

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

SAMSARA INC.
Petitioner

v.

MOTIVE TECHNOLOGIES, INC.
Patent Owner

Case IPR2025-00574
Patent No. 11,875,580

**DECLARATION OF TREVOR J. DARRELL, PH.D.,
IN SUPPORT OF PETITION FOR *INTER PARTES* REVIEW OF
U.S. PATENT NO. 11,875,580**

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

TABLE OF CONTENTS

I.	Introduction.....	1
II.	Qualifications.....	3
III.	Understanding of Applicable Patent Law.....	6
	A. Claim Construction.....	6
	B. Obviousness.....	6
IV.	Technology Overview	8
	A. History of Predictive Models for Object Detection and Horizon Estimation.....	9
	B. Prior Art Relied on for the Grounds of Unpatentability	13
	C. '580 Patent.....	28
V.	Prosecution History	31
VI.	Level of Ordinary Skill in the Art	34
VII.	Grounds of Unpatentability	34
VIII.	Ground 1: Choe renders obvious claims 1-4.	35
	A. Independent Claim 1.....	35
	B. Dependent Claim 2	49
	C. Dependent Claim 3	50
	D. Dependent Claim 4	51
IX.	Ground 2: Choe in view of Khan and further in view of Ali renders obvious claims 1-5.....	52
	A. Motivation to Combine.....	52
	B. Claims 1-4.....	58
	C. Dependent Claim 5	61

X.	Ground 3: Choe in view of Workman renders obvious claims 1-4.....	64
A.	Motivation to Combine.....	64
B.	Claims 1-4.....	66
XI.	Ground 4: Choe in view of Westmacot renders obvious claims 5-7.....	70
A.	Motivation to Combine.....	70
B.	Dependent Claim 5	78
C.	Dependent Claim 6	80
D.	Dependent Claim 7	88
XII.	Ground 5: Choe in view of Westmacot, further in view of Khan and Ali renders obvious claims 1-7.....	91
XIII.	Ground 6: Choe in view of Westmacot, further in view of Workman renders obvious claims 1-7.....	92
XIV.	No Objective Indicia of Nonobviousness.....	93
XV.	The art cited in the Grounds are substantively different than and not cumulative to the references cited during prosecution.....	94
XVI.	Conclusion	1

I. Introduction

I, Trevor J. Darrell, declare as follows:

1. I have been retained on behalf of Samsara Inc. for the above-captioned *inter partes* review proceeding. I understand that this proceeding involves U.S. Patent No. 11,875,580 (EX1001, “the ’580 patent”), which is titled “Camera Initialization for Lane Detection and Distance Estimation Using Single-View Geometry.”

2. I understand the ’580 patent is assigned to Motive Technologies, Inc. I also understand that the ’580 patent names three inventors: Ali Hasaan, Ahmed Ali, and Syed Wajahat Ali Shah Kazmi. EX1001, front page, (72).

3. I reviewed and am familiar with the specification and the claims of the ’580 patent. The ’580 patent was filed on October 4, 2021 as U.S. Patent Application No. 17/493,011. EX1001, front page, (22).

4. I understand the ’580 patent has been provided as Exhibit 1001. I will cite to the specification using the following format: ’580 patent, 1:1-10. This citation points to the specification of the ’580 patent at column 1, lines 1-10.

5. I have reviewed and am familiar with the file history of the ’580 patent. The file history of the ’580 patent has been provided as Exhibit 1002. A summary of my understanding of the prosecution history is in Section V.

6. My opinions are based on more than 30 years of working with computer vision and pattern recognition, which includes decades of experience teaching and researching predictive models for computer vision applications. My opinions are also based on investigation and study of the relevant materials, including the patent at issue and its file history, and the prior art.

7. For example, I have reviewed and am familiar with the following prior art used in the petition for *inter partes* review of the '580 patent:

- U.S. Patent Application Publication No. 2020/0410704 to Choe *et al.* (EX1005) and its filed figures (EX1006)
- Khan *et al.*, Lane Detection Using Lane Boundary Marker Network with Road Geometry Constraints (EX1007)
- Ali *et al.*, Real-Time Vehicle Distance Estimation using Single View Geometry (EX1008)
- PCT Application Publication No. PCT/EP2019/056356 to Westmacot *et al.* (EX1010)
- Scott Workman *et al.*, Horizon Lines in the Wild (EX1011)

8. I have been asked to provide my technical review, analysis, insights, and opinions about the '580 patent and the above-listed references that form the basis for the grounds of unpatentability.

9. A full list of documents I considered is in my Materials Considered, which is attached to this declaration as Appendix A.

10. To the best of my knowledge, the accompanying exhibits are true and accurate copies of what they purport to be, and that an expert in the field would reasonably rely on them to formulate opinions such as those in this Declaration.

11. I may rely upon these materials and/or additional materials to rebut arguments raised by the Patent Owner. Further, I may also consider additional documents and information in forming any necessary opinions, including documents that may not yet have been provided to me.

12. My analysis of the materials relevant to this proceeding is ongoing, and I will continue to review any new material as it is provided. This report presents only those opinions I have formed to date. I reserve the right to revise, supplement, and/or amend my opinions stated herein based on new information and on my continuing analysis of the materials already provided.

13. I am being compensated at \$810 per hour for my work in this case. My compensation does not depend on the outcome of the case. The fact that I am being compensated has not altered the opinions in this declaration.

II. Qualifications

14. I am currently a Professor in the Computer Science Division (EECS) at UC Berkeley. I am also currently a member of the steering committee for the Berkeley Artificial Intelligence Research Center (BAIR). I was previously a Professor in the Department of Electrical Engineering and Computer Science at

MIT from 1999 to 2008 and a Group Leader at the International Computer Science Institute (ICSI) from 2008 to 2014.

15. I have been researching and/or teaching in the area of computer science, with a particular focus on computer vision, machine learning, computer graphics, and perception-based human-computer interfaces for over 30 years. I currently lead a laboratory where our current research focuses on algorithms to enable visual recognition across a variety of platforms and applications.

16. I am a past or current member of a number of scientific societies that focus on computer vision and machine learning, including the DARPA Information Science and Technology Advisory Study Group (2003-2014) and various National Science Foundation (NSF) Panels (2002-present). I have been invited to chair numerous conferences related to computer vision and machine learning. For example, I have served as both the Area Chair and Program Chair for the Conference on Computer Vision and Pattern Recognition (CVPR), the Area Chair for the International Conference on Computer Vision (ICCV), the Area Chair and General Chair of the International Conference on Multimodal Interfaces (ICMI), the Area Chair for European Conference on Computer Vision (ECCV), the Area Chair for the International Conference on Machine Learning (ICML), the Area Chair for the International Conference on Learning Representations (ICLR), and the Program Chair for the Conference on Perceptual User Interfaces (PUI).

17. I have been an editor for a number of leading journals in the field of computer vision and artificial intelligence, such as the Artificial Intelligence Journal (2004-2011) and the IEEE Transactions Pattern Analysis and Machine Intelligence Journal (2005-2011, 2015-2017).

18. I have published over 300 articles and conference papers related to computer algorithms, computer vision, and image processing. My publications have won several awards and honors, including the International Conference on Machine Learning (ICML) Test of Time Award in 2024, the Computer Vision Foundation Longuet-Higgins Prize in 2024, the Association for Computer Machinery SIGMM Test of Time Paper Award in 2024, the North American Associate for Computation Linguistics Best Paper award in 2016, the Computer Vision and Pattern Recognition Best Paper Honorable Mention in 2015, the ACM Multimedia Best Open Source Software Award in 2014, and the Internet Vision Workshop Best Paper Award in 2008. In 2016, my team also won the Everingham Prize at the International Conference on Computer Vision, recognizing our work in providing a significant contribution to the computer vision community. To date, my publications have been cited by other scholars over 269,000 times.

19. I am a named inventor on several patents related to object recognition and image processing. I have also served as a consultant in the area of computer vision for a number of start-up and established companies, including Pinterest

(IPO), IQ Engines (acquired by Yahoo), and BotSquare/Flutter (acquired by Google). I continue to advise a number of startup companies including Nexar, Grabango, and SafelyYou. I also have previously served as a consulting and/or testifying expert witness in a number of patent infringement cases, including two ITC investigations, as described more fully in my CV (EX1004).

III. Understanding of Applicable Patent Law

20. I am not a lawyer. However, I have been advised that technical experts should apply certain legal standards in forming their opinions.

A. Claim Construction

21. I understand that claim terms should be interpreted under a claim construction standard called *Phillips*. I understand that, according to the *Phillips* standard, claim terms should be construed according to the terms' meaning as understood by a person having ordinary skill in the art (POSA) at the time of the invention. I understand that the claim terms should be construed in light of the specification. I also understand that the prosecution history of the patent should be consulted in determining claim meaning. In this declaration, I have construed the terms according to their ordinary and customary meanings as understood by a POSA in view of the specification and prosecution history.

B. Obviousness

22. I understand a patent claim is unpatentable if the claimed invention would have been obvious to a person of ordinary skill in the art (POSA) at the time

of the alleged invention. This means that a patent claim can be unpatentable even if all of the claim elements cannot be found in a single prior-art reference.

23. I understand an invention is obvious when the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious to a POSA at the time of the invention.

24. I understand that to prove that prior art renders a patent claim obvious, it is necessary to: (1) identify the particular references that, singly or in combination, teach or suggest every element of the claim; (2) identify which claim elements appear in which reference; and (3) explain why a POSA would have sought to combine the prior-art references.

25. I understand that when considering whether a patent claim is obvious, one should consider whether at the time of the alleged invention a teaching, suggestion, or motivation to combine the references existed so as to avoid impermissibly applying hindsight. I understand this test should not be rigidly applied, but that the test can be important to avoid such hindsight.

26. I understand the following rationales can support a motivation to combine prior-art references:

- (A) Combining prior art elements according to known methods to yield predictable results;
- (B) Simple substitution of one known element for another to obtain predictable results;

- (C) Use of known technique to improve similar devices (methods, or products) in the same way;
- (D) Applying a known technique to a known device (method, or product) ready for improvement to yield predictable results;
- (E) Choosing from a finite number of identified, predictable solutions, with a reasonable expectation of success; and
- (F) Known work in one field of endeavor may prompt variations of it for use in either the same field or a different one based on design incentives or other market forces if the variations are predictable to one of ordinary skill in the art.

27. I understand objective indicia (also called secondary considerations) can be important evidence as to whether a patent claim is obvious. Such indicia include: (1) commercial success of products covered by the patent claim; (2) a long-felt need for the invention; (3) failed attempts by others to make the invention; (4) copying of the invention by others; (5) unexpected results achieved by the invention; (6) praise of the invention by the infringer or others in the field; (7) taking of licenses under the patent by others; (8) expressions of surprise by experts and those skilled in the art at the making of the invention; and (9) the patentee proceeded contrary to the accepted wisdom of the prior art.

IV. Technology Overview

28. The '580 patent relates to “automotive systems and to machine learning systems for analyzing sensor data in such systems.” '580 patent, 1:7-9. In

particular, the '580 patent is directed to “identif[ying] one or more lines in the video using a predictive model (e.g., a convolutional neural network)” “*includ[ing] a horizon line*” and “comput[ing] some camera parameters based on the horizon line.” '580 patent, 1:38-44.¹ The methods described by the '580 patent were well known in the art at the time of filing.

A. History of Predictive Models for Object Detection and Horizon Estimation

29. Using predictive models such as a convolutional neural network (CNN) to detect objects and estimate the horizon line had been known well before the inventors of the '580 patent filed their patent application in October 2021. These predictive models had been known for *decades*, and applying them to traffic videos to perform the same functions—e.g., identifying the horizon—was not new.

30. Applying CNNs to images and video frames has been going on since the late 1980s. Rawat, p. 1 (EX1015). “Image classification . . . is a fundamental problem in computer vision.” Rawat, p. 1. CNNs with multiple layers of nonlinear information processing has been shown to overcome accuracy problems with traditional feature extraction techniques. Rawat, pp. 1-2. In 1979, a novel multilayer neural network was proposed, based on advances in neuroscience. Rawat, p. 7. Later, in 1989, LeCun “proposed the first multilayer CNNs and

¹ All emphasis added unless stated otherwise.

successfully applied these large-scale networks, to real . . . image classification problems.” Rawat, p. 8. Because CNNs require fewer free parameters and consider the spatial topology of the input data, CNNs were better able to deal with the variability of 2D shapes in images. Rawat, p. 9.

31. In the late 1990s and early 2000s, neural network research diminished but later experienced a resurgence in 2006. Rawat, pp. 9-10. In 2006, researchers experienced successes applying CNNs to various image processing problems including off-road obstacle avoidance and generic object classification. Rawat, pp. 10-11. Further, finding that CNNs require less supervised learning than other neural networks, interest in CNNs grew, creating more research in different applications of CNNs in computer vision. Rawat, p. 15. Some of these applications included vision for autonomous off-road driving. Rawat, p. 15.

32. CNNs were also widely used for object detection for autonomous driving in a city environment. Jiao, p. 1 (EX1012); He, p. 9 (EX1009). For example, a predictive model called “Mask R-CNN” was a predictive model that used a CNN to detect objects in cityscape images, including detecting cars and

pedestrians. He, p. 9.² The Mask R-CNN model used both a *backbone* architecture for feature detection and a network *head* for bounding-box recognition.

Network Architecture: To demonstrate the generality of our approach, we instantiate Mask R-CNN with multiple architectures. For clarity, we differentiate between: (i) the convolutional *backbone* architecture used for feature extraction over an entire image, and (ii) the network *head* for bounding-box recognition (classification and regression) and mask prediction that is applied separately to each RoI.

He, p. 4.

33. To train CNNs such as the ones described above, there was a need to quickly generate a database of annotated images. Russell, Abstract (EX1013); Workman, pp. 1-2 (EX1011). Hand-collecting data is cumbersome and it is difficult to obtain the same variety of images that can be found on the Internet. Russell, 1. To remedy this problem, researchers such as Russell proposed techniques to annotate images by exploiting labels that humans have placed on said images. Russell, 1. These labels included “labeling horizon lines, junctions, and edge types.” Russell, 1. Instead of having human users label everything in an

² The Mask R-CNN paper derives from earlier work which I was a co-author, which introduced the "R-CNN" family of methods: Rich feature hierarchies for accurate object detection and semantic segmentation, by Ross Girshick, Jeff Donahue, Trevor Darrell, Jitendra Malik, published at CVPR 2014.

image, a mixed method was used, which had a computer infer image information and a human user provide labels of object names and location in the image.

Russell, 1-2.

34. CNNs were also used to determine camera parameters and the horizon line. For example, Russell described an approach that determined the pitch of the camera based on the camera's focal length, the center of the image, and the image location of the horizon line. Russell, 3. The horizon line was determined by finding the height of the camera from the height of various objects in the image. Russell, 5-6. Using the horizon line, Russell's method could then ignore detected objects that lie above the horizon line without contact to the detected ground plane. Russell, 6.

35. Bartl is another example that used predictive models to determine the horizon line. Bartl's approach was used as a "Fully Automatic Horizon Estimation for Surveillance Cameras." Bartl, Title (EX1014). "The proposed algorithm is fully automatic in the sense that no user input is needed per-camera and it works with various scenes (indoor, outdoor, traffic, pedestrian, livestock, etc.)." Bartl, Abstract. However, Bartl's approach used humans to "annotate[] ground truth horizons" after the horizon was predicted. Bartl, Abstract. The horizon line was then used to calibrate the camera, such as to correct its capture of the scene it is recording. Bartl, 1.

36. Additional prior art, including patents and research publications, further described how predictive models could identify the horizon line in traffic images. The prior art shows how it was known for many years how human annotators could adjust the predicted line for higher accuracy. And the art shows that the horizon line was then used to calibrate cameras in autonomous vehicles.

B. Prior Art Relied on for the Grounds of Unpatentability

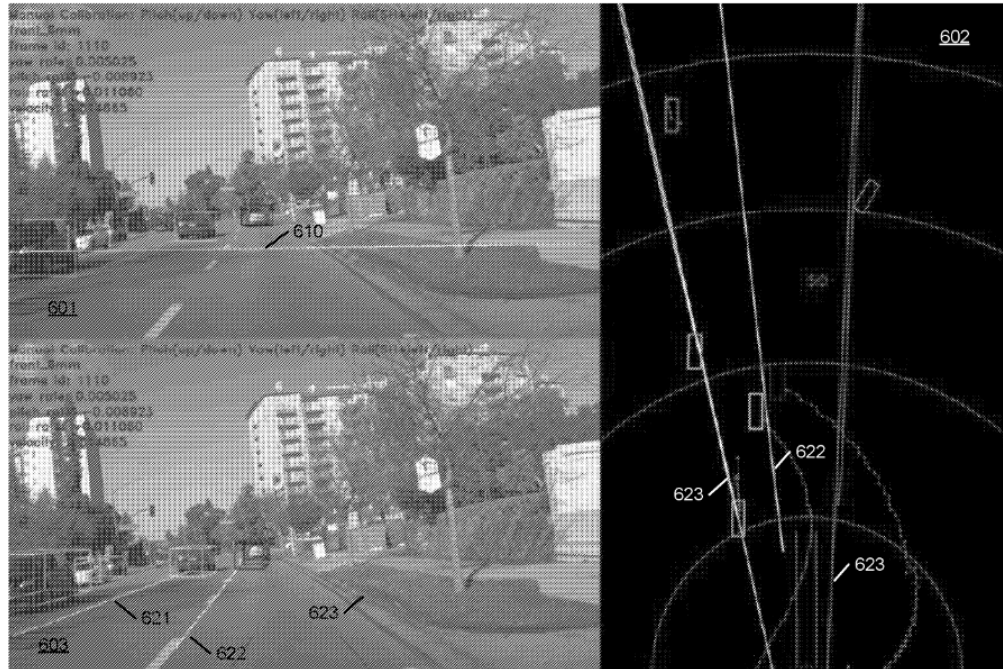
37. The following prior art used in the Grounds of unpatentability show that using predictive models to estimate the horizon line and an annotator to manually adjust the predicted horizon line was well known before the '580 patent.

