

Description

A vehicle positioning method based on deep neural network image recognition

5 **Technical Field**

The present invention relates to a technical field of image recognition and positioning, in particular to a vehicle positioning method based on deep neural network image recognition, and a training method of the deep neural network.

10 **Background**

Currently, before a public transportation vehicle enters a station, a driver only relies on a visual judgment to determine a distance between the vehicle and the station, which is impossible to achieve an accurate inbound route planning and speed controlling. In order to enable the vehicle to achieve the accurate inbound route planning and speed controlling before entering the station, it is necessary to accurately determine the distance between the vehicle and the station. Currently, the vehicle positioning technology mainly adopts GPS technology and high-precision map matching and positioning.

There are several problems in the use of GPS technology: when using an ordinary GPS mode for positioning, the positioning error may reach a meter level, which cannot meet the accuracy requirement of the vehicle inbound; if a RTK mode of GPS is used, it is necessary to obtain satellite information and ground reference positioning information at the same time, which needs to deploy reference positioning communication equipment along the roads, so that the equipment cost and the using cost are relatively high; when the vehicle enters a road with poor satellite signal, such as a dense forest or a tunnel, the GPS signal is easily lost, thereby losing the positioning information.

The high-precision map matching and positioning generally adopts point cloud data matching or stereo vision matching, which needs to establish map data and store it on the vehicle in advance. When the vehicle is running, the point cloud data

or image data of the current environment of the vehicle is obtained through an external lidar or camera device, and is matched with the pre-stored map data. The map production cost and the hardware and software costs of matching calculation of the positioning method are relatively high.

Therefore, a low-cost and high-precision vehicle positioning method is needed to provide a reliable data support for vehicle positioning, inbound route planning and speed controlling.

Summary

The present invention provides a vehicle positioning method based on deep neural network image recognition, and a training method of the deep neural network. The present invention can increase the training accuracy of the deep neural network by increasing the amount of training samples and optimizing the parameters of the neural network, so that the positioning accuracy of the vehicle is improved, and the required equipment cost and using cost are lower.

The first aspect of the present invention provides a training method of a deep neural network for road mark recognition, comprising the following steps:

(1) a step of road marking graph configuration: configuring a road marking graph on a road surface in an inbound direction of a station, a distance between a marking point of the road marking graph and an edge of the station in the inbound direction is L ;

(2) a step of photographing device configuration: installing a photographing device on a vehicle, an optical axis of the photographing device coincides with a longitudinal symmetry centreline of a vehicle body of the vehicle, and a distance from an optical centre of a lens of the photographing device to a ground is H ;

(3) a step of image sample collection: under different illuminating or weather conditions, using the photographing device to photograph the road marking graph, and changing an angle between the optical axis of the photographing device and the road surface, respectively in a forward direction of the vehicle and a direction that is perpendicular to the forward direction of the vehicle, thereby making the photographing device photograph one image sample of the road marking graph by

every certain angle, within a certain range of the angle between the optical axis and the road surface;

(4) a step of training sample preparation: calculating a position coordinates of the marking point of the road marking graph, in each image sample, in an image coordinate system to make a label set, and pairing each image sample with a corresponding label set to form a training sample;

(5) a step of deep neural network construction: on a basis of a deep neural network for target recognition and classification, modifying a final classification output layer of the network to an output layer composed of two nodes to output the position coordinates of the marking point of the road marking graph;

(6) a step of deep neural network training: inputting the training sample to the deep neural network for training.

Preferably, in the step of image sample collection, a photographing time is selected at a sunny noon and a sunny night.

Preferably, in the step of image sample collection, a photographing time is selected at a rainy noon and a rainy night.

Preferably, in the step of image sample collection, a photographing time is selected at a foggy noon and a foggy night.

Preferably, in the step of image sample collection, the photographing device photographs one image sample of the road marking graph by every 5° , within the range of the angle between the optical axis and the road surface from 5° to 180° .

