IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received October 27, 2020, accepted November 9, 2020, date of publication November 16, 2020,
date of current version November 24, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3037692

Visibility Detection Based on the Recognition of the Preceding Vehicle's Taillight Signals

YANJUN GUO[1], QINGJIN XU[2], YANQI SU[2], AND SIYANG JIANG[2]
[1]Liaoning Provincial College of Communications, Shenyang 110122, China
[2]School of Automobile, Chang'an University, Xi'an 710064, China

Corresponding author: Qingjin Xu (qingjinxu@outlook.com)

This work was supported in part by the Natural Science Foundation of Liaoning Province under Grant 20180551142, in part by the Key Research and Development Program of Shaanxi under Grant 2020GY-163 and Grant 2019ZDLGY03-09-02, in part by the Fundamental Research Funds for the Central Universities under Grant CHD 300102220202, and in part by the Key Laboratory of Shaanxi Province for Development and Application of New Transportation Energy under Grant 300102220507.

ABSTRACT This paper proposes a method for visibility detection based on the recognition of the preceding vehicle's taillight signals using in-vehicle camera and millimeter-wave (mm-W) radar. First, we design two methods of vehicle identification. One is to use Haar-like features and an AdaBoost algorithm to train the vehicle classifier, which is mainly used to identify vehicles without turning on the taillights. The other is to identify vehicles with taillights on by means of taillight segmentation. The two identification methods are combined with a Discriminative Scale Space Tracker (DSST) to track the vehicle in the image acquired by vehicle camera and to measure anthropic visibility with mm-W radar. In addition, we drove a test vehicle on a foggy highway and collected experimental data through in-vehicle camera and mm-W radar. The experimenter observed the movement of the vehicle in front until it disappeared from the field of vision and recorded the distance of the vehicle in front measured by radar at that time as human visibility, which was also used as the ground truth to verify the accuracy of the proposed visibility detection method. The experimental results show that the visibility measured by the proposed algorithm is essentially consistent with the visibility obtained by human eyes, that is, the visibility of vehicles with no taillights, clearance lamp, emergency flasher, or fog lamp tends to rise, with an average accuracy of 88%, 91%, 90%, and 95%, respectively. In contrast to the traditional visibility measurement, this method mainly measures the maximum distance that the driver can observe when the front vehicle is not turned on or different taillights are turned on.

INDEX TERMS Visibility detection, sensor fusion, foggy highway, adaboost algorithm, target tracking, taillight signals.

I. INTRODUCTION

Visibility is the maximum horizontal distance a person with normal vision can see a target clearly from the background in current weather conditions. The study of visibility detection in foggy conditions has received considerable critical attention. Existing visibility detection methods on foggy days mainly include visual measurement and instrument measurement. Among them, the former has high requirements for surveyors and is difficult to popularize, and its accuracy cannot be guaranteed. The instrument measuring method is mainly used to monitor the atmospheric visibility with transmission or scattering measuring instruments, but both have limitations in different application fields, so they cannot be popularized. Moreover, due to the problems of high cost, short service life and high maintenance costs, it is difficult to apply the method to highway systems. Foggy days can accompany mass fog, which is a kind of fog with lower visibility that appears in a local range of tens of meters to hundreds of meters in heavy fog. The line of sight outside it is good, but inside it is hazy. Mass fog is highly regional and difficult to predict, especially on highways. It can cause sudden changes in visibility, which is extremely harmful to highway traffic safety and can easily lead to major traffic accidents, which also leads to higher requirements for real-time visibility measurement. With the popularization of vehicle-mounted sensors and the development of video image processing technology as well as the continuous improvement of video surveillance systems along highways, researchers have gradually shifted their research emphasis to video visibility detection methods. At present,

The associate editor coordinating the review of this manuscript and approving it for publication was Chao Yang.

visibility research based on image processing focuses on daytime visibility detection. MIT [1] proposed a visibility calculation method based on logo images, which compared and analyzed the images of the detection scene with the meteorological images of the predicted visibility to obtain the relative visibility. This method does not require auxiliary facilities such as flags to be set up, but using the camera is difficult, and it is easily affected by the occlusion of moving objects in the scene.