1. Choe discloses a system that calibrates a camera for lane detection using a predicted horizon line and annotations.

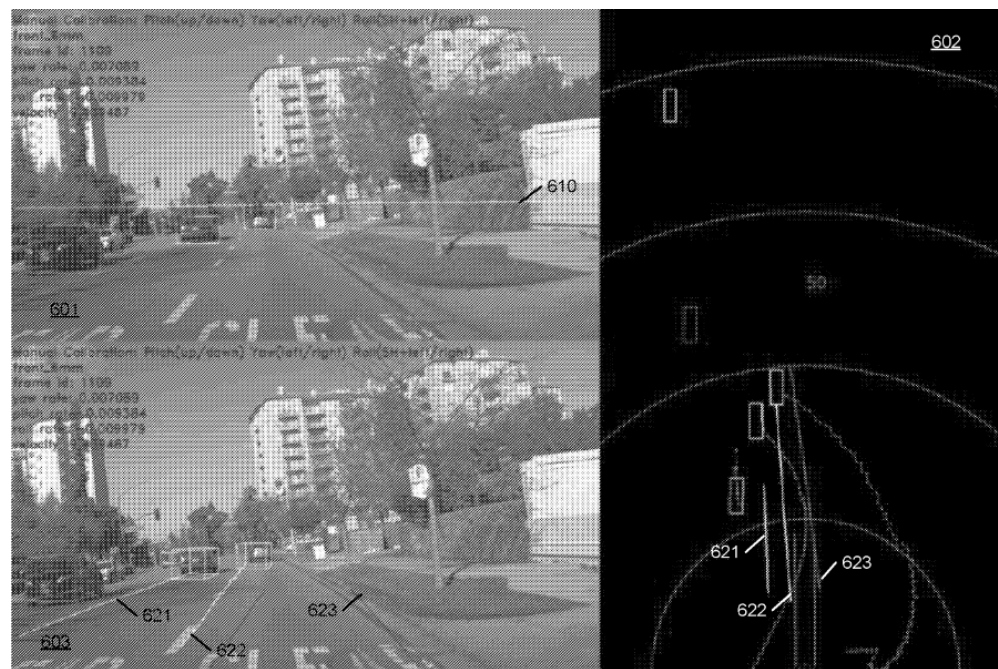
38. Choe is directed towards “calibrating a sensor system of an autonomous driving vehicle.” Choe, [0001] (EX1005). As Choe explains, sensors in autonomous vehicles need to be periodically calibrated to ensure that they accurately detect and recognize objects and lane configuration of the driving environment. Choe, [0003]. At the time of Choe’s invention, however, there was “a lack of efficient sensor calibration systems.” Choe, [0003]. One part of Choe’s sensor calibration system involves calibrating a camera for lane detection.

39. To calibrate its camera, Choe describes processing logic that receives images from the camera over a network. *See* Choe, [0018]. The camera is described as “still cameras and/or video cameras.” Choe, [0022]. When using a

video camera, the “images” are video frames, as seen in Choe’s Figures 6A and 6B:



Choe, FIG. 6A.



Choe, FIG. 6B.

40. After receiving the video frames, Choe predicts the horizon line and lane lines. Choe, [0015]-[0016]. “[A] horizon line is determined based on the camera’s hardware settings.” Choe, [0015]. Additionally, “[o]ne or more lane lines are determined based on the first image via a perception process performed on the first image.” Choe, [0015]; *see also* Choe, [0056], [0059], FIG. 7. The prediction of horizon and lane lines is done by a “machine learning engine 122 [that] generates or trains a set of rules, algorithms, and/or predictive models 124.” Choe, [0030].

41. Choe further describes superimposing lines for the horizon line and lane lines on the frames of the video. Choe, [0016], [0050]. After determining the horizon line within a video frame, the horizon line is superimposed on the frame. Choe, [0050]. Then, “a second image is rendered” and “includes the lane lines superimposed thereon.” Choe, [0050].



EX1006, FIG. 6A (excerpt, annotated); *see also* Choe, FIG. 6A.

42. After superimposing the lines, Choe sends the video to a user to review the predicted lines. *See* Choe, [0014]-[0016]. Choe describes having a “visualizer” which “includes a user interface to visually display a two-dimensional (2D) image captured by a camera.” Choe, [0014]. On this visualizer, a user views the lines and manually lines them up with the real lane lines and horizon line. Choe, [0055]-[0057]. A user may do this manual correction using an “input device (e.g., keyboard, joystick, voice interactive command).” Choe, [0015]. “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.” Choe, [0015]. The user manipulates the lines until they are satisfied with the results. *See* Choe, [0014].

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].

44. Once the user is satisfied with the line placement, Choe transmits updated camera parameters back to the vehicle's camera. *See* Choe, [0014]-[0016]. After a user has moved 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, [0015]. “Similarly, the yaw angle and the roll angle of a camera can also be calibrated using the techniques as described above.” Choe, [0057].

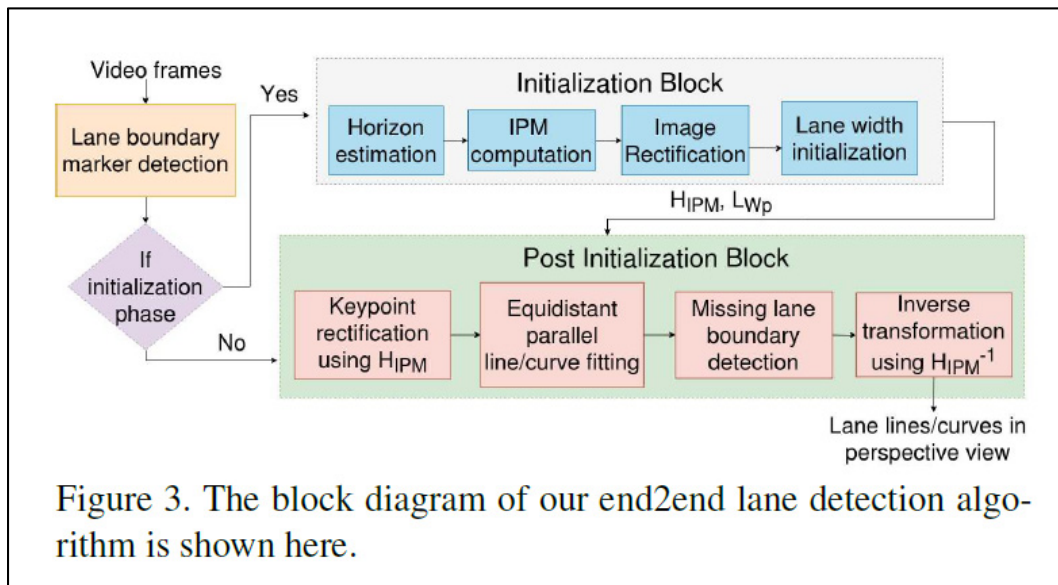
2. Khan discloses predictive models to identify lane lines, the horizon line, and initial camera parameters.

45. Khan discloses “a lane boundary marker network to detect keypoints along the lane boundaries.” Khan, Abstract (EX1007). Khan explains how lane detection for autonomous vehicles is “of critical importance,” Khan, p. 1823, and current solutions “suffer[ed] from the problem of occlusion and struggle[d] to detect lanes with low or no evidence on the road.” Khan, Abstract. At the time, Khan claimed that its solution outperformed other predictive models in the industry. *See* Khan, pp. 1828-30 (showing detailed results).

46. Khan uses a predictive model called “Mask-RCNN” to identify lane lines. Khan, p. 1825 (citing He, EX1009). Khan and the other authors did not invent this model. Rather, Mask-RCNN was a well-known model developed at Facebook AI Research. *See* He, p. 1. As Facebook explained in its publication (He), a Mask-RCNN has “(i) the convolutional backbone architecture used for

feature extraction” and “(ii) the network head for bounding-box recognition.” He, p. 4. Facebook also published that a Mask-RCNN could be used effectively for boundary detection in cityscapes. *See* He, p. 9. Khan merely repurposed the Mask-RCNN to “predict [] lane markers.” Khan, p. 1825.

47. Khan’s lane detection algorithm is shown in Figure 3 (below).



Khan, p. 1825.

48. Shown above, if the camera is being initialized (i.e., calibrated), the algorithm estimates the horizon line after the lanes are predicted (shown above in the top left blue box). Khan, pp. 1826-27. Khan explains that the horizon is detected based on the lane lines and states that the horizon prediction algorithm is further disclosed in a companion paper. *See* Khan, p. 1827 (citing Ali, EX1008).

49. The next step of Khan’s algorithm is called Inverse Perspective Mapping (IPM) computation. Khan, p. 1825. Using the predicted lane and horizon

lines, the IPM computation “use[s] road geometry to find camera pose.” Khan, p. 1826. This camera pose includes the camera tilt, pan angle, and road plane normal. *See* Khan, p. 1826. Instead of using the camera’s known parameters to determine the road geometry, Khan describes using “road geometry to find camera pose and in-turn find the IPM.” Khan, p. 1826. By using the lane lines and horizon line, Khan “comput[es] plane normal $\hat{n}$ ” to determine how to rotate the camera view to “align its Z-axis with the road’s normal.” Khan, p. 1827.

50. Khan’s algorithm then includes additional steps such as lane width initialization and post-initialization blocks. Khan, p. 1825. Although Khan states that “[t]he main contribution of this paper is the post processing pipeline, shown in Figure 3, following the lane boundary markers detection,” Khan, p. 1825, these steps are less relevant to the claims of the ’580 patent.

3. Ali discloses in more detail Khan’s predictive model for identifying the horizon line.

51. Like Khan, Ali estimates lane lines and horizon lines for running autonomous vehicles. Ali, p. 1100. Specifically, Ali explains that “[d]istance estimation is required for advanced driving assistance systems” and that it “uses geometric features of the road lane markings for distance estimation.” Ali, p. 1100.

52. Ali, like Khan, uses a Mask-RCNN to predict lane lines and the horizon line. *See* Ali, p. 1100, 1102. Ali cites to Khan to describe how it estimates lane lines. Ali, p. 1102. Then, Ali describes how it uses the lane lines to estimate a

horizon line. *See* Ali, p. 1103. Ali provides pseudocode for how its algorithm estimates the best fitting horizon (l_h).

