

Samsara Inc.
v.
Motive Technologies, Inc.

IPR2025-00574

U.S. Patent No. 11,875,580



Petitioner's Demonstratives

May 27, 2026

Samsara EX1088
Samsara v. Motive
IPR2025-00574

1

Obviousness Grounds

Claims	References
1-4	Choe (EX1005, EX1006)
1-5	Choe, Khan (EX1007), and Ali (EX1008)
1-4	Choe and Workman (EX1011)
5-7	Choe and Westmacot (EX1010)
1-7	Choe, Westmacot, Khan, and Ali
1-7	Choe, Westmacot, and Workman

Agenda

'580 Patent

- Overview
- POSA Definitions
- Claim Construction – “predictive model”

Grounds

- Choe (Claims 1-4)
- Choe in view of Khan/Ali (Claims 1-5)
- Choe in view of Workman (Claims 1-4)
- Choe in view of Westmacot (Claims 5-7)
- Choe in view of Westmacot and Khan/Ali (Claims 1-7)
- Choe in view of Westmacot and Workman (Claims 1-7)

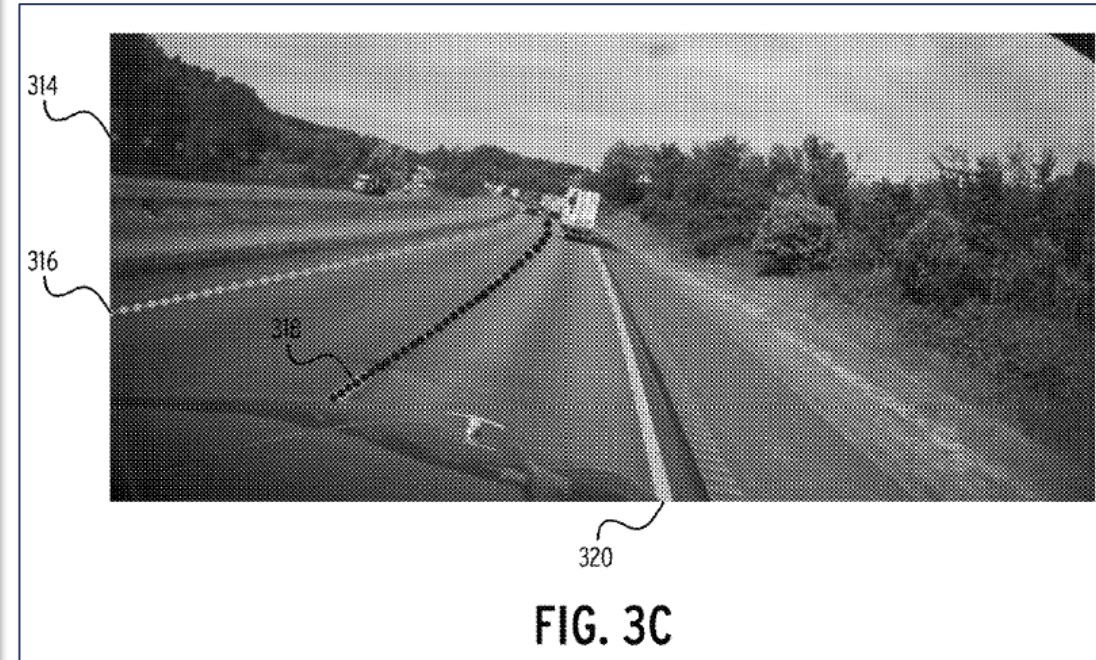
The '580 Patent



(12) United States Patent
Hassan et al.

(10) Patent No.: US 11,875,580 B2
(45) Date of Patent: Jan. 16, 2024

In process 206, the camera initialization service 232 receives the video transmitted in transmission 204. In an embodiment, the camera initialization service 232 can comprise a remote endpoint configured to receive network requests and process the request body (i.e., the video). In an embodiment, the camera initialization service 232 detects lane and horizon lines in the video using a predictive model, such as a deep learning or other AI model. In an embodiment, process 206 can comprise inputting the video into a predictive model, such as a CNN or other type of deep learning model to identify lane and horizon markers in a given video file automatically. Details of this type of predictive model are provided in the description of FIG. 6 and are not repeated herein. Additionally, examples of such models are provided in FIGS. 3A and 3B. In general, the model can output a set of lane and horizon lines identified by key points or markers.



EX1001, FIG. 3C.

EX1001, 5:21-37.

The '580 Patent: Independent Claim 1

1. A method comprising:
 - [1A] receiving, over a network from a camera device, a video comprising a set of image frames;
 - [1B] identifying one or more lines in the video using a predictive model, the one or more lines including a horizon line;
 - [1C] computing at least one camera parameter based on the one or more lines;
 - [1D] overlaying the one or more lines on the video to generate an overlaid video;
 - [1E] transmitting the overlaid video to an annotator device for manual review;
 - [1F] receiving a confirmation from the annotator device, the confirmation indicating that the one or more lines are accurate; and
 - [1G] transmitting data representing the at least one camera parameter to the camera device.

District Court Claim Constructions

Term	Samsara's Construction	Motive's Construction
predictive model	plain and ordinary meaning – a model that makes predictions	model that predicts values using non-geometric calculation
camera parameter	property of the camera itself	value(s) associated with one or more operating characteristics of a camera
order of steps	Step [b] must be performed after step [a] Step [c] must be performed after step [b] Step [d] must be performed after step [b] Step [e] must be performed after step [d] Step [f] must be performed after step [e] Step [g] must be performed after steps [c] and [f]	plain and ordinary meaning

EX1086, 9, 12, 15; EX1087, 2, 10, 14.

POSA Definitions

Samsara's Definition	Motive's Definition
Someone with a Bachelor's in Computer Science with a focus on computer/machine vision, and at least two years of experience with each of: (1) machine learning (including neural network) methods as applied to machine/computer vision and (2) classical machine/computer vision algorithms (e.g., edge detection, line identification, computation of camera parameters from horizon line, etc.). Additional education could serve as a substitute for the experience requirement.	A bachelor's degree in engineering or computer science with one- or two-years training and/or experience in design of software for camera/sensor systems in the automotive industry.

Pet., 7-8.

POR, 8-9.

Dr. Darrell

8. However, if the Board adopts Motive's definition of a POSA, that definition would not materially impact my opinions. As I explain in Section II.A, an engineer with one- or two-years training and/or experience in the design of software for camera/sensor systems in the automotive industry would have been familiar with computer vision, including the machine learning methods and machine/computer vision algorithms I specified in my POSA definition.

example, at Mobileye, an entry level position requiring a Bachelor's of Science in Computer Science or similar field (e.g., an Algorithm Developer) would have been prototyping solutions for complex computer vision problems. EX1083. A POSA with two years of experience would have also gained knowledge about machine learning, deep neural networks (DNNs), and simultaneous localization and mapping (SLAM), which involves edge detection and line identification. EX1083.

Dr. Darrell, Petitioner's expert, exceeds the level of skill of a POSA.



Trevor J. Darrell, Ph.D.

Professional Experience

- UC Berkeley

Professor in the Computer Science Division (EECS)

Member of the Steering Committee for Berkeley Artificial Research Center (BAIR)

- Massachusetts Institute of Technology

Professor in the Dep't of Electrical Engineering and Computer Science (1999-2008)



Industry Experience with Automotive Industry

- Chief Science Advisor of Nexar (2016-2024)
- Collaborated with Nvidia on deep learning for autonomous driving
- Affiliated Director of California's Partners for Advanced Transit and Highways (PATH)
- Worked on BDD100K, at the time the largest open-source video dataset for autonomous driving
- Many academic papers and invited talks on computer vision for autonomous vehicles
- Advised several PhD dissertations related to autonomous driving

“predictive model”

Samsara’s Construction	Motive’s Construction
Plain and Ordinary Meaning – a model that makes predictions	a model that predicts values using non-geometric calculation

Pet., 8; EX1087, 10.

EX1086, 12.

- The '580 Patent does not define “predictive model.”
- The '580 Patent does not disavow geometric calculations from the scope of the claims.
- The prosecution history does not support Motive’s disavowal argument.

The '580 Patent does not disavow geometric calculations.



“To constitute disclaimer, there must be a clear and unmistakable disclaimer.”

Thorner v. Sony Computer Entertainment Am. LLC, 669 F.3d 1362, 1366-67 (Fed. Cir. 2012).

➤ Statements in the '580 patent do not clearly and unmistakably exclude geometric calculations.

In an embodiment, the neural network is trained to predict lane and horizon key points 308. During testing, image frames are fitted with lane lines and horizon lanes based on estimated key points. As described, in some embodiments, human annotators can be used to perform this fitting, as described in the description of FIG. 2. In other embodiments, a geometric process can be performed to identify lane and horizon lines. As discussed in previous figures, once the horizon line is determined, downstream processes can use the horizon line to determine the height of a camera and its viewing angle (i.e., road plane normal) with respect to the ground. Further, from these, downstream processes can determine the road-plane normal vector.

EX1001, 8:6-19.

requests and process the request body (i.e., the video). In an embodiment, the camera initialization service 232 detects lane and horizon lines in the video using a predictive model, such as a deep learning or other AI model. In an embodiment, process 206 can comprise inputting the video into a predictive model, such as a CNN or other type of deep learning model to identify lane and horizon markers in a given video file automatically. Details of this type of predictive model are provided in the description of FIG. 6 and are not repeated herein. Additionally, examples of such models are provided in FIGS. 3A and 3B. In general, the model can output a set of lane and horizon lines identified by key points or markers.

EX1001, 5:22-37.

The '580 Patent does not disavow geometric calculations.

- Relaxing requirements of having three lane lines ≠ disavowing geometric calculations altogether.

For example, the use of the neural network can relax constraints that the vehicle needs to be driving on a highway having three or more lane lines. Indeed, just two lane-lines with a known lane-width would be sufficient for camera initialization. That is, since the model is capable of predicting the horizon line directly, previous requirements of having three lane lines in frame (e.g., to geometrically calculate a horizon line) can be relaxed.

EX1001, 8:39-46.

- The '580 Patent's title of the invention uses geometry.

(54) **CAMERA INITIALIZATION FOR LANE
DETECTION AND DISTANCE ESTIMATION
USING SINGLE-VIEW GEOMETRY**

EX1001, (54).



"We have used the title of a patent to aid in claim construction."

Ruckus Wireless, Inc. v. Innovative Wireless Solutions, LLC,
824 F.3d 999, 1003 n.2 (Fed. Cir. 2016).

The prosecution history does not support Motive's construction.

In response to a rejection, the applicant argued:

- **Step 1B:** “Kwant1 does not disclose using a *predictive model* to detect a horizon line. Instead, Kwant1 describes predicting *lane lines*. Kwant1 then describes extending these lines and, using a geometric approach, identifying the horizon lines.”
- **Steps 1C and 1G:** “[Kwant1’s] ‘parameters’ are not parameters related to a *camera*, they are of a ‘parametric representation of a single lane line.’”
- **Step 1D:** “Kwant1 also fails to disclose ‘overlaying’ lines on a video to generate an overlaid video.”
- **Step 1E:** “Kwant1 also fails to describe transmitting an overlaid video ton an ‘annotator device for manual review.’”
- **Step 1F:** “Kwant2 fails to describe a ‘confirmation’ from an ‘annotator device.’”

The prosecution history does not support Motive's construction.

