

ESTIMATING OBJECT PROPERTIES USING VISUAL IMAGE DATA

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application is a continuation of, and claims priority to, U.S. Patent App. No. 16/279,657 titled “ESTIMATING OBJECT PROPERTIES USING VISUAL IMAGE DATA” and filed on February 19, 2019, the disclosure of which is hereby incorporated herein by reference in its entirety.

BACKGROUND OF THE INVENTION

[0002] Autonomous driving systems typically rely on mounting numerous sensors including a collection of vision and emitting distance sensors (e.g., radar sensor, lidar sensor, ultrasonic sensor, etc.) on a vehicle. The data captured by each sensor is then gathered to help understand the vehicle's surrounding environment and to determine how to control the vehicle. Vision sensors can be used to identify objects from captured image data and emitting distance sensors can be used to determine the distance of the detected objects. Steering and speed adjustments can be applied based on detected obstacles and clear drivable paths. But as the number and types of sensors increases, so does the complexity and cost of the system. For example, emitting distance sensors such as lidar are often costly to include in a mass market vehicle. Moreover, each additional sensor increases the input bandwidth requirements for the autonomous driving system. Therefore, there exists a need to find the optimal configuration of sensors on a vehicle. The configuration should limit the total number of sensors without limiting the amount and type of data captured to accurately describe the surrounding environment and safely control the vehicle.

SUMMARY

[0002a] According to an aspect, the present disclosure refers to a system, comprising: one or more processors configured to: receive image data based on an image captured using a camera of a vehicle, the image data depicting an object within the vehicle's surrounding environment;

and utilize the image data as input to a trained machine learning model, the trained machine learning model outputting a distance from the vehicle to the object, wherein the trained machine learning model is configured to output the distance using only the image data; ~~and~~ wherein the trained machine learning model was trained using a training set, the training set comprising a plurality of images captured over a period of time and ~~a correlated~~ outputs of an emitting distance sensor corresponding to each image of the plurality of images, and wherein, for each image of the plurality of images, the output of the emitting distance sensors indicates at least one distance from the emitting distance sensor to at least one object represented by the image of the plurality of images; and a memory coupled to the one or more processors.

[0002b] According to another aspect, the present disclosure refers to a computer program product, the computer program product being embodied in a non-transitory computer readable storage medium and comprising computer instructions for: receiving image data based on an image captured using a camera of a vehicle, the image data depicting an object within the vehicle's surrounding environment; and utilizing the image data as input to a trained machine learning model, the trained machine learning model outputting a distance from the vehicle to the object, wherein the trained machine learning model is configured to output the distance using only the image data, ~~and~~ wherein the trained machine learning model was trained using a training set, the training set comprising a plurality of images captured over a period of time and ~~a correlated~~ outputs of an emitting distance sensor corresponding to each image of the plurality of images, and wherein, for each image of the plurality of images, the output of the emitting distance sensors indicates at least one distance from the emitting distance sensor to at least one object

represented by the image of the plurality of images.

[0002c] According to another aspect, the present disclosure refers to a method, comprising: receiving a selected image based on an image captured using a camera of a vehicle; receiving distance data based on an emitting distance sensor of the vehicle; identifying an object using the selected image as an input to a trained machine learning model; extracting a distance estimate of the identified object from the received distance data; creating a training image by annotating the selected image with the extracted distance estimate; training a second machine learning model to predict a distance measurement using a training data set that includes the training image; and providing the trained second machine learning model to a second vehicle equipped with a second camera.

[0002d] According to another aspect, the present disclosure refers to a method implemented by a processor included in a vehicle, the method comprising: receiving image data based on an image captured using a camera of the vehicle, the image data depicting an object within the vehicle's surrounding environment; and utilizing the image data as input to a trained machine learning model, the trained machine learning model outputting a distance from the vehicle to the object, wherein the trained machine learning model is configured to output the distance without using distance information from a first emitting distance sensor, ~~and~~ wherein the trained machine learning model was trained using a training set, the training set comprising a plurality of images captured over a period of time and ~~a correlated~~ outputs of a second emitting distance sensor corresponding to each image of the plurality of images, and wherein, for each image of the plurality of images, the output of the emitting distance sensors indicates at least one distance from the emitting distance sensor to at least one object represented by the image of the plurality of image.