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
6     for  $j \leq {}^B C_2$  do
7       forward_vpij = crossprod(line_1, line_2)
8     end
9   end
10  if  $B \geq 3$  then
11    for  $k \leq {}^B C_3$  do
12      Get lateral_vpik using cross ratios as described
        in Section 3.2.2
13    end
14  end
15 end
16 RANSAC for best fitting horizon across video:
17 initialize;
18 s = random_int between 1 to num_forward_vps;
19 t = random_int between 1 to num_lateral_vps;
20 while iterations  $\leq t$  OR num_inliers  $\leq \text{max\_inliers}$  do
21   Fit line  $l_h$  to forward_vp(s) and lateral_vp(t);
22   if  $\text{dist}(l_h, \text{forward\_vp}(s)) \leq \text{threshold}_s$  &
       $\text{dist}(l_h, \text{lateral\_vp}(t)) \leq \text{threshold}_t$  then
23     record and count inlier points
24   end
25 end
```

Ali, p. 1104.

53. Using the horizon line, Ali explains that it may “find[] camera height and road plane normal.” Ali, p. 1103. As Ali describes determining the camera height, Ali also describes “comput[ing] [the camera’s] viewing angle α alpha.” Ali, p. 1105. Once the viewing angle is determined, Ali explains that the “camera height h_c ” is given by:

$$h_c = \frac{\frac{Lwr}{2}}{\tan\left(\frac{\alpha}{2}\right)} \quad (3)$$

54. Using the camera height, road plane normal, and horizon line, Ali then determines the distance between the camera and a car in front of it. *See* Ali, p.

1104.

4. Workman discloses a CNN for horizon prediction.

55. Workman uses CNNs to predict horizon lines without requiring any use of geometric constraints or spatial cues. Workman, Abstract. Workman acknowledges that CNNs had been used to predict horizon lines many times prior for “a wide variety of applications.” *See* Workman, pp. 1-2. One known application was “pedestrian and vehicle detection.” Workman, p. 1.

56. Workman states that it provides three main contributions over other predictive models at the time:

Our main contributions are: 1) a novel approach for using structure from motion to automatically label images with a horizon line, 2) a large evaluation dataset of images with labeled horizon lines, 3) a CNN-based approach for directly estimating the horizon line in a single image, and 4) an extensive evaluation of a variety of CNN design choices.

Workman, p. 2.

57. Workman provides a “CNN-based approach for directly estimating [a] horizon line” Workman, 2. As part of this approach, Workman discusses determining a horizon line, “manually validat[ing] a global horizon line for each model, and then project[ing] the horizon line back into each image.” Workman, 4.

5. Westmacot discloses an annotator that adds lane lines.

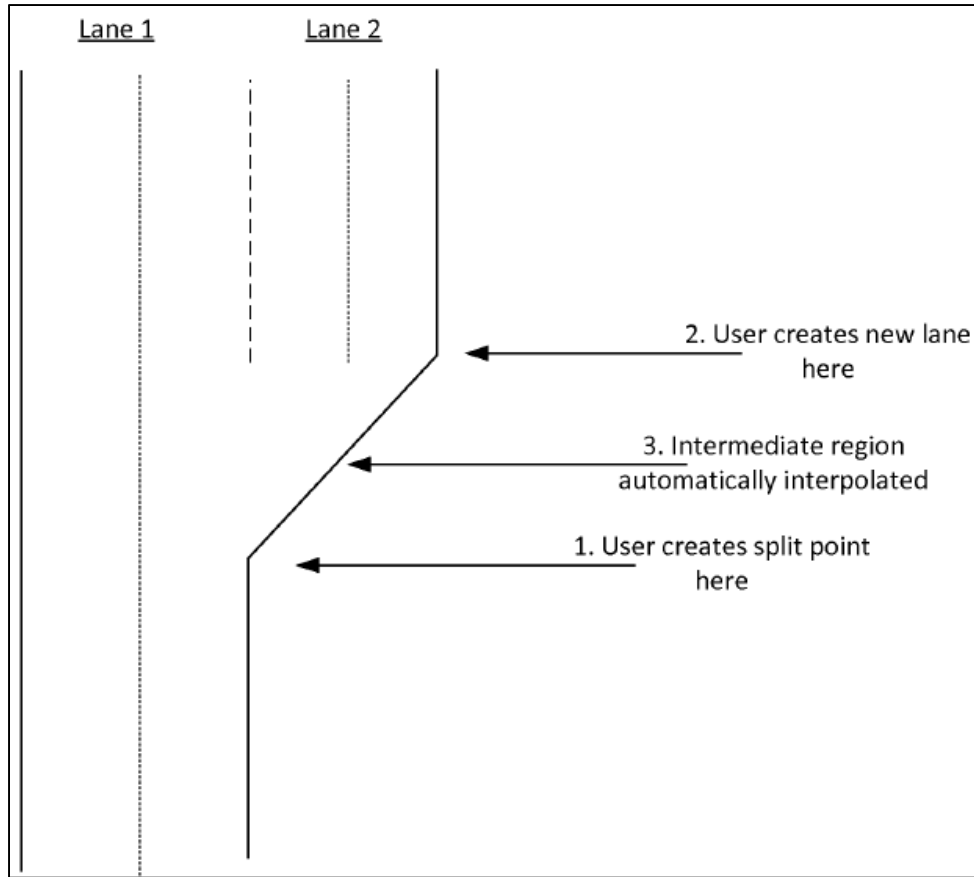
58. Westmacot is directed towards “image annotation,” specifically for “the creation of annotated street road images which can be used (among other things) for training a machine learning road detection component.” Westmacot, 1:05-07 (EX1010). As Westmacot explains, “[convolutional neural networks] require large numbers of training images.” Westmacot, 1:29. At the time of Westmacot, however, “[t]he conventional technique for creating the annotated training images is to have humans manually hand annotate the images.” Westmacot, 2:06-07. Westmacot describes a process for “automatic [and] semi-automatic image annotation technology” to “allow[] large numbers of training images to be generated more efficiently and quickly.” Westmacot, 3:01-05.

59. Westmacot describes how it annotates the road images to show the “road structure.” *See* Westmacot, 3:15-24. The road structure is the “real-world structure, such as road or lane structure, e.g. which part of the image is road surface, which part of the image makes up lanes on the road, etc.” Westmacot, 1:22-26. Westmacot explains the following method for annotating road images:

receiving a time sequence of two dimensional images as captured by an image capture device of travelling vehicle; processing the images to reconstruct, in three-dimensional space, a path travelled by the vehicle; using the reconstructed vehicle path to determine expected road structure extending along the reconstructed vehicle path; and generating road annotation data for marking at least one of the images with an expected road structure location, by performing a geometric projection of the expected road structure in three-dimensional space onto a two- dimensional plane of that image.

Westmacot, 3:17-24.

60. Westmacot describes that, after determining lane lines, a horizon line, and camera parameters, a manual reviewer may adjust the lines and the camera parameters. *See* Westmacot, 34:05-16. Westmacot describes this as the “human fixer stage.” *See* Westmacot, 19:17-18. “Figure 12 shows an example of a manual adjustment at a human fixer stage.” Westmacot, 17:14.



Westmacot, FIG. 12.

61. Using Figures 18 to 24, Westmacot illustrates “an annotation tool user interface for annotating images based on the techniques disclosed herein.”

Westmacot, 17:23-24. “[T]he annotator is provided with controls to manipulate currently rendered lanes (narrow, widen, move to the left or right, move the boundaries of the lane etc.) and add new lanes.” Westmacot, 34:07-09. Further, “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 vehicle’s dash or bonnet.” Westmacot, 34:09-11. “All annotations are performed in the

estimated 3D road plane, but immediate feedback is provided via projection in the 2D camera view.” Westmacot, 34:11-13.



Fig. 18

Westmacot, FIG. 18.

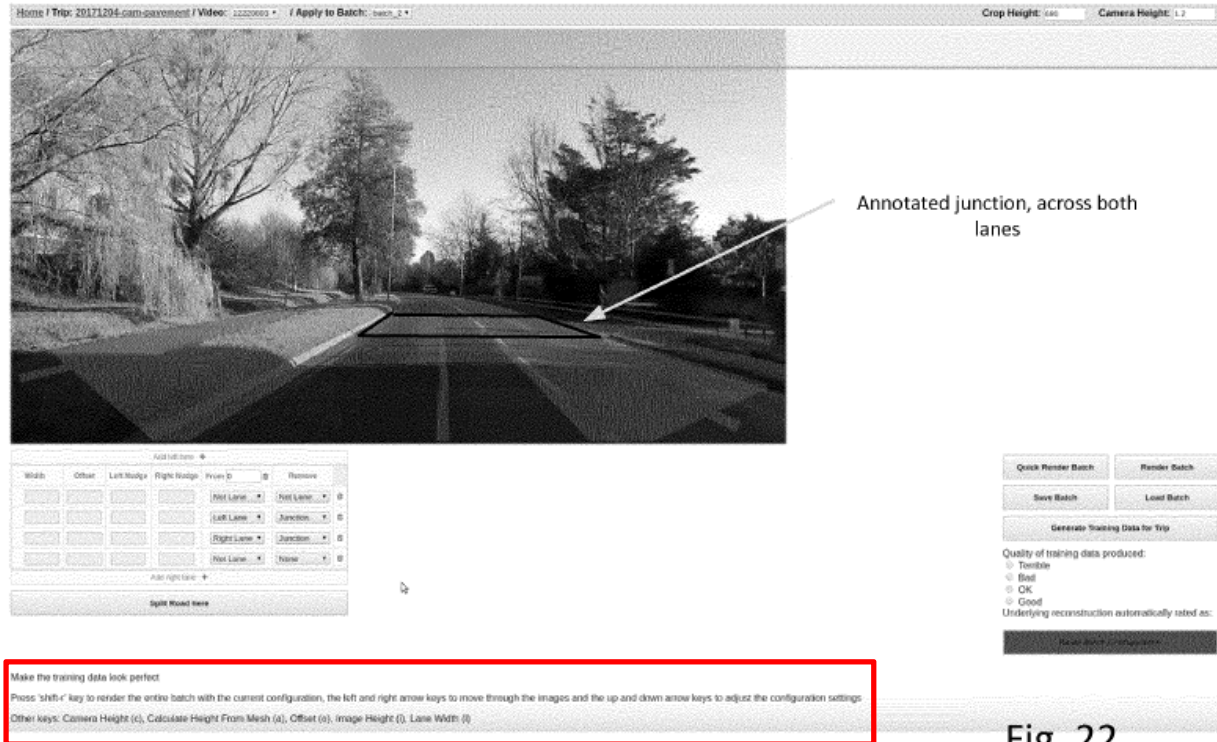


Fig. 22

Westmacot, FIG. 22 (annotated).

Other keys: Camera Height (c), Calculate Height From Mesh (a), Offset (o), Image Height (i), Lane Width (l)

62. During the process described above, Westmacot determines camera parameters from the image. Westmacot, 3:26-34, 6:10-11, 7:03-04. Westmacot explains that “one or more parameters of the image capture device . . . may be derived from (only) the image data captured by the image capture device.” Westmacot, 3:31-34. “A camera pose R_i may be determined for each image i in performing the SLAM processing.” Westmacot, 6:10-11. “The [image capture device’s] height may be estimated based on an average height of a reconstructed path of the image capture device and a 3D road structure mesh computed from the images.” Westmacot, 7:03-04. Westmacot further describes methods for measuring

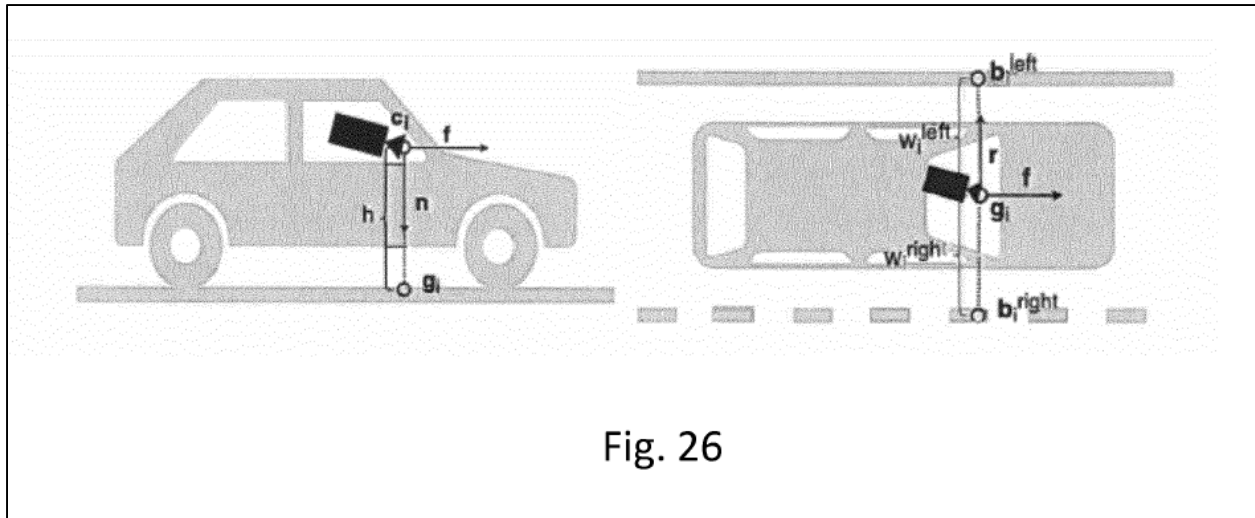
the angular orientation and rotational orientation of the image capture device from the images. *See* Westmacot, 7:10-22, 9:4-17, 9:27-33.

63. Westmacot further describes that camera parameters include the road plane normal (n). “The expected road structure may be determined based on *a road normal vector*.” Westmacot, 5:1. The road normal vector n is “the surface normal of the road relative to the camera,” 25:1-2, and may be estimated from the image as:

$$\mathbf{n} = \frac{1}{\sum_{i=2}^{N-2} \|\mathbf{n}_i\|} \sum_{i=2}^{N-2} \mathbf{n}_i$$
$$\mathbf{n}_i = \mathbf{R}_i^{-1}(\mathbf{m}_{i-1,i} \otimes \mathbf{m}_{i,i+1}),$$

Westmacot, 5:2-4; *see also* Westmacot, 24:1-6, 26:9-27:5, 39:31-40:2, FIGs. 8-9, 16, 27 (additional details on the road plane normal).

64. Westmacot’s Figure 26, shown below, further illustrates the camera height (h) and road normal (n) in relation to the determined lane lines.



Westmacot, FIG. 26.

C. '580 Patent

65. The '580 patent describes well-known, conventional algorithms to detect lanes and objects in images. In general, the '580 patent describes “automotive systems and . . . machine learning systems for analyzing sensor data in such systems.” '580 patent, 1:1-9. In more detail, the '580 patent describes “a network-based initialization system” for “receiv[ing] a video comprising a set of image frames and identif[ying] one or more lines in the video using a predictive model” to “compute[] some camera parameters.” '580 patent, 1:36-44. The '580 patent also describes a “neural network” embodiment that “records a video comprising one or more image frames,” “generates a set of intermediate features representing the video,” “predict[s] a horizon line,” and “compute[s] a camera parameter . . . based on the horizon line.” '580 patent, 1:53-64. Finally, the '580 patent “records a video comprising one or more image frames,” “generates a set of

intermediate features representing the video,” “predict[s] camera parameters using the output of the camera estimation head, the camera parameter estimation head trained using an output of the backbone network, and an output of a lane and horizon line prediction head.” ’580 patent, 1:65-2:08.

66. The ’580 patent starts by describing at a high level the device 230, camera initialization service 232, annotator device 234, and initialization service 236. *See* ’580 patent, 4:58-7:10. Device 230 is “a camera and processor installed in a vehicle” that “captures a video, represented as, for example, a set of one or more images.” ’580 patent, 4:60-62. After capturing a video, called a camera initialization event, device 230 “transmits a camera initialization event to a camera initialization service 232.” ’580 patent, 5:13-15. The camera initialization service 232 then “detects lane and horizon lines in the video using a predictive model.” ’580 patent, 5:25-28. The ’580 patent defines a predictive model as “a deep learning or other AI model” or “a CNN or other type of deep learning model.” ’580 patent, 5:25-32. A CNN is a convolutional neural network. The ’580 patent then describes “if process 206 can detect at least two lanes . . . , process 206 can further include computing camera parameters for the device 230 that recorded the video.” ’580 patent, 5:38-41. The process can predict “camera height, viewing angle, and road plane normal.” ’580 patent, 5:45-47. The process then “overlays the computed lane and horizon lines on the original video file, generating an overlaid video file.”

'580 patent, 5:48-50. The overlaid video is then sent to an annotator device 234.

'580 patent, 5:54-56.

67. The annotator device 234 is merely described as “a workstation or web-based application allowing for human review of automatically detected horizon and lane lines.” '580 patent, 5:65-68. After receiving the overlaid video, the annotator device 234 “presents the overlaid video file to the human annotator” and “requests that the human annotator confirm the lane lines (if present) and horizon in the overlaid video file.” '580 patent, 6:01-05. “If the annotator confirms the overlaid lines, the annotator can accept the lines and transmit its acceptance back to the camera initialization service 232.” '580 patent, 6:05-08. If the annotator does not confirm the overlaid lines, the annotator can “manually add a horizon line and/or lane lines.” '580 patent, 6:31-33. The annotator then transmits the overlaid video to initialization service 236, which may use “a deep learning or AI model” to “predict[] camera parameters (e.g., height, IPM, etc.) based on” the overlaid video. '580 patent, 6:44-46. It is unclear whether initialization service 236 is different from camera initialization service 232. The '580 patent describes that, in some embodiments, camera initialization service 232 can comprise two different models: “one predicting lane lines and one predicting camera parameters based on lane lines.” '580 patent, 6:47-49. The patent further describes how “these two models

may comprise heads of a multi-headed model sharing a common backbone neural network.” ’580 patent, 6:52-55.

68. Later in the patent, “heads” are described as “a neural network.” ’580 patent, 2:03. For example, the ’580 patent describes the lane-horizon marking head as being a CNN; a fully convolutional network (FCN), which can be based on Unet, DeepLab, PSPNet, SCNN, or EnetSAD; feature pyramid network (FPN); spatial pyramid pooling (SPP); recurrent neural network (RNN); or long short-term memory (LSTM). ’580 patent, 7:24-40. Further, “a backbone neural network” is described as “an EfficientNet backbone network.” ’580 patent, 7:58-59. The backbone network is further described as “a plurality of convolutional layers followed by one or more LSTM layers.” ’580 patent, 7:61-64.

V. Prosecution History

69. During the prosecution history of the ’580 patent, there was only a single office action and response before allowance, with no claim amendments.

70. In the single office action, the Examiner rejected claims 1-3, 5, 8-12, and 14-15 over prior art and objected to claims 4, 6-7, 13, and 17-20 for being dependent upon the rejected base claims. EX1002, p. 129. The cited prior art included: (1) U.S. Patent Application Publication No. 2019/0034740 A1 to Kwant *et al.* (“Kwant1”) and (2) U.S. Patent Application Publication No. 2019/0102674 A1 to Kwant *et al.* (“Kwant2”). EX1002, p. 131.

71. In response to the rejections, the Applicant submitted five arguments without amending the claims. First, the Applicant argued that Kwant1 did not disclose “a predictive model, the one or more lines including a horizon line” (claim 1) or “predicting a horizon line using the lane or horizon prediction head” (claim 8). EX1002, p. 145. According to the Applicant, Kwant1 did not disclose a predictive model to detect a *horizon* line but instead predicted *lane* lines and then used a geometric approach to identify the horizon line. EX1002, pp. 145-46.

72. Second, the Applicant argued that neither Kwant1 nor Kwant2 disclose “at least one camera parameter” (claims 1 and 8). EX1002, pp. 146-47. According to the Applicant, the parameters in Kwant1 and Kwant2 “have no relation to the camera used to generate the image but rather describe the lines—they are a pure mathematical representation of a line.” *See* EX1002, pp. 146-47. In the response, “Applicant note[d] 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, p. 146.

73. Third, the Applicant argued that Kwant1 did not disclose overlaying lines on a video to generate an overlaid video as recited in claim 1. EX1002, pp. 147-48. According to the Applicant, “while the figure in Kwant1 may arguably illustrate an image with polylines overlaid (a point not conceded), Kwant1 does not

describe a method step for doing this (i.e., the explicit step of ‘overlaying’ as claimed).” EX1002, p. 147. According to the Applicant, the alleged overlay image “[wa]s a construction of the patent attorney or agent and not of a process.” EX1002, p. 147. To satisfy the claim element, the Applicant noted that the reference “must describe a *process* for generating an *image* and the failure to do so renders Kwant1 deficient.” EX1002, p. 147 (emphasis in original).

74. Fourth, the Applicant argued that Kwant1 did not disclose transmitting an overlaid video to an “annotator device for manual review” as claim 1 recites. According to the Applicant, even if Kwant1 displayed an image with overlaid lines, “the ultimate recipient in Kwant1 is a driver, not an ‘annotator device’ as claimed.” EX1002, p. 148.

75. Finally, the Applicant argued that the cited prior art failed to describe a “confirmation” from an “annotator device for manual review” as claim 1 recites. EX1002, p. 148. The Applicant explained:

The Office Action cites a “learner module” of Kwant2 as meeting this element. However, this learning module is not part of a “manual review.” Instead, it is an automated machine learning technique to adjust weights of a model based on a loss function. This is a fundamentally different process than using manual review to confirm the accuracy of a horizon detection.

EX1002, p. 148.

76. The Examiner then allowed all claims.

VI. Level of Ordinary Skill in the Art

77. A person of ordinary skill in the art (“POSA”) in the technical field of the ’580 patent at the relevant time would typically have had 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.

VII. Grounds of Unpatentability

78. The six grounds of unpatentability in this declaration are shown below. This declaration explains why each ground renders obvious the claims.

Ground	Reference(s)	Claims
1	Choe	1-4
2	Choe, Khan, and Ali	1-5
3	Choe and Workman	1-4
4	Choe and Westmacot	5-7

5	Choe, Westmacot, Khan, and Ali	1-7
6	Choe, Westmacot, and Workman	1-7

VIII. Ground 1: Choe renders obvious claims 1-4.

79. Choe discloses every element in claims 1-4. However, if Choe does not explicitly disclose “video” and “overlaid video,” these features would have been at least obvious in view of Choe’s teachings about processing “frames” (e.g., Figures 6A and 6B, disclosing frame ids 1109 and 1110).

80. Additionally, while Choe teaches that predictive models are used for Choe’s sensor calibration ([0030]), I understand that the Patent Owner may allege that Choe does not explicitly use a predictive model to identify horizon and lane lines. However, a POSA reading Choe would have understood that using predictive models to identify the horizon and lane lines was at least obvious.

A. Independent Claim 1

81. Choe discloses every element in claim 1.

1. [1Pre] “A method comprising”

82. Choe discloses this element, as shown in the following excerpt:

The processes *or methods* depicted in the preceding figures may be performed by processing logic that comprises hardware (e.g. circuitry, dedicated logic, etc.), software (e.g., embodied on a non-transitory computer readable medium), or a combination of both.