REASONS FOR ALLOWANCE

Prior art was found for the claims as follows:

Re. Claim 1, Kwantl et al., [US Pub. No.: 2019/0034740 A1] discloses:

. . .

identifying one or more lines in the video using a predictive model [the output of a lane feature detection process include lane models representing the detected lanes or features, and associated prediction confidence values indicating a probability that the processed features of the input image are predictive of the detected lane features] 0046], the one or more lines including a horizon line [determining a horizontal line in the image at which a maximum number of the one or more extended road lanes crosses over a minimum horizontal extent of the horizontal line [Abstract];

Whereas the Examiner maintained the rejection for Step 1B, the Examiner did not maintain the rejection for:

- Steps 1C and 1G
- Step 1D
- Step 1E

No Disclaimer



"[I]t is not clear from the record that allowance was based on this understanding, or that the patentee disclaimed this claim scope."

Convolve, Inc. v. Compaq Computer Corp.,
812 F.3d 1313, 1322 (Fed. Cir. 2016)

EX1002, 0169-70; Pet. Reply, 4-5.

The Petition argued only that step [1C] does not need to occur before step [1F].

Sur-Reply

Similarly, with respect to the order of the steps in claim 1, while in the Petition Samsara argued that “neither the claims themselves, nor the specification require a specific order for this element [1C]” (Pet., p.28), in district court Samsara argues that “[t]he steps of the claim must be performed in the following order ... Step [c] must be performed after step [b].” EX_2007, p.4-5. These positions are diametrically opposed.

. . .

downstream. Reply, pp.18-19. Further, in the Petition Samsara argued that the claimed steps did not need to be performed in any order. Pet., p.28. Thus, this too is contrary to Samsara’s positions on Reply.

Sur-Reply, 5, 22.

Petition

4. [1C] “computing at least one camera parameter based on the one or more lines”

Choe discloses this element. EX1003, ¶98.

Choe computes camera parameters (e.g., calibration factors or parameters for calibrating the camera’s pitch, yaw, and roll) based on the horizon line and/or the lane lines. EX1003, ¶98. Choe states: “[a] first calibration factor or first correction value is determined for calibrating a pitch angle of the camera based on a difference between the initial horizon line and the updated horizon line.”

EX1005, ¶[0015]. Choe also computes other parameters based on the lane lines. EX1005, ¶¶[0014] (“A user can calibrate any of the sensors such as a pitch angle, a yaw angle, and/or a roll angle....”), [0049], [0052]-[0053], [0057], [0059], FIG. 7.

PO may allege that Choe computes these camera parameters *after* the manual calibration. Even if true, Choe still satisfies [1C] because the computation is nonetheless “based on the one or more lane lines.” EX1003, ¶98. And neither the claims themselves, nor the specification require a specific order for this element.

Pet., 28.

Ground 1: Choe (Claims 1-4)

Choe discloses Step 1A:

“receiving, over a network from a camera device, a video comprising a set of image frames;”

Choe's sensor calibration system 125 receives video frames over a network.

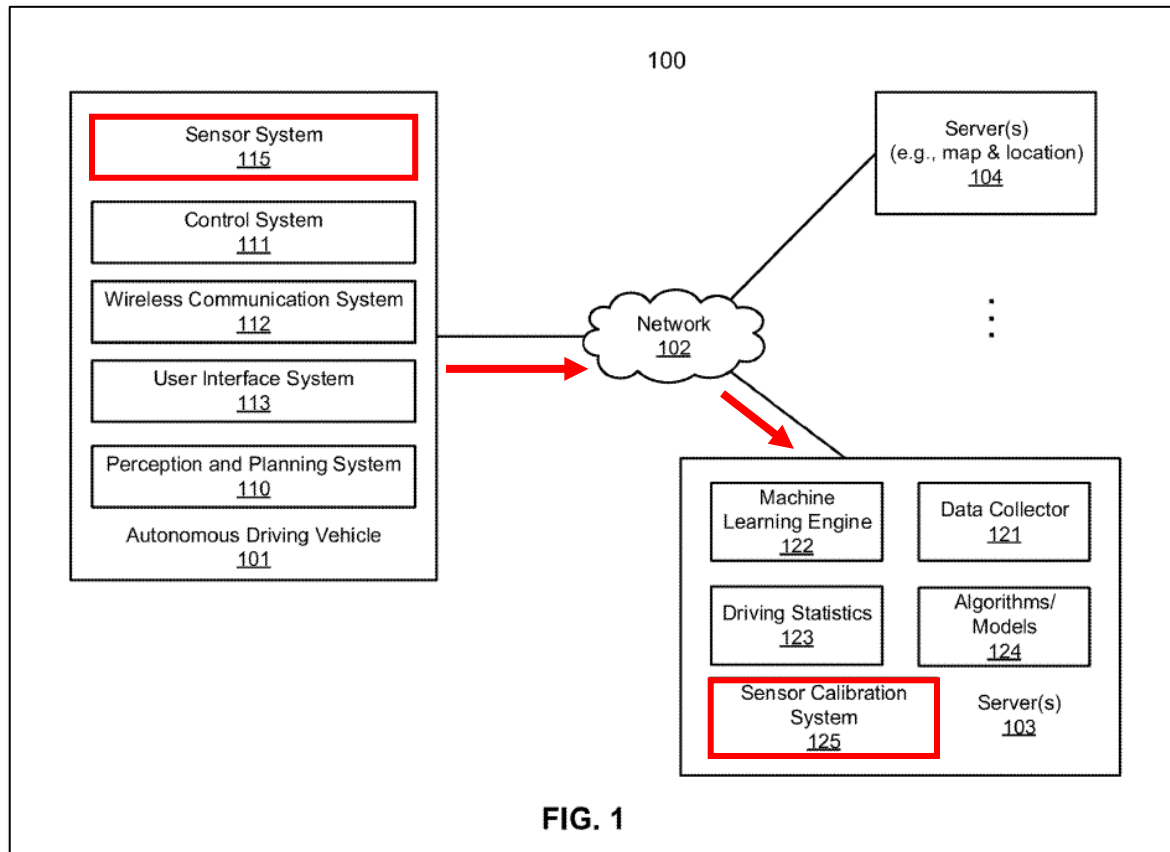


FIG. 1

EX1005, FIG. 1; Pet., 26; see also EX1005, [0031].

[0022] Referring now to FIG. 2, in one embodiment, sensor system 115 includes, but it is not limited to, one or more cameras 211, global positioning system (GPS) unit 212, inertial measurement unit (IMU) 213, radar unit 214, and a light detection and range (LIDAR) unit 215. GPS

...

Cameras 211 may be still cameras and/or video cameras.

EX1005, [0022]; Pet., 25.



EX1005/EX1006, FIGs. 6A, 6B; Pet., 25.

Choe discloses Step 1B: “identifying one or more lines in the video . . . , the one or more lines including a horizon line;”



EX1005/EX1006, FIG. 6B; Pet., 27.

[0015] According to one embodiment, in response to a first image captured by a camera of an ADV, a horizon line is determined based on the camera's hardware settings. The horizon line represents a vanishing point from a view point of the ADV based on an initial or default pitch angle of the camera. One or more lane lines are determined based on the first image via a perception process performed on the first image.

EX1005, [0015]; Pet., 27.

In addition, horizon line 610 is determined and placed on a location within image 601 based on the hardware setting of the targeted camera. Image 601 may be rendered

In addition, optional image 603, referred to herein as a 2D re-projected view, is generated based on 2D view image 601 and 3D view image 602, including projecting lane lines 621-623 onto the 2D view.

EX1005, [0055], [0056]; Pet., 27.

Choe discloses Step 1B: “identifying one or more lines in the video using a predictive model, the one or more lines including a horizon line;”

Choe

[0030] Based on driving statistics 123, machine learning engine 122 generates or trains a set of rules, algorithms, and/or predictive models 124 for a variety of purposes. In one embodiment, algorithms 124 may include algorithms for autonomous driving a vehicle as described further below. Algorithms 124 may further include sensor calibration algorithms, which will be described further below. Algorithms 124 can then be uploaded on ADVs to be utilized during autonomous driving in real-time.

EX1005, [0030]; Pet., 27.

Dr. Darrell

94. In view of Choe’s express teachings of using predictive algorithms for camera calibration and the industry’s wide use of predictive algorithms to identify horizon and lane lines, *see* Section IV.B.1, a POSA would have understood that Choe’s predictive models 124 would have predicted these lines. Once the initial horizon and lane lines are identified, the user would then adjust the position of the lines. For example, “[t]he input signal may represent an incremental adjustment for adjusting the position of the horizon line.” Choe, [0015].

EX1003, [0094]; Pet. Reply, 8.

Mr. Andrews

46. The predictive models 124 of Choe are generated by a machine learning engine 122 “based on driving statistics 123.” *Id.* at [0044]. Thus, these “predictive models 124” are based on “driving statistics 123” and are used for sensor calibration “to determine a set of calibration parameters.” Nothing in Choe teaches or suggests “identifying one or more lines in the video using a predictive model, the one or more lines including a horizon line.”

EX2003, [0046] (POPR Declaration); Pet. Reply, 7.

Choe discloses Step 1D: “overlaying the one or more lines on the video to generate an overlaid video;”



EX1005/EX1006, FIG. 6B; Pet., 27, 34.

[0016] In one embodiment, the first image with the horizon line superimposed thereon is displayed within a first display area of a display device. When the position of the horizon line is updated, the first image is updated on the first display area to update the position of the horizon line to give a visual feedback to a user. In one embodiment, a second image is displayed within a second display area of the display device. The second image includes the lane lines superimposed thereon from a top view or a 3D view based on the 2D view of the first image. When the horizon line is

EX1005, [0016]; Pet., 29, 33.

Cameras 211 may be still cameras and/or video cameras.

EX1005, [0022]; Pet., 25; Pet. Reply, 12.

Choe discloses Step 1D: “overlaying the one or more lines on the video to generate an overlaid video;”

Choe:

frame id: 1109

frame id: 1110



EX1005/EX1006, FIGs. 6A, 6B; Pet., 25.

POR:

But the claim requires generating an “overlaid video,” which must include “a set of image frames,” not just a single frame.

POR, 18-19.

Petitioner Reply:

The user interface is overlaying frames 1109 and 1110. Two image frames is a “set” of image frames. EX1061, 103:19-20.

