction condition.

[0186] In some embodiments, the one or more defined lane prediction conditions comprise that multiple substantially parallel edges are detected, in the multiple images, on the roadway on one side of a lane on the roadway. In some embodiments, image analysis techniques applied by the system to the multiple images received may detect edges in one or more of the multiple images received; in some embodiments, the detection of edges may be used to recognize objects and markings in the environment around the vehicle. In some embodiments, the detection of edges in the multiple images may be used to recognize lane lines or markings or the edges or roadways, and to thereby determine the location of lane constraints. In some embodiments, including those in which a guard rail shadow is cast onto a roadway or lane in which the vehicle is traveling, the system may detect the edge of a guard rail shadow as an edge in the multiple images. In some other embodiments, other objects or markings may cause the system to detect multiple edges on the same side of one lane. In these embodiments, determination of which edge corresponds to the correct lane constraint may be difficult, slow, or unreliable. For this reason, it may be advantageous in some embodiments to refrain from attempting to determine, based on the multiple images, which of several detected edges corresponds to a lane constraint, and instead to simply predict the location of the lane constraint using lane constraint prediction when

Tesla EX1002 Page 1644 IPR2025-01034

multiple edges are detected in the multiple images as discussed above. In the example of FIGS. 10A and 10B, system 100 detects, in the multiple images, three different edges on the left side of lane 1002, indicated by the bolded lines marked 1004, 1005a, and 1005b. In the example of FIGS. 10A and 10B, as explained above, system 100 may determine that multiple edges are detected on the left side of lane 1002, and this detection may constitute a lane prediction condition.

[0187] In some embodiments, if it is determined, at block 1206, that one or more defined lane prediction conditions are not satisfied, then method 1200 may return to step 1202 and may continue to acquire additional images of the roadway as time progresses.

[0188] In some embodiments, if it is determined, at block 1206, that one or more defined lane prediction conditions are satisfied, then method 1200 may proceed to step 1208. At step 1208, in some embodiments, in accordance with a determination that one or more defined lane prediction conditions are satisfied, the system predicts the location of a second roadway lane constraint. In some embodiments, the system may predict a lane constraint on the opposite side (the vehicle being a reference as to direction of the opposite side) of a lane from a detected lane constraint. Thus, in some embodiments, if one lane constraint is recognizable from lane lines or other visible indicia in the received images, and lane prediction conditions such as any of the conditions discussed above are determined to be satisfied, then the system may predict the location of the second lane constraint on the opposite side of the lane from the first lane constraint.

[0189] In some embodiments, the predicted lane constraint may be predicted to extend for a predetermined distance along the roadway. In some embodiments, the predicted lane constraint may be predicted to be parallel to the recognized lane constraint opposite it. In some embodiments, the predicted lane constraint may be predicted to continue along the same angle, path, or location as the same lane constraint was previously detected by the system. In some embodiments, the predicted lane constraint may be predicted based on the location of another lane constraint, based on the location of other discernible features of a roadway, or based on the location of the vehicle associated with the system and/or on surrounding vehicles. In some embodiments, the predicted lane constraint may be predicted as being located a predetermined distance from a recognized lane constraint or from any other object recognized or predicted based on the multiple images of the roadway. In some embodiments, the predicted lane constraint may be predicted to be a lane constraint type that matches or is determined in accordance with other lane constraint types detected in the multiple images, or determined in accordance with a lane constraint type determined for the predicted lane constraint based on previous images of the roadway from a previous time (e.g., if a system detects a solid single white line on the right side of the vehicle, then ceases to detect any lane line, the system may predict that a solid single white lane constraint continues on the right side of the vehicle; alternately or additionally, a system may predict that a solid single white line is located on the opposite side of a lane from a solid single yellow line).

[0190] In the example of FIGS. 8A-8C, system 100 recognizes lane constraint 804 based on the detection of lane line 803a in the multiple images. When vehicle 200 reaches its position in FIG. 8B and does not detect any lane lines on the right side of lane 802, then a lane prediction condition may be satisfied. Accordingly, in some embodiments, system 100 may predict lane constraint 806 on the right side of lane

Tesla EX1002 Page 1645 IPR2025-01034

802, as shown in FIG. 8B. In some embodiments, lane constraint 806 may be predicted as a single dashed line lane constraint, based on the previous recognition, in FIG. 8A based on lane lines 803b, of lane constraint 806 as a single dashed line lane constraint. In some embodiments, the position, angle, type, and/or any other characteristic of lane constraint 806 may be predicted based on any stored or otherwise determined information accessible by system 100 about vehicle 200, roadway 800, or any other surrounding environmental factors.

[0191] In some embodiments, predicting the location of a lane constraint comprises predicting a width of a lane based on information about a previously determined width of the lane. In some embodiments, a width of a lane may be predicted based on historical information about lane width or cached information about the width of a specific lane in the recent past (e.g., a few seconds before lane constraint prediction was performed). In some embodiments, when a system is detecting both constraints of a lane and may thereby determine the width of that lane, and the system then ceases to detect or recognize one of the lane constraints, then the system may predict that the lane continues to be the same width. The system may accordingly predict the location of the predicted lane constraint such that the lane is predicted to remain constant or substantially constant in width. In the example of FIGS. 8A-8C, when system 100 is predicting the location of lane constraint 806 in FIG. 8B, constraint 806 is predicted as being a distance 812 from lane constraint 804. This distance 812 may be the same distance 812 that was detected between lane constraint 806 and lane constraint 804 when vehicle 200 was in the position shown in FIG. 8A. System 100 may, in some embodiments, log or cache historical information about lane width, such that it can quickly recall lane width information in order to predict the location of a predicted lane constraint.