Choe, [0064].

83. Choe's specification and figures generally also describe its method.

2. [1A] “receiving, over a network from a camera device, a video comprising a set of image frames”

84. Choe discloses this element.

85. In Choe, an autonomous driving vehicle includes cameras 211. Choe, [0020]-[0022], FIGs. 1-2. Choe states that cameras 211 may be video cameras:

Cameras 211 may be still cameras and/or *video cameras*. A camera may be mechanically movable, for example, by mounting the camera on a rotating and/or tilting a platform.

Choe, [0022].

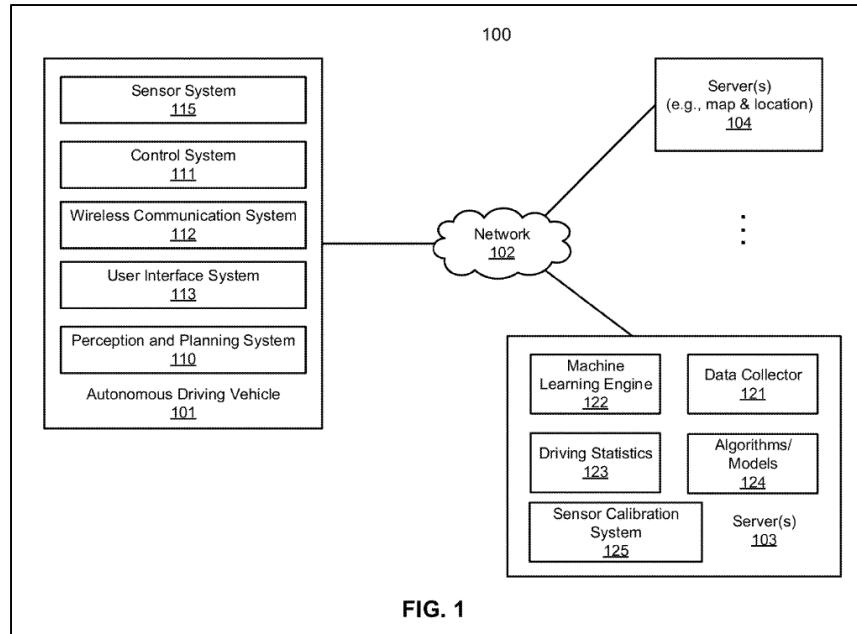
86. A POSA would have understood from Choe that the video captured by a video camera “compris[es] a set of image frames” as recited in claim 1 of the '580 patent. For example, Choe's Figures 6A and 6B show the video frames with frame ids 1109 (FIG. 6B) and 1110 (FIG. 6A, below). In the context of Choe, which uses a video camera, a sequence of over a thousand image frames is a video.

frame id: 1110



EX1006, FIG. 6A (excerpt, annotated); *see also* Choe, FIGs. 6A-6B.

87. Choe also discloses that its system sends information “over a network” because, “[r]eferring to FIG. 1, network configuration 100 includes autonomous vehicle 101 that may be communicatively coupled to one or more servers 103-104 *over a network 102.*” Choe, [0018]. The video captured by the video camera (sensor of sensor system 115) is received, over the network, by the sensor calibration system 125 on server(s) 103. *See* Choe, [0031], FIG. 1.



Choe, FIG. 1.

88. I understand that the Patent Owner may allege that receiving video frames does not necessarily mean that Choe’s sensor calibration system 125 receives a video. However, as I explain above, Choe discloses frames, which collectively form a video. To the extent Choe does not explicitly disclose video, combining the frames to form a video would have been obvious to a POSA. It would have also enabled the system to form and transmit a single, compressed file to the sensor calibration system in an efficient manner.

89. Additionally, arguing that processing video frames does not satisfy the claim elements deviates from what the Applicant argued during prosecution. During prosecution, the Applicant argued that the prior art “must describe a *process* for generating an *image* and the failure to do so renders [the prior art]

deficient.” EX1002, p. 147 (emphasis in original). The Applicant therefore argued that the claims required processing an *image*, not necessarily an entire video. Choe does, in fact, disclose a process for generating images (video frames) of a video captured by its video camera 211.

3. [1B] “identifying one or more lines in the video using a predictive model, the one or more lines including a horizon line”

90. Choe discloses (or at least suggests) this element.

91. Choe’s method “identif[ies] one or more lines . . . including a horizon line” because Choe determines a horizon and lane lines. *See* Choe, [0015], [0045], [0048], [0055], [0056], [0059], FIG. 7. For example, Choe states:

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.

Choe, [0015].

92. In Choe, the horizon and lane lines are identified in video frames because, as I explain for [1A], Choe captures video. *See also* Choe, FIGs. 6A-6B; EX1006, FIGs. 6A-6B (identifying horizon and lane lines on video frames). The identified lines would therefore be “in the video.”

93. A POSA would have understood that Choe discloses (or at least suggests) the initial horizon and lane lines being identified “using a predictive model” as recited in claim 1 of the ’580 patent. Choe states that it uses machine learning engine 122 and predictive models 124 for a variety of purposes. Choe, [0030]. One of these purposes is camera calibration. Choe, [0030]; *see also* Choe, FIG. 1 (showing algorithms/models 124 with the sensor calibration system 125).

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.

Choe, [0030].

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].

95. After a predictive model 124 predicts the initial horizon and lane lines and a user manually adjusts the lines to ground truth, the lines are then identified again and used by additional predictive models. Choe states:

The sensor data is then processed by perception module 302 such as image processing, including detecting an object within the image and determining a horizon line representing a vanishing point of a road. *The result of the perception processing can be utilized by other modules such as prediction module 303 and planning module 305 for autonomous driving.*

Choe, [0045]. This post-processing step provides additional support for [1B].

96. A POSA would have therefore understood that Choe discloses (or at least suggests) “identifying one or more lines in the video using a predictive model, the one or more lines including a horizon line” as recited in claim 1.

97. If Choe is found to not identify the horizon and lane lines using a predictive model, a POSA would have found it obvious because predictive models were widely used to identify these lines in cityscape and traffic images. For example, the references in Grounds 2 and 3 (Khan, Ali, and Workman) explicitly disclose predictive models to identify the horizon and lane lines. Indeed, this claim element was well known and performed in many applications, including for autonomous driving applications.

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

98. Choe discloses this element because it 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. For example, Choe states:

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.*

Choe, [0015].

At block 702, processing logic determines one or more lane lines based on the first image based on a perception process on the first image. In response to a first input signal received from an input device, at block 703, the position of the horizon line is updated based on the first input signal. At block 704, a position of at least one of the lane lines is updated based on the updated horizon line. *At block 705, processing logic determines a calibration factor or parameter for calibrating a pitch angle of the camera based on the difference between the initial horizon line and the updated horizon line.*

Choe, [0059]; *see also* Choe, [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], FIG. 7 (support for computing camera parameters based on lines).

5. [1D] “overlaying the one or more lines on the video to generate an overlaid video”

99. Choe discloses this element.

100. In Choe, the horizon line is superimposed on the video frame. As a POSA would have understood, superimposing a line on a video frame “overlay[s]” the line on the video frame. Choe states that the initial horizon line is superimposed on the video frame:

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.

Choe, [0016]; *see also* Choe, [0046], [0055]-[0057], FIGs. 6A-6B.

101. Shown below, the horizon line is superimposed on a video frame with “frame id: 1110.” Choe, FIG. 6A; EX1006, FIG. 6A. Shown below, this line is “too low” and will need to be adjusted by manual calibration. *See* Choe, [0057].



EX1006, FIG. 6A (excerpt, annotated); *see also* Choe, FIG. 6A.

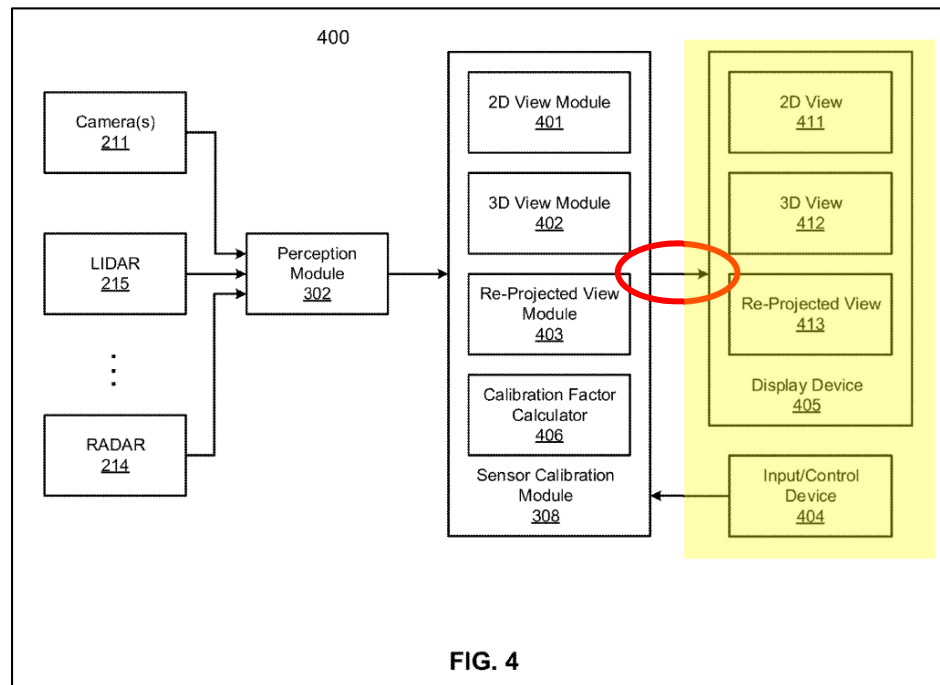
102. A video having the horizon line superimposed thereon (even if the overlay is on a single frame) is “an overlaid video.”

103. If the Patent Owner argues that overlaid video frames are not an overlaid video, I would disagree. Indeed, during prosecution, the Applicant argued that this element required “a *process* for generating an *image*” and “‘overlay[ing]’ data on an image.” EX1002, pp. 147, 148 (emphasis in original). A video frame is an image, and Choe overlays lines on the images. However, to the extent a collection of overlaid video frames is not an overlaid video, it would have been obvious to combine the frames to form an overlaid video.

6. [1E] “transmitting the overlaid video to an annotator device for manual review”

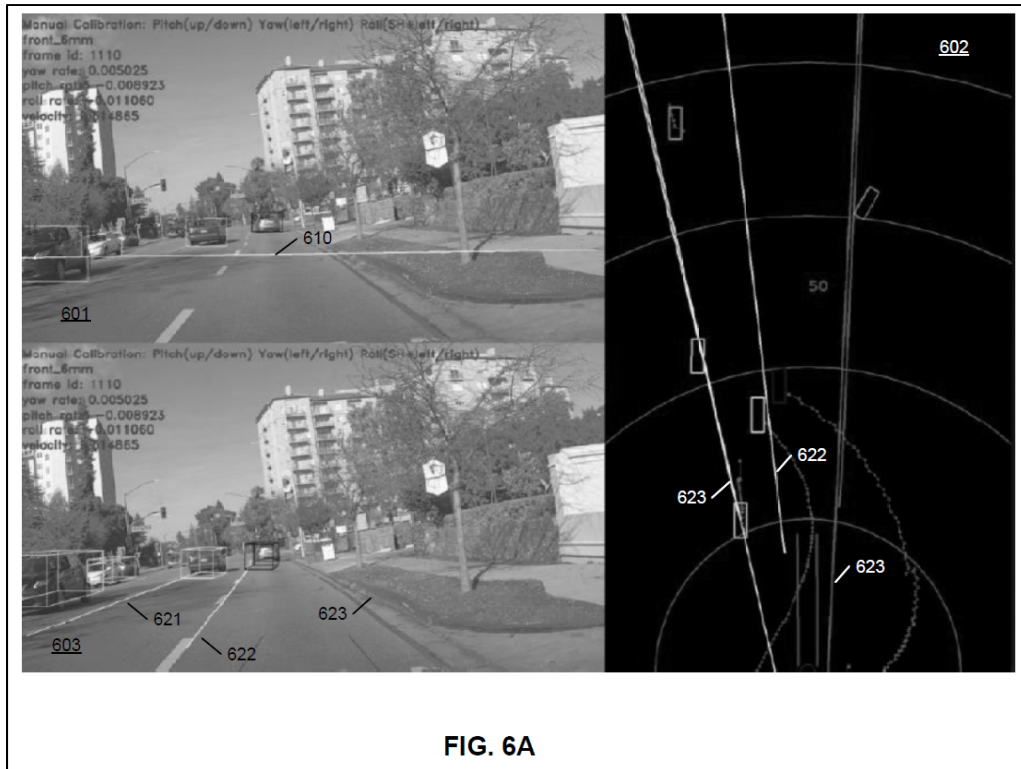
104. Choe discloses this element.

105. Shown below in Choe's Figure 4, the video with the initial overlaid horizon and lane lines is first transmitted to the sensor calibration module 308. Choe, [0048]. The overlaid video is then transmitted (shown below with the circled arrow) to the annotator device, e.g., elements 404, 405, 411, 412, and/or 413 (highlighted below in yellow). Choe, [0045]-[0054].



Choe, FIG. 4 (annotated); *see also* Choe, [0046]-[0047] (explaining FIG. 4).

106. Choe discloses that this annotator device is “for manual review” as recited in claim 1 of the ’580 patent. Specifically, a user uses the annotator device’s input device 404 (e.g., a computer station’s keyboard) to review, annotate, and reposition the overlaid lines manually on display 405. Choe, [0045]-[0054]. Figure 6A shows the annotator device’s graphical user interface (GUI), which would be on display 405. Choe, [0010], [0046], [0055]-[0058].



EX1006, FIG. 6A; *see also* Choe, FIGs. 6A-6B; EX1006, FIG. 6B.

107. Shown above in Figure 6A and as Choe describes, the GUI on display 405 instructs the user how to perform manual calibration with a keyboard:

Manual Calibration: Pitch(up/down) Yaw(left/right) Roll(SHift/left/right)

EX1006, FIG. 6A (excerpt); *see also* Choe, FIG. 6A.

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

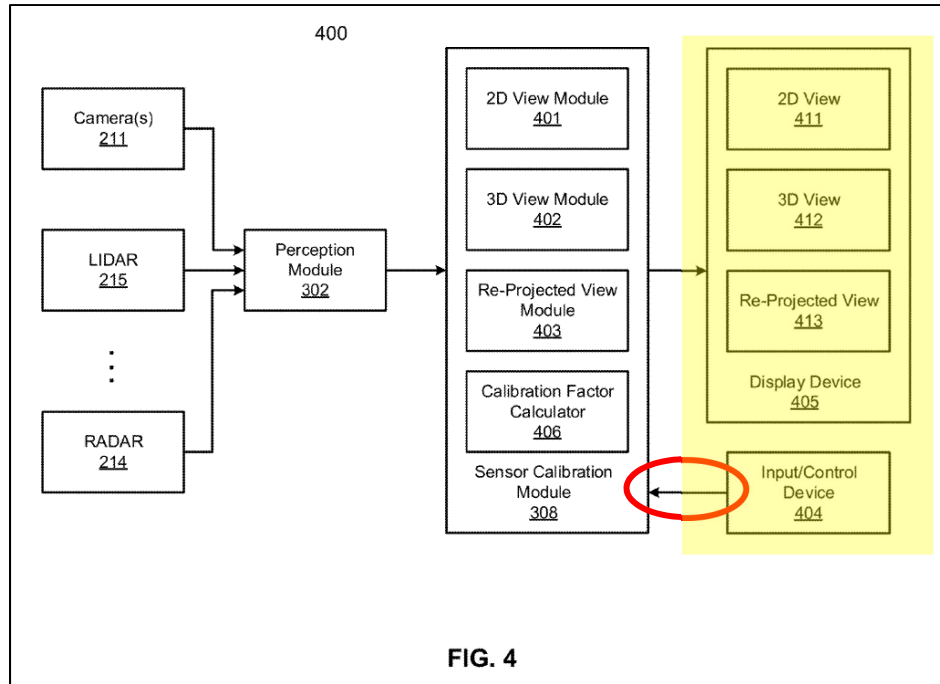
calibrating a pitch angle of the camera based on a difference between the initial horizon line and the updated horizon line.

Choe, [0015]; *see also* Choe, [0016]-[0017], [0047], [0049], [0052], [0055]-[0057], (explaining how a user manually adjusts the horizon and lane lines).

7. [1F] “receiving a confirmation from the annotator device, the confirmation indicating that the one or more lines are accurate”

108. Choe discloses this element.

109. Once the annotator device, e.g., elements 404, 405, 411, 412, and/or 413 (highlighted below in yellow), receives the video with overlaid image frames, the manual reviewer either adjusts the lines or confirms that they are accurate. *See* Choe, [0047]. Shown below in Figure 4, the annotator device’s input/control device 404 sends confirmation indicating that the lines and camera parameters are accurate. “[T]he user can calibrate the sensors until the calibration results as shown on display device 405 are satisfied.” Choe, [0047]. The sensor calibration module 308 receives this confirmation. *See* Choe, FIG. 4.



Choe, FIG. 4 (annotated).

8. [1G] “transmitting data representing the at least one camera parameter to the camera device”

110. Choe discloses this element.

111. In Choe, after the manual calibration, the data representing at least one camera parameter (e.g., calibration factors or parameters for calibrating the pitch, yaw, and roll of camera 211) is transmitted to the camera 211. Choe states:

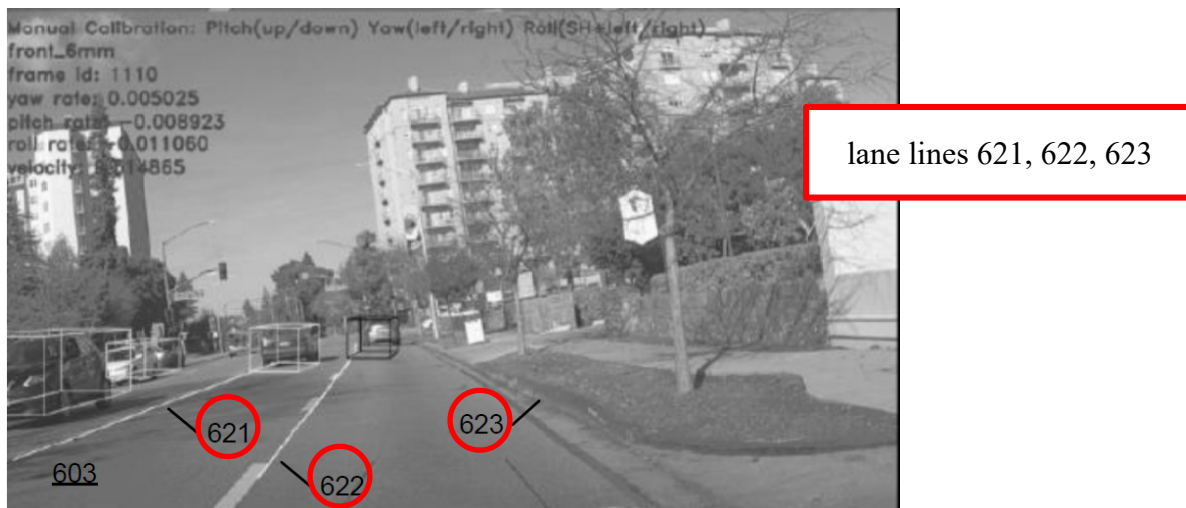
The calibration parameters can then be uploaded to the corresponding vehicles and utilized by autonomous driving software to calibrate and compensate the corresponding sensors.

Choe, [0031].

112. The camera parameters are transmitted to software modules of Choe's camera 211 to "perform proper adjustment at the software level to compensate the hardware settings of the sensors." Choe, [0047]; *see also* Choe, [0044].

B. Dependent Claim 2

113. Claim 2 depends from claim 1 and further recites "wherein the one or more lines includes at least one lane line of a roadway." Choe discloses this element because, as I show below, the lines 621, 622, 623 superimposed on the video frames include lane lines of a roadway Choe, [0056]; *see also* Choe, [0015]-[0017], [0048]-[0053], [0057], [0059], FIGs. 6B, 7.



EX1006, FIG. 6A (excerpt, annotated); *see also* Choe, FIG. 6A.

114. These lane lines are identified "using a predictive model" as recited in [1B] because the lane lines are determined and positioned based on the horizon line, which is identified using a predictive model. Choe states:

Image 402 includes one or more *lane lines 621-623 positioned based on the position of horizon line 610.*

Choe, [0056].

One or more lane lines are determined based on the first image via a perception process performed on the first image. 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.*

Choe, [0015]; *see also* Choe, [0016], [0048]-[0053], [0059].

C. Dependent Claim 3

115. Claim 3 depends from claim 1 and further recites “wherein computing the at least one camera parameter comprises predicting the at least one camera parameter using the predictive model.”

116. Choe discloses this element because the camera parameters (e.g., calibration factors or parameters for calibrating the camera’s pitch, yaw, and roll) are predicted “using a predictive model.” Choe states that it uses a “machine learning engine 122 [that] generates or trains a set of rules, algorithms, *and/or predictive models 124* for a variety of purposes.” Choe, [0030]. One of these purposes is camera calibration:

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.

Choe, [0030]; *see also* Choe, FIG. 1 (showing algorithms/models 124 with the sensor calibration system 125).

117. These sensor calibration algorithms/models predict initial camera parameters and updated camera parameters. Choe, [0015], [0017], [0047], [0049], [0052]-[0054], [0059].

D. Dependent Claim 4

118. Claim 4 depends from claim 1 and further recites “wherein computing the at least one camera parameter comprises predicting the at least one camera parameter using a second predictive model.” Choe discloses this element.

119. As I explain for [1B], Choe discloses a predictive model for identifying a horizon line in a video frame (a first predictive model). Additionally, as I explain for claim 3, Choe discloses a predictive model for predicting a camera parameter (a second predictive model). Choe states that it uses *multiple* predictive models: “machine learning engine 122 generates or trains a set of rules, algorithms,

and/or predictive models 124 [(plural)] for a variety of purposes.” Choe, [0030].

Choe further describes two different kinds of algorithms/models. Choe, [0030].

IX. Ground 2: Choe in view of Khan and further in view of Ali renders obvious claims 1-5.

120. Choe in view of Khan and further in view of Ali discloses (or at least suggests) every element in claims 1-5.

A. Motivation to Combine

1. Choe

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.

122. I understand that Khan and Ali are papers that include the inventors as authors. However, I understand that Khan and Ali are prior art to the ’580 patent because they were published and publicly available over a year before the effective filing date of the ’580 patent. Although Khan and Ali include the inventors, many of the concepts described in the papers are borrowed from prior work of many

others before them, based on their broad citing from related publications. Thus, the reliance on Khan and Ali is really a reliance on their summaries of the work of others. So this declaration is not crediting them with discovering the teachings being applied to the claims, but rather showing the inventors knew about other scientists using well-known predictive models to identify horizon and lane lines.

2. Khan

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").

124. A POSA would have also been motivated to use Khan's predictive models for Choe's camera calibration because both use similar methods. If the predicted horizon line is too low, the predicted lane lines diverge. Choe, [0057],

FIG. 6A. If the horizon line is too high, the predicted lane lines converge. Choe, [0057], FIG. 6B. Accordingly, during calibration, Choe manually sets the horizon line to maintain the parallelism of the lane lines. Choe, [0057]. Similarly, Khan relies on the parallelism of the lane lines when calibrating the camera. Khan states: “We fit lines to predicted keypoints and initialize lane width and camera rotation *using the equidistant and parallelism properties of the road lanes.*” Khan, pp. 1825-26. Additionally, like Choe’s use of a top-down view, Choe, [0057], to visualize the parallel property of lane lines, Khan’s model also “synthetically rotates the camera into a top (bird’s eye) view thus rendering the intersecting lines as parallel; the way they actually are in the real world.” Khan, p. 1826. Given these similarities in the method of predicting the lines and camera parameters, a POSA would have understood that Khan’s predictive models would have been compatible with Choe. Additionally, because predictive models such as the ones in Khan were so well known, a POSA would have had a reasonable expectation of success when combining Khan’s predictive models with Choe’s system.

3. Ali

125. A POSA would have also been motivated to use Ali’s horizon estimation algorithm in the Choe-Khan combination because Khan states that it uses Ali’s horizon estimation algorithm to find the best horizon l_h :

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 .*

Khan, pp. 1826-27.

126. Ali discloses the algorithm to find the best horizon l_h , which uses a predictive model using RANSAC and predicted keypoints for each video frame, as shown in the pseudocode below.