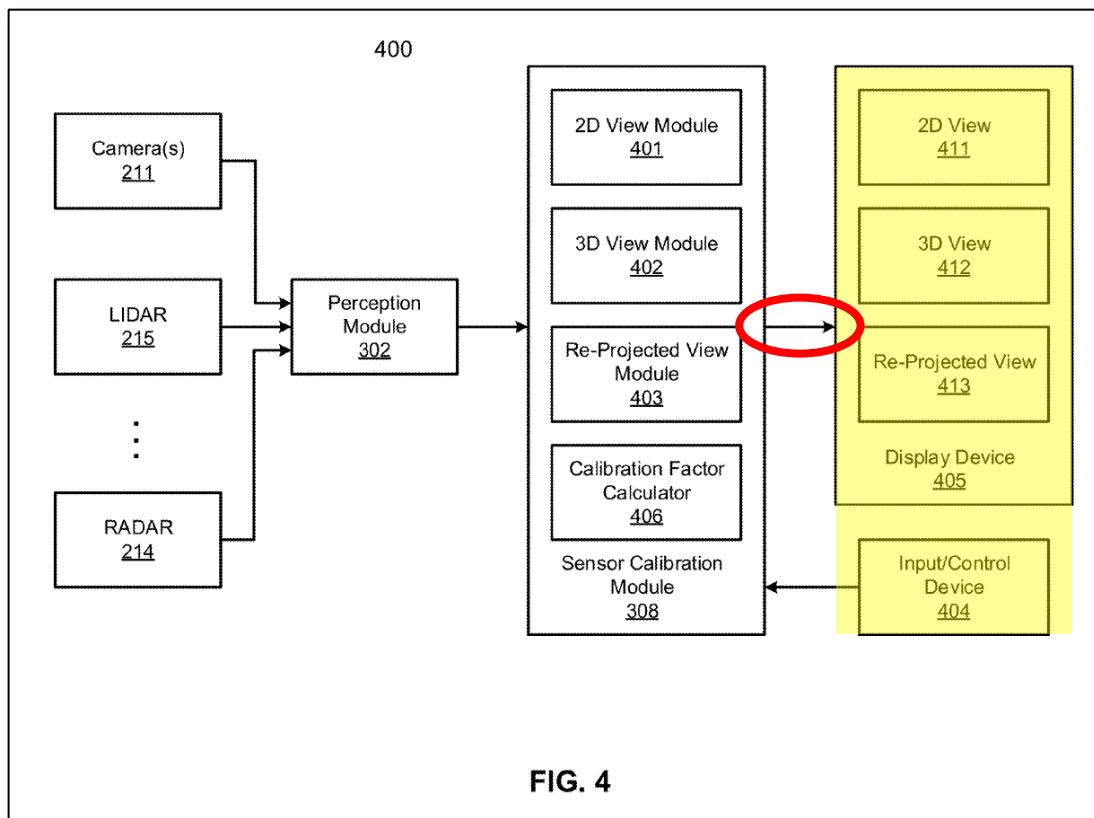
Pet. Reply, 12.

'580 Patent: A video comprises “one or more frames.”

In a second embodiment, an on-vehicle neural network is disclosed. In this embodiment, a camera records a video comprising one or more image frames.

EX1001, 1:53-55; Pet., 2.

Choe discloses Step 1E: “transmitting the overlaid video to an annotator device for manual review;”

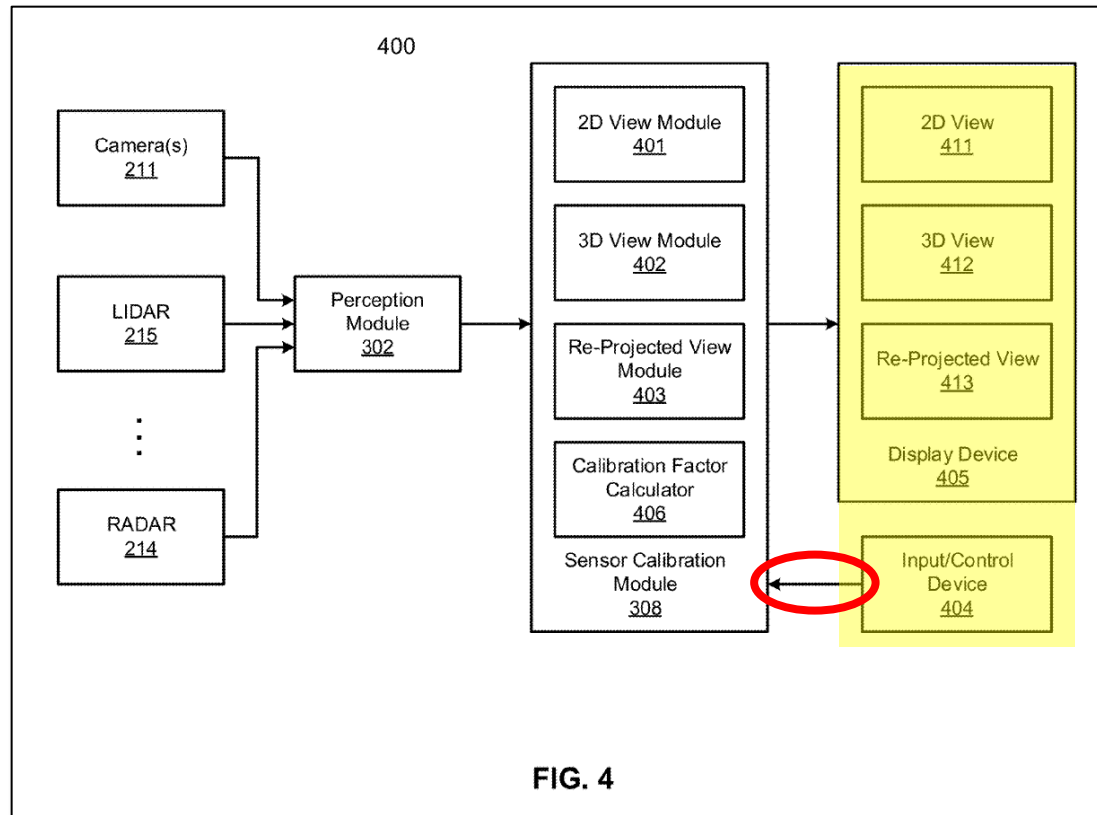


EX1005, FIG. 4; Pet., 31.

[0047] The re-projected 2D view is then displayed on display device 405 as part of 2D re-projected view 413. A user can then enter feedback or adjustment using input device 404 such as a keyboard or joystick. In response to the input provided by the user, modules 411-413 adjust certain parameters, render the images again, and display the updated images on display device 405. As a result, the user can calibrate the sensors until the calibration results as shown on display device 405 are satisfied. Once the calibration results are satisfied, calibration factor calculator 406 determines the calibration factor 410 and feeds back to perception module 302 or other software modules such that the software modules can perform proper adjustment at the software level to compensate the hardware settings of the sensors.

[0049] In response to a first input signal (e.g., an up or down arrow keypress) received from an input device (e.g., a keyboard, a joystick, a microphone to receive voice interactive commands), a position of the horizon line is updated based on the first input signal and a position of at least one of the lane lines is updated based on the updated horizon line. The input signal may represent an incremental value for adjusting the position of the horizon line. A first calibration factor or first correction value is determined for calibrating a pitch angle of the camera based on a difference between the initial horizon line and the updated horizon line.

Choe discloses Step 1F: “receiving a confirmation from the annotator device, the confirmation indicating that the one or more lines are accurate; and”



EX1005, FIG. 4; Pet., 32.

[0047] The re-projected 2D view is then displayed on display device 405 as part of 2D re-projected view 413. A user can then enter feedback or adjustment using input device 404 such as a keyboard or joystick. In response to the input provided by the user, modules 411-413 adjust certain parameters, render the images again, and display the updated images on display device 405. As a result, the user can calibrate the sensors until the calibration results as shown on display device 405 are satisfied. Once the calibration results are satisfied, calibration factor calculator 406 determines the calibration factor 410 and feeds back to perception module 302 or other software modules such that the software modules can perform proper adjustment at the software level to compensate the hardware settings of the sensors.

EX1005, [0047]; Pet., 32.

Not Disputed

Choe discloses Step 1C: “computing at least one camera parameter based on the one or more lines;”

Choe

In response to a first input signal (e.g., an up or down arrow key) received from an input device (e.g., keyboard, joystick, voice interactive command), a position of the horizon line is updated based on the first input signal and a position of at least one of the lane lines is updated based on the updated horizon line. The input signal may represent an incremental adjustment for adjusting the position of the horizon line. A first calibration factor or first correction value is determined for calibrating a pitch angle of the camera based on a difference between the initial horizon line and the updated horizon line.

[0017] According to one embodiment, in response to a second input signal received from the input device (e.g., left or right arrow key), the position of at least one of the lane lines is modified based on the second input signal. The positions of the lane lines are updated on the second image displayed on the second display area. A second calibration factor for calibrating a yaw angle of the camera is determined based on the modification of the positions of the lane lines. Similarly, in response to a third signal received from the input device (e.g., shift-left or shift-right arrow key), the lane lines are modified and a third calibration factor for calibrating a roll angle of the camera is determined.

EX1005, [0015], [0017]; Pet., 28; Pet. Reply, 16.

Dr. Darrell

43. Based on determining the lines, Choe then predicts camera parameters, such as the pitch angle of the camera. Choe, [0015]. For example, “[a] first calibration factor or first correction value is determined for calibrating a pitch angle of the camera based on a difference between the initial horizon line and the updated horizon line.” Choe, [0015]. Choe also specifies that the system uses the visualizer to manually calibrate other camera parameters, including the camera’s “yaw angle, and/or a roll angle.” Choe, [0014].

EX1003, [0043]; Pet., 10; Pet. Reply, 11-12.

Mr. Andrews

Q So you agree that pitch, yaw, and roll are examples of camera parameters?

A They're some examples, yes.

EX1061, 10:1-3; Pet. Reply, 11-12.

Neither party alleges that Step 1C must occur before Steps 1D, 1E, 1F.

**Choe discloses Step 1C:
“computing at least one camera parameter based on the one or more lines;”**

Samsara’s Construction	Motive’s Construction
property of the camera itself	value(s) associated with one or more operating characteristics of a camera

EX1087, 2.

EX1086, 9.

Prosecution History

Applicant notes that the specification is clear that a camera parameter is a property of the camera itself (“(e.g., height, viewing angle, road plane normal, etc.)”, ¶ 5, and 6–7) and not a property of a lane line.

EX1002, 0146; Pet. Reply, 4.

Dr. Darrell

60. Motive argues that Choe does not compute a “camera parameter” because Choe only computes “calibration factors.” POR, 18. Motive’s argument is flawed because Choe’s “calibration factors” are camera parameters. Choe’s calibration factors are camera parameters because they relate to the camera itself and are associated with one or more operating characteristics of the camera.

61. After Choe computes calibration factors, Choe then uses the calibration factors to compute the camera’s pitch, yaw, and roll.

EX1085, [0060]-[0061]; Pet. Reply, 11.

Choe discloses Step 1G: “transmitting data representing the at least one camera parameter to the camera device.”

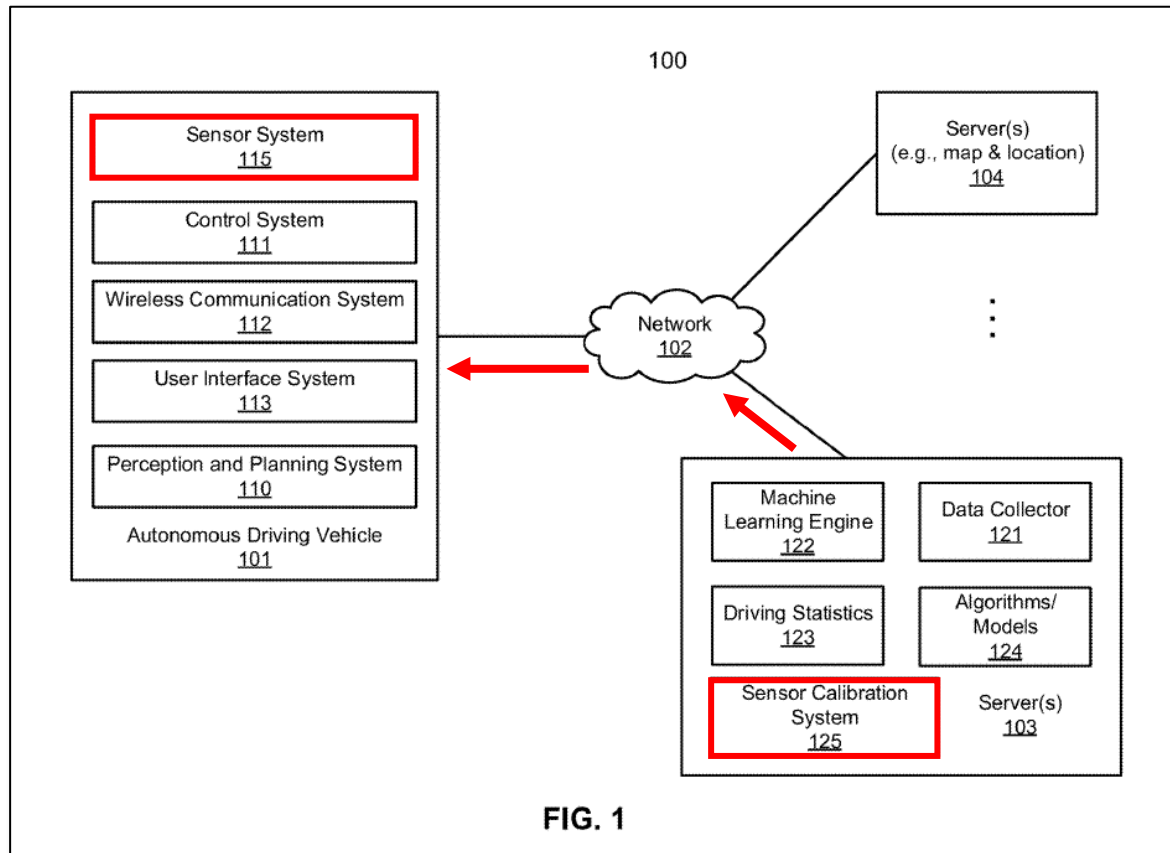


FIG. 1

EX1005, FIG. 1; Pet., 26, 33.