Preferably, a lens parameter of the photographing device is selected as that when the road marking graph totally appears in a photograph frame, the road marking graph can occupy more than 20% area of the photograph frame.

Preferably, the photographing device is installed on a front roof of the vehicle and points to the forward direction of the vehicle

Preferably, the road marking graph is a triangle, or a rectangle, or an arc, or a combination of other geometric elements that are easy to recognize.

Preferably, the road marking graph is a barcode or a two-dimensional code.

Preferably, the marking point of the road marking graph is a geometric centre thereof.

Preferably, the deep neural network is a ResNet50 network, replacing a final classification output layer of the network with two fully connected layers with 1024 nodes, and an output layer with two node outputs is connected after the fully connected layers.

5 Preferably, the deep neural network is a ResNet50 network, replacing a final classification output layer of the network with two fully connected layers with 2048 nodes, and an output layer with two node outputs is connected after the fully connected layers.

10 Preferably, a floating point data output by the two nodes belong to a closed interval of $[0, 1]$, and a pixel coordinates can be obtained by multiplying the output floating point data with a corresponding image width and height.

The second aspect of the present invention provides a vehicle positioning method, which uses the above-mentioned training method of a deep neural network, comprising:

15 (1) a step of road marking graph recognition: using a deep neural network after training to recognize a road marking graph that is photographed during an actual inbound process of a vehicle, and obtain a position coordinates (u, v) of a marking point P thereof in an image coordinate system;

20 (2) a step of road marking graph positioning: calculating a coordinates (X_w, Y_w, Z_w) of the marking point P of the road marking graph in a world coordinate system by a transformation relationship between the image coordinate system and the world coordinate system, so as to obtain a distance between the marking point P of the road marking graph and the photographing device;

25 (3) a step of vehicle positioning: determining a distance between the photographing device and an edge of a station in an inbound direction, according to the obtained distance between the marking point P of the road marking graph and the photographing device, and then determining a distance between the vehicle and the edge of the station in the inbound direction in connection with an installation position of the photographing device on the vehicle.

30 Preferably, the marking point P of the road marking graph is on an optical centre axis of a lens of the photographing device;

an origin of a coordinate system of the photographing device is set at a position of a imaging aperture of the photographing device, and a horizontal distance between the optical centre of the lens of the photographing device and the marking point P of the road marking graph is Z_c ;

5 a positive direction of a Z axis of the coordinate system of the photographing device is selected as a forward direction of the vehicle, a positive direction of a Y axis of the coordinate system of the photographing device is selected as a downward direction of the vehicle, and a positive direction of a X axis of the coordinate system of the photographing device is selected as a right direction of the vehicle;

10 the world coordinate system coincides with the coordinate system of the photographing device;

an origin of the image coordinate system is on the Z axis of the coordinate system of the photographing device, and a X axis and a Y axis of the image coordinate system are respectively parallel to the X axis and the Y axis of the coordinate system of the photographing device;

15 according to a formula:

$$Z_c = \frac{f \cdot H}{v \cdot dy}$$

the horizontal distance Z_c between the photographing device and the marking point P of the road marking graph is obtained;

20 and then according to a formula:

$$L_{cz} = Z_c + L$$

a horizontal distance L_{cz} between the photographing device and the edge of the station in the inbound direction is obtained.

The advantages of the present invention are as following.

25 (1) The collection process of the image samples is carried out in a plurality of time periods under different illuminating and weather conditions, which reduces the impact of environmental factors on the training results and improves the environmental adaptability of the deep neural network.

(2) One sampled image is taken by every certain angle in a forward direction
30 of the vehicle and a direction that is perpendicular to the forward direction,

obtaining a large amount of training sample data, which can improve the training accuracy of the deep neural network, thereby improving the accuracy of subsequent vehicle positioning.

5 (3) A trained deep neural network is used to recognize the road marking graph configured in front of a station. A spatial position of an on-board photographing device is calculated by a conversion of the image coordinate system and the world coordinate system, thereby positioning the vehicle. The above-mentioned method can provide a distance data between the vehicle and the station, and provide a data support for vehicle positioning, inbound route planning, and speed controlling,
10 having the advantages of simple operation, low cost, and high reliability.