```
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
6     for  $j \leq {}^B C_2$  do
7        $\text{forward\_vp}i_j = \text{crossprod}(\text{line}_1, \text{line}_2)$ 
8     end
9   end
10  if  $B \geq 3$  then
11    for  $k \leq {}^B C_3$  do
12      Get  $\text{lateral\_vp}i_k$  using cross ratios as described
        in Section 3.2.2
13    end
14  end
15 end
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  $\text{dist}(l_h, \text{forward\_vp}(s)) \leq \text{threshold}_s$  &
      $\text{dist}(l_h, \text{lateral\_vp}(t)) \leq \text{threshold}_t$  then
23     record and count inlier points
24   end
25 end
```

Ali, p. 1104.

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 high-way 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.

4. Additional KSR Rationales

128. Using predictive models for Choe's lane and horizon prediction would have combined known elements according to known methods to yield predictable results. Predictive models, including convolutional neural networks (CNNs) and region-based CNNs (R-CNNs), were known and were already being used to predict lane lines and horizon lines in traffic images and cityscapes. Khan, pp. 1825-26; Ali, p. 1102; He, p. 9. Implementing these known mathematical models in Choe's calibration system, according to their known functions, would have yielded predictable results. Indeed, Choe is already designed to implement predictive models in its camera calibration system. Choe, [0030], [0037]. Khan also states that its predictive model is "simple." Khan, p. 1824.

129. Using predictive models for Choe's lane and horizon prediction would have used a known technique (predictive models) to improve similar devices in the same way. Choe and Khan/Ali identify and annotate video frames with lane lines. Choe, [0057], FIG. 6B; Khan, pp. 1823, 1826; Ali, pp. 1102-03. Both Choe and Khan/Ali compute and annotate the horizon line. Choe, [0055], FIG. 6A; Khan, pp. 1823, 1826-27; Ali, pp. 1100, 1103-04. They also determine camera parameters (e.g., calibration factors or parameters for calibrating the camera's pitch, yaw, and roll) based on those lines. Choe, [0014], [0015], [0049]; Khan, p. 1826; Ali, p. 1103. Applying Khan/Ali's predictive models to determine the lane and horizon lines, something Choe is already intended to do, would have improved Choe in the same way.

130. Using predictive models for Choe's lane and horizon prediction would have also applied a known technique (predictive models) to a known device (a manual annotator device) ready for improvement. As I just explained, Khan/Ali's predictive models and Choe's manual calibration system were known. Choe was also ready for improvement. Choe states:

In the foregoing specification, embodiments of the disclosure have been described with reference to specific exemplary embodiments thereof. It will be evident that *various modifications may be made* thereto without departing from the broader spirit and scope of the disclosure as set forth in the following claims. The specification and drawings are,

accordingly, to be regarded in an illustrative sense rather than a restrictive sense.

Choe, [0066].

131. Accordingly, a POSA would have been motivated to combine Choe and Khan/Ali to identify lane and horizon lines with predictive models.

B. Claims 1-4

132. As I explain for Ground 1, Choe discloses (or at least suggests) every element of claims 1-4. However, if Patent Owner alleges that Choe fails to adequately describe the “predictive models” as claimed, Khan in view of Ali provides this teaching.

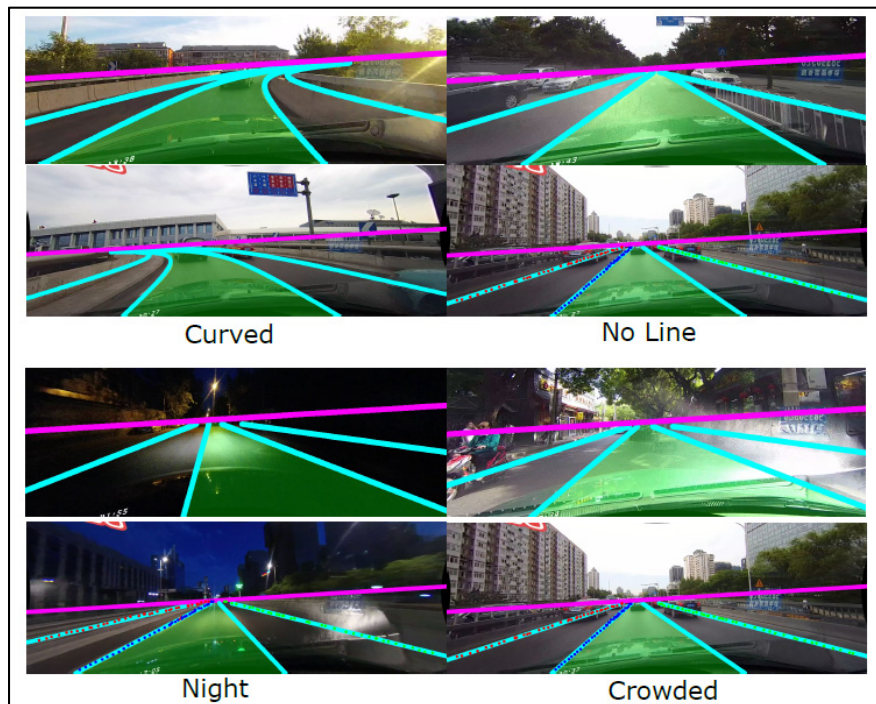
133. Khan in view of Ali “identif[ies] one or more lines in the video using a predictive model, the one or more lines including a horizon line” as recited in [1B]. In Khan, the predictive model used to identify lane and horizon lines is called a “Mask-RCNN.” *See* Khan, p. 1825. Khan did not invent this model. *See* Khan, p. 1825 (citing He). Rather, a Mask-RCNN was a well-known model that combined (i) a convolutional *backbone* architecture for feature detection and (ii) a network *head* for classification and regression. He, p. 4. A Mask-RCNN is similar to the predictive model recited in other claims of the ’580 patent (e.g., claim 8), which likewise combines a backbone network and a head.

134. Khan in view of Ali uses this predictive model to identify lane lines and horizon lines. Specifically, the Mask-RCNN predicts lane boundary markers

and then fits lines to the predicted keypoints. Khan, pp. 1825-26; *see also* Khan, p. 1828 (Algorithm 1). Khan describes: “*We fit lines to predicted keypoints and initialize lane width and camera rotation using the equidistant and parallelism properties of the road lanes.*” Khan, pp. 1825-26.

135. After lane lines are fit to the detected markers, Khan uses RANSAC (a predictive model) to predict the horizon line. *See* Khan, pp. 1826-27. This predictive model for the horizon is described more fully in Ali. Ali, p. 1104; *see also* Khan, p. 1827 (citing Ali).

136. Once Khan in view of Ali identifies the horizon and lane lines, they overlay the horizon line and/or lane lines on the video, as shown below.

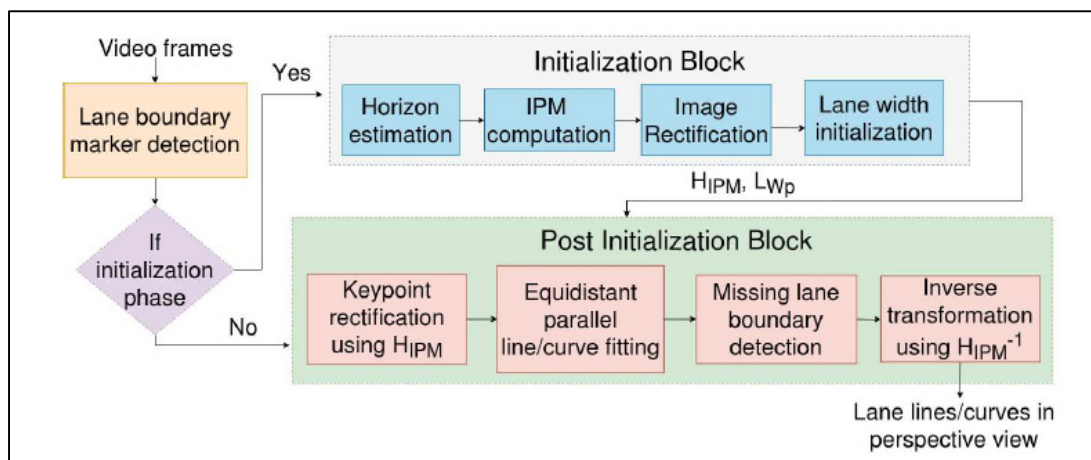


Khan, p. 1823; *see also* Khan, pp. 1824, 1828, 1829 (additional figures).



Ali, p. 1100; *see also* Ali, pp. 1103-06 (additional figures).

137. Using these lines, Khan in view of Ali then “predict[s] the at least one camera parameter using [a] predictive model” as recited in claim 3. In Khan and Ali, camera parameters are predicted during its Inverse Perspective Mapping (IPM) computation step. Khan, p. 1826; Ali, p. 1103. As Khan states: “We instead use road geometry to *find camera pose* and in-turn find the IPM.” Khan, p. 1826. Other camera parameters such as camera height are also predicted. *See* Khan, p. 1827; Ali, p. 1105.



Khan, p. 1825; *see also* Ali, p. 1102.

138. If Khan's entire end2end lane detection algorithm (e.g., Khan, FIG. 3) is "the predictive model," the IPM computation is part of "the predictive model" as recited in claim 3. One could also view the lane boundary marker detection, the horizon estimation, and IPM computation as different predictive models because all three use well-known models to predict different features. With this understanding, the lines are predicted with a first predictive model and the camera parameters are predicted with "a second predictive model" as recited in claim 4.

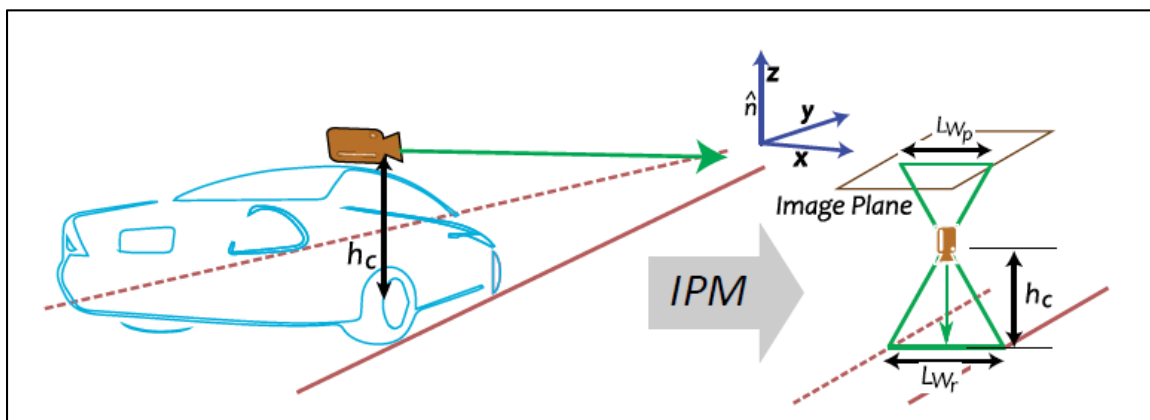
139. For these reasons, using the predictive models in Khan in view of Ali for Choe's camera calibration system would have disclosed (or at least suggested) every element of claims 1-4 of the '580 patent.

C. Dependent Claim 5

140. Claim 5 depends from claim 1 and further recites "wherein the at least one camera parameter comprises a camera height and road plane normal." Choe in view of Khan, further view of Ali discloses this element. As a POSA would have understood, Choe would compute the road plane normal and camera height as part of its calibration process because the camera's pitch, roll, and yaw are algebraically defined based on the camera height and the road plane normal. However, to the extent Choe does not disclose these elements, Khan and Ali do.

141. As I explain in Section IX.A, a POSA would have been motivated to use Khan's predictive models and Ali's horizon predictive model in Choe's camera

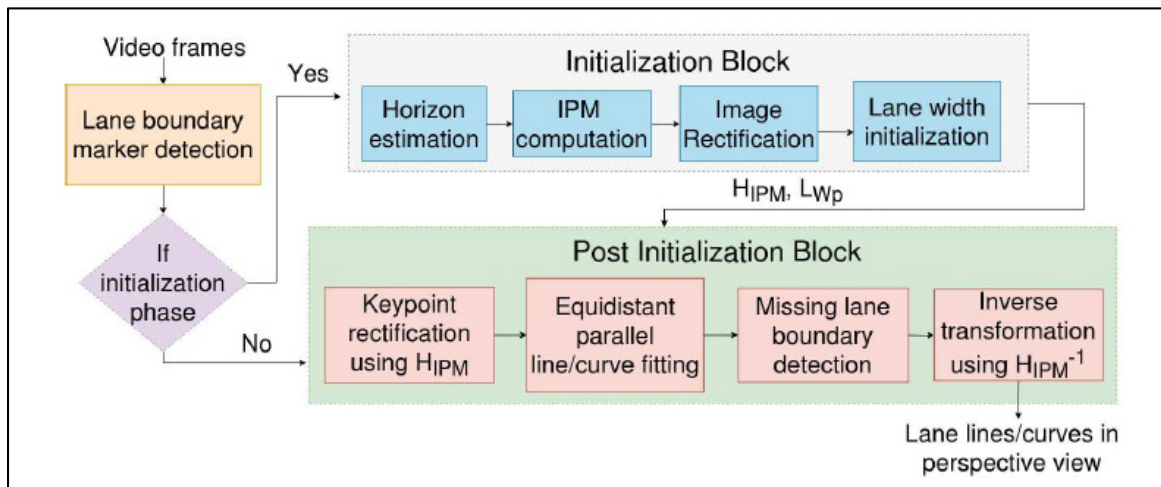
calibration system. These predictive models compute “a camera height” and “road plane normal” based on the lane and horizon lines. Shown below, the predictive model predicts the camera height (h_c) based on a lane line and the lane width. *See* Khan, p. 1826 (“ h_c is the camera height”); Ali, p. 1100 (“[I]t determines an Inverse Perspective Mapping (IPM) and camera height from a known lane width and the detected horizon.”), p. 6 (additional detail on how to calculate camera height). Also shown below, the predictive models predict the road plane normal ($\hat{n}$) based on the lines. Khan, p. 1826. Khan explains: “After estimating horizon, we compute plane normal $\hat{n}$ using camera intrinsic matrix K and horizon l_h [11] by $\hat{n} = K^T l_h$.” Khan, p. 1827; *see also* Ali, p. 1103 (“Camera initialization involves finding camera height and road plane normal.”). “The road plane normal is along the Z-axis (upwards)” Khan, p. 1826.



Khan, p. 1826; *see also* Ali, p. 1103.

142. Shown above with the arrow labeled “IPM” (Inverse Perspective Mapping), these parameters are calculated as part of Khan’s predictive model.

They are also calculated based on one or more of the lines because, as shown in the figure above and the figure below, the camera parameters are calculated during the IPM computation *after* and based on the lane boundary marker detection and horizon estimation. *See* Khan, pp. 1826-27; Ali, pp. 1103-04 (using the lane and horizon lines to determine the camera parameters).



Khan, p. 1825; *see also* Ali, p. 1102.

143. As I explain in Section IX.A, a POSA would have been motivated to use Khan's predictive models, which predict these camera parameters.

Additionally, a POSA would have been motivated to calculate the camera height and road plane normal based on the lines, as taught in Khan in view of Ali, to compensate for variability in the vehicle height, the camera height relative to the dashboard, and the camera's pitch, yaw, and roll. As a POSA would have appreciated, calculating using these camera parameters to calibrate the camera would have improved accuracy and consistency of lane detection.

144. For these reasons, a POSA would have been motivated to combine Choe, Khan, and Ali in a manner that satisfied every element of claim 5.

X. Ground 3: Choe in view of Workman renders obvious claims 1-4.

145. Choe in view of Workman discloses (or at least suggests) every element in claims 1-4. As Ground 2 explains, a POSA would have been motivated to use a predictive model in Choe's camera calibration system to identify the horizon line and lane lines. To show that predictive models were already used to identify these lines, I relied on Khan and Ali. These publications include the inventors as authors. However, the inventors were not the first to use predictive models for horizon and lane detection. As Workman shows, predictive models (e.g., CNNs) were used to identify the horizon line years prior.

A. Motivation to Combine

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.

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:

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.

Workman, p. 1; *see also* Workman, pp. 2 (reporting “significantly improve[d] performance for vanishing point based horizon line estimation”), 8-10 (experiment results showing “state-of-the-art results on all datasets”).

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

The resulting network implicitly combines both geometric and semantic cues, makes no explicit assumptions about the contents of the underlying scene, and is *several orders of magnitude faster than current state-of-the-art methods* which focus on vanishing points.

Workman, p. 1.

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.

Workman, p. 2.

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.

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.*

Workman, p. 10.

150. Using Workman's predictive model to identify the horizon line would have therefore combined known elements (CNNs and annotation for camera calibration) according to known methods to yield predictable results and used known techniques (CNNs) to improve similar devices in the same way. Additionally, because these predictive models had been so well known, a POSA would have had a reasonable expectation of success.

B. Claims 1-4

151. As I explain for Ground 1, Choe discloses (or at least suggests) every element of claims 1-4. However, if the Patent Owner alleges that Choe fails to adequately describe the "predictive model," Workman provides this teaching.

152. Workman “identif[ies] one or more lines in the video using a predictive model, the one or more lines including a horizon line” as recited in [1B]. Specifically, Workman uses a CNN to directly estimate the horizon line:

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.

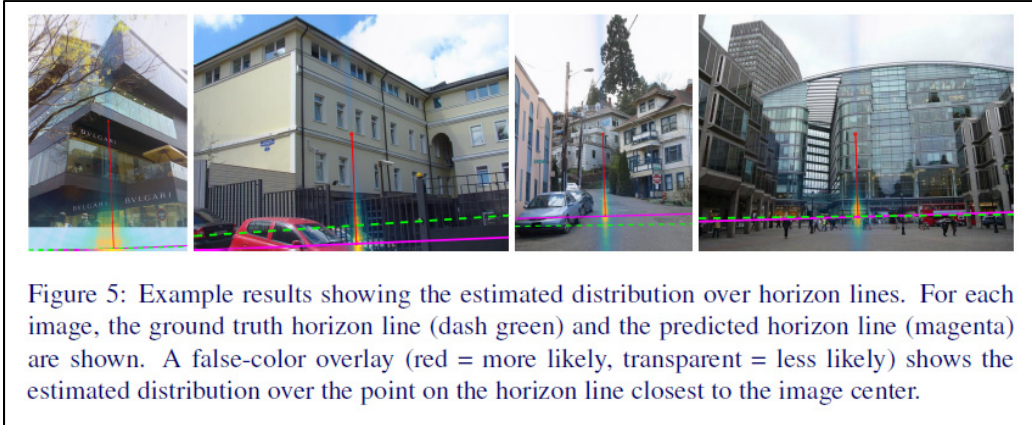
Workman, p. 1 (original emphasis removed and new emphasis added).

Instead of focusing on a particular geometric cue, we take a learning-based approach and propose to *use a deep convolutional neural network (CNN) to directly estimate the horizon line*.

Workman, p. 1.

We propose to use *convolutional neural networks (CNNs) to estimate the location of the horizon line* from raw pixel intensities. This approach is fast and does not require extensive manual tuning of parameters.

Workman, p. 6.



Workman, p. 9.

Using HLW, we investigated methods for *directly estimating the horizon line using convolutional neural networks*, including both classification and regression formulations.

Workman, p. 10.

153. When Choe uses Workman's predictive model to identify the horizon line, the combination also identifies the lane lines as recited in claim 2 because the lane lines are identified based on the identified horizon line. Choe states:

Image 402 includes one or more *lane lines 621-623 positioned based on the position of horizon line 610*.

Choe, [0056].

One or more lane lines are determined based on the first image via a perception process performed on the first image. 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.

Choe, [0015]; *see also* Choe, [0016], [0048]-[0053], [0059].

154. When using Workman's predictive model, the lane lines would therefore also be identified "using a predictive model."