[0031] In one embodiment, sensor calibration system 125 may be hosted by server 103 to calibrate sensors such as cameras, LIDAR, and/or RADAR devices offline based on the images and/or point clouds captured by the sensors. The sensor calibration process may be performed offline based on the previously captured sensor data captured by the targeted sensors of ADVs. The calibration parameters can then be uploaded to the corresponding vehicles and utilized by autonomous driving software to calibrate and compensate the corresponding sensors. The calibration process will be described in details further below.

EX1005, [0031]; Pet., 33.

Not Disputed

Choe discloses Claim 2: “wherein the one or more lines includes at least one lane line of a roadway”



EX1005/EX1006, FIG. 6B; Pet., 27, 33-34.

[0015] According to one embodiment, in response to a first image captured by a camera of an ADV, a horizon line is determined based on the camera's hardware settings. The horizon line represents a vanishing point from a view point of the ADV based on an initial or default pitch angle of the camera. One or more lane lines are determined based on the first image via a perception process performed on the first image.

In addition, optional image 603, referred to herein as a 2D re-projected view, is generated based on 2D view image 601 and 3D view image 602, including projecting lane lines 621-623 onto the 2D view.

EX1005, [0015], [0056]; Pet., 33.

Ground 1 – Claims 3 and 4

Claim 3	Claim 4
“wherein computing the at least one camera parameter comprises predicting the at least one camera parameter using the predictive model”	“wherein computing the at least one camera parameter comprises predicting the at least one camera parameter using a second predictive model”

Mr. Andrews

46. The predictive models 124 of Choe are generated by a machine learning engine 122 “based on driving statistics 123.” *Id.* at [0044]. Thus, these “predictive models 124” are based on “driving statistics 123” and are used for sensor calibration “to determine a set of calibration parameters.” Nothing in Choe teaches or suggests “identifying one or more lines in the video using a predictive model, the one or more lines including a horizon line.”

EX2003, [0046] (POPR Dec); Pet. Reply, 7.

First Predictive Model: identifies the lines (see Step 1B)
Lines are then used to determine camera parameters (see Step 1C)

Pet., 34-35; Pet. Reply, 14-15.

Choe

[0030] Based on driving statistics 123, machine learning engine 122 generates or trains a set of rules, algorithms, and/or predictive models 124 for a variety of purposes. In one embodiment, algorithms 124 may include algorithms for autonomous driving a vehicle as described further below. Algorithms 124 may further include sensor calibration algorithms, which will be described further below. Algorithms 124 can then be uploaded on ADVs to be utilized during autonomous driving in real-time.

EX1005, [0030]; Pet., 27.

First Predictive Model: identifies the lines (see Step 1B)
Second Predictive Model: determines camera parameters

Pet., 35; Pet. Reply, 14-15.

Ground 2: Choe and Khan/Ali (Claims 1-5)

Ground 2: Khan and Ali are prior art to the '580 Patent.

'580 Patent

US 11,875,580 B2

(12) United States Patent Hassan et al. (10) Patent No.: US 11,875,580 B2 (45) Date of Patent: Jan. 16, 2024

(54) CAMERA INITIALIZATION FOR LANE DETECTION AND DISTANCE ESTIMATION USING SINGLE-VIEW GEOMETRY

(71) Applicant: MOTIVE TECHNOLOGIES, INC.

(22) Filed: Oct. 4, 2021

(65) Prior Publication Data US 2023/0106961 A1 Apr. 6, 2023

(51) Int. Cl. G06T 7/80 (2017.01) G06T 7/50 (2017.01) G06T 7/75 (2017.01) G06T 7/59 (2017.01) G06T 18/214 (2023.01) G06N 3/044 (2023.01)

(52) U.S. Cl. CPC: G06V 20/588 (2022.01); G06F 18/214 (2023.01); G06N 3/044 (2023.01); G06T 7/59 (2017.01); G06T 7/75 (2017.01); G06T 2207/30256 (2013.01)

(58) Field of Classification Search CPC: G06V 20/588; G06V 18/214; G06N 3/044; G06N 3/0442; G06N 3/045; 20 Claims, 10 Drawing Sheets

(21) Appl. No.: 17/493,011

(22) Filed: Oct. 4, 2021

(65) Prior Publication Data US 2023/0106961 A1 Apr. 6, 2023

(51) Int. Cl. G06T 7/80 (2017.01) G06T 7/50 (2017.01) G06T 7/75 (2017.01) G06T 7/59 (2017.01) G06T 18/214 (2023.01) G06N 3/044 (2023.01)

(52) U.S. Cl. CPC: G06V 20/588 (2022.01); G06F 18/214 (2023.01); G06N 3/044 (2023.01); G06T 7/59 (2017.01); G06T 7/75 (2017.01); G06T 2207/30256 (2013.01)

(58) Field of Classification Search CPC: G06V 20/588; G06V 18/214; G06N 3/044; G06N 3/0442; G06N 3/045; 20 Claims, 10 Drawing Sheets

302

Backbone Network 304

Lane-Marking Horizon Head 306

Camera Initialization Head 310

Lane and Horizon Key Points 308

Camera Parameters 312

Samsara EX1001 U.S. Patent No. 11,875,580

Dr. Hall-Ellis

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

SAMSARA EX1016

MOTIVE TECHNOLOGIES, INC.

Case: Patent

DECLARATION OF SAMSARA EX1016 UNDER 37 C.F.R. § 1.601 FOR INTERVIEW

Mail Stop "PATENT BOARD"
Patent Trial and Appeal Board
U.S. Patent & Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450

Exhibit	Publication	Publicly Available
1007	Khan, Hussam Ullah, Afsheen Rafaqat Ali, Ali Hassan, Ahmed Ali, Wajahat Kazmi, and Aamer Zaheer. "Lane Detection Using Lane Boundary Marker Network with Road Geometry Constraints." <i>2020 IEEE Winter Conference on Applications of Computer Vision</i> (pp. 1834-1843). WACV. Piscataway, NJ: IEEE, 2020.	May 14, 2020
1008	Ali, Ahmed, Ali Hassan, Afsheen Rafaqat Ali, Hussam Ullah Khan, Wajahat Kazmi, and Aamer Zaheer. "Real-Time Vehicle Distance Estimation Using Single View Geometry." <i>2020 IEEE Winter Conference on Applications of Computer Vision</i> (pp. 1111-1120). WACV. Piscataway, NJ: IEEE, 2020.	May 14, 2020

Samsara EX1016
U.S. Patent No. 11,875,580

EX1016, [0079]; Pet., 9.

Not Disputed

Ground 2: A POSA would have been motivated to use Khan/Ali's "predictive model" in Choe.

There is no dispute that Khan uses Ali's horizon estimation algorithm.

Khan

Lane detection using lane boundary marker constraints

Hussam Ullah Khan, Afsheen Rafiqat Ali, Ali Hassan, Ahmed Ali, Waqar Khan
KeepTruckin Inc. (KTI)
Lahore, Pakistan
(hussam.khan, afsheen.ali, ali.hassan, ahmed.ali, waqar.khan)

Abstract

Lane detection is of critical importance to both the self-driving cars as well as advanced driver assistance systems. While current methods use a range of features from low-level to deep features extracted from convolutional neural networks, they all suffer from the problem of occlusion and struggle to detect lanes with low or no evidence on the road. In this paper, we use a lane boundary marker network to detect keypoints along the lane boundaries. An inverse perspective mapping is estimated using road geometry which is then applied to the detected markers and line/curves are fitted jointly on the rescaled points. Finally, missing lane boundaries are predicted using lane geometry constraints i.e., equidistance and parallelism. Reciprocal weighted averaging ensures lane boundaries with strong evidence dominate their predicted alternatives. The results show a significant improvement of +7.8%, +6.8% and +1.2% of F1 scores over the state-of-the-art on CULane, Caltech and TuSimple datasets, respectively. This proves our algorithm's robustness against both occluded and missing lanes cases. Furthermore, we also show that our algorithm can be combined with other lane detectors to improve their lane retrieval potential.

1. Introduction

Fully autonomous vehicles in the form of self-driving cars are gradually appearing on the scene and we have Google Waymo [17], Tesla's Auto-pilot [19], Cruise AI's OpenPilot kit [16] and Intel's Mobileye solutions [18] just to name a few. Along with self-driving cars, Advanced Driver Assistance Systems (ADAS) have also captured a surge in the interest. Both self-driving cars and the ADAS require different levels of autonomy but some aspects of automation are still common. One of them is road line detection and segmentation. Lane detection is critical in identifying and ensuring safe driving practices and an on-board



Figure 1: A sequence of four images showing a road with lane boundaries. The first image shows a clear road with lane markings. The second image shows the road with some occlusion. The third image shows the road with more occlusion. The fourth image shows the road with significant occlusion, where the lane boundaries are mostly hidden.

3.2.1 Horizon Estimation

The estimation of IPM assumes a planar road scene having no elevation and more than one lanes (3 or more lane boundaries), such as a typical highway scenario. We start by fitting lines to the detected markers. A best fit line must have a minimum number of inlier markers (i.e., 25% of total number of markers for each lane boundary). Inlier markers are those having horizontal distance of under 8 pixels to the best 1st degree line fit through RANSAC [8].

Critical part of initialization is the estimation of a vanishing line (horizon) from these lines. This requires two sets of vanishing points (VPs), forward and lateral VPs. Forward VPs can be found from the cross product of the world parallel lines whereas the process of finding lateral vanishing point is a bit involved. For that we use cross-ratios of three real-world parallel lines which are the lane lines as used by Ali *et al.* [2]. Once a set of forward and lateral VPs is found, we use their horizon estimation algorithm to find the best fitting horizon l_h .

Mr. Andrews

58. Thus, Khan does not disclose a specific approach to determine the horizon line but rather refers to the approach disclosed in Ali (Ex. 1008) to estimate the horizon.

EX2003 (revised), [0058]; see also POR, 25.

Q And I'll -- you reference that Khan and Ali work together; is that fair?
A Yeah, that's true.

EX1061, 27:1-3; see also POR, 25.

EX1007, 1826-27; Pet., 37-38.

Ground 2: A POSA would have been motivated to use Khan/Ali's "predictive model" in Choe.