[0192] At step 1210, in some embodiments, the system enables the moving vehicle to avoid the first roadway lane constraint and the second roadway lane constraint. Thus, in some embodiments, the system may cause the vehicle to travel in a lane defined on one side by a detected/recognized lane constraint and on the opposite side by a predicted lane constraint. In some embodiments, the system may cause any suitable steering, accelerating, decelerating, or other control of the vehicle to cause it to travel within the lane defined by the predicted and detected lane constraints. In the example of FIGS. 8A-8C, when system 100 detects lane constraint 806 in FIG. 8B, system 100 then causes vehicle 200 to travel along lane 802 between detected lane constraint 804 and predicted lane constraint 806 to travel from its position in FIG. 8B to its position in FIG. 8C.

[0193] At step 1212, in some embodiments, the system determines whether one or more defined lane prediction cancellation conditions are satisfied. In some embodiments, a lane prediction cancellation condition may be any predetermined condition by which the system may determine whether to cease predicting the location of a lane constraint. In some embodiments, the location of a lane constraint may continue to be predicted once the system begins predicting the location of the lane constraint until one or more predefined lane prediction cancellation conditions are satisfied. In some embodiments, only one lane prediction cancellation condition need be satisfied in order for lane prediction to cease. In some embodiments, a predefined number or lane prediction conditions greater than one must be satisfied in order for lane prediction to be ceased. Optionally, a confidence score may be assigned to a lane prediction condition according to predefined criteria. Optionally, some lane prediction condition can have a greater weight than others. Still further by way of example, an

IPR2025-01034
Tesla EX1002 Page 1646

IPR2025-01034
Tesla EX1002 Page 1647

overall score can be computed to determine whether a lane prediction cancellation condition is met, and such a calculation can take into account confidence scores and weights which were assigned to lane prediction conditions.

[0194] In some embodiments, the one or more defined lane prediction cancellation conditions comprise that, after ceasing to be recognized, the second lane constraint is recognized based on the one or more images. For example, in some embodiments, as the vehicle is traveling along a roadway, a lane constraint may cease to be recognized (as discussed above), such as when a roadway is unmarked or lane lines are not detectable in the multiple images; lane prediction may be accordingly performed, as discussed above. In some embodiments, thereafter, for example as the vehicle continues to travel along the roadway, the system may once again recognize the lane constraint in the multiple images as the system continues to acquire new images. For example, if a vehicle reaches a portion of the roadway where the lane lines are no longer missing, or if the lane lines otherwise become detectable by the system once more, the system may recognize the lane constraint based on the updated images of the roadway. This recognition of the second lane constraint—the lane constraint that was being predicted—may in some embodiments constitute a lane prediction cancellation condition that may cause the system to cease to predict the location of the second lane constraint and to instead return to reliance on the detected/recognized location of the lane constraint based on the multiple images.

[0195] In the example of FIGS. 8A-8C, while system 100 may be predicting the location of lane constraint 806 when vehicle 200 is located at the position shown in FIG. 8B, as vehicle 200 continues to the position shown in FIG. 8C, the dashed lane lines 803b on the upper half of the diagram may once again be visible in the images acquired by system 100. Based on these new images, system 100 may once again detect/recognize the presence of lane constraint 806 based on lane line 803b, and this recognition may constitute a lane prediction cancellation condition that triggers the cessation of the prediction of the location of lane constraint 806.

[0196] In some embodiments, the one or more defined lane prediction cancellation conditions comprise that a driver is moving the vehicle between lanes. In some embodiments, the system may determine based on the multiple images that the driver is steering or merging the vehicle to another lane. In some embodiments, this detection may satisfy a lane prediction cancellation condition and may trigger the cessation of lane constraint location prediction.

[0197] In some embodiments, the one or more defined lane prediction cancellation conditions comprise that the vehicle is traversing a curve that exceeds a predefined sharpness threshold. In some embodiments, the system may access a predefined or dynamically determined (e.g., determined in accordance with vehicle speed, posted speed limit, road conditions, weather conditions, environment brightness, the vicinity of other vehicles, etc.) sharpness threshold for curves, and may determine whether a curve in a roadway or a lane exceeds the sharpness threshold (e.g., is sharper than the maximum permissible sharpness). If it is determined that a curve being traversed or approached by the vehicle exceeds the maximum permissible sharpness, then the system may determine that a lane prediction cancellation condition is satisfied, and may accordingly cease lane constraint prediction.

[0198] In some embodiments, determining whether one or more defined lane prediction cancellation conditions are satisfied comprises selecting a maximum lane prediction

distance setting, and determining whether a distance traveled by the vehicle while predicting the location of the second roadway lane constraint is greater than a maximum lane prediction distance associated with the maximum lane prediction distance setting. In some embodiments, a maximum lane prediction distance may be predefined, set, or dynamically determined by the system. In some embodiments, the system may monitor the distance traveled by the vehicle while the system is predicting the location of a lane constraint, and may cease to predict the location of the lane constraint if the distance traveled exceeds the maximum permissible lane prediction distance. In some embodiments, implementation of a maximum permissible lane prediction distance may improve safety and functionality by ensuring that the system does not predict the location of a lane constraint for an overly-long period of time or over an overly-long distance, over which distance or time one or more characteristics of the lane or the roadway may change or vary. As characteristics of lanes and roadways can be expected, in some embodiments, to change or vary over space and time, the system may improve safety by ensuring that lane constraint reference points or indicators, such as lane lines or other lane markers, are detected based on the multiple images acquired by the system at least once over a maximum lane prediction distance (or, in some embodiments, a maximum lane prediction time).

[0199] In some embodiments, the maximum lane prediction distance may be varied by the system in different situations. In some embodiments, a maximum lane prediction distance may be associated with one or more maximum lane prediction distance settings, such that the system may select one maximum lane prediction distance setting and implement the associated maximum lane prediction distance. In some embodiments, different maximum lane prediction distance settings may be implemented in different driving scenarios, as detected and determined by the system.