155. Choe in view of Workman also discloses "predicting the at least one camera parameter using [a] predictive model" as recited in claim 3. As explained for Ground 1, Choe uses the predicted horizon line to predict camera parameters. The Choe-Workman combination likewise performs this step.

156. If Workman's algorithm and Choe's camera parameter prediction are collectively "the predictive model," the camera parameter prediction is part of "the predictive model" as recited in claim 3. One could also view the horizon line prediction and the camera parameter prediction as different predictive models because each predict different features. With this understanding, the lines are predicted with a first predictive model and the camera parameters are predicted with "a second predictive model" as recited in claim 4.

157. For these reasons, using Workman's predictive model in Choe's camera calibration system would have disclosed (or at least suggested) every element of claims 1-4 of the '580 patent.

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

158. Choe in view of Westmacot discloses (or at least suggests) every element in claims 5-7.

A. Motivation to Combine

159. A POSA would have found it obvious to combine Choe and Westmacot for at least two reasons. First, a POSA would have been motivated to compute and calibrate the camera height and road plane normal in Choe's camera calibration system. Second, a POSA would have been motivated to allow the user to add lane lines during Choe's manual calibration.

1. *How and Why* it would have been obvious to compute and calibrate the camera height and the road plane normal

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.

161. In Choe's camera calibration system, a POSA would have been motivated to calculate "a camera height and road plane normal" as taught in Westmacot because the left and right lane lines are derived from equations that require the camera height (h) and the road plane normal (n). Westmacot, 36:26-37:9, 65:1-4. For example, the ground point g_i on the road below the camera at frame i can be calculated with the following equation, where h is the height of the

camera above the road and n is the surface normal of the road relative to the camera. Westmacot, 24:24-25:2.

$$g_i = c_i + hR_i n$$

Westmacot, 24:26. The ground point g_i is then used to determine the left and right lane borders, as Westmacot's equation shows:

$$\begin{aligned} b_i^{\text{left}} &= g_i + w_i^{\text{left}} R_i r \\ b_i^{\text{right}} &= g_i + w_i^{\text{right}} R_i r \end{aligned}$$

Westmacot, 25:2-4.

162. Additionally, a POSA would have been motivated to compute the camera height and the road plane normal because these parameters are used to determine the rotational orientation of the camera. *See* Westmacot, 9:19-22. Since Choe's camera calibration requires calculating the camera's rotational orientation (pitch, yaw, and roll), a POSA would have understood that the system would need to compute these camera parameters.

163. A POSA would have also been motivated to enable the user to manually adjust the camera height during calibration to correct errors. For example, Westmacot explains:

The fully automated annotation system described produces usable annotation, however *it is still subject to a number of errors tha[t] can be fixed up with very little effort by a human annotator*. For example:

- *the calculated camera height may not be exactly correct*
- *the left/right position within the lane may not be exactly right*
- *the width of the lane may not be exactly right*

Westmacot, 47:20-26.

164. For example, “[i]n Figure 1[8], the annotator is provided with means to adjust the camera height” 34:9-10.³ This figure shows a “simple” user interface that has an option on the upper right that allows the user to adjust the camera height. *See* Westmacot, 58:6-7, FIG. 18.

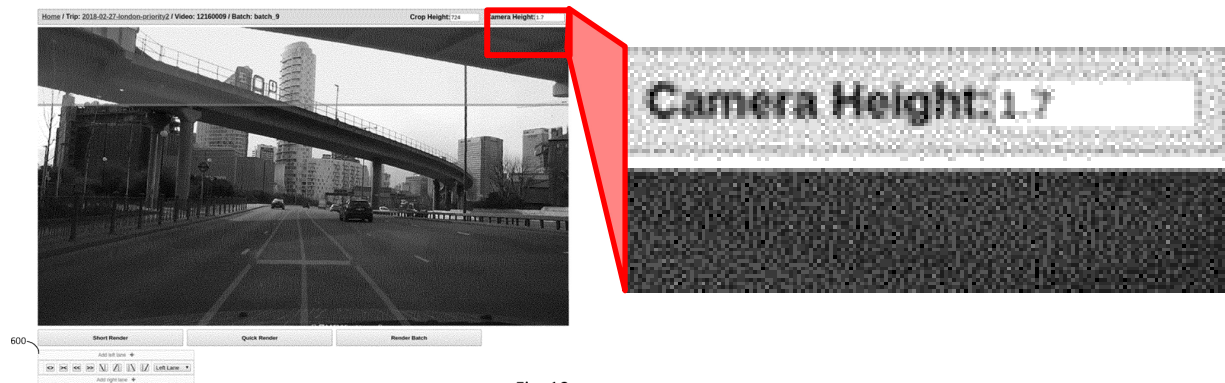


Fig. 18

Westmacot, FIG. 18 (excerpted and annotated).

³ Westmacot cites Figure 19 at 34:9-10 but meant to cite Figure 18.

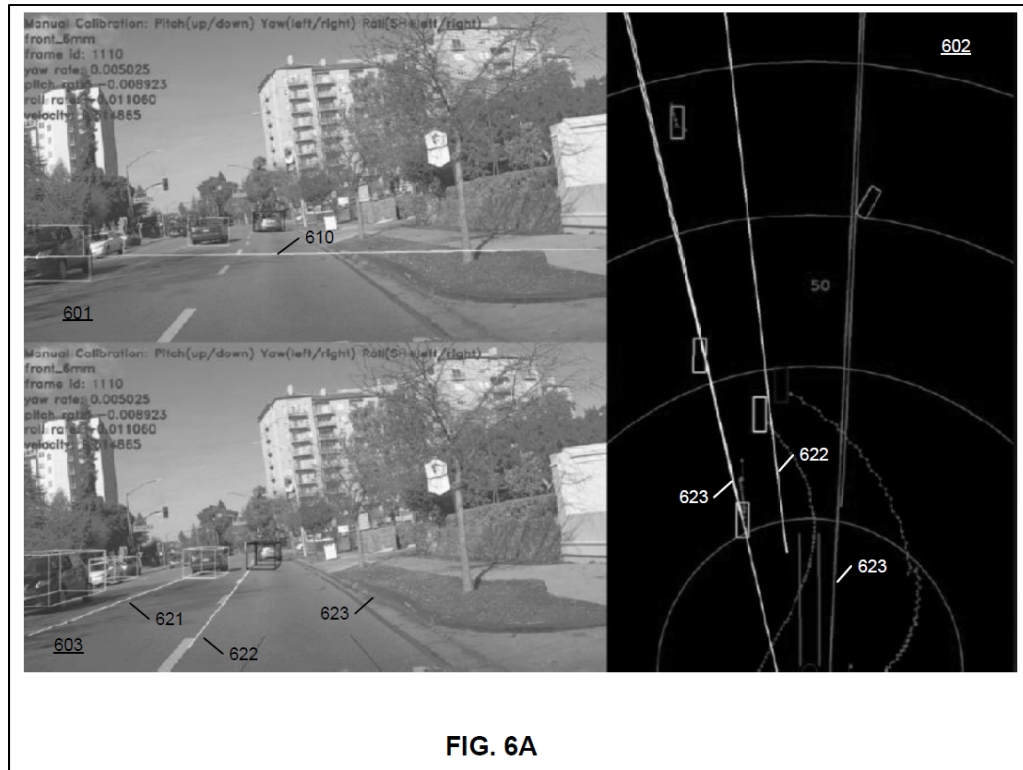
165. Computing and calibrating the camera height and the road plane normal would have combined known elements according to known methods to yield predictable results. Specifically, the camera height and road plane normal were known calculations or measurements that used known methods (e.g., multiplication and division). *See* Westmacot, 5:2-4 (equation for the road plane normal), 26:4 (“The camera height above the road ft 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.

166. Computing and calibrating the camera height and road plane normal would have used a known technique to improve similar devices (annotator devices) in the same way. Choe and Westmacot are both annotator devices. Choe, [0055], FIG. 6A; Westmacot, 33:20-25, 34:5-11. Both rely on manual annotations. Choe, [0014], [0057]; Westmacot, 33:18-34:16. Both annotate video frames with lane lines. Choe, [0057], FIG. 6B; Westmacot, 34:5-9. Both also compute and annotate the horizon line. Choe, [0055], FIG. 6A; Westmacot, 37:3-4, 44:14-18. Applying

Westmacot's option to adjust the camera height during its manual calibration would have improved Choe's annotator in the same way—by increasing the accuracy of the annotated lane lines.

2. *How and why* it would have been obvious to enable a user to add lane lines during manual calibration

167. As I explain for Ground 1, Choe discloses an annotator device that enables a user to manually adjust lane lines. For example, Figure 6A shows the annotation device's graphical user interface (GUI). Choe, [0010], [0055].



EX1006, FIG. 6A; *see also* Choe, FIGs. 6A-6B; EX1006, FIG. 6B.

168. Shown in the top left corner of the video frames in Figure 6A, the GUI instructs the user how to perform manual calibration with a keyboard:

“Manual Calibration: Pitch (up/down) Yaw (left/right) Roll (SH+left/right).”

EX1006, FIG. 6A; Choe, FIG. 6A. Choe then explains that a user may update the position of the horizon line, resulting in the position of the lane lines to be updated.

See Choe, [0015]-[0017], [0047], [0049], [0052], [0055]-[0057], FIGs. 6A-6B.

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.

Choe, [0015].

169. Although Choe allows users to adjust lane lines, I understand that the Patent Owner may allege that Choe does not explicitly allow users to *add* lane lines. A POSA would have understood that lane prediction models would not identify all the lane lines that exist in the real world. Khan, pp. 1827-28. For example, some lane lines may be faded and others may not have lane markings at all. Khan, p. 1824 (“fading color of the lane boundary over time . . . is quite common”). Without sufficient lane markings, Choe’s lane detection algorithm, in certain circumstances, would be unable to detect lane features.

170. A POSA would have therefore searched for references that disclose how to remedy the problem of missing lane lines. This would have led the POSA to Westmacot, which discloses an annotator that allows users to add missing lane lines. Westmacot states:

In the bottom-left, the annotator is provided with controls to manipulate currently rendered lanes (narrow, widen, move to the left or right, move the boundaries of the lane etc.) and *add new lanes*.

Westmacot, 34:7-9.

171. Westmacot also shows the option on the GUI to add lines in its figures, including Figure 23 (below).



Westmacot, FIG. 23 (annotated).

172. A POSA would have been motivated to have Choe's annotator device enable a user to add lane lines, as taught in Westmacot. Not only does it solve the problem of not properly identifying faded and/or missing lane lines, Westmacot's annotation is quick and efficient. Westmacot, 4:28-30, 19:17-24, 47:28-48:10, 55 (Table 1). Indeed, Westmacot's annotation requires "100 times less human effort [than past annotation methods], as measured in terms of annotation time per image." Westmacot, 19:17-24. The feature of adding lane lines also requires "very low effort to add but provide[s] significant benefit." Westmacot, 47:28-30.

173. Adding lane lines during manual calibration would have combined known elements according to known methods to yield predictable results. Specifically, manually adding a lane line would have used known methods (e.g., an annotator and overlaid annotations). *See* Choe, [0016], [0057], FIG. 6A; Westmacot, 34:5-11, 49:2-5. Using this known annotation option (new lane lines) in Choe's system would have yielded predictable results because Choe is already configured to annotate images with lane lines. Additionally, Westmacot's annotations, which include adding lane lines, are "simple." Westmacot, 48:6-7. Accordingly, a POSA would have had a reasonable expectation of success when combining Westmacot's option to add lane lines with Choe's manual calibration.

174. Adding lane lines during manual calibration would have used a known technique (a conventional GUI with an option to add a line into a video frame) to

improve similar devices (annotators) in the same way. Choe and Westmacot are both annotator devices. Choe, [0055], FIG. 6A; Westmacot, 33:20-25, 34:5-11. Both rely on manual annotations. Choe, [0014], [0057]; Westmacot, 33:18-34:16. Both also annotate video frames with lane lines. Choe, [0057], FIG. 6B; Westmacot, 34:5-9. Adding the option for the user to manually add lane lines would have improved Choe's annotator in the same way—e.g., by adding missing lane lines and/or adding bike/bus lanes.

B. Dependent Claim 5

175. Claim 5 depends from claim 1 and further recites “wherein the at least one camera parameter comprises a camera height and road plane normal.” Choe in view of Westmacot discloses this element.

176. As I explain in Ground 1 (Section VIII.A.4), Choe “comput[es] at least one camera parameter based on the one or more lines.” I also explain why a POSA would have understood that Choe would compute the road plane normal and camera height as part of its calibration process. However, to the extent the Patent Owner alleges that Choe does not explicitly disclose which parameters are required for accurate calibration, Westmacot discloses that the important camera parameters include “a camera height and road plane normal.”

177. Westmacot shows the camera height (h) and road plane normal (n) in Figure 26 (below) in relation to the left and right lane lines.

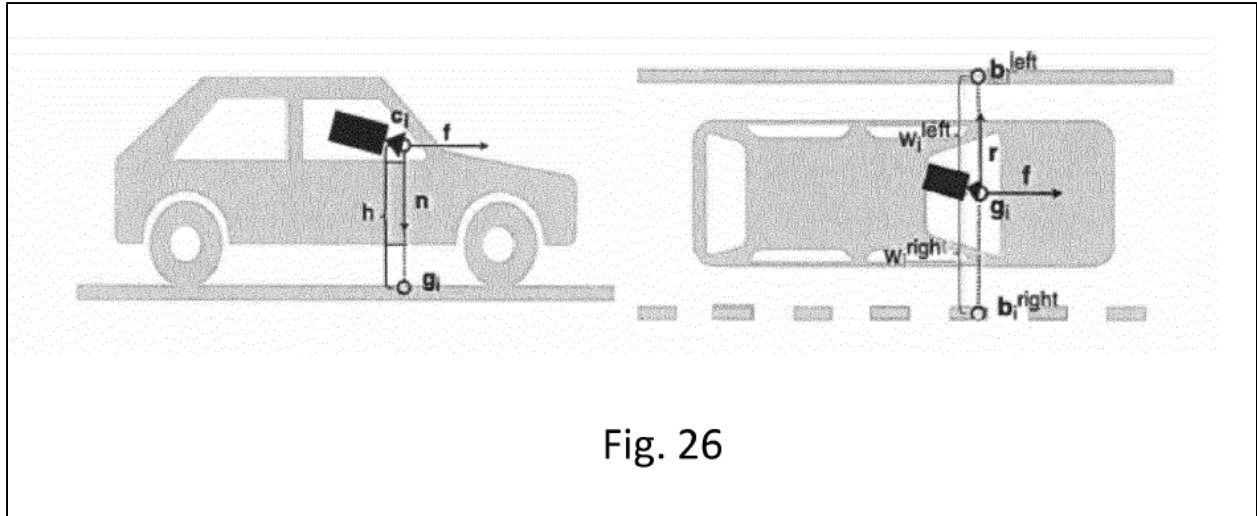


Fig. 26

Westmacot, FIG. 26.

178. As Westmacot explains, “[t]he projection may be performed based on *[the] height of the image capture device in the vehicle.*” Westmacot, 7:1-2; *see also* Westmacot, 24:4 (“*h* is the height of the camera above the road”), 6:5-7, 26:4-8, 30:24-28, 45:20-31, FIGs. 3, 3A, 26 (additional details on the camera height).

179. Westmacot also discloses that the camera parameters include the road plane normal (*n*). Westmacot states: “The expected road structure may be determined based on *a road normal vector.*” Westmacot, 5:1. The road normal vector *n* is “the surface normal of the road relative to the camera,” 25:1-2, and may be estimated from the image as:

$$\mathbf{n} = \frac{1}{\sum_{i=2}^{N-2} \|\mathbf{n}_i\|} \sum_{i=2}^{N-2} \mathbf{n}_i$$

$$\mathbf{n}_i = \mathbf{R}_i^{-1}(\mathbf{m}_{i-1,i} \otimes \mathbf{m}_{i,i+1}),$$

Westmacot, 5:2-4; *see also* Westmacot, 24:1-6, 26:9-27:5, 39:31-40:2, FIGs. 8-9, 16, 27 (additional details on the road plane normal).

180. Accordingly, Westmacot discloses that the camera parameters needed to determine the camera position include “a camera height and road plane normal.”

181. As I explain in Section XI.A, a POSA would have been motivated to combine Choe and Westmacot to compute these variables. The combination therefore also discloses this claim element.

C. Dependent Claim 6

182. Claim 6 depends from claim 1 and recites the following elements.

- 1. [6A] “receiving, over the network from the camera device, a second video comprising a second set of images”**

183. Choe discloses this element.

184. As I explain for Ground 1 (Section VIII.A.2), Choe’s sensor calibration system “receiv[es], over a network from a camera device, a video comprising a set of image frames.” Choe performs the same step with “a second video comprising a second set of images” because it calibrates the camera multiple times. Choe calibrates the camera multiple times. Choe states:

In order to accurately autonomously drive the vehicle, the system relies heavily on the sensor data obtained from the sensors such as cameras, LIDAR, and RADAR devices, etc. that are mounted on the vehicle. However, such sensors require *periodically* calibrated [sic] to

make sure the accuracy of the sensors, such that the perception system of an autonomous driving system can accurately detect and recognize the objects and lane configuration of the driving environment.

Choe, [0003].

185. Additionally, performing the same function (receiving a video, over a network, from the camera device) twice instead of once does not produce new and unexpected results. I understand that it is well settled law that mere duplication of parts has no patentable significance unless a new and unexpected result is produced. Here, repeating Choe's process would not have any new or unexpected result. It would merely perform the same function (recalibration) that Choe's system performed the first time (a calibration).

2. [6B] “transmitting the second video to the annotator device for manual review”

186. Choe discloses this element.

187. As I explain for Ground 1 (Section VIII.A.6), Choe's sensor calibration system transmits video to an annotator device for manual review. Choe performs the same step with “the second video” because it calibrates the camera multiple times. As the prior section explains, Choe performs its process multiple times because the camera must be “periodically” calibrated. Choe, [0003].