For Ground 2, both experts agree there is a shortcoming in Choe.

Dr. Darrell

121. As I explain for Ground 1, Choe discloses a “machine learning engine
122 [that] generates or trains a set of rules, algorithms, and/or predictive models,”
including for its sensor calibration algorithms. Choe, [0030]. However, I
understand that the Patent Owner may allege that Choe does not explicitly disclose
which predictive models it uses or whether it uses predictive models to identify the
horizon line. A POSA would have therefore searched for references that use
predictive models and how such models were used. This would have led to Khan
and Ali, which disclose predictive models for horizon and lane detection.

EX1003, [0121]; Pet., 36.

Mr. Andrews

90. Choe then explains how the “set of rules” are used. For example,
“Decision module 304 may make such decisions according to a set of rules such as
traffic rules or driving rules 312, which may be stored in persistent storage device
352.” EX 1004 at [0038]. It also explains how “algorithms” are used for “
autonomous driving a vehicle,” for “sensor calibration,” and how they can be
uploaded and “utilized during autonomous driving in real-time.” *Id.* at [0030].
Choe does not actually describe how “predictive models” are used for any purpose,
how they may be trained, or provide any specifics about them. More importantly,
nothing in Choe teaches or suggests the use of a predictive model to identify a
horizon line.

EX2003 (revised), [0090]; Pet. Reply, 15.

Ground 2: A POSA would have been motivated to use Khan/Alī's "predictive model" in Choe.

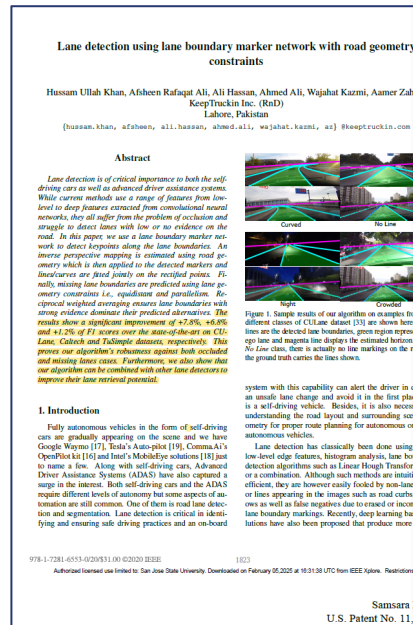
Dr. Darrell explained some advantages of using Khan/Alī's predictive models.

Dr. Darrell

123. A POSA would have been motivated to use Khan's predictive models in Choe's camera calibration system because Khan states that the models provide "a robust and efficient lane detection algorithm." Khan, p. 1830. Khan's predictive models also outperformed other predictive models in the industry for this use case. Specifically, "[t]he results show a significant improvement of +7.8%, +6.8% and +1.2% of F1 scores over the state-of-the-art on CULane, Caltech and TuSimple datasets, respectively." Khan, p. 1823; *see also* Khan, pp. 1828-30 (detailed results). Khan's predictive models also improved occluded and missing lanes cases. Khan, p. 1823; *see also* Khan, p. 1824 ("Our approach significantly improved the results on public datasets, specifically for crowded or missing lane scenarios."), 1826 ("This pipeline greatly improves the lane detection accuracy").

EX1003, [0123]; Pet., 36.

Khan



Abstract

Lane detection is of critical importance to both the self-driving cars as well as advanced driver assistance systems. While current methods use a range of features from low-level to deep features extracted from convolutional neural networks, they all suffer from the problem of occlusion and struggle to detect lanes with low or no evidence on the road. In this paper, we use a lane boundary marker network to detect keypoints along the lane boundaries. An inverse perspective mapping is estimated using road geometry which is then applied to the detected markers and lines/curves are fitted jointly on the rectified points. Finally, missing lane boundaries are predicted using lane geometry constraints i.e., equidistant and parallelism. Reciprocal weighted averaging ensures lane boundaries with strong evidence dominate their predicted alternatives. The results show a significant improvement of +7.8%, +6.8% and +1.2% of F1 scores over the state-of-the-art on CULane, Caltech and TuSimple datasets, respectively. This proves our algorithm's robustness against both occluded and missing lanes cases. Furthermore, we also show that our algorithm can be combined with other lane detectors to improve their lane retrieval potential.

EX1007, 1823; Pet., 36-37.

Ground 2: A POSA would have been motivated to use Khan/Ali's “predictive model” in Choe.

Dr. Darrell explained some advantages of using Khan/Ali's predictive models.

Dr. Darrell

127. A POSA would have been motivated to use Ali's horizon estimation with Khan's predictive models also because it could efficiently predict the horizon line. Ali, pp. 1106-07 (detailed results). It was “computationally inexpensive and is data-domain invariant.” Ali, p. 1100. It was “well suited” for autonomous highway driving. Ali, p. 1101. It was also, according to Ali, “simple, intuitive, independent of data and performs almost equally on all the tested datasets” Ali, p. 1107. A POSA would have therefore had a reasonable expectation of success when using Ali's horizon prediction in the Choe-Khan system.

Ali

Real-time vehicle distance estimation using single view geometry

Ahmed Ali, Ali Hassan, Ahsan Ali, and Ali Hassan

(Ahmed.Ali, Ali.Hassan, Ahsan.Ali, and Ali.Hassan)

Abstract

Distance estimation is required for autonomous driving systems (ADS) as well as for obstacle avoidance, and is crucial for accident prevention. Currently, radar is used for this purpose which offers poor resolution. Deep learning based distance estimation techniques require large amount of data and ensuring domain invariance is a challenge. In this paper, we propose a single view distance estimation method which is lightweight and uses geometric line markings for distance estimation with the lane and vehicle direction in ADS. Our system introduces novel use cross-ratio of lane boundaries to determine an inverse Perspective camera height from a known lane width. Distances of the vehicles are calculated by back projecting image points to the reconstructed road plane. For radar data as ground truth and compare our algorithm with radar as well as the learning based monocular depth prediction results on three public datasets (Kitti level 5) showed that the proposed system RMSE between 6.10 to 7.31. algorithms on two of the datasets while one supervised deep learning method. Furthermore, it is computationally inexpensive and is data-domain invariant.

1. Introduction

Human error is one of the major causes of traffic accidents. At the heart of safe driving is the judgment to estimate the vehicle's position relative to other road users. For this reason, Autonomous Systems (ADS) have evolved in the last few years and features such as

Abstract

. . . .

the reconstructed road plane. For evaluation, we used lidar data as ground truth and compare the performance of our algorithm with radar as well as the state-of-the-art deep learning based monocular depth prediction algorithms. The results on three public datasets (Kitti, nuScenes and Lyft level 5) showed that the proposed system maintains a consistent RMSE between 6.10 to 7.31. *It outperforms other algorithms on two of the datasets while on KITTI it falls behind one (supervised) deep learning method. Furthermore, it is computationally inexpensive and is data-domain invariant.*

EX1003, [0127]; Pet., 39.

EX1008, 1100; Pet., 39.

Ground 2: A POSA would have been motivated to use Khan/Ali's "predictive model" in Choe.

Dr. Darrell explained some advantages of using Khan/Ali's predictive models.

69. I was also not persuaded by Mr. Andrews' claim that Choe would not care about accuracy of the initial horizon line when calibrating its camera system.

During deposition, Mr. Andrews alleged that "Choe doesn't care how accurate that candidate horizon line is." EX1061, 79:5-6. However, I find this statement to be inconsistent with basic design principles for autonomous vehicles.

. . .

73. A POSA would have understood that having an accurate initial horizon line would not only result in more accurate calibration, it would also result in faster calibration. If the predicted horizon line is completely wrong, the reviewer of Choe's system would have to spend significant time manually correcting the horizon line. In real-time use, this would be cumbersome. Conversely, by accurately predicting the initial horizon line, the reviewer could quickly assess and correct minor inaccuracies. This would better enable real-time calibration.

Ground 2: Khan/Ali discloses the “predictive model” recited in Step 1B.

Khan/Ali discloses Step 1B in at least three ways:

1 Ali’s horizon estimation algorithm (I_h) is a “predictive model” that identifies the horizon line.

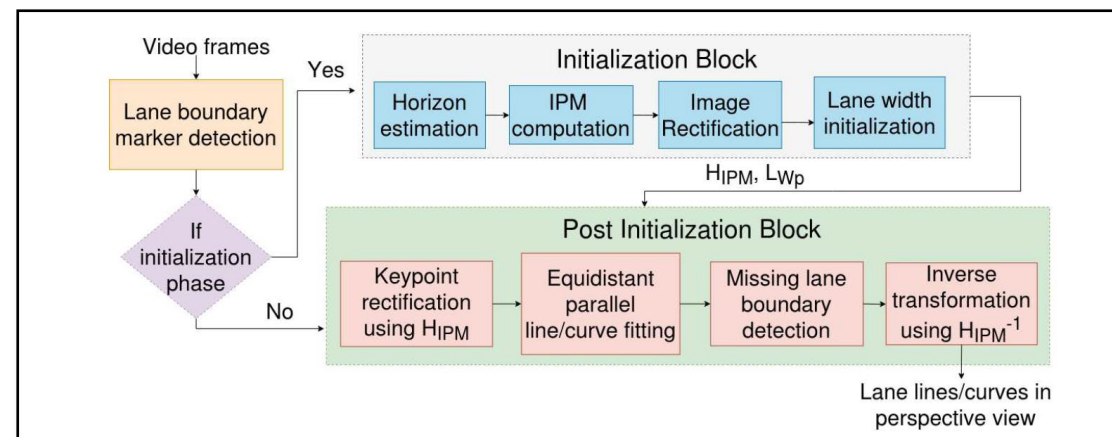
Pet., 41; EX1003, [0135]; Pet. Reply, 19-20; EX1085, [0094]-[0096].

2 Khan’s CNN for lane detection (a predictive model) is used in Ali’s horizon estimation algorithm (I_h) to identify the horizon line.

Pet., 41; EX1003, [0133]-[0134]; Pet. Reply, 17-19; EX1085, [0079]-[0093].

3 Khan/Ali’s end-2-end algorithm is a “predictive model” that identifies the horizon line.

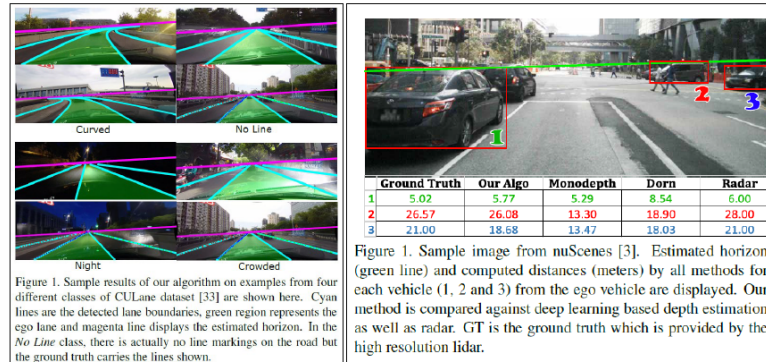
Pet., 43; EX1003, [0138]; Pet. Reply, 20; EX1085, [0097]-[0105].



Ground 2: Khan/Alì's horizon estimation algorithm is a "predictive model."