Brief Description of the Drawings

The above content and the following specific embodiments of the present invention will be better understood when read in conjunction with the
15 accompanying drawings. It should be noted that the drawings are only examples of the claimed invention. In the drawings, the same reference numerals represent the same or similar elements.

Fig. 1 is a flowchart of a training method of a deep neural network for road mark recognition;

20 Fig. 2 is a schematic diagram of photographing in the step of image sample collection;

Fig. 3 is a flowchart of a vehicle positioning method based on a deep neural network after training;

Fig. 4 is a side view of the vehicle in the process of entering a station;

25 Fig. 5 is a top view of the vehicle in the process of entering a station;

Fig. 6 is a schematic diagram of calculating the distance between the vehicle and the edge of the station in the inbound direction.

Detailed Description of Embodiments

30 The present invention will be described in further detail in conjunction with the accompanying drawings and embodiments as following.

Fig. 1 is a flowchart of a training method of a deep neural network for road marking recognition, according to the present invention, comprising a step 101 of road marking graph configuration, a step 102 of photographing device configuration, a step 103 of image sample collection, a step 104 of training sample preparation, a step 105 of deep neural network construction, and a step 106 of deep neural network training.

The step 101 of road marking graph configuration: configuring a road marking graph on a road surface of a station in an inbound direction. A distance between the marking point of the road marking graph and an edge of the station in the inbound direction is L . The road marking graph can be, but not limited to, a triangle, a rectangle, an arc, or a combination of other geometric element that is easy to recognize, or a text graph, or a bar code or two-dimensional code that incorporates relative information of the station. The marking point of the road marking graph may be a geometric centre point, a vertex or other geometric characteristic point of the road marking graph.

The step 102 of photographing device configuration: installing a photographing device on a vehicle. The lens of the photographing device points to a forward direction of the vehicle, installed on a front roof of the vehicle or other position where the road marking graph can be photographed, and making an optical axis of the photographing device coincides with a longitudinal symmetry centreline of a vehicle body of the vehicle. A distance from an optical centre of a lens of the photographing device to a ground is recorded as H . A lens parameter of the photographing device is selected as that when the road marking graph totally appears in a photograph frame, the road marking graph can occupy more than 20% area of the photograph frame. The larger the occupied area, the more accurate the positioning of the marking point of the road marking graph in the photograph frame.

The step 103 of image sample collection: under different illuminating or weather conditions, such as sunny noon and sunny night, rainy noon and rainy night, foggy noon and foggy night, using the photographing device to photograph the road marking graph. The photographing angle is shown in Fig. 2, wherein the letter A represents the road marking graph, and the angle between the optical axis of the

photographing device and the road surface is changed, respectively in the forward direction of the vehicle and the direction perpendicular to the forward direction of the vehicle, so that the photographing device photographs one image sample of the road marking graph by every 5° , within the range of the angle between its optical axis and the road surface from 5° to 180° .

The step 104 of training sample preparation: calculating a position coordinates of the marking point of the road marking graph in each image sample in an image coordinate system to make a label set, and pairing each image sample with a corresponding label set to form a training sample for subsequent input to the deep neural network for training.

The step 105 of deep neural network construction: using a deep neural network for target recognition and classification, while modifying the final classification output layer of the network to an output layer composed of two nodes. The values output by these two nodes are the coordinates of the marking point of the road marking graph in an image frame. More specifically, a ResNet50 network can be used, by removing the final classification output layer, and using two fully connected layers with 1024 nodes or 2048 nodes instead, according to the required recognition effect. An output layer with two node outputs is connected after the fully connected layers. The floating point data output by the two nodes belong to a closed interval of $[0,1]$. A pixel coordinates can be obtained by multiplying the output floating point data with a corresponding image width and height.

The step 106 of deep neural network training: inputting the above-mentioned training sample to the optimized deep neural network for training. After training, the deep neural network can be used to recognize the road marking graph and obtain the position coordinates of the geometric centre thereof.