The popularization of vehicle-mounted radar provides more ideas for visibility measurement. Mori *et al.* [2] obtained the distance and azimuth of the car in front through the vehicle-mounted mm-W radar, used the azimuth to determine and intercept the vehicle position in the image obtained from the camera, and obtained the visibility by analyzing the power spectrum of the pixels of the vehicle. However, this method is susceptible to interference from vehicle taillights. Gabb [3] makes use of the fact that mm-W radar is not easily disturbed by adverse weather conditions and combines with the identification and tracking of vehicles in front of the camera to measure the visibility of the road. However, due to the lack of ground truth visibility range, the accuracy of the method must be further verified.

Machine learning and big data provide new ideas for visibility estimation. At this stage, image-based visibility estimation is based on the classic Koschmieder model [4]. By extracting image features and training, the required parameters of the model are obtained to realize visibility prediction.

Zhang *et al.* [5] focus on the visibility changes caused by factory emissions, and use the regression method of XGBoost and LightGBM fusion to train the features in the satellite image to predict the future visibility range, but it is difficult to obtain large numbers of high-precision satellite images. Kim [6] compare the visibility recognition method based on satellite image with the visibility based on human eye observation and optical measurement, and find that the method is similar to the concept of visual distance and the visibility error is about 15% without the influence of precipitation and other weather, but there are unavoidable operating errors. Based on the in-vehicle camera method, Hautiere *et al.* [7] take the lead in studying the driving assistance system that can measure visibility using the atmospheric scattering principle and V-parallax algorithm, combined with the original local comparison algorithm and improved parallax. He also proposes collecting a large number of scene images as samples for machine learning, studying the physical characteristics of scenes, and addressing the visibility by means of nonlinear data regression. Guo *et al.* [8] select the region of interest(ROI) in the image to obtain the extinction coefficient required by Koschmieder's law to calculate visibility. Through comparison and quantitative analysis, it is concluded that the method has better measurement accuracy and faster computing speed, and it is easily affected by local changes of fog density. In addition, the effectiveness of the classic Koschmieder's formula is verified for the first time through a large number of road traffic surveillance video sets [9]. The method also provides the required extinction coefficient for Koschmieder's formula to measure visibility by fitting a piecework function of observed brightness curve. On the basis of the same video, Cheng *et al.* [10] propose another visibility measurement method. This method uses spectrum features to filter out images with visibility less than 300 meters. Then, based on the relationship between total bounded variation and visibility, machine learning is used to establish a Piecewise stationary Time Series equation to estimate visibility. Both methods achieve an error rate of less than 10%, but the effect of headlights on visibility estimation is not taken into account.

Image defogging algorithm proposed by He *et al.* [11] provides a new idea for visibility estimation in foggy days. By analyzing a large number of fog-free images, it is found that in most outdoor fog-free images (excluding the sky), there are always some (at least one) pixels, and the intensity value of one or several color channels of them is very low and close to zero, which is the dark channel prior (DCP) theory. Li *et al.* [12], [13] propose Inflection point estimation (IPE) based on DCP and Transmission Refining (TR) models and apply them respectively to detect the Inflection Point and refine the transmission Map. The position of the inflection point, which changes along the vertical direction, is necessary to calculate visibility through Koschmieder model. However, it is not suitable for the image of large amount of vegetation on the road side. Ye *et al.* [14] extend the application of DCP and propose a real-time highway visibility measurement scheme, but this method needs to eliminate the interference of vehicles, road signs and other non-road objects on the road that lead to abrupt changes in environmental visibility. Qin *et al.* [15] proposed a method for estimating traffic visibility at night. This method combines DCP and image brightness contrast to analyze images and uses Laser Atmospheric Transmission Theory to estimate visibility, which has the effect of reducing estimation errors, but limited to a large and pure black target. Yang *et al.* [16] use improved DCP combined with weighted image entropy (WIE) to analyze foggy images, and train the results to obtain a visibility indicator through support vector machine (SVM). Because this method only classifies the visibility, the estimation accuracy and running speed are improved.