[0003] One example includes a system. The system comprises one or more processors configured to: receive image data based on an image captured using a camera of a vehicle; and utilize the image data as a basis of an input to a trained machine learning model to at least in part identify a distance of an object from the vehicle; wherein the trained machine

WHAT IS CLAIMED IS:

1. A system, comprising:
one or more processors configured to:
receive image data based on an image captured using a camera of a vehicle, the image data depicting an object within the vehicle's surrounding environment; and
utilize the image data as input to a trained machine learning model, the trained machine learning model outputting a distance from the vehicle to the object, wherein the trained machine learning model is configured to output the distance using only the image data;
~~and~~ wherein the trained machine learning model was trained using a training set, the training set comprising a plurality of images captured over a period of time and a correlated outputs of an emitting distance sensor corresponding to each image of the plurality of images, and
wherein, for each image of the plurality of images, the output of the emitting distance sensors indicates at least one distance from the emitting distance sensor to at least one object represented by the image of the plurality of images; and
a memory coupled to the one or more processors.
2. The system of claim 1, wherein the trained machine learning model outputs a direction of the object with respect to the vehicle.
3. The system of claims 1-2, wherein the one or more processors are further configured to identify a velocity vector of the object with respect to the vehicle.
4. The system of claims 1-3, wherein the emitting distance sensor includes a radar sensor.
5. The system of claims 1-3, wherein the emitting distance sensor includes an ultrasonic sensor.
6. The system of claims 1-3, wherein the emitting distance sensor includes a lidar sensor.
7. The system of claims 1-6, wherein the object is one of the following: a pedestrian, a vehicle, an obstacle, a barrier, or a traffic control object.
8. The system of claims 1-7, wherein the training image and the correlated output of the

emitting distance sensor were captured on a training vehicle.

9. The system of claim 8, wherein the output of the emitting distance sensor includes an estimated distance of a second object from the training vehicle and an estimated direction of the second object with respect to the training vehicle, and wherein the second object is identified in

the training image.

10. The system of claim 9, wherein the training image is one of a time series of images captured using an image sensor of the training vehicle.

11. The system of claim 10, wherein the second object is identified at least in part by analyzing a plurality of images included in the time series of images.

12. The system of claim 10, wherein the time series of images is used to disambiguate an output portion of the emitting sensor corresponding to the second object among a plurality of objects detected by the emitting distance sensor.

13. The system of claims 1-12, wherein the identified distance of the object from the vehicle is used to predict a semantic label associated with the object.

14. The system of claims 1-13, wherein the trained machine learning model was trained to predict a distance output of the emitting distance sensor based on an input image.

15. The system of claims 1-14, wherein the distance of the object identified using the trained machine learning model is used in combination with an output of an installed emitting distance sensor included on the vehicle to determine a relative location of the object with respect to the vehicle.

16. A computer program product, the computer program product being embodied in a non-transitory computer readable storage medium and comprising computer instructions for:

receiving image data based on an image captured using a camera of a vehicle, the image data depicting an object within the vehicle's surrounding environment; and

utilizing the image data as input to a trained machine learning model, the trained machine learning model outputting a distance from the vehicle to the object, wherein the trained machine learning model is configured to output the distance using only the image data, ~~and~~ wherein the trained machine learning model was trained using a training set, the training set comprising a plurality of images captured over a period of time and a correlated-outputs of an emitting distance sensor corresponding to each image of the plurality of images, and wherein, for each image of the plurality of images, the output of the emitting distance sensors indicates at least one distance from the emitting distance sensor to at least one object

represented by the image of the plurality of images.

17. A method, comprising:
receiving a selected image based on an image captured using a camera of a vehicle;
receiving distance data based on an emitting distance sensor of the vehicle;
identifying an object using the selected image as an input to a trained machine learning

model;
extracting a distance estimate of the identified object from the received distance data;
creating a training image by annotating the selected image with the extracted distance estimate;
training a second machine learning model to predict a distance measurement using a training data set that includes the training image; and
providing the trained second machine learning model to a second vehicle equipped with a second camera.

18. The method of claim 17, wherein the captured image is part of a time series of captured images and the distance data is part of a time series of captured distance data.

19. The method of claims 17-18, further comprising tracking the identified object across the time series of captured images and wherein the extracted distance estimate is based on correlating the tracked identified object to the received distance data.

20. The method of claims 17-19, wherein the identified object is a detected vehicle or a pedestrian.

21. The system of claim 1, wherein the trained machine learning model outputs a plurality of distances corresponding to points on a bounding box associated with the object as depicted in the image data.

22. The computer program product of claim 16, wherein the trained machine learning model outputs a plurality of distances corresponding to points on a bounding box associated with the object as depicted in the image data.

23. A method implemented by a processor included in a vehicle, the method comprising:
receiving image data based on an image captured using a camera of the vehicle, the image data depicting an object within the vehicle's surrounding environment; and
utilizing the image data as input to a trained machine learning model, the trained machine learning model outputting a distance from the vehicle to the object, wherein the trained machine

learning model is configured to output the distance without using distance information from a first emitting distance sensor,

~~and~~ wherein the trained machine learning model was trained using a training set, the training set comprising a plurality of images captured over a period of time and a ~~correlated~~ outputs of a second emitting distance sensor corresponding to each image of the plurality of images, and

wherein, for each image of the plurality of images, the output of the emitting distance sensors indicates at least one distance from the emitting distance sensor to at least one object represented by the image of the plurality of image.