[0200] In some embodiments, a standard or default maximum lane prediction distance setting may be associated with a standard or default maximum lane prediction distance. In some embodiments, the standard or default maximum lane prediction distance may be 18 meters.

[0201] In the example of FIGS. 8A-8C, vehicle 200 has traveled distance 814 from its position in FIG. 8B, at which point lane constraint prediction was commenced, to its position in FIG. 8C, at which point lane prediction is ceased. This distance 814 may be less than a standard or default maximum lane prediction distance of 18 meters, such that lane constraint prediction may have been performed continuously between the positions of vehicle 200 in FIGS. 8B and 8C.

[0202] In some embodiments, selecting a maximum lane prediction distance setting comprises, in accordance with a determination that a first lane is merging with a second lane in the vicinity of the vehicle, selecting a maximum lane prediction distance setting such that the maximum lane prediction distance is greater than a default maximum lane prediction distance. In some embodiments, the system may thus implement an extended or increased maximum lane prediction distance. For example, in some embodiments, the system may determine, based on the multiple images, that a lane is merging with the lane in which the vehicle is traveling, or that the lane in which the vehicle is traveling is merging with another lane. The system may determine that a lane merge is occurring in the vicinity of the vehicle by detecting a lane constraint (e.g., by recognizing a lane line) on the far side of the merging lane adjacent to the vehicle's lane, and by determining that the lane constraint on the opposite side of the adjacent lane is approaching the lane in which the vehicle is traveling at above a predefined or

Tesla EX1002 Page 1648 IPR2025-01034

dynamically determined threshold rate of approach. That is, in some embodiments, the system may recognize the far side of an adjacent lane approaching the lane in which the vehicle is traveling as the vehicle progresses along a roadway, and may accordingly determine that the adjacent lane is merging with the lane in which the vehicle is traveling.

[0203] In some embodiments, roadways may be expected to feature long distances of unmarked lane constraints when two lanes are merging with one another, so it may be advantageous for the system to be configured to predict the location of a lane constraint for a greater distance and/or a greater time than a default maximum lane prediction distance or time. In some embodiments, the increased maximum lane prediction distance associated with lane merge situations may be 1000 meters.

[0204] In the example of FIGS. 9A-9C, vehicle 200 has traveled distance 914 from its position in FIG. 9B, at which point lane constraint prediction was commenced, to its position in FIG. 9C, at which point lane prediction is ceased. This distance 914 may be greater than a standard or default maximum lane prediction distance, such that lane constraint prediction may have been ceased before vehicle 200 reached its position in FIG. 9C if the default maximum lane prediction distance setting was selected. However, as indicated by the bolded line in FIG. 8B, system 100 recognized lane constraint 922 on the right side of lane 916 approaching lane 902 as vehicle 200 progressed, and may have accordingly determined that the lanes were merging and that a maximum lane prediction distance setting associated with lane merges should be selected. The selected setting may be associated with a maximum lane prediction distance of 1000 meters, greater than the default maximum lane prediction distance of 18 meters. Thus, in some embodiments, distance 914 may be less than the increased maximum lane prediction distance of 1000 meters, such that lane constraint prediction may have been performed continuously between the positions of vehicle 200 in FIGS. 9B and 9C.

[0205] In some embodiments, selecting a maximum lane prediction distance setting comprises, in accordance with a determination that defined guard-rail conditions are satisfied, selecting a maximum lane prediction distance setting having no maximum lane prediction distance. That is, in some embodiments, when system 100 detects that guard-rail conditions are satisfied, system 100 may predict the location of a lane constraint indefinitely, without regard for any maximum lane prediction distance. In this way, in some embodiments, the maximum lane prediction distance associated with guard-rail conditions may be thought of as an infinite maximum distance, or as a non-existent maximum distance.

[0206] In some embodiments, guard-rail conditions may be satisfied in any situation where the system detects the presence of one or more guard rails, guard rail supports, and/or guard rail shadows in the vicinity of the vehicle. For example, guard-rail conditions may be satisfied in any or all of the situations discussed above in which the system may determine that a guard rail and/or a guard rail shadow are present in the vicinity of the vehicle and/or on a roadway or lane in which the vehicle is traveling. In some situations, guard-rail conditions may be satisfied when the presence, location, and/or orientation of a guard rail and/or guard rail shadow causes the system to engage in lane constraint prediction, as discussed above.

[0207] In some embodiments, it may be advantageous for no maximum lane prediction distance (or an infinite maximum lane prediction distance) to be applied in situations in which a guard rail and/or guard-rail shadow is present, because guard rails and their

Tesla EX1002 Page 1649 IPR2025-01034

associated shadows can be expected to persist continuously for very long distances along roadways (often greater than a default maximum lane prediction distance or even an increased lane prediction distance associated with lane merges). In the example of FIGS. 10A and 10B, when system 100 begins predicting the location of lane constraint 1004 due to the detection of guard rail shadow 1016, system 100 may select a maximum lane prediction distance setting associated with guard rail conditions; the selection of a maximum lane prediction distance setting associated with guard rail conditions may cause the system to perform lane prediction of lane constraint 1004 without being subject to any maximum lane prediction distance.

[0208] In some embodiments, if it is determined, at block 1212, that one or more defined lane prediction cancellation conditions are not satisfied, then method 1200 may return to step 1202 and may continue to acquire additional images of the roadway as time progresses.