Dr. Darrell

95. Khan describes using Ali's horizon estimation algorithm to find the horizon line. See EX1007, 1827 ("Once a set of forward and lateral VPs is found, we use [Ali's] horizon estimation algorithm to find the best fitting horizon l_h "). Ali describes that the horizon estimation model finds the best fitting horizon across the video using RANSAC. EX1008, 1104, Algorithm 1. RANSAC is a well-known machine learning model that is used as a tool for predictive modeling. The best fit horizon line is then overlaid onto the video as shown in Khan and Ali's figures:



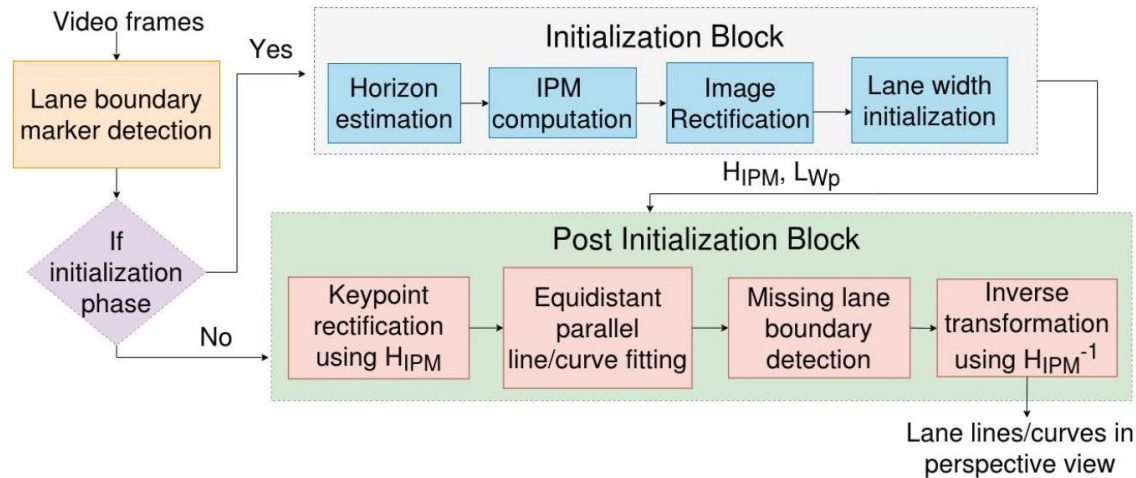
Left: EX1007, FIG. 1; Right: EX1008, FIG. 1

96. The '580 patent similarly determines markers (also called key points) for the lane and horizon lines and then "fits one or more lines to the markers." EX1001, 13:20-21 ("Next, in step 614, the method fits one or more lines to the markers."). The '580 patent does not exclude algorithms from the scope of "predictive model." Thus, a POSA would have understood that an algorithm that predicts the horizon line is a model that identifies a horizon line. Because Khan/Alì describes a prediction model that similarly estimates a best fit horizon line based on lane markers, a POSA would have understood that Khan/Alì's horizon line estimation is a predictive model.

Ground 2: Khan/Alì's CNN ("a predictive model") is used to identify lane lines.

Motive's expert admitted Khan's CNN is a "predictive model."

Khan



We use a simple encoder-decoder based Convolutional Neural Network (CNN) to detect keypoints or markers (used interchangeably) along the lane boundaries. Inverse Perspective Mapping (IPM) estimated using scene geometry is then applied to the detected markers and lines/curves are fitted on the rectified points.

Mr. Andrews

Q Right. So a person of ordinary skill in the art would have understood that Khan's CNN would be a predictive model as that term is used in the '580 patent, right?

A It's an example of a predictive model. It's not the only example.

Q I -- understood.

It's not the only example of predictive model, but a person of ordinary skill in the art would have understood that Khan's CNN is an example of a predictive model as that term is used in the '580 patent, right?

MR. RIBERA: Objection; vague.

THE WITNESS: I think that they would have understood that a CNN, if they had read the '580 patent and understood how the '580 patent described its predictive model, they would understand that Khan's CNN is an example of a predictive model.

Ground 2: Khan/Ali's CNN ("a predictive model") is used to identify a horizon line.

Motive's expert admitted Ali's horizon estimation algorithm "uses" Khan's CNN.

Ali

Algorithm 1: Best fitting Horizon (l_h) estimation

Input: Detected markers by lane detector
Output: Estimated Horizon line
1 Find forward and lateral VP for each frame in a video:
2 initialize;
3 for $i \leq \text{num_frames in video}$ do
4 Fit lines to keypoints ($B = \text{num_lines}$);
5 if $B \geq 2$ then

Algorithm 1: Best fitting Horizon (l_h) estimation

Input: Detected markers by lane detector

Output: Estimated Horizon line

16 RANSAC for best fitting horizon across video:
17 initialize;
18 $s = \text{random_int between 1 to num_forward_vps}$;
19 $t = \text{random_int between 1 to num_lateral_vps}$;
20 while $\text{iterations} \leq t$ OR $\text{num_inliers} \leq \text{max_inliers}$ do
21 Fit line l_h to $\text{forward_vp}(s)$ and $\text{lateral_vp}(t)$;
22 if $\text{dist}(l_h, \text{forward_vp}(s)) \leq \text{threshold}_s$ &
23 $\text{dist}(l_h, \text{lateral_vp}(t)) \leq \text{threshold}_t$ then
24 record and count inlier points
25 end

EX1008, 1104; Pet., 38, 41; Pet. Reply, 18-19.

Mr. Andrews

To be clear, Algorithm 1 of Ali would use the CNN of Khan or the CNN of Khan plus the missing boundaries approach of Khan, right?

MR. RIBERA: Objection; vague, misstates the witness.

THE WITNESS: Could you read that back?

(Record read.)

THE WITNESS: Yes.

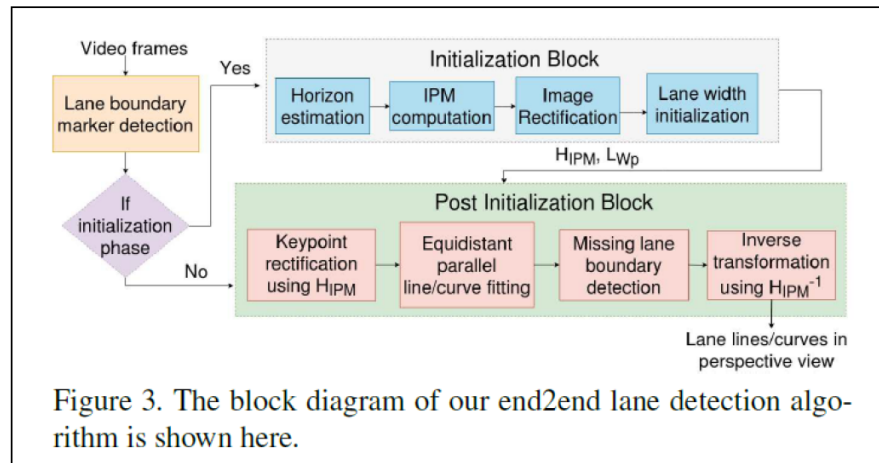
EX1061, 44:16-22; Pet. Reply, 19.

Step 1B: "identifying one or more lines in the video using a predictive model, the one or more lines including a horizon line"

Ground 2: Khan/Ali's end-2-end algorithm is a “predictive model.”

Dr. Darrell

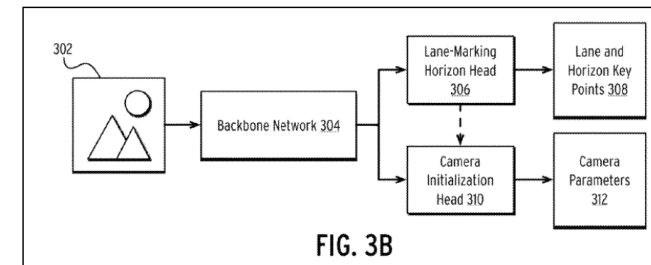
98. Taken as a whole, the end2end lane detection algorithm described by Khan and Ali would be considered a “predictive model.” As I described above, Khan and Ali describe an algorithm that uses a CNN to detect lane markers, which then is used to estimate a horizon line. Khan and Ali’s horizon line estimation could not happen without the use of the CNN.



EX1007, FIG. 3.

99. Mr. Andrews argues that the end2end algorithm is not a predictive model because it includes pre- and post-initialization steps. See EX2003, ¶95. However, a predictive model can have multiple steps.

100. In fact, the '580 patent itself describes its predictive model as having multiple steps to determine the lane lines, determine the horizon line, and initialize the camera. For example, “FIG. 3B is a block diagram illustrating a neural network for predicting lane and horizon lines and camera parameters according to some embodiments.” EX1001, 2:21-23. In Figure 3B, the predictive model is shown to have separate steps such as the Lane-Marking Head 306 and Camera initialization Head 310:



EX1001, FIG. 3B

Khan

Ali

EX1008, 1104; Pet., 38; Pet. Reply, 21.

Ground 2 – Claim 2: “wherein the one or more lines includes at least one lane line of a roadway.”

Khan

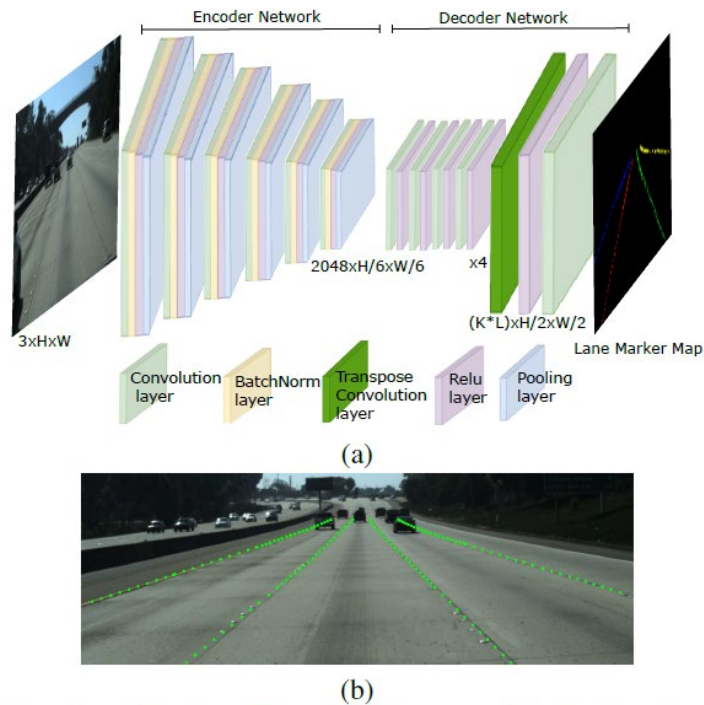


Figure 2. (a) The figure shows the architecture of the lane boundary marker network. (b) The sampled keypoints from the ground truth lane line are shown here.

EX1007, FIG. 2; Pet., 41-42; Pet. Reply, 21.