The improvement of computer performance once again highlights the advantages of neural network, especially in the field of image processing. Neural network models used to be discouraged by the huge amount of calculation. Now, it has been tried in image-based visibility estimation. Kwon [17] propose a detection method based on infrared video visibility. By extracting image edges of different visibility, a neural network algorithm is used to classify and convert them into corresponding visibility levels. Infrared camera has low noise, but its price is high, so it is difficult to lay densely along the road. After filtering out the low-level features in the image, Palvanov and Cho [18] use the VisNet model, which is composed of three-stream parallel convolutional

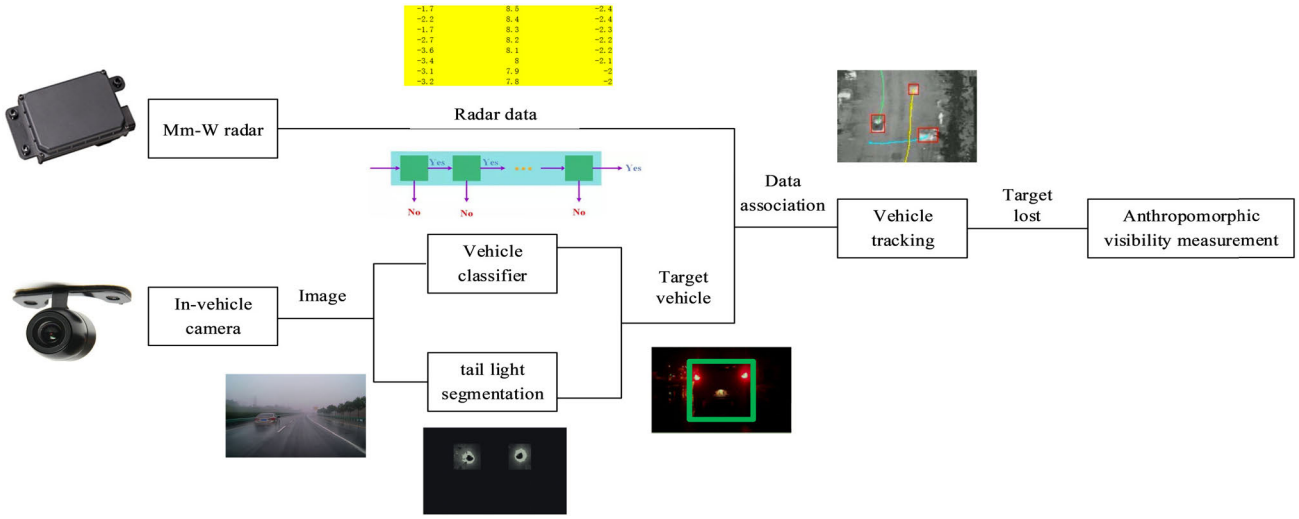

FIGURE 1. Flow chart of anthropomorphic visibility measurement.

neural network, to estimate the visibility of different scenes. Through comparison, it is found that the accuracy of this method has an advantage over the existing neural network model. Chaabani *et al.* [19] measure the visibility of the images acquired by the camera based on a neural network with a hybrid global feature descriptor. This method can be applied to both the roadside and the on-board with an accuracy rate of 90.2%. You *et al.* [20] propose a relative CNN-RNN model to simulate the visual perception visibility of human eyes, in which CNN structure is used to learn the overall visibility of the whole picture, and RNN is used to simulate the visual attention transfer from coarse to fine of human eyes. The error rate of this method is maintained at 19% without a large amount of training data, which is better than the existing image processing technology. In a word, the contradiction between precision and training amount of deep learning has always existed.

This paper introduces a method of visibility detection based on mm-W radar and camera fusion. The method flow chart is shown in Fig.1. The aim is to ensure that the visibility measured by the proposed method is consistent with the maximum distance a driver can detect in front of a vehicle in a driving situation. First, we use Haar-AdaBoosting combined with vehicle taillight segmentation to anthropomorphize the image obtained by the camera and use the DSST to track the identified vehicles. After that, the mm-W radar is used to measure the distance between the identified vehicles in real time. Under the condition that the target recognition algorithm is consistent with the human visual characteristics, the distance measured by the radar at the time at which the target vehicle disappears is recorded as the visibility.