[0209] In some embodiments, if it is determined, at block 1212, that one or more defined lane prediction cancellation conditions are satisfied, then method 1200 may proceed to step 1214. At step 1212, in some embodiments, in accordance with a determination that one or more lane prediction cancellation conditions are satisfied, the system ceases to predict the location of the second lane constraint. In the example of FIGS. 8A-8C, when vehicle 200 reaches its position depicted in FIG. 8C, a lane prediction cancellation condition may be satisfied by the recognition of lane constraint 806 based on lane line 803b, and system 100 may therefore cease to predict the location of lane constraint 806. Instead, system 100 may resume determining the position of lane constraint 806 based on the detected location of lane line 803b, as it did previously in FIG. 8A.

[0210] Construction Zone Detection and Response

[0211] System 100 may provide vehicle 200 with construction zone detection, such that the system recognizes the presence of construction zones on roadways and takes action in accordance therewith. In some embodiments, system 100 may recognize various elements in the one or more images that it receives and may determine, in accordance with the recognition of those one or more elements, that vehicle 200 is approaching or inside a construction zone. In some embodiments, any automated or driver-assist action may be taken in accordance with the detection of a construction zone, including but not limited to changing the speed of vehicle 200, steering vehicle 200, engaging or disengaging one or more lights of vehicle 200, etc. In some embodiments, in accordance with the detection of a construction zone, a signal may be outputted by system 100 that indicates that a construction zone has been detected and/or that a construction zone mode is activated. In some embodiments, in accordance with the detection of a construction zone (and/or in accordance with the detection of a complex construction zone), automated or driver-assist control of vehicle 200 by system 100 may be disabled, such that control of vehicle 200 may be returned to the driver.

[0212] FIG. 13 illustrates an embodiment in which vehicle 200 is traveling on roadway 1300 and approaching a construction zone. Roadway 1300 has two lanes 1302 and 1304. Lane 1302 is defined by lane constraints 1306 and 1308, while lane 1304 is defined by lane constraints 1310 and 1312. These lane constraints are detected by system 100, as shown by the bolded lines indicating lane constraints 1306, 1308, 1310, and 1312. Lane constraints 1306 and 1308 are marked on roadway 1300 by painted

Tesla EX1002 Page 1650 IPR2025-01034

lane lines 1320 and 1322, respectively; while lane constraints 1310 and 1312 are marked by on roadway 1300 by painted lane lines 1322 and 1324, respectively.

[0213] Roadway 1300 includes several indicators of a construction zone or construction area, including painted construction zone lane line 1326. In some embodiments, construction zone lane lines may be painted in a different color and/or different style than conventional lane lines. For example, in Germany, yellow lane lines are painted on the road to indicate a construction zone, while conventional lane lines are painted in white. As indicated by the bolded line, system 100 detects construction zone lane constraint 1314 based on painted construction zone lane line 1326. In accordance with the detection of construction zone lane constraint 1314, system 100 may cause vehicle 200 to drive within the lane defined by construction zone lane constraint 1314 rather than the conventional lane constraints. In some embodiments, when system 100 has detected that vehicle 200 is in a construction zone, construction zone lane lines may be given precedence over conventional lane lines.

[0214] Roadway 1300 may include additional construction zone indicators, including cones 1316, some of which are located on the surface of roadway 1300 and inside lane 1304. In response to detecting cones 1316, system 100 may cause vehicle 200 to alter its path so as to avoid cones 1316, changing a position in a lane of travel or changing lanes if necessary. In response to detecting cones 1316, system 100 may determine that vehicle 200 is in or near a construction zone; activate a construction zone mode; output a construction zone signal; disable automated or driver-assist control of vehicle 200; and/or control vehicle 200 by changing its speed, steering it, or taking any other appropriate action (such as to avoid cones 1316).

[0215] Roadway 1300 may include additional construction zone indicators, including construction road signs 1318a and 1318b, which may indicate the presence of a construction zone and/or indicate an instruction for drivers to follow while in the construction zone (e.g., slow, stop, merge, turn, etc.). System 100 may detect construction road signs 1318a and 1318b through the manner in which the system may detect any or all other traffic signs, and may take any appropriate action in response to detection of construction road signs 1318a and 1318b, including determining that vehicle 200 is in or near a construction zone; activating a construction zone mode; outputting a construction zone signal; disabling automated or driver-assist control of vehicle 200; and/or controlling vehicle 200 by changing its speed, steering it, or taking any other appropriate action (such as in accordance with an instruction on a construction road sign).

[0216] FIG. 14 is an exemplary block diagram of memory 140 or 150, which may store instructions for performing one or more operations consistent with disclosed embodiments. As illustrated in FIG. 14, memory 140 or 150 may store one or more modules for performing construction-zone detection as described herein. For example, memory 140 or 150 may store image obtaining module 1400, construction zone indicator recognizing module 1402, determining module 1404, signal outputting module 1406, vehicle moving enabling module 1408, and vehicle control disabling module 1410. As would be appreciated by a person of ordinary skill in the art, any of the previously recited modules may be combined with each other or with other modules, interfaced with one another or with other modules, and/or subdivided into further modules and/or sub-modules.

[0217] In some embodiments, processing unit 110 may execute instructions stored on

Tesla EX1002 Page 1651 IPR2025-01034

the one or more modules in memory 140 or 150, and the execution by processing unit 110 may cause system 100 to perform the methods discussed herein.

[0218] In some embodiments, processing unit 110 may cause system 100 to obtain (e.g., with image obtaining module 1400) multiple images from a camera (e.g., image capture devices 122, 124, and/or 126), wherein the camera is adapted to capture, when mounted on a vehicle (e.g., vehicle 200), and while the vehicle is in motion, images of a roadway within a field of view of the camera, apply image processing to the multiple images to recognize (e.g., with construction zone indicator recognizing module 1402) one or more indicators of a construction zone, and, in accordance with recognizing the one or more indicators of a construction zone, output (e.g., with signal outputting module 1406) a signal indicating that the vehicle is proximate to a construction zone.