188. Additionally, performing the same function (transmitting video to an annotator device) twice instead of once does not produce new and unexpected

results. I understand that it is well settled law that mere duplication of parts has no patentable significance unless a new and unexpected result is produced. Here, repeating Choe's process would not have any new or unexpected result. It would merely perform the same function (recalibration) that Choe's system performed the first time (a calibration).

3. [6C] “adding, by the annotator device, one or more lines to the second video to generate a second overlaid video”

189. Choe in view of Westmacot discloses this element.

190. As I explain for Ground 1 (Section VIII.A.6), Choe discloses an “annotator device” that modifies or confirms lane and horizon lines in video frames of a video. However, if the Patent Owner alleges that Choe does not explicitly disclose “adding” a line to the video, Westmacot provides this teaching.

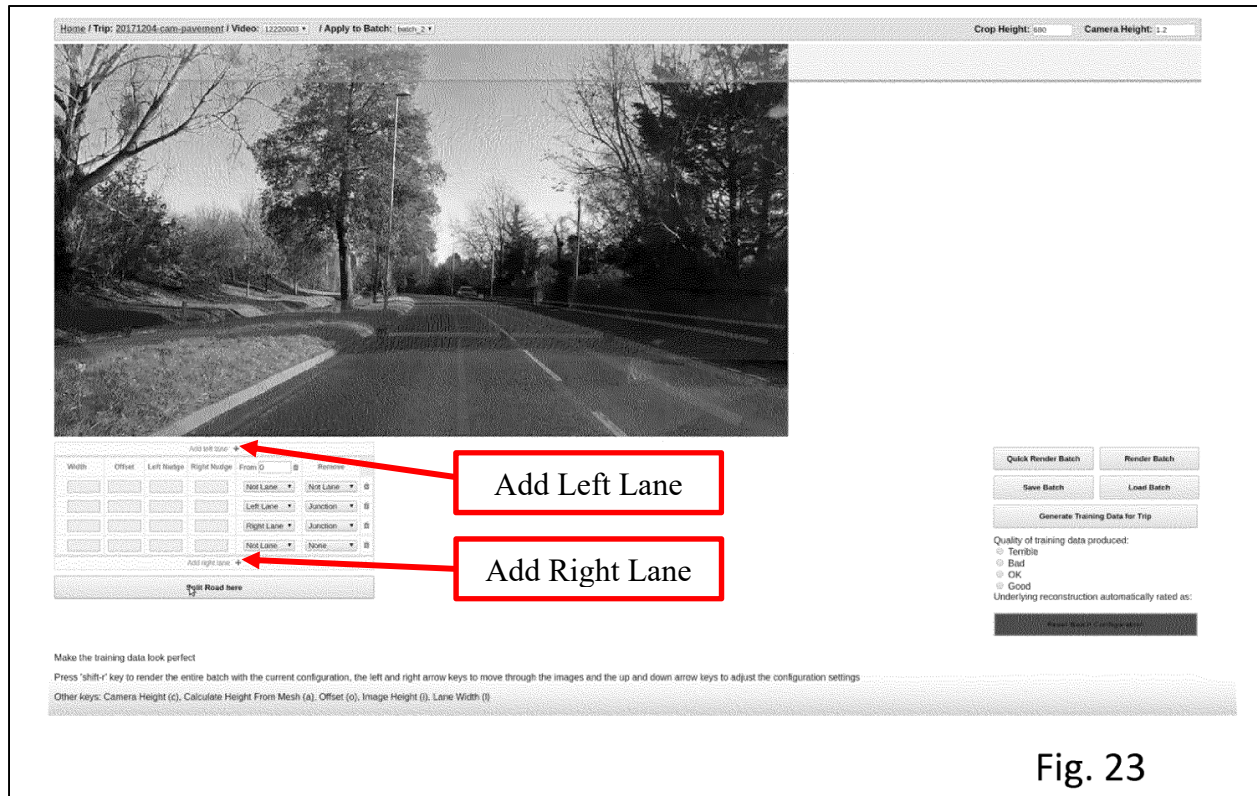
191. Westmacot discloses an “annotator device” (e.g., a computer with an annotation interface) that adds lines to a video. Specifically, “[a]n annotation interface provides a human annotator with the ability to view the frames in a batch individually with the currently rendered lanes projected into the images.”

Westmacot, 33:20-23. Westmacot further describes “a manual correction (‘human fixer’) stage” using the annotator device. Westmacot, 19:17.

192. These described images are video frames. Westmacot, 19:26-28. The annotator device can then add lane lines. Westmacot states:

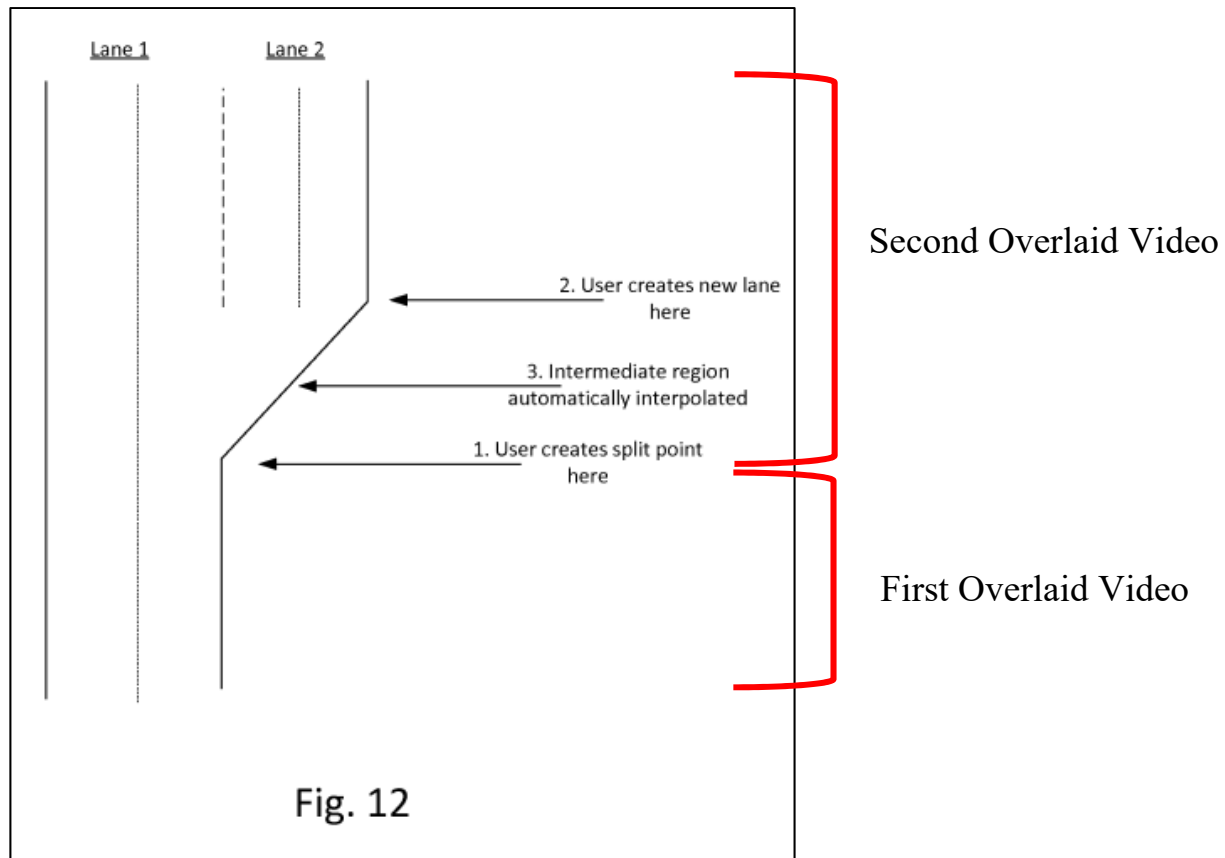
In the bottom-left, the annotator is provided with controls to manipulate currently rendered lanes (narrow, widen, move the left or right, move the boundaries of the lane etc.) and *add new lanes*.

Westmacot, 34:7-9; *see also* Westmacot, 51:25-30. The user can add lines for adjacent lanes (47:31-33), cycle lanes (52:18-21), and bus/taxi lanes (48:1-3).



Westmacot, FIG. 23 (annotated).

193. Once added, Westmacot's annotator device generates "a second overlaid video." As Westmacot explains, a user can "split" the video at the junction where a second lane line is added. *See* Westmacot, 50:11-17, 51:18-52:4. A first overlaid video includes the frames before the split point and the second overlaid video includes the frames after the split point.



Westmacot, FIG. 12 (annotated).

194. Even if Westmacot did not “split” the video, it would still disclose a “second overlaid video” because it performs its process multiple times.

Westmacot, 19:26-20:4 (annotating “batches” of video frames), FIGs. 18, 22 (two different overlay videos). The second time Westmacot performed its process would generate a “second” overlaid video. Westmacot therefore “add[s], by the annotator device, one or more lines to the second video to generate a second overlaid video.”

195. Because a POSA would have combined Choe and Westmacot to enable Choe’s annotator device to add lines to the videos (Section XI.A.2), the combination also discloses this element.

4. [6D] “predicting, using the one or more lines in the second overlaid video, at least one camera parameter”

196. Choe in view of Westmacot discloses this element. As I explain in Section XI.C.3, Choe in view of Westmacot discloses a second overlaid video. I also explain in Sections VIII.A.4 and XI.B ([1C], claim 5) how both Choe and Westmacot predict camera parameters using the lines in the overlaid video. For example, Choe states:

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.*

Choe, [0015].

At block 702, processing logic determines one or more lane lines based on the first image based on a perception process on the first image. In response to a first input signal received from an input device, at block 703, the position of the horizon line is updated based on the first input signal. At block 704, a position of at least one of the lane lines is updated based on the updated horizon line. *At block 705, processing logic determines a calibration factor or parameter for calibrating a*

pitch angle of the camera based on the difference between the initial horizon line and the updated horizon line.

Choe, [0059]; *see also* Choe, [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], FIG. 7 (support for computing camera parameters based on lines).

197. Westmacot also shows how a camera height (h) mathematically relates to the left and right lane lines. *See* Westmacot, 24:1-29:2.

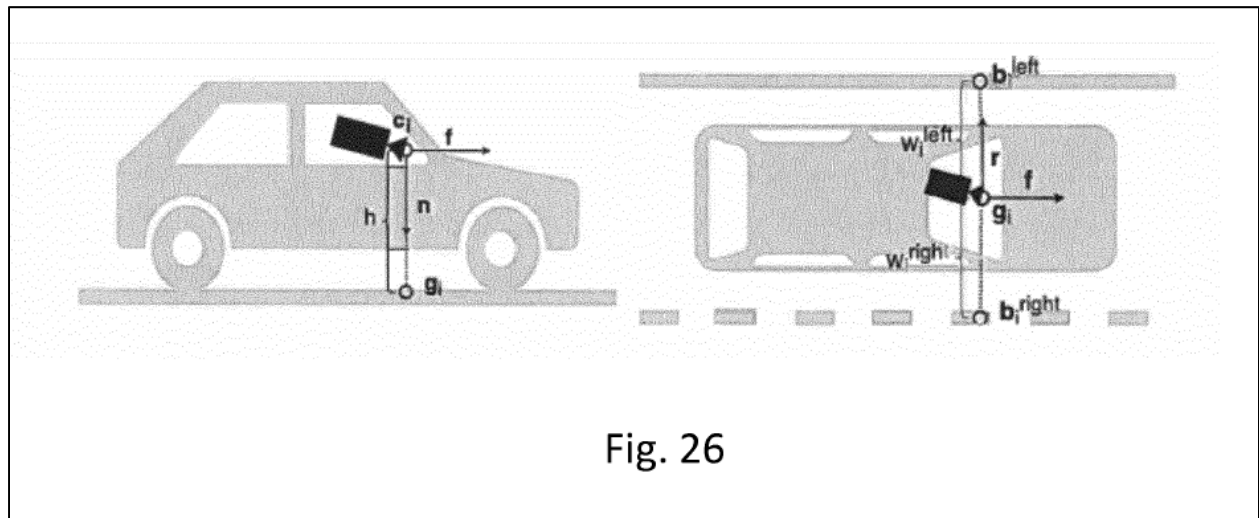


Fig. 26

Westmacot, FIG. 26.

198. Accordingly, in the Choe-Westmacot combination, Choe would calculate camera parameters (e.g., parameters for the camera’s pitch, yaw, roll, and height) based on the annotated lines, including lines added during manual calibration, as taught in Westmacot.

5. [6E] “transmitting the at least one camera parameter to the annotator device”

199. Choe in view of Westmacot discloses this element.

200. Choe “transmit[s] the at least one camera parameter to the annotator device.” As I explain for Ground 1(Section VIII.A.6), Choe’s sensor calibration module 308 “transmit[s] the overlaid video to an annotator device for manual review.” This transmission includes “the at least one camera parameter” because, as Choe shows, the GUI of the annotator device includes the camera parameters.



EX1006, FIG. 6A (excerpt, annotated); *see also* Choe, FIG. 6A-6B.

201. Westmacot also transmits camera parameters to the annotator device because, as I show below, the annotator device’s GUI includes camera height.



Fig. 18

Westmacot, FIG. 18 (excerpted and annotated).

202. Choe in view of Westmacot therefore transmits the camera parameters to an annotator device, which receives and displays them to a user.

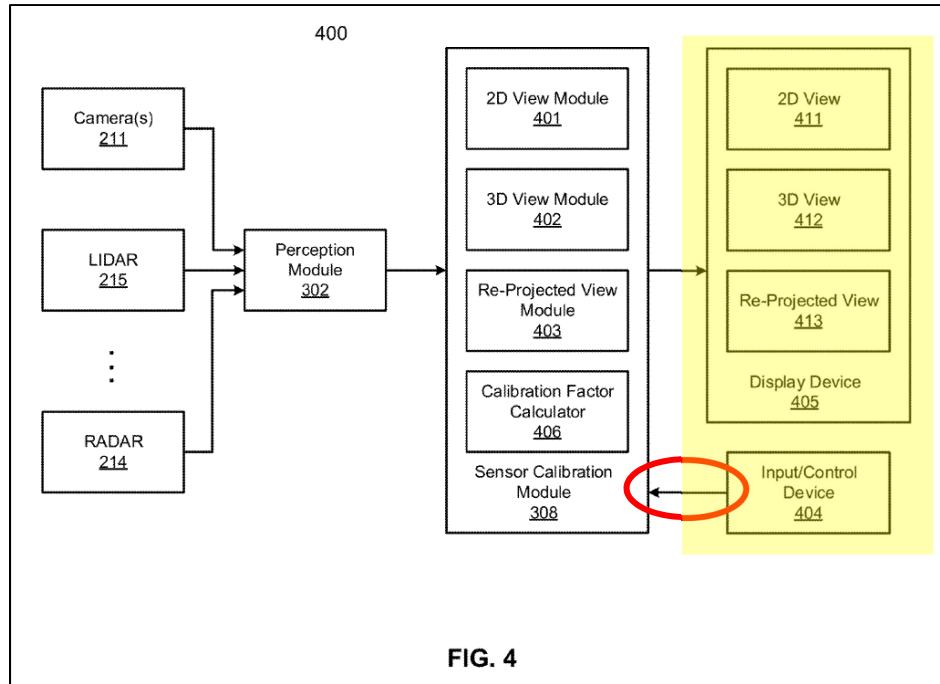
D. Dependent Claim 7

203. Claim 7 depends from claim 6 and recites the following elements.

1. **[7A] “receiving a second confirmation from the annotator device, the second confirmation indicating that the at least one camera parameter are accurate”**

204. Choe in view of Westmacot discloses this element.

205. Choe discloses this claim element. In Choe, once the annotator device receives the video with overlaid image frames, the reviewer either adjusts the lines or confirms that they are accurate. *See* Choe, [0047]. Shown below, the annotator device’s input/control device 404 sends confirmation indicating that the lines and camera parameters are accurate. *See* Choe, [0047] “[T]he user can calibrate the sensors until the calibration results as shown on display device 405 are satisfied.” Choe, [0047]. The sensor calibration module 308 receives this confirmation. *See* Choe, FIG. 4.



Choe, FIG. 4 (annotated).

206. I understand that the Patent Owner may allege that claim 7 requires two *separate* confirmations: (1) a first confirmation that confirms that the *lines* are accurate and (2) a second, separate confirmation that confirms that a *camera parameter* is accurate. Even with this understanding, Choe in view of Westmacot discloses this element.

207. As I explain for [1F], Choe discloses a first confirmation indicating that the lines are accurate. When combined with Westmacot, which computes and calibrates camera height, the combination also separately confirms that a camera parameter is accurate.

208. Shown below, during manual calibration, the user either adjusts the camera height or confirms that this parameter is accurate. Westmacot, 34:9-11.



Fig. 18

Westmacot, FIG. 18 (excerpted and annotated).

209. A POSA would have understood that this decision (either adjust or confirm) would likewise get transmitted to and received by Choe’s sensor calibration module 308 to calibrate the camera accordingly.

210. Because a POSA would have been motivated to combine Choe and Westmacot to compute and calibrate the camera height (*see* Section XI.A.1), Choe in view of Westmacot also discloses [7A].

2. [7B] “transmitting data representing the one or more lines of the second video to the camera device”

211. Choe in view of Westmacot discloses this element.

212. In Choe, after the manual calibration, “[t]he calibration result can then be uploaded onto the vehicle to be utilized online during the image processing as a part of the perception process.” Choe, [0045]. In the Choe-Westmacot combination, modifying the perceived lane lines and adding new lane lines is part of the calibration result. Choe, [0047]. In Choe’s process, the result (e.g., a new lane line) would therefore be transmitted to the camera’s perception module. *See*

Choe, FIG. 4 (showing perception module 302 associated with camera(s) 211).

This understanding is consistent with Choe's disclosure, which uses the camera's perception software to identify lane features:

The perception can include the lane configuration, traffic light signals, a relative position of another vehicle, a pedestrian, a building, crosswalk, or other traffic related signs (e.g., stop signs, yield signs), etc., for example, in a form of an object. The lane configuration includes information describing a lane or lanes, such as, for example, a shape of the lane (e.g., straight or curvature), a width of the lane, how many lanes in a road, one-way or two-way lane, merging or splitting lanes, exiting lane, etc.

Choe, [0035].

213. To the extent data representing a line is not sent to the camera, doing so would have been obvious such that a passenger viewing the camera's video feed can "visually 'feel' the calibration result based on user feedback." Choe, [0014], [047]; *see also* Choe, [0027] (the user can be a passenger).

214. For at least these reasons, Choe in view of Westmacot discloses (or at least suggests) every element of claims 5-7.

XII. Ground 5: Choe in view of Westmacot, further in view of Khan and Ali renders obvious claims 1-7.

215. As I explain for Ground 1, Choe discloses (or at least suggests) every element in claims 1-4. I also explain for Ground 4 why Choe in view of Westmacot discloses (or at least suggests) every element in claims 5-7. However, I understand

Patent Owner may argue that Choe-Westmacot does not explicitly disclose the predictive models, Khan and Ali provide this teaching.

216. To the extent the Choe-Westmacot combination does not explicitly disclose which predictive models it uses, a POSA would have searched for references that use predictive models for horizon and lane detection. As I explain for Ground 2, this would have led to Khan and Ali, which disclose predictive models for horizon and lane detection. *See* Section IX.A.

217. As I explain for Ground 2, a POSA would have been motivated to use Khan and Ali's predictive models in Choe's system. *See* Section IX.A. The same motivations to combine apply equally to the Choe-Westmacot combination.

218. I also explain for Ground 2 why applying Khan and Ali's predictive models satisfy the claim elements. Indeed, Khan and Ali are publications from the '580 patent inventors (and other authors) disclosing to the public the specific predictive models used in the '580 patent.

XIII. Ground 6: Choe in view of Westmacot, further in view of Workman renders obvious claims 1-7.

219. As I explain for Ground 1, Choe discloses (or at least suggests) every element in claims 1-4. I also explain for Ground 4 why Choe in view of Westmacot discloses (or at least suggests) every element in claims 5-7. However, I understand Patent Owner may argue that Choe-Westmacot does not explicitly disclose the predictive models. Workman provides this teaching.

220. To the extent the Choe-Westmacot combination does not explicitly disclose which predictive models it uses, a POSA would have searched for references that use predictive models for horizon and lane detection. As I explain for Ground 3, this would have led to Workman, which disclose predictive models for horizon detection. *See* Section X.A.

221. As I explain for Ground 3, a POSA would have been motivated to use Workman's predictive models in Choe's system. *See* Section X.A. The same motivations to combine apply equally to the Choe-Westmacot combination.

222. I also explain for Ground 3 why applying Workman's predictive models satisfy the claim elements. Indeed, Workman shows that the inventors were far from the first people to use predictive models to identify the horizon line in an image/video. Nor were they the first to use manual annotations to adjust the lines. These functions were well known and used throughout the industry well before the inventors disclosed their methods to the public.

XIV. No Objective Indicia of Nonobviousness

223. In reaching my opinions regarding obviousness in this declaration, I also considered whether there were any objective indicia of nonobviousness. The basis for my understanding of obviousness and objective indicia of nonobviousness is provided above in Section III.

224. I understand that any objective indicia of nonobviousness must have a nexus, or a close connection, to the novel aspects of the claims. I understand that objective indicia of nonobviousness include the following: commercial success, long-felt but unmet need, failure of others, copying, praise in the industry, and unexpectedly superior results as compared with the closest prior art.

225. Considering the information I have reviewed, I am not aware of any objective evidence that might support the patentability of the claims of the '580 patent. Further, from my review of the file history (EX1002) of the '580 patent, I did not see that the applicant raised any objective indicia of nonobviousness during prosecution. I am not otherwise aware of any publicly available evidence of objective indicia of nonobviousness.