Dr. Darrell

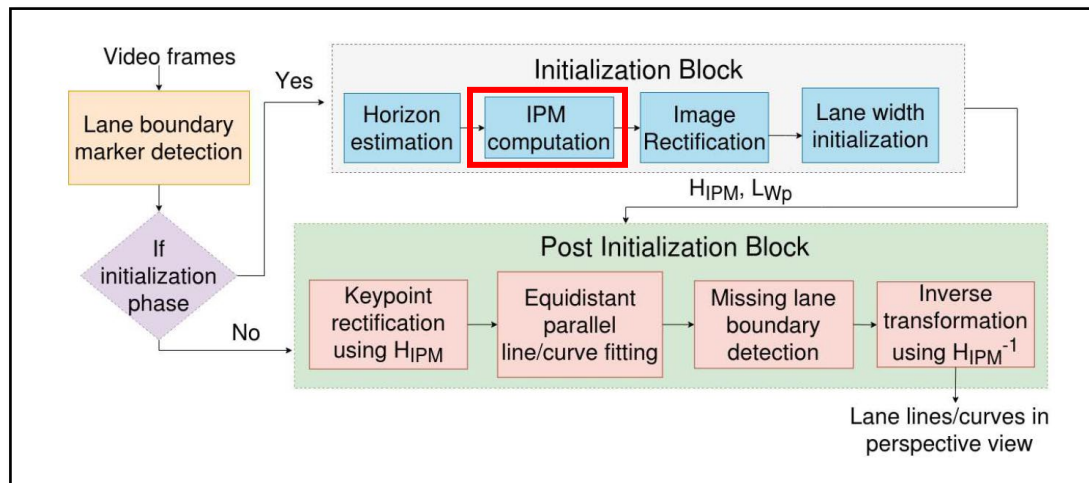
83. This model uses an architecture with multiple convolution layers, batch normalization layers, a transpose convolution layer, ReLu (Rectified Linear Unit) layers, and pooling layers. See EX1007, 1826; EX1008, 1102. This model is a deep learning AI model trained to predict the lane lines based on identifying keypoints of the lane boundaries. See EX1008, 1102 (Khan uses “a fully convolutional deep network”).

EX1085, [0083]; Pet. Reply, 17.

Ground 2 – Claims 3 and 4

Claim 3	Claim 4 – Not Disputed (see POR, 35-39)
“wherein computing the at least one camera parameter comprises predicting the at least one camera parameter using the predictive model”	“wherein computing the at least one camera parameter comprises predicting the at least one camera parameter using a second predictive model”

Khan



EX1007, FIG. 3; Pet., 42-43; Pet. Reply, 23-24.

Mr. Andrews

Q The '580 patent, though, refers to IPM as a camera parameter, right?

A It does, yes.

Q And Khan calculates the IPM based on the horizon line and other variables, right?

A Yes.

Q So the '580 patent uses the same equations as Khan and Ali to determine the IPM, correct?

MR. RIBERA: Objection; vague.

THE WITNESS: It uses the same equations that are described in Khan. 3.2.2 of Khan.

EX1061, 58:8-13, 61:9-13; Pet. Reply, 21.

Ground 2 – Claim 5: “wherein the at least one camera parameter comprises a camera height and a road plane normal.”

Mr. Andrews

Q Okay. So the '580 patent uses the same equation to determine the road plane normal as Khan and Ali, right?

A Yes.

MR. RIBERA: Objection; vague.

THE WITNESS: Yes, it does.

BY MR. DUTTON:

Q And if we go to Equation 2, it's also correct that the '580 patent uses the same equation to determine H as Khan and Ali, right?

MR. RIBERA: Objection; vague.

THE WITNESS: Yes.

EX1061, 60:10-20; Pet. Reply, 25.

Khan

3.2.2 IPM and Image Rectification

After estimating horizon, we compute plane normal $\hat{n}$ using camera intrinsic matrix K and horizon l_h [11] by $\hat{n} = K^T l_h$. The IPM is then computed using $H_{IPM} = K R K^{-1}$, where $R = [l_h \times \hat{n}; (l_h \times \hat{n}) \times -\hat{n}; -\hat{n}]$ is the rotation matrix and H_{IPM} is the rectification homography which will rotate the camera view to align its Z-axis with the road's normal.

EX1007, 1827; Pet., 44-45; Pet. Reply, 25.

'580 Patent

In step 704, the method computes the IPM for a camera and rectifies the sample images. In an embodiment, step 704 comprises computing a road plane normal vector ($\hat{n}$) using an intrinsic camera matrix (K) and the detected horizon (l_h) according to:

$$\hat{n} = K^T l_h \quad (\text{Equation 1})$$

where K^T comprises the transpose of the intrinsic camera matrix. In an embodiment, the intrinsic camera matrix comprises the geometric properties of the camera. After computing the plane normal ($\hat{n}$), the method computes the IPM (H) using:

$$H = K R K^{-1} \quad (\text{Equation 2})$$

where H is the rectification homography which will rotate the camera view to align its Z-axis with the road plane normal, and R is the rotation matrix defined by:

$$R = [l_h \times \hat{n}; (l_h \times \hat{n}) \times -\hat{n}; -\hat{n}] \quad (\text{Equation 3})$$

EX1001, 14:15-33; Pet. Reply, 25.

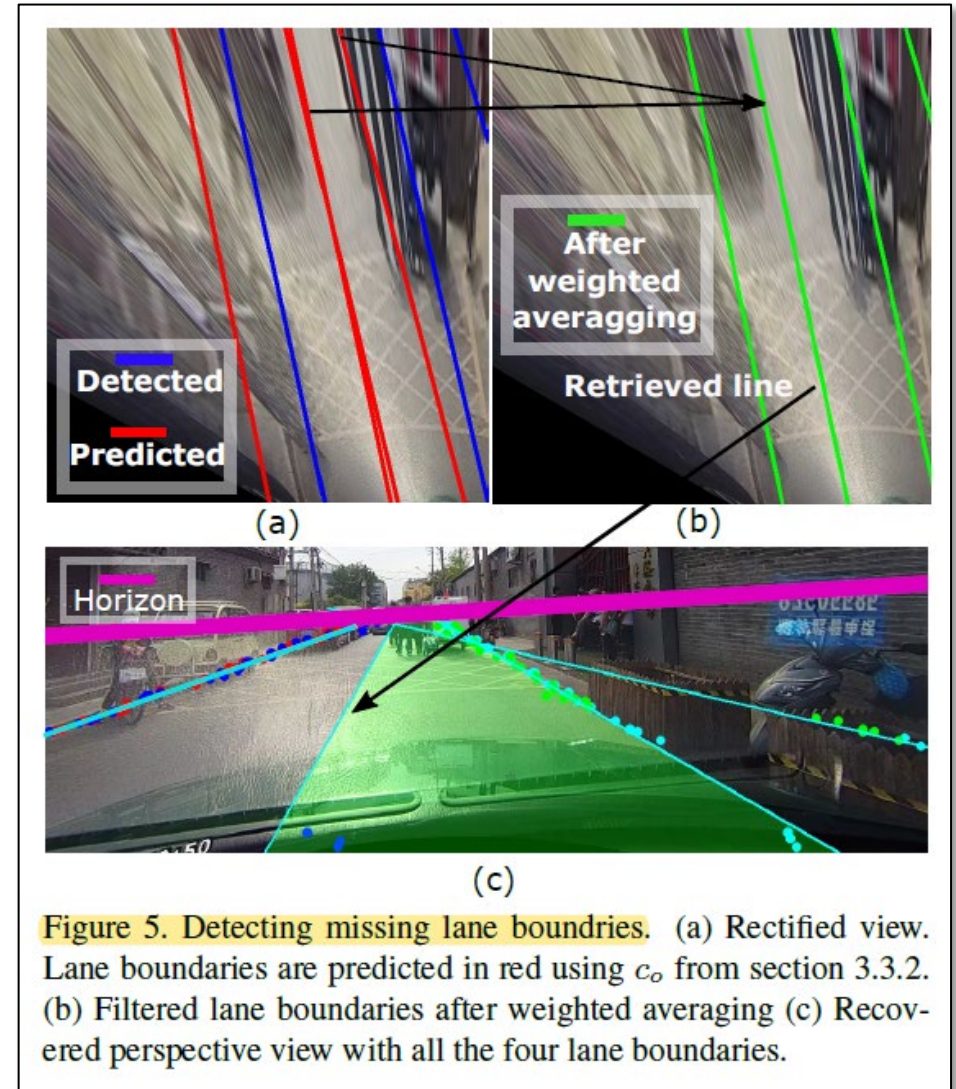
Ground 2: Khan/Ali does not require three lane lines to identify the horizon line.

3.3.2 Detecting Weak or Missing Lane Boundaries

In order to estimate the missing lanes, we used the initialized lane width in pixels (L_{Wp}). We calculate the lane boundary offset for the missing or low confidence lanes from Equation 1 as:

$$c_o = L_{Wp} \sqrt{a^2 + b^2} \quad (2)$$

EX1007, 1827; Pet. Reply, 21.



EX1007, 1828; Pet., 42.

Ground 2: Khan/Ali does not require three lane lines to identify the horizon line.

Mr. Andrews agrees Khan/Ali determines the horizon line even if only two lane lines are visible.

Q So Khan and Ali would determine the horizon line if only two lane lines were visible on the input image to the CNN predictive model, right?

MR. RIBERA: Objection; vague.

THE WITNESS: They would determine a horizon line if only two lines were visible on the image and they synthetically derived a third line.

BY MR. DUTTON:

Q And Khan describes how it synthetically derives a third line when that's the case, right?

A Yes.

EX1061, 47:10-20.

Q And Khan discloses how you would use that approach to infer a third line in Section 3.3.2, right?

A Yes.

EX1061, 41:2-4.

Ground 3: Choe and Workman (Claims 1-4)

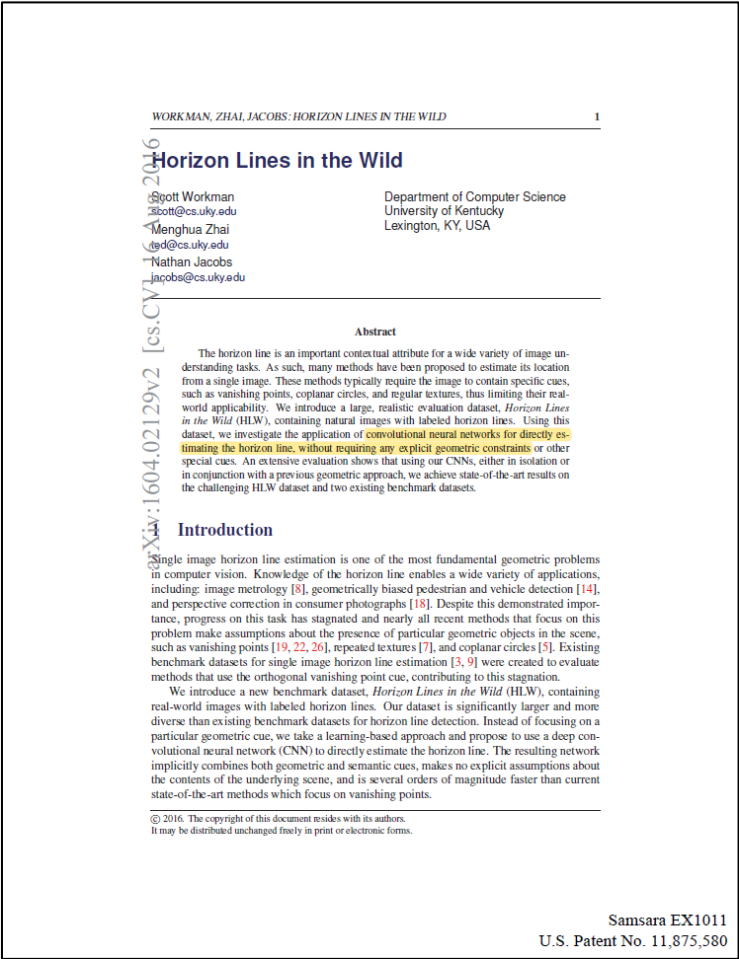
'580 Patent

Dr. Hall-Ellis

Not Disputed

Ground 3: Workman identifies a horizon line using a “predictive model.”

Workman



Workman discloses a “predictive model” that identifies a horizon line.