The main contribution of this paper lies in that compared with traditional visibility estimation method for natural environment, we focus on the visibility of vehicles in front of ego vehicle from the driver's perspective, which takes into account the visibility difference caused by the vehicle in front when the vehicle is not turned on or different taillights are turned on. This is a new visibility estimation method for traffic safety. After the visibility measurement is completed, each vehicle traveling in the road network will transmit the measured visibility to other vehicles in the same road network by using the intelligent connected environment in the future, so as to achieve more targeted visibility prediction. Moreover, visibility measurement only by identifying the vehicle in front eliminates the interference brought by other information in the image, improves the accuracy of measurement and reduces the requirement of computing capability. In addition, compared with the common use of public data sets or virtual images as experimental verification data in previous studies, this paper collects foggy images by experimental vehicles and makes real-time observation by vehicle-mounted radar combined with human eye observation, providing reliable real visibility values for the experiment.

The rest of the paper is organized as follows: in Section 2, the vehicle identification algorithm is presented. In Section 3, we describe how to track the identified vehicle. We also introduce the design of the experiment and present the obtained data in Section 4. Section 5 evaluates the results and compares the data to the ground truth information. Section 6 concludes the paper.

II. VEHICLE IDENTIFICATION

A. IDENTIFICATION OF VEHICLE WITH TAILLIGHTS OFF

Although most vehicles turn on their taillights in foggy day, there are still vehicles that do not. Therefore, in order to ensure that the visibility measurement method in this paper is effective for all vehicles, it is necessary to identify vehicles whose taillights are off. Among the many vehicle recognition algorithms, we choose the recognition method based on feature data learning. This is because this method has the characteristics of small sample demand and high identification accuracy, which is especially suitable for the identification of vehicles in front only under the expressway scene in this paper. At present, there are many target recognition methods based on feature data learning, among which there are three main target image features: Histogram of Oriented Gradient

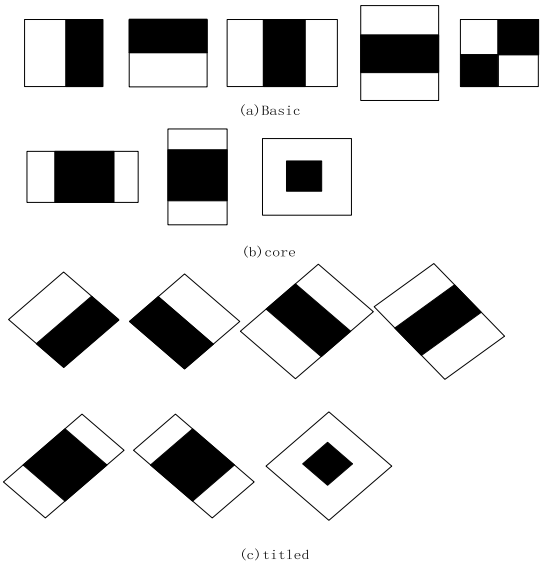

FIGURE 2. Haar-like features.

(HOG) [21], Local Binary Pattern (LBP) [22] and haar-like. Among them, haar-like features are more suitable for this particular scene in terms of both computing speed and representational capability, so we rely on Haar-like features to train a cascade classifier to identify them in this case. It should be emphasized that the vehicle detector based on Haar-like features trained in this section is mainly for the detection of vehicles with taillights off, but it still has a certain ability to identify vehicles with taillights on.

1) HAAR-LIKE FEATURES

A Haar-like feature, first proposed by Viola and Jones [23], was first applied to face representation. This is a feature that uses black and white rectangles to divide image pixels into modules to calculate the difference to reflect the grayscale changes of the image. The main reason for choosing the feature(modules) of the image rather than choosing the pixel to recognize the target lies in that features can be used as coding for specific domain knowledge, which is difficult to learn using a limited amount of training data. Haar features are divided into three categories: edge features, linear features and diagonal features, which are combined into feature templates, as shown in Fig.2. There are white and black rectangles in the feature template, and the feature value of the template is defined as the sum of white rectangle pixels and minus black rectangle pixels. The Haar eigenvalue reflects the gray level changes of the image. Lienhart *et al.* [24] Extended the initial Haar features by a set of 45° rotated features, which add additional domain knowledge to the learning framework and which is otherwise difficult to learn. These novel features can be computed rapidly at all scales in constant time.