226. Therefore, it is my opinion that no objective indicia of nonobviousness support the patentability of the claims of the '580 patent.

XV. The art cited in the Grounds are substantively different than and not cumulative to the references cited during prosecution.

227. As Section V explains, the Applicant raised five arguments during prosecution to distinguish the prior art, including Kwant1 and Kwant2. The prior-art references cited in the Grounds disclose all five features in a manner different than the prior art cited during prosecution.

228. First, the Applicant argued that the prior art cited during prosecution (Kwant1) did not predict a horizon line. EX1002, p. 145. According to the

Applicant, Kwant1 predicted *lane lines* and then used a geometric approach to identify the horizon line. EX1002, p. 145. I disagree that the claims necessarily exclude such an embodiment. Nor did the Examiner make such a finding.

However, the references in the grounds identify the horizon *directly* with a predictive model. For example, Khan/Ali identifies the horizon using RANSAC, a predictive model. *See* Ali, p. 1104; *see also* Khan, pp. 1826-27 (relying on Ali's algorithm). Workman also uses a CNN as a predictive model to directly identify the horizon line using pixel intensities rather than a geometric approach. Workman states:

Instead of focusing on a particular geometric cue, we take a learning-based approach and propose to *use a deep convolutional neural network (CNN) to directly estimate the horizon line.*

Workman, p. 1.

We propose to use convolutional neural networks (CNNs) to *estimate the location of the horizon line from raw pixel intensities.*

Workman, p. 6.

229. Second, the Applicant argued that the prior art cited during prosecution did not compute a camera parameter. EX1002, p. 146. According to the Applicant, Kwant1 did not disclose parameters related to a camera but instead disclosed “parametric representation of a single lane line.” EX1002, p. 146. In contrast, the art used in the Grounds all describe computing parameters related to

the camera itself. Choe discloses calibration factors or parameters for calibrating the camera's pitch, yaw, and roll. *See* Choe, [0014], [0015], [0049], [0052]-[0053], [0057], [0059], FIG. 7. Westmacot describes how a camera height (h) can be determined based on a mathematical relationship to the left and right lane lines. *See* Westmacot, 24:1-29:2. Khan describes how it “use[s] road geometry to *find camera pose*” and other parameters such as camera height. *See* Khan, pp. 1826-27. Similarly, Ali describes estimating the camera height. Ali, p. 1105. Each of these references describe calculating parameters related to the camera itself.

230. Third, the Applicant argued that the prior art cited during prosecution did not disclose “overlaying the one or more lines on the video to generate an overlaid video.” EX1002, p. 147. According to the Applicant, the overlay in Kwant1 was not actually overlaid but instead was merely an annotation added by a patent attorney or agent. *See* EX1002, p. 147. In contrast, the horizon and lane lines are *actually* superimposed/overlayed on the video frames of Choe, Khan, Ali, Workman, and Westmacot. *See, e.g.*, Choe, [0016]-[0017], FIGs. 6A-6B; Khan, p. 1823; Ali, FIGs. 1, 3; Workman, FIG. 5; Westmacot, FIGs. 13, 18-24.

231. Fourth, the Applicant argued that the prior art cited during prosecution did not disclose “an annotator device for manual review.” EX1002, pp. 147-48. In contrast, both Choe and Westmacot explicitly disclose an annotator to manually

adjust the predicted lane and horizon lines. *See, e.g.*, Choe, [0016], FIGs. 4, 6A-6B; Westmacot, 33:20-25, FIGs. 18-24.

232. Finally, the Applicant argued that the prior art cited during prosecution did not “receiv[e] a confirmation from the annotator device, the confirmation indicating that the one or more lines are accurate.” EX1002, p. 148. Similar to their fourth argument, the Applicant argued that no manual reviewer confirmed that predicted lines were accurate. *See* EX1002, p. 148. Choe, however, does disclose this element because “the user can calibrate the sensors until the calibration results as shown on display device 405 are satisfied.” Choe, [0047]. The sensor calibration module 308 then receives this confirmation. *See* Choe, FIG. 4. Westmacot similarly confirms that predicted lines and the camera height are accurate. *See, e.g.*, Westmacot, FIGs. 18, 22-24.

233. Accordingly, the prior-art references in the Grounds are substantively different than the references cited and overcome during prosecution.

XVI. Conclusion

In signing this declaration, I recognize that the declaration will be filed as evidence in a contested case before the Patent Trial and Appeal Board of the United States Patent and Trademark Office. I also recognize that I may be subject to cross-examination in the case and that cross-examination will take place within the United States. If cross-examination is required of me, I will appear for cross-examination within the United States during the allotted time.

I declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code.

Respectfully submitted,

A handwritten signature in black ink, appearing to read "Trevor J. Darrell", written in a cursive style.

Trevor J. Darrell, Ph.D.

Date: February 14, 2025

Appendix A

MATERIALS CONSIDERED

Exhibit	Description
1001	U.S. Patent No. 11,875,580 to Hassan <i>et al.</i> (“Hassan”)
1002	Prosecution History of U.S. Patent No. 11,875,580
1004	<i>Curriculum Vitae</i> of Dr. Trevor Darrell
1005	U.S. Publication No. 2020/0410704 to Choe <i>et al.</i> (“Choe”)
1006	Figures of U.S. Patent Application Publication No. 2020/0410704 to Choe <i>et al.</i> , <i>available at</i> the USPTO SCORE Database via Patent Center
1007	Khan <i>et al.</i> , “Lane Detection Using Lane Boundary Marker Network with Road Geometry Constraints,” <i>available at</i> https://ieeexplore.ieee.org/document/9093296 (May 14, 2020) and https://openaccess.thecvf.com/content_WACV_2020/html/Khan_Lane_detection_using_lane_boundary_marker_network_with_road_geometry_WACV_2020_paper.html (Mar. 2020) (“Khan”)
1008	Ali <i>et al.</i> , “Real-Time Vehicle Distance Estimation using Single View Geometry,” <i>available at</i> https://ieeexplore.ieee.org/document/9093634 (May 14, 2020) and https://openaccess.thecvf.com/content_WACV_2020/papers/Ali_Real-time_vehicle_distance_estimation_using_single_view_geometry_WACV_2020_paper.pdf (Mar. 2020) (“Ali”)
1009	He <i>et al.</i> , “Mask R-CNN” Version 3, <i>available at</i> https://arxiv.org/abs/1703.06870 (Jan. 24, 2018) (“He”)
1010	PCT Application Publication No. PCT/EP2019/056356 to Westmacot <i>et al.</i> (“Westmacot”)
1011	Scott Workman and Menghua Zhai and Nathan Jacobs, “Horizon Lines in the Wild,” <i>available at</i> https://arxiv.org/abs/1604.02129 (“Workman”)
1012	L. Jiao, <i>et al.</i> , “A Survey of Deep Learning-Based Object Detection,” IEEE Access vol. 7, p 128,837-128,868, 2019, <i>available at</i> https://arxiv.org/abs/1907.09408 . (“Jiao”)

Exhibit	Description
1013	B. C. Russell and A. Torralba, "Building a database of 3D scenes from user annotations," 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 2009, p 2711-2718 ("Russell")
1014	V. Bartl and A. Herout, "Fully Automatic Horizon Estimation for Surveillance Cameras," 2017 International Conference on Digital Image Computing: Techniques and Applications (DICTA), Sydney, NSW, Australia, 2017, p 1-8 ("Bartl")
1015	W. Rawat and Z. Wang, "Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review," in Neural Computation, vol. 29, no. 9, p 2352-2449, Sept. 2017 ("Rawat")
1016	Declaration of Sylvia D. Hall-Ellis, Ph.D.
1017	Sylvia D. Hall-Ellis, Ph.D. Curriculum Vitae
1018	Attachment 1A to the Hall-Ellis Declaration – MARC Record for Ali
1019	Attachment 1B to the Hall-Ellis Declaration – MARC Record for Ali
1020	Attachment 1C to the Hall-Ellis Declaration – LOC Entry for Computer Vision Congresses
1021	Attachment 1D to the Hall-Ellis Declaration – LOC Entry for Optical Pattern Recognition Congresses
1022	Attachment 1E to the Hall-Ellis Declaration – IEEE Xplore Page for Ali, <i>available at</i> https://ieeexplore.ieee.org/document/9093634
1023	Attachment 1F to the Hall-Ellis Declaration – IEEE Digital Library Page for Ali, <i>available at</i> https://www.computer.org/csdl/proceedings-article/wacv/2020/09093634/1jPbBO7T7ji
1024	Attachment 1G to the Hall-Ellis Declaration – ResearchGate Page for Ali, <i>available at</i> https://www.researchgate.net/publication/341400645_Real-time_vehicle_distance_estimation_using_single_view_geometry
1025	Attachment 1H to the Hall-Ellis Declaration – SemanticScholar Page for Ali, <i>available at</i> https://www.semanticscholar.org/paper/Real-time-vehicle-distance-estimation-using-single-Ali-Hassan/1d48563e016ef0932c49a977944fafe61c4dea77

Exhibit	Description
1026	Attachment 1I to the Hall-Ellis Declaration – SJSU Library Page for Ali, <i>available at</i> https://csu-sjsu.primo.exlibrisgroup.com/discovery/fulldisplay?docid=alma991072196389102901&context=L&vid=01CALS_SJO:01CALS_SJO&lang=en&search_scope=MyInst_and_CI&adaptor=Local%20Search%20Engine&tab=Everything&query=any,contains,2020%20IEEE%20Winter%20Conference%20on%20Applications%20of%20Computer%20Vision
1027	Attachment 1J to the Hall-Ellis Declaration – WorldCat for Ali, <i>available at</i> https://search.worldcat.org/title/1183029756
1028	Attachment 2A to the Hall-Ellis Declaration – MARC Record for He
1029	Attachment 2B to the Hall-Ellis Declaration – MARC Record for He
1030	Attachment 2C to the Hall-Ellis Declaration – IEEE Xplore Page for He, <i>available at</i> https://ieeexplore.ieee.org/document/8237584
1031	Attachment 2D to the Hall-Ellis Declaration – arXiv Page for He, <i>available at</i> https://arxiv.org/abs/1703.06870
1032	Attachment 2E to the Hall-Ellis Declaration – SJSU Library Page for Ali, <i>available at</i> https://csu-sjsu.primo.exlibrisgroup.com/discovery/fulldisplay?docid=alma991068786465502901&context=L&vid=01CALS_SJO:01CALS_SJO&lang=en&search_scope=MyInst_and_CI&adaptor=Local%20Search%20Engine&tab=Everything&query=any,contains,2017%20IEEE%20International%20Conference%20on%20Computer%20Vision%20Workshops
1033	Attachment 2F to the Hall-Ellis Declaration – WorldCat for He, <i>available at</i> https://search.worldcat.org/title/1048182166
1034	He <i>et al.</i> , “Mask R-CNN,” <i>available at</i> https://ieeexplore.ieee.org/document/8237584
1035	Attachment 3A to the Hall-Ellis Declaration – MARC Record for Khan
1036	Attachment 3B to the Hall-Ellis Declaration – MARC Record for Khan

Exhibit	Description
1037	Attachment 3C to the Hall-Ellis Declaration – IEEE Xplore Page for Khan, <i>available at</i> https://ieeexplore.ieee.org/document/9093296
1038	Attachment 3D to the Hall-Ellis Declaration – ResearchGate for Khan, <i>available at</i> https://www.researchgate.net/publication/341400531_Lane_detection_using_lane_boundary_marker_network_with_road_geometry_constraints
1039	Attachment 3E to the Hall-Ellis Declaration – SJSU Library Page for Khan, <i>available at</i> https://csu-sjsu.primo.exlibrisgroup.com/discovery/fulldisplay?docid=alma991072196389102901&context=L&vid=01CALS_SJO:01CALS_SJO&lang=en&search_scope=MyInst_and_CI&adaptor=Local%20Search%20Engine&tab=Everything&query=any,contains,2020%20IEEE%20Winter%20Conference%20on%20Applications%20of%20Computer%20Vision
1040	Attachment 3F to the Hall-Ellis Declaration – WorldCat for Khan, <i>available at</i> https://search.worldcat.org/title/1183029756
1041	Attachment 4A to the Hall-Ellis Declaration – MARC Record for Workman
1042	Attachment 4B to the Hall-Ellis Declaration – LOC Entry for Image Processing Congresses
1043	Attachment 4C to the Hall-Ellis Declaration – LOC Entry for Artificial Intelligence Congresses
1044	Attachment 4D to the Hall-Ellis Declaration – Collection of Abstracts from the British Machine Vision Conference 2016
1045	Attachment 4E to the Hall-Ellis Declaration – Proceedings of the British Machine Vision Conference 2016, <i>available at</i> https://bmva-archive.org.uk/bmvc/2016/toc.html
1046	Attachment 4F to the Hall-Ellis Declaration – Proceedings of the British Machine Vision Conference 2016 Foreword, <i>available at</i> https://bmva-archive.org.uk/bmvc/2016/index.html
1047	Attachment 4G to the Hall-Ellis Declaration – British Machine Vision Conference 2016 Page for Workman, <i>available at</i> https://bmva-archive.org.uk/bmvc/2016/papers/paper020/index.html

Exhibit	Description
1048	Attachment 4H to the Hall-Ellis Declaration – arXiv Page for Workman, <i>available at</i> https://arxiv.org/abs/1604.02129
1049	Attachment 4I to the Hall-Ellis Declaration – Multimodal Vision Research Laboratory Page for Workman, <i>available at</i> https://mvrl.cse.wustl.edu/datasets/hlw/
1050	Attachment 4J to the Hall-Ellis Declaration – ResearchGate Page for Workman, <i>available at</i> https://www.researchgate.net/publication/317193732_Horizon_Lines_in_the_Wild
1051	Attachment 4K to the Hall-Ellis Declaration – BMVC History and Description, <i>available at</i> https://www.bmva.org/bmvc
1052	Attachment 4L to the Hall-Ellis Declaration – WorldCat for Workman, <i>available at</i> https://search.worldcat.org/title/34231812
1053	Khan YouTube Video (Mar. 20, 2020), <i>available at</i> https://www.youtube.com/watch?v=Lj-FylnEboQ&t=3280s
1054	Khan YouTube Webpage (Mar. 20, 2020), <i>available at</i> https://www.youtube.com/watch?v=Lj-FylnEboQ&t=3280s
1055	Ali YouTube Video (Mar. 20, 2020), <i>available at</i> https://www.youtube.com/watch?v=0RYqfwR5YNk
1056	Ali YouTube Webpage (Mar. 20, 2020), <i>available at</i> https://www.youtube.com/watch?v=0RYqfwR5YNk
1057	U.S. Patent Application Publication No. 2019/0034740 to Kwant <i>et al.</i> (“Kwant1”)
1058	U.S. Patent Application Publication No. 2019/0102674 to Kwant <i>et al.</i> (“Kwant2”)