Abstract

The horizon line is an important contextual attribute for a wide variety of image understanding tasks. As such, many methods have been proposed to estimate its location from a single image. These methods typically require the image to contain specific cues, such as vanishing points, coplanar circles, and regular textures, thus limiting their real-world applicability. We introduce a large, realistic evaluation dataset, *Horizon Lines in the Wild* (HLW), containing natural images with labeled horizon lines. Using this dataset, we investigate the application of **convolutional neural networks for directly estimating the horizon line, without requiring any explicit geometric constraints** or other special cues. An extensive evaluation shows that using our CNNs, either in isolation or in conjunction with a previous geometric approach, we achieve state-of-the-art results on the challenging HLW dataset and two existing benchmark datasets.

EX1011, 1; Pet., 19.

Not Disputed

Ground 3: A POSA would have been motivated to use Workman’s “predictive model” in Choe.

For Ground 3, both experts agree there is a shortcoming in Choe.

Dr. Darrell

146. As I explain in Ground 1, Choe discloses a “machine learning engine 122 [that] generates or trains a set of rules, algorithms, and/or predictive models,” including for its sensor calibration algorithms. Choe, [0030]. However, I understand that the Patent Owner may allege that Choe does not explicitly disclose which predictive models it uses or whether it uses predictive models to identify the horizon line. A POSA would have therefore searched for references that use predictive models and how such models were used. This would have led to Workman, which discloses predictive models to identify the horizon line.

EX1003, [0146]; Pet., 47.

Mr. Andrews

90. Choe then explains how the “set of rules” are used. For example, “Decision module 304 may make such decisions according to a set of rules such as traffic rules or driving rules 312, which may be stored in persistent storage device 352.” EX 1004 at [0038]. It also explains how “algorithms” are used for “autonomous driving a vehicle,” for “sensor calibration,” and how they can be uploaded and “utilized during autonomous driving in real-time.” *Id.* at [0030]. Choe does not actually describe how “predictive models” are used for any purpose, how they may be trained, or provide any specifics about them. More importantly, nothing in Choe teaches or suggests the use of a predictive model to identify a horizon line.

EX2003 (revised), [0090]; Pet. Reply, 15.

Ground 3: A POSA would have been motivated to use Workman's "predictive model" in Choe.

Dr. Darrell explained advantages of using Workman's predictive model.

Dr. Darrell

147. A POSA would have been motivated to use Workman's predictive models to identify the horizon line because it achieved state-of-the-art results:

. . .

148. A POSA would have also been motivated to use Workman's predictive model because it identifies the horizon line faster than other models:

. . .

149. A POSA would have understood that Workman's state-of-the-art predictive model, in addition to outperforming all existing methods, was intended to be used in other methods.

EX1003, [0147]-[0149]; Pet., 47-49.

Workman

WORKMAN, ZHAI, JACOBS: HORIZON LINES IN THE WILD 1
Horizon Lines in the Wild
Siyun Workman
zhai@cs.uky.edu
Mengzha Zhai
zhai@cs.uky.edu
Zhaojun Jacobs

Department of Computer Science
University of Kentucky
Lexington, KY, USA

We propose to use a CNN to directly estimate the horizon line location. However, we show that by using our CNN as context for their method, replacing the one they proposed, significantly improves performance for vanishing point based horizon line estimation. Extensive experiments demonstrate that our CNN-based approach is fast, requiring only milliseconds per image, and accurate, achieving state-of-the-art performance on two popular datasets designed to showcase purely geometric methods, and the challenging HLW dataset.

significantly outperform their geometric methods, and several orders of magnitude faster than current state-of-the-art methods which focus on vanishing points.
© 2016. The copyright of this document rests with its authors.
It may be distributed unchanged freely in print or electronic form.

EX1011, 0002; Pet., 47-48.

Despite this generality, the performance of our methods is competitive, achieving state-of-the-art results on two existing benchmark datasets designed for geometric methods, and outperforming all existing methods on the challenging real-world imagery contained in HLW. Our method is fast, works in natural environments, and can provide a prior over horizon line location that can be used as input to other methods.

EX1011, 0010; Pet., 47-48.

Ground 3: Motive's arguments disputing the motivation to combine lack merit.

1 Ground 3 identified a shortcoming in Choe – that Choe's predictive models are not described fully.

Pet., 47; EX1003, [0146]; EX2003 (revised); Pet. Reply, 15; EX1085, [0116]-[0117].

2 Dr. Darrell did not use the '580 Patent as a roadmap – his MTC analysis relies on Choe and Workman.

EX1003, [0146]-[0150]; Pet., 47-49; Pet. Reply, 26-27.

3 A POSA would care about speed, especially when Choe's system calibrates the camera while the vehicle is driving.

EX1085, [0124], [0126] (citing EX1005, [0017]); Pet. Reply, 27.

4 Workman's performance improvement is relative to vanishing point estimation – the type of system in Choe.

Pet. Reply, 27-28; Pet., 47-48; EX1003, [0147].

5 Workman is analogous art and relevant to identifying a horizon line in automobile applications.

EX1085, [0119]-[0122]; Pet. Reply, 28-29. 51

Ground 4: Choe and Westmacot (Claims 5-7)

Ground 4 – Choe in view of Westmacot renders obvious claims 5-7.

Patent Owner Argues:

1 A POSA would not have been motivated to compute “a camera height and a road plane normal” (claim 5).

POR, 53-56.

2 Westmacot teaches away from camera calibration.

POR, 56.

3 Westmacot does not compute the camera height and road plane normal based on the one or more lines.

POR, 57-58

If the Board determines that Choe satisfies claim 1, Ground 4 renders obvious claims 6 and 7.

Ground 4 – A POSA would have been motivated to combine Choe and Westmacot.

Dr. Darrell

160. Choe's camera calibration system predicts lane lines. Choe, [0015].

To the extent Choe does not explicitly disclose the mathematics for predicting these lane lines, a POSA would have sought other references. This would have led to Westmacot, which discloses equations for determining the lane lines.

. . .

163. A POSA would have also been motivated to enable the user to manually adjust the camera height during calibration to correct errors.

EX1003, [0160], [0163]; Pet., 53.

165. Computing and calibrating the camera height and the road plane normal would have combined known elements according to known methods to yield predictable results. Specifically, the camera height and road plane normal were known calculations or measurements that used known methods (e.g., multiplication and division). See Westmacot, 5:2-4 (equation for the road plane normal), 26:4 ("The camera height above the road fit is *easy* to measure"), FIGs. 3, 3A, 26; see also Ali, pp. 1104 (computing road plane normal), 1105 (equation for camera height). Performing known mathematical calculations in Choe's camera calibration system would have yielded predictable results because, given a set of rules and known variables, a POSA could consistently calculate the camera height and the road plane normal. Accordingly, a POSA would have had a reasonable expectation of success when computing the road plane normal and camera height in Choe's camera calibration system.

EX1003, [0165]; Pet., 55.

Ground 4 – A POSA would have been motivated to combine Choe and Westmacot.

POR

In fact, a POSITA would not be motivated to combine Westmacot with Choe's camera calibration system because Westmacot teaches away for doing any camera calibration: "Preferably the expected road structure is determined using only information derived from image data captured by the image capture device ... without the use of any other vehicle sensor and **without any requiring any calibration of the image capture device.**" Ex_1010, p.5:25-31. This is precisely

POR, 56.



"[A] reference does not teach away if it merely expresses a general preference for an alternative invention but does not criticize, discredit or otherwise discourage investigation into the claimed invention."

UCB, Inc. v. Actavis Labs, UT, Inc.,
65 F.4th 679, 692 (Fed. Cir. 2023).

Ground 4 –Westmacot computes camera parameters based on the lines.

Westmacot explains the mathematical relationship between camera height (h), road plane normal (n) and the lane boundaries (b_i^{left} and b_i^{right}).

$$g_i = c_i + hR_i n$$

$$b_i^{left} = g_i + w_i^{left} R_i r$$

$$b_i^{right} = g_i + w_i^{right} R_i r$$

EX1010, 24:26, 25:2-4; Pet., 54.

g_i = road point below camera

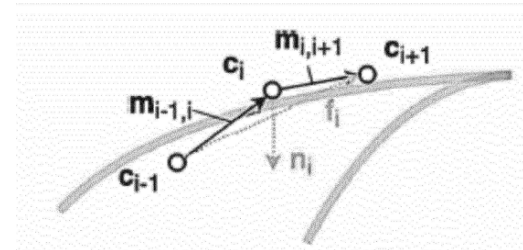
c_i = camera's 3D position

R_i = camera's pose

w_i = width left or right

r = vector within the road plane

Fig. 27 shows the principles of the estimation of the road normal n .



EX1010, 26:9, FIG. 27; Pet., 62.

The camera height above the road h is easy to measure manually. However, in case this cannot be done (e.g. for dash-cam videos downloaded from the web) it is also possible to obtain the height of the camera using the estimated mesh of the road surface obtained from OpenSfM. A rough estimate for h is sufficient, since it is corrected via manual annotation, see the following section.

...

In Figure 19, the annotator is provided with means to adjust the camera height, to match the reconstruction to the road surface, and the crop height, to exclude the vehicles dash or bonnet.

EX1010, 26:4-8, 34:9-11; Pet., 62.

Ground 5: Choe, Khan/Ali, and Westmacot (Claims 1-7)

Ground 5 – A POSA would have been motivated to combine Choe, Khan/Ali, and Westmacot.

POR:

E. Ground 5 – The Choe/Westmacot/Khan/Ali Combination Does Not Render the Challenged Claims Unpatentable.

. . .

Petitioner fails to provide any motivation of why a POSITA would combine Choe, Westmacot, Khan, and Ali all together. “Petitioner’s proposed combination of teachings plainly takes into account knowledge gleaned only from applicant’s disclosure, i.e., impermissible hindsight.” *See* IPR2023-00656, Paper No. 33, p.52 (citing *In re McLaughlin*, 443

POR, 58.

Pet. Reply:

VIII. GROUND 5

Motive alleges “Petitioner fails to provide any motivation of why a POSA would combine Choe, Westmacot, Khan, and Ali all together.” POR, 58. Not so.

The motivation to use Khan/Ali’s predictive model in Choe is articulated in Ground 2, Pet., 36-40. And the motivation to compute the camera height and road plane normal in Choe’s system, as described in Westmacot, is articulated in Ground 4, Pet., 53-60. These rationales are independent and relate to different parts of Choe’s system. EX1085, ¶¶141-143. If a POSA were to make both combinations, the result would disclose every element of the challenged claims. EX1085, ¶¶141-143.

Pet. Reply, 32-33.

Ground 6: Choe, Workman, and Westmacot (Claims 1-7)

Ground 6 – A POSA would have been motivated to combine Choe, Workman, and Westmacot.

POR:

F. Ground 6 – Petitioner Does Not Identify Any Additional Motivation to Combine Workman or Westmacot with Choe.

. . .

Petitioner

fails to provide any motivation of why a POSITA would combine Choe with both Westmacot and Workman. Again, “Petitioner’s proposed combination of teachings plainly takes into account knowledge gleaned only from applicant’s disclosure, i.e., impermissible hindsight.” *See* IPR2023-00656, Paper No. 33, p.52

POR, 59.

Pet. Reply:

IX. GROUND 6

Motive alleges “Petitioner fails to provide any motivation of why a POSITA would combine Choe with both Westmacot and Workman.” POR, 59. Not so. The motivation to use Workman’s predictive model in Choe is articulated in Ground 3. Pet., 47-49. And the motivation to compute the camera height and road plane normal in Choe’s system, as described in Westmacot, is articulated in Ground 4. Pet., 53-60. These rationales are independent and relate to different parts of Choe. EX1085, ¶¶141-143. If a POSA were to make both combinations, the result would disclose every element of the challenged claims. EX1085, ¶¶141-143.

Pet. Reply, 33.

Thank You