[0219] In some embodiments, processing unit 110 may cause system 100 to determine (e.g., with determining module 1404), based on the one or more indicators, that the vehicle is proximate to a construction zone, and in accordance with a determination that the vehicle is proximate to a construction zone, output (e.g., with signal outputting module 1406) the signal indicating that the vehicle is proximate to a construction zone.

[0220] In some embodiments processing unit 110 may cause system 100 to, in accordance with the determination that the vehicle is proximate to a construction zone, cause the moving vehicle (e.g., with vehicle moving enabling module 1408) to move in accordance with the construction zone indicators.

[0221] In some embodiments processing unit 110 may cause system 100 to determine (e.g., with determining module 1404), based on the one or more indicators, that the construction zone is a complex construction zone, and in accordance with a determination that the construction zone is a complex construction zone, disable (e.g., with vehicle control disabling module) automated vehicle control functionality.

[0222] FIG. 15 illustrates a process 1500 for implementing lane constraint prediction consistent with the disclosed embodiments. In some embodiments, the steps in process 1500 may be illustrated by the situations illustrated in FIG. 13. According to some embodiments, process 1500 may be implemented by one or more components of navigation system 100, such as at least one processing unit 110. According to some embodiments, process 1500 may be implemented by a system 100 associated with a vehicle 200.

[0223] At step 1502, in some embodiments, a system receives multiple images of a roadway in the vicinity of a moving vehicle associated with the system. In some embodiments, the one or more images are acquired using at least one image capture device, such as image capture devices 122, 124, and/or 126. The plurality of images may be received by the system through a data interface such as data interface 128. In the example illustrated in FIG. 13, system 100 may acquire a plurality of images of roadway 1300 in the vicinity of vehicle 200.

[0224] At step 1504, in some embodiments, the system recognizes, in the multiple images, one or more indicators of a construction zone. In some embodiments, the system may use image analysis techniques to process the one or more images received to detect signs, objects, other vehicles, marked lane lines, and any other objects or characteristics of a surrounding environment. In some embodiments, certain

signs, markings, lane lines, objects, and/or combinations of any such detected aspects may be predetermined by the system to be indicators of a construction zone. In some embodiments, cones, barrels, flashing lights, poles, barriers, and/or markers present in the vicinity of the roadway may be predetermined by the system to be indicators of a construction zone. In some embodiments, predefined traffic signs, such as those bearing text, symbols, or images known to indicate a construction zone, may be predetermined by the system to be indicators of a construction zone. In some embodiments, lane lines may be determined by the system to be indicators of a construction zone. For example, in Germany, construction zones may be marked with special construction zone lane lines painted on roadways, the construction zone lane lines possibly being of a different color or a different appearance than conventional lane lines. The system may detect lane lines of a color, appearance, orientation, and/or location that indicates to the system that the lines are construction zone lane lines, and the system may recognize such construction zone lane lines as indicators of a construction zone.

[0225] In the example of FIG. 13, system 100 may recognize cones 1316, construction road signs 1318a and 1318b, and/or painted construction zone lane line 1326. In some embodiments, any or all of those elements may be predetermined by system 100, or dynamically determined by system 100, to be construction zone indicators.

[0226] At step 1506, in some embodiments, the system determines, based on the one or more indicators, whether the vehicle is proximate to a construction zone. In some embodiments, a system may determine based on the proximity (to the vehicle, roadway, or lane) of construction zone indicators that the vehicle is proximate to a construction zone. In some embodiments, a system may assume based on the orientation or placement with respect to a roadway or a lane that the vehicle is proximate to a construction zone. In some embodiments, a system may determine from the content (e.g., the images, symbols, and/or text) of one or more road signs that the vehicle is proximate to a construction zone. In some embodiments, a minimum threshold number of construction zone indicators must be detected within a maximum threshold time of one another in order for a system to determine that the vehicle is proximate to a construction zone. In the example of FIG. 13, system 100 may determine based on the simultaneous detection of cones 1316, construction road signs 1318a and 1318b, and/or painted construction zone lane line 1326 that vehicle 200 is proximate to a construction zone.

[0227] In some embodiments, if it is determined, at block 1506, that the vehicle is not proximate to a construction zone, then method 1500 may return to step 1502 and may continue to acquire additional images of the roadway as time progresses.

[0228] In some embodiments, if it is determined, at block 1506, that the vehicle is proximate to a construction zone, then method 1500 may proceed to step 1508. At step 1508, in some embodiments, in accordance with a determination that the vehicle is proximate to a construction zone, the system outputs a signal indicating that the vehicle is proximate to a construction zone. In some embodiments, the signal output by system 100 may be an electronic signal or an electronic message that communicates to any component of the system or to any associated system that a construction zone is detected. For example, a signal may be sent to a display system associated with the system, or a signal may be sent to any other local or remote system communicatively coupled with the system. In some embodiments, the signal may be perceptible by a driver of the vehicle. For example, the signal may be a visible signal (such as a lighted

indicator or a displayed element on a computer or entertainment system display), an audible signal emitted by a speaker system, and/or a haptic signal such as a vibration.