2) ADABOOST ALGORITHM

After completing the calculation of image features, the next step needs to classify the obtained feature values. In this paper, adaboost algorithm proposed by Freund and Schapire in 1995 [25] is used to train the classifier of small feature sets. They proved that the error of the classifier trained by this method approaches zero exponentially with the increase of rounds. More importantly, the classifier can achieve large margins quickly and acquire strong generalization performance accordingly. The trained classifier is a strong classifier composed of a series of weak classifiers, which mainly performs to match and associate the weight type with the classification function. For example, the larger weight is associated with a better classification effect, and vice versa. The main form of the AdaBoost algorithm is to use a weighted majority voting method to aggregate a large number of weak classification functions. There are many kinds of AdaBoost algorithms, such as discrete AdaBoost (DAB) and real AdaBoost (RAB) [7], and Gentle AdaBoost (GAB) has been proved to be superior to other AdaBoost algorithms in experiments, because GAB has lower computational complexity and requires fewer characteristics to achieve the same performance under the same conditions. Therefore, in this paper, we choose GAB combined with Haar-like features to identify vehicles.

We define N given training data $(x_1,y_1),(x_2,y_2),\ldots,(x_n,y_n)$ with $x \in \Re^k$ and $y_i \in \{-1, 1\}$. In our case, the input K-component vector x_i is one Haar-like feature. $y_i = 1$ indicates that the input pattern contains a complete instance of the object class of interest. Each component encodes a feature relevant for the learning task at hand. The desired two-class output is encoded as -1 and $+1$. In the case of object detection, the input component is one Haar-like feature $y_i = 1$. An output of $+1$ and -1 indicates whether the input pattern does contain a complete instance of the object class of interest. The principle of classifier is described below:

(a) let $f_1,f_2,\ldots,f_T$ represent a series of weak classifiers, and the strong classifier can be expressed as

$$F(x) = E[y \mid x] = \sum_{t=1}^{T} f_t(x). \tag{1}$$

where E stands for mathematical expectation.

(b) Given the current estimate F, the estimate error is obtained

$$J(F + f) = E\left[e^{-y(F(x)+f(x))}\right]. \tag{2}$$

(c) Use adaptive Newton iteration to optimize the function J(F+f).

$$\begin{aligned} F(x) &\leftarrow F(x) + \frac{E\left[e^{-yF(x)}y\right]}{E\left[e^{-yF(x)}\right]} \\ &= F(x) + E_\omega[y|x]. \end{aligned} \tag{3}$$

where $E_\omega[y|x]$ is the probability distribution with weight.

(d) Update the weight through (4).

$$\omega \leftarrow \omega \cdot e^{-yf(x)}. \tag{4}$$

(e) After completing each stage of learning, the weak classifier f(x) can be expressed as

$$f(x) = E_\omega[y|x] = \frac{E\left[e^{-yF(x)}y\right]}{E\left[e^{-yF(x)}\right]}. \tag{5}$$

(f) In order to optimize f(x), it is necessary to make the second derivative of J(F+f) and f(x) zero at the same time, and then minimize f(x) point by point to obtain the final weak classifier.

$$\hat{f}(x) = \arg\min_{f} E_{\omega}\left[(y - f(x))^2 | x\right]. \tag{6}$$

Accordingly, strong classifier F(x) can be obtained by combining (1).

3) CASCADE OF CLASSIFIERS

According to Haar-like features, we conduct rough vehicle detection in a cascade manner. The overall form of this stage is a degenerate tree that integrates the various stages of the classifier trained by AdaBoost algorithm. Cascade classifiers are characterized by higher detection performance and lower calculation cost. In our paper, the goal of the classifier at each stage is to detect almost all vehicles while eliminating non-vehicles. A large number of negative samples is eliminated without complex processing by the initial classifier. Positive samples from the initial classifier go into subsequent classifiers to eliminate the remaining negative samples in the same way.