Fig. 3 is a flowchart of a vehicle positioning method, using the above-mentioned training method of a deep neural network, comprising a step 201 of road marking graph recognition, a step 202 of road marking graph positioning, and a step 203 of vehicle positioning.

The step 201 of road marking graph recognition: using a deep neural network after training to recognize a road marking graph that is photographed during an

actual inbound process of a vehicle, and obtain a position coordinates (u, v) of a marking point P thereof in an image coordinate system.

The step 202 of road marking graph positioning: calculating a coordinates (X_w , Y_w , Z_w) of the marking point P of the road marking graph in a world coordinate system by a transformation relationship between the image coordinate system and the world coordinate system, so as to obtain a distance between the marking point P of the road marking graph and the photographing device. The transformation between the above-mentioned image coordinate system and world coordinate system can be described by a pinhole imaging model. For a point P_w with a coordinates (X_w , Y_w , Z_w) in the world coordinate system that is imaged, by a lens, to a point P_i with a coordinates (u, v) in the two-dimensional image coordinate system, the coordinates of point P_w and point P_i can be converted by using formula (1):

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{dx} & 0 & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (1)$$

in formula (1), Z_c represents a horizontal distance between the marking point P of the road marking graph and the optical centre of the lens of the photographing device; d_x , d_y , u_0 , v_0 , f are lens internal parameters related to the lens of the photographing device. Specifically:

d_x and d_y respectively represents a physical length of a unit pixel in the X direction and Y direction of the image coordinate system; u_0 and v_0 respectively represents an offset of the origins of the image coordinate system and the coordinate system of the photographing device, in the X and Y directions, in the image coordinate system; f represents the imaging focal length of the lens.

In formula (1), R represents a rotation relationship between the world coordinate system and the coordinate system of the photographing device, which is calculated by formula (2):

$$R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha \\ 0 & -\sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{bmatrix} \begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2)$$

wherein α , β and γ respectively represents the angle need to be rotated around the X axis, Y axis, and Z axis to make the world coordinate system coincides with the coordinate system of the photographing device.

In formula (1), T represents a translational relationship between the world coordinate system and the coordinate system of the photographing device, which is calculated by formula (3):

$$T = \begin{bmatrix} t_x & t_y & t_z \end{bmatrix}^T \quad (3)$$

wherein t_x , t_y and t_z respectively represents the translation amount between the world coordinate system and the coordinate system of the photographing device on the X axis, Y axis, and Z axis.

During the specific implementation, the above parameters d_x , d_y , u_0 , v_0 , f , α , β , γ , t_x , t_y , t_z can be calibrated by using, but not limited to, the following conditions.

As shown in Fig. 4 and Fig. 5, the photographing device is installed on the front roof of the vehicle and points to the forward direction. The optical centre axis of the lens of the photographing device coincides with the geometrical symmetry centreline of the longitudinal axis of the vehicle. The marking point P of the road marking graph, on the road in front of the vehicle, is on the optical centre axis of the lens of the photographing device, the horizontal distance from which to the optical centre of the lens of the photographing device is Z_C . The origin of the coordinate system of the photographing device is set at the position of the imaging aperture of the photographing device. To simplify the calculation, the world coordinate system is assumed to coincide with the coordinate system of the photographing device, the forward direction of the vehicle is selected as the positive direction of the Z axis, the downward direction of the vehicle is selected as the positive direction of the Y axis, and the right direction of the vehicle is selected as the positive direction of the X axis. The origin of the image coordinate system is on the Z axis of the coordinate system of the photographing device, and the X axis and the Y axis of the image coordinate system are respectively parallel to the X axis and the Y axis of the coordinate system of the photographing device.