[0229] At step 1510, in some embodiments, in accordance with the determination that the vehicle is proximate to a construction zone, the system enables the moving vehicle to move in accordance with the construction zone indicators. In some embodiments, in response to determining that the vehicle is proximate to a construction zone, the system may cause the vehicle to be accelerated, decelerated, steered, or otherwise controlled and moved in accordance with the detected construction zone indicators. In some embodiments, the vehicle may be automatically slowed to a reduced speed to increase safety. In some embodiments, the vehicle may be steered or caused to change lanes so as to follow construction zone lane lines, avoid cones/barrels/poles/barriers or other construction zone indicators, or follow instructions indicated on road signs indicating or associated with the construction zone. In the example of FIG. 13, system 100 may modify a course of travel of vehicle 200 such that vehicle 200 follows detected lane constraint 1314 in accordance with construction zone lane line 1326 (rather than lane constraint 1312 detected in accordance with lane line 1324), avoids cones 1316, and obeys traffic sign 1318b (which includes a symbol instructing vehicle 200 to merge left). In some embodiments, vehicle 200 may be caused to accordingly merge from lane 1304 left into lane 1302.

[0230] At step 1512, in some embodiments, the system determines, based on the one or more indicators, whether the construction zone is a complex construction zone. In some embodiments, a complex construction zone may be a construction zone having one or more predefined or dynamically determined characteristics associated by the system with a construction zone in which autonomous and/or driver-assist control of a vehicle may be unsafe, in which autonomous and/or driver assist control of a vehicle may be unreliable or undependable, or in which control of a vehicle should be returned to a driver of the vehicle. In some embodiments, a complex construction zone may be a construction zone where an obstruction exists in a lane or on a roadway on or in which the vehicle is traveling. In some embodiments, a complex construction zone may be a construction zone in which a predefined or dynamically determined threshold number of cones/barrels/poles/barriers or other construction zone indicators are detected within a threshold distance or threshold time of one another. That is, in some embodiments, when a construction zone is densely populated with construction zone indicators, it may be determined to be a complex construction zone. In the example of FIG. 13, in some embodiments, roadway 1300 may be determined by system 100 to contain a complex construction zone, based on the presence of cones 1316 on roadway 1300, cones 1316 in lane 1304, a minimum threshold number of cones 1316 being detected, and or a minimum threshold number of total construction zone indicators being detected.

[0231] In some embodiments, if it is determined, at block 1512, that the construction zone is not a complex construction zone, then method 1500 may return to step 1502 and may continue to acquire additional images of the roadway as time progresses.

[0232] In some embodiments, if it is determined, at block 1512, that the construction zone is a complex construction zone, then method 1500 may proceed to step 1514. At step 1514, in some embodiments, in accordance with a determination that the construction zone is a complex construction zone, the system may disable automated vehicle control functionality. In some embodiments, driver assist functionality may be disabled. In some embodiments, fully automated driving functionality may be disabled. In some embodiments, full control (e.g., acceleration, deceleration, steering, etc.) of

vehicle 200 may be returned by system 100 to the driver of vehicle 200.

[0233] 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.

[0234] 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. Further, it is to be understood that the functionality of the program modules described herein is not limited to the particular type of program module named herein, but rather can be programmed in any suitable manner across among one or more modules or other software constructs.

[0235] 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.

IPR2025-01034
Tesla EX1002 Page 1655

Espacenet

Claims: EP3183688 (A1) — 2017-06-28

RECOGNITION AND PREDICTION OF LANE CONSTRAINTS AND CONSTRUCTION AREAS IN NAVIGATION

Claims not available for EP3183688 (A1)
Claims of corresponding document: US2016046290 (A1)

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

CN107111742 (A) US2016046290 (A1) WO2016027270 (A1) CN113158820 (A) US2019077400 (A1) US2021291827 (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 device mountable on a vehicle, comprising:
a camera adapted to capture, when mounted on the vehicle, and while the vehicle is in motion, multiple images of a roadway within a field of view of the camera;
a processor configured to:
obtain the multiple images from the camera;
apply image processing to the multiple images to recognize a first roadway lane constraint appearing in at least one of the multiple images;
determine, based on the multiple images, that one or more defined lane prediction conditions are satisfied;
in accordance with a determination that one or more defined lane prediction conditions are satisfied, predict a location of a second roadway lane constraint;
issue a notification or a control signal to control a motion of the vehicle according to the predicted location of the second roadway lane constraint.

2. The device according to claim 1, wherein the processor is configured to issue a notification or a control signal to direct a motion of the vehicle according to a location of the first roadway lane constraint.

3. The device according to claim 1, wherein the processor is configured to issue a notification or a control signal to direct a motion of the vehicle along a lane or an area between the first roadway lane constraint and the second roadway lane constraint.

4. The device according to claim 1, wherein the processor is configured to issue a

notification or a control signal for moving the vehicle issue a notification or a control signal to direct a motion of the vehicle to avoid the first roadway lane constraint and the second roadway lane constraint.

5. The device of claim 1, wherein the one or more defined lane prediction conditions comprise that the second roadway lane constraint ceases to be recognized, based on the multiple images, as the vehicle moves on the roadway.

6. The device of claim 1, wherein the one or more defined lane prediction conditions comprise that the first roadway lane constraint has been recognized for at least a threshold time or a threshold distance.

7. The device of claim 1, wherein the one or more defined lane prediction conditions comprise that a guard rail is present in the vicinity of the vehicle.

8. The device of claim 7, wherein determining that a guard rail is present in the vicinity of the vehicle comprises detecting, in the one or more images, a pattern created by posts supporting the guard rail.

9. The device of claim 7, wherein the one or more defined lane prediction conditions comprise that the guard rail is casting a shadow onto the roadway.

10. The device of claim 1, wherein the one or more defined lane prediction conditions comprise that multiple substantially parallel edges are detected, in the multiple images, on one side of a lane on the roadway.

11. The device of claim 1, wherein predicting the location of the second roadway lane constraint comprises accounting for a location of the first roadway lane constraint.

12. The device of claim 1, wherein predicting the location of the second roadway lane constraint comprises predicting a width of a lane based on information about a previously determined width of the lane.