According to the above conditions, it can be known that the coordinates of

point P in the world coordinate system are (X_w, Y_w, Z_w) , the coordinates thereof in the coordinate system of the photographing device are (X_c, Y_c, Z_c) , and $X_w=X_c=0$, $Y_w=H$, $Z_w=Z_c$. After pinhole imaging by the photographing device, the coordinates of the point P in the image coordinate system are (u, v) , and $u=0$, $u_0=0$, $v_0=0$. The translation parameters, between the coordinate system of the photographing device and the world coordinate system, are $t_x=t_y=t_z=0$. Therefore, formula (1) can be simplified to:

$$Z_c \begin{bmatrix} 0 \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{f}{dy} \cdot Y_w \\ Z_w \end{bmatrix} \Rightarrow Z_c = \frac{f \cdot Y_w}{v \cdot dy} \quad (4)$$

that is

$$Z_c = \frac{f \cdot H}{v \cdot dy} \quad (5)$$

The step 203 of vehicle positioning: as shown in Fig. 6, after obtaining the horizontal distance Z_c between the marking point P of the road marking graph and the optical centre of the lens of the photographing device, and combined with the distance L between the marking point P of the road marking graph and the edge of the station in the inbound direction, the horizontal distance L_{cz} between the optical centre of the lens of the photographing device and the edge of the station in the inbound direction can be calculated:

$$L_{cz} = Z_c + L \quad (6)$$

then combining the installation position of the photographing device on the vehicle, the distance between the vehicle and the edge of the station in the inbound direction is determined, thereby realizing vehicle positioning.

The terms and expressions based here are only used for description, and the present invention should not be limited to these terms and expressions. The use of these terms and expressions does not mean to exclude any equivalent features of the illustration and description (or part of them), and it should be recognized that various modifications that may exist should also be included in the scope of the

claims. Other modifications, changes and replacements may also exist. Accordingly, the claims should be regarded as covering all these equivalents.

5 Similarly, it should be pointed out that although the present invention has been described with reference to the current specific embodiments, those of ordinary skill in the art should recognize that above embodiments are only used to illustrate the present invention, and various equivalent changes or substitutions can be made without departing from the spirit of the present invention. Therefore, as long as the changes and modifications of the above-mentioned embodiments are within the scope of the essential spirit of the present invention, they will fall within the scope
10 of the claims of the present application.

Claims

1. A training method of a deep neural network for road mark recognition, wherein the method comprises:

5 a step of road marking graph configuration: configuring a road marking graph on a road surface in an inbound direction of a station, a distance between a marking point of the road marking graph and an edge of the station in the inbound direction is L ;

10 a step of photographing device configuration: installing a photographing device on a vehicle, an optical axis of the photographing device coincides with a longitudinal symmetry centreline of a vehicle body of the vehicle, and a distance from an optical centre of a lens of the photographing device to a ground is H ;

15 a step of image sample collection: under different illuminating or weather conditions, using the photographing device to photograph the road marking graph, and changing an angle between the optical axis of the photographing device and the road surface, respectively in a forward direction of the vehicle and a direction that is perpendicular to the forward direction of the vehicle, thereby making the photographing device photograph one image sample of the road marking graph by every certain angle, within a certain range of the angle between the optical axis and
20 the road surface;

a step of training sample preparation: calculating a position coordinates of the marking point of the road marking graph in each image sample in an image coordinate system to make a label set, and pairing each image sample with a corresponding label set to form a training sample;

25 a step of deep neural network construction: on a basis of a deep neural network for target recognition and classification, modifying a final classification output layer of the network to an output layer composed of two nodes to output the position coordinates of the marking point of the road marking graph;

30 a step of deep neural network training: inputting the training sample to the deep neural network for training.

2. The training method of a deep neural network for road mark recognition of claim 1, wherein in the step of image sample collection, a photographing time is selected at a sunny noon and a sunny night.

5 3. The training method of a deep neural network for road mark recognition of claim 1, wherein in the step of image sample collection, a photographing time is selected at a rainy noon and a rainy night.

10 4. The training method of a deep neural network for road mark recognition of claim 1, wherein in the step of image sample collection, a photographing time is selected at a foggy noon and a foggy night.