13. The device of claim 1, wherein the processor is configured to:
determine whether one or more defined lane prediction cancellation conditions are satisfied, and
in accordance with a determination that one or more defined lane prediction cancellation conditions are satisfied, cease to predict the location of the second roadway lane constraint.

14. The device of claim 13, wherein the one or more defined lane prediction cancellation conditions comprise that, after ceasing to be recognized, the second roadway lane constraint is recognized based on the one or more images.

15. The device of claim 13, wherein the one or more defined lane prediction cancellation conditions comprise that a driver is moving the vehicle between lanes.

16. The device of claim 13, wherein the one or more defined lane prediction cancellation conditions comprise that the vehicle is traversing a curve that exceeds a predefined sharpness threshold.

17. The device of claim 13, wherein determining whether one or more defined lane

Tesla EX1002 Page 1657 IPR2025-01034

prediction cancellation conditions are satisfied comprises:
selecting a maximum lane prediction distance setting, and
determining whether a distance traveled by the vehicle while predicting the location of the second roadway lane constraint is greater than a maximum lane prediction distance associated with the maximum lane prediction distance setting.

18. The device of claim 17, wherein selecting a maximum lane prediction distance setting comprises, in accordance with a determination that a first lane is merging with a second lane in the vicinity of the vehicle, selecting a maximum lane prediction distance setting such that the maximum lane prediction distance is greater than a default maximum lane prediction distance.

19. The device of claim 17, wherein selecting a maximum lane prediction distance setting comprises, in accordance with a determination that defined guard-rail conditions are satisfied, selecting a maximum lane prediction distance setting having no maximum lane prediction distance.

20. The device of claim 17, wherein selecting a maximum lane prediction distance setting comprises, in accordance with a determination that defined guard-rail conditions are satisfied, selecting a maximum lane prediction distance setting such that the maximum lane prediction distance is greater than a default maximum lane prediction distance.

21. A computer system comprising:
a memory that stores instructions, and
a processor that executes the instructions to cause the system to:
obtain multiple images from a camera, the camera being adapted to capture, when mounted on the vehicle, and while the vehicle is in motion, images of the roadway within a field of view of the camera;
apply image processing to the multiple images to recognize a first roadway lane constraint appearing in at least one of the multiple images;
determine, based on the multiple images, that one or more defined lane prediction conditions are satisfied;
in accordance with a determination that one or more defined lane prediction conditions are satisfied, predict a location of a second roadway lane constraint;
issue a notification or a control signal to direct a motion of the vehicle according to the predicted location of the second roadway lane constraint.

22. A method comprising:
at a system comprising a memory that stores instructions and a processor that executes the instructions:
obtaining multiple images from a camera, the camera being adapted to capture, when mounted on a vehicle, and while the vehicle is in motion, images of a roadway within a field of view of the camera;
applying image processing to the multiple images to recognize a first roadway lane constraint appearing in at least one of the multiple images;
determining, based on the multiple images, that one or more defined lane prediction conditions are satisfied;
in accordance with a determination that one or more defined lane prediction conditions are satisfied, predicting a location of a second roadway lane constraint;
issuing a notification or a control signal to direct a motion of the vehicle according to the predicted location of the second roadway lane constraint.

Tesla EX1002 Page 1658 IPR2025-01034

23. A vehicle comprising:
a body;
a camera adapted to capture, when mounted on the vehicle, and while the vehicle is in motion, multiple images of a roadway within a field of view of the camera;
a processor configured to:
obtain the multiple images from the camera;
apply image processing to the multiple images to recognize a first roadway lane constraint appearing in at least one of the multiple images;
determine, based on the multiple images, that one or more defined lane prediction conditions are satisfied;
in accordance with a determination that one or more defined lane prediction conditions are satisfied, predict a location of a second roadway lane constraint;
issue a notification or a control signal to control a motion of the vehicle according to the predicted location of the second roadway lane constraint.

24. A device mountable on a vehicle, comprising:
a camera adapted to capture, when mounted on the vehicle, and while the vehicle is in motion, multiple images of a roadway within a field of view of the camera, and
a processor configured to:
obtain the multiple images from the camera;
apply image processing to the multiple images to recognize, in the multiple images, one or more indicators of a construction zone, and
in accordance with recognizing the one or more indicators of a construction zone, output a signal indicating that the vehicle is proximate to a construction zone.

25. The device of claim 24, wherein the processor is configured to:
determine, based on the one or more indicators, that the vehicle is proximate to a construction zone, and
in accordance with the determination that the vehicle is proximate to a construction zone, output the signal indicating that the vehicle is proximate to a construction zone.

26. The device of claim 24, wherein the output signal is perceptible by a driver of the vehicle.

27. The device of claim 24, wherein the output signal is a control signal that controls a motion of the vehicle.

28. The device of claim 24, wherein the one or more indicators of a construction zone comprise one of: cones present in the vicinity of the roadway, barrels present in the vicinity of the roadway, poles present in the vicinity of the roadway, and barriers present in the roadway.

29. The device of claim 24, wherein the one or more indicators of a construction zone comprise predefined traffic signs.

30. The device of claim 24, wherein the one or more indicators of a construction zone comprise construction zone lane lines.

31. The device of claim 24, wherein the instructions cause the system to:
in accordance with the determination that the vehicle is proximate to a construction zone, cause the moving vehicle to move in accordance with the construction zone

indicators.

32. The device of claim 31, wherein causing the moving vehicle to move in accordance with the construction zone indicators comprises one of: causing reducing a speed of the vehicle; causing changing a lane in which the vehicle is traveling; enabling avoiding cones, barrels, poles, and/or barriers; and causing following construction zone lane constraints rather than conventional lane constraints.

33. The device of claim 24, wherein the instructions cause the system to:
determine, based on the one or more indicators, that the construction zone is a complex construction zone, and
in accordance with a determination that the construction zone is a complex construction zone, disable automated vehicle control functionality.

34. The device of claim 33, wherein the determination that the construction zone is a complex construction zone comprises the recognition of one of: cones present in a lane in which the vehicle is present, barrels present in a lane in which the vehicle is present, poles present in a lane in which the vehicle is present, and barriers present in a lane in which the vehicle is present.

35. The device of claim 33, wherein the determination that the construction zone is a complex construction zone comprises the recognition of a number of cones, barrels, poles, and/or barriers in the vicinity of the vehicle exceeding a predefined threshold number.

36. A computer system comprising:
a memory that stores instructions, and
a processor that executes the instructions to cause the system to:
obtain multiple images from a camera, the camera being adapted to capture, when mounted on the vehicle, and while the vehicle is in motion, images of the roadway within a field of view of the camera;
apply image processing to the multiple images to recognize, in the multiple images, one or more indicators of a construction zone, and
in accordance with recognizing the one or more indicators of a construction zone, output a signal indicating that the vehicle is proximate to a construction zone.

37. A method comprising:
at a system comprising a memory that stores instructions and a processor that executes the instructions:
obtaining multiple images from a camera, the camera being adapted to capture, when mounted on a vehicle, and while the vehicle is in motion, images of a roadway within a field of view of the camera;
applying image processing to the multiple images to recognize, in the multiple images, one or more indicators of a construction zone, and
in accordance with recognizing the one or more indicators of a construction zone, output a signal indicating that the vehicle is proximate to a construction zone.

38. A vehicle, comprising:
a body;
a camera adapted to capture, when mounted on the vehicle, and while the vehicle is in motion, multiple images of a roadway within a field of view of the camera, and
a processor configured to:

IPR2025-01034
Tesla EX1002 Page 1660

obtain the multiple images from the camera;
apply image processing to the multiple images to recognize, in the multiple images, one or more indicators of a construction zone, and
in accordance with recognizing the one or more indicators of a construction zone, output a signal indicating that the vehicle is proximate to a construction zone.

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

(11) Veröffentlichungsnummer:
(11) Publication number: **EP 3 183 688 A0**
(11) Numéro de publication:

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

WO 2016/027270 (Art. 153(3) EPÜ).

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

WO 2016/027270 (Art. 153(3) EPC).

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

WO 2016/027270 (art. 153(3) CBE).

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

(11) EP 3 514 032 A1

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
24.07.2019 Bulletin 2019/30

(21) Application number: **19160915.5**

(22) Date of filing: **04.12.2014**

(51) Int Cl.:
B60W 30/14 *(2006.01)* ***G06K 9/00*** *(2006.01)*
B60K 31/00 *(2006.01)* *G06T 7/00* *(2017.01)*

(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

(30) Priority: **04.12.2013 US 201361911490 P**
14.05.2014 US 201461993084 P
14.05.2014 US 201461993111 P
23.06.2014 US 201461015524 P
09.07.2014 US 201462022221 P
21.08.2014 US 201462040224 P
21.08.2014 US 201462040269 P

(62) Document number(s) of the earlier application(s) in accordance with Art. 76 EPC:
14843184.4 / 3 092 599

(71) Applicant: **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)

Remarks:
This application was filed on 05-03-2019 as a divisional application to the application mentioned under INID code 62.

(54) ADJUSTING VELOCITY OF A VEHICLE FOR A CURVE

(57) Systems and methods use cameras to provide autonomous navigation features. In one implementation, 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: 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.

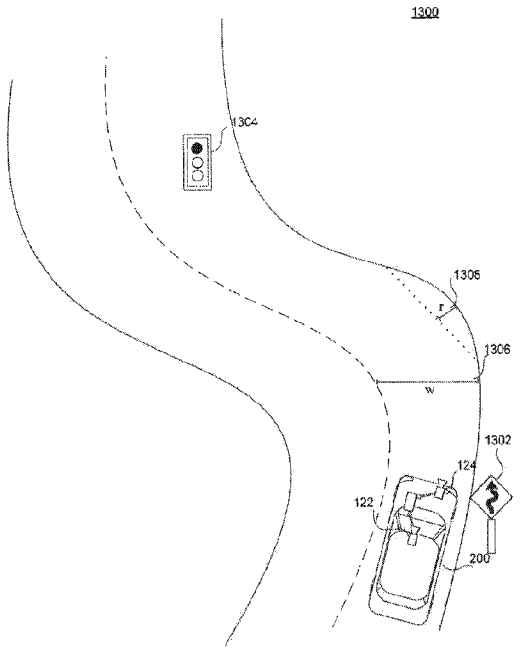


FIG. 13A

EP 3 514 032 A1

IPR2025-01034
Tesla EX1002 Page 1663

Description

[0001]

Cross References to Related Applications

[001] This application claims the benefit of priority of United States Provisional Patent Application No. 61/911,490, filed on December 4, 2013; United States Provisional Patent Application No. 61/993,084, filed on May 14, 2014; United States Provisional Patent Application No. 61/993,111, filed on May 14, 2014; United States Provisional Patent Application 62/015,524, filed on June 23, 2014; United States Provisional Patent Application 62/022,221, filed on July 9, 2014; United States Provisional Patent Application No. 62/040,224, filed on August 21, 2014; and United States Provisional Patent Application 62/040,269, filed on August 21, 2014. All of the foregoing applications are incorporated herein by reference in their 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 provide autonomous vehicle navigation features.

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, 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

IPR2025-01034
Tesla EX1002 Page 1664

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.

[006] 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

IPR2025-01034
Tesla EX1002 Page 1665

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.

[007] 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.

[008] 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.

[009] 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

IPR2025-01034
Tesla EX1002 Page 1666