15 5. The training method of a deep neural network for road mark recognition of any one of claims 1~4, wherein in the step of image sample collection, the photographing device photographs one image sample of the road marking graph by every 5°, within the range of the angle between the optical axis and the road surface from 5° to 180°.

20 6. The training method of a deep neural network for road mark recognition of claim 1, wherein a lens parameter of the photographing device is selected so that when the road marking graph totally appears in a photograph frame, the road marking graph can occupy more than 20% area of the photograph frame.

25 7. The training method of a deep neural network for road mark recognition of claim 6, wherein the photographing device is installed on a front roof of the vehicle and points to the forward direction of the vehicle.

30 8. The training method of a deep neural network for road mark recognition of claim 1, wherein the road marking graph is a triangle, or a rectangle, or an arc, or a combination of other geometric element that is easy to recognize.

9. The training method of a deep neural network for road mark recognition of claim 1, wherein the road marking graph is a barcode or a two-dimensional code.

10. The training method of a deep neural network for road mark recognition of claim 1, wherein the marking point of the road marking graph is a geometric centre thereof.

11. The training method of a deep neural network for road mark recognition of claim 1, wherein the deep neural network is a ResNet50 network, replacing a final classification output layer of the network with two fully connected layers with 1024 nodes, and an output layer with two node outputs is connected after the fully connected layers.

12. The training method of a deep neural network for road mark recognition of claim 1, wherein the deep neural network is a ResNet50 network, replacing a final classification output layer of the network with two fully connected layers with 2048 nodes, and an output layer with two node outputs is connected after the fully connected layers.

13. The training method of a deep neural network for road mark recognition of any one of claims 11~12, wherein a floating point data output by the two nodes belong to a closed interval of $[0,1]$, and a pixel coordinates can be obtained by multiplying the output floating point data with a corresponding image width and height.

14. A vehicle positioning method using the training method of a deep neural network of claim 1, wherein the method comprises:

a step of road marking graph recognition: using a trained deep neural network to recognize a road marking graph that is photographed during an actual inbound

process of a vehicle, and obtaining a position coordinates (u, v) of a marking point P of the road marking graph in an image coordinate system;

a step of road marking graph positioning: calculating a coordinates (X_w , Y_w , Z_w) of the marking point P of the road marking graph in a world coordinate system by a transformation relationship between the image coordinate system and the world coordinate system, so as to obtain a distance between the marking point P of the road marking graph and the photographing device;

a step of vehicle positioning: determining a distance between the photographing device and an edge of a station in an inbound direction, according to the obtained distance between the marking point P of the road marking graph and the photographing device, and then determining a distance between the vehicle and the edge of the station in the inbound direction in connection with an installation position of the photographing device on the vehicle to.

15. The vehicle positioning method of claim 14, wherein the marking point P of the road marking graph is on an optical centre axis of a lens of the photographing device;

an origin of a coordinate system of the photographing device is set at a position of a imaging aperture of the photographing device, and a horizontal distance between the optical centre of the lens of the photographing device and the marking point P of the road marking graph is Z_C ;

a positive direction of a Z axis of the coordinate system of the photographing device is selected as a forward direction of the vehicle, a positive direction of a Y axis of the coordinate system of the photographing device is selected as a downward direction of the vehicle, and a positive direction of a X axis of the coordinate system of the photographing device is selected as a right direction of the vehicle;

the world coordinate system coincides with the coordinate system of the photographing device;

an origin of the image coordinate system is on the Z axis of the coordinate system of the photographing device, and a X axis and a Y axis of the image coordinate system are respectively parallel to the X axis and the Y axis of the

coordinate system of the photographing device;

according to a formula:

$$Z_c = \frac{f \cdot H}{v \cdot dy}$$

5 the horizontal distance Z_C between the photographing device and the marking point P of the road marking graph is obtained;

and then according to a formula:

$$L_{cz} = Z_c + L$$

a horizontal distance L_{CZ} between the photographing device and the edge of the station in the inbound direction is obtained.

10