4) DATA TRAINING

Since the cascades classifier training program is integrated into OpenCV, we will only introduce how to use OpenCV (version 3.4.0) to train the classifier we need to identify the vehicle. First, we collected the required vehicle and non-vehicle samples in foggy environments with a sample size of 20 × 20. The samples are mainly from the videos we collected during the driving experiment on the expressway, and the experimental details can be referred to the experimental section. In this paper, the rear and sides of the vehicle are selected as positive samples, while highways or areas with no vehicles in the approximate scene are used as negative samples. In order to improve the robustness of the trained classifier, we additionally added some undetected vehicle samples and some non-vehicle samples from the public data set as supplements. Considering the impact of the number of positive and negative samples on the training performance, the number of positive and negative samples collected is 2580 and 8433 respectively.

The next step is to use opencv_createsamples.exe to generate description files in vec format from the collected positive and negative samples. In addition, negative samples need to have a txt file describing the size and path of the image.

Finally, we take the generated vec file, txt file and a series of preset parameters (picture size, number of samples, maximum number of training stages(numstages), etc.) as input, call opencv_traincascade.exe to complete the training of cascade classifier.

During the training, when falsealarm is less than the preset maxfalsealarm, each stage of training is completed. Further, when the number of training stages reaches numstages (set as 20 in this paper) or the falsealarmrate reaches maxfalsealarm$^{\text{numstages}}$, the training is completed.

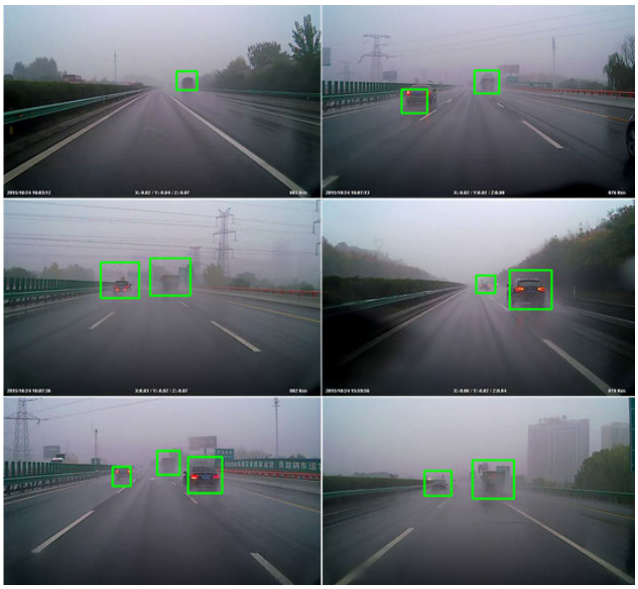

FIGURE 3. Vehicle identification results.

Among them, falsealarm is the ratio of the number of falsely detected samples in negative samples, and falsealarmrate is the accumulation of false alarms in each stage. Finally, the cascade classifier is trained to 14 stages to complete the training, whose minimum detection rate of each stage classifier was 99.75%, and highest false detection rate was 50%. Some of the detection results are shown in Fig.3.

B. IDENTIFICATION OF VEHICLE WITH TAILLIGHTS ON

1) TAILLIGHT SEGMENTATION

In foggy conditions, the driver can observe a vehicle with the lights on at a greater distance than a vehicle without the lights on. In order to ensure that the recognition method is consistent with human vision characteristics, the taillight segmentation method is considered here to supplement the aforementioned vehicle recognition method. After comparing several existing taillight segmentation algorithms, we chose a Corona segmentation algorithm proposed by Alpar [26], which has the advantages of being able to accurately segment vehicle taillights under low visibility and not being disturbed by noise. The specific implementation steps are as follows:

The first step is to convert RGB images to grayscale images, including the 345,600 pixels p_{ij}, which preserves the necessary gradient features while reducing the computational complexity.

$$G_{i,j}(i \in [1:720], j \in [1:480]). \tag{7}$$

The second is to extract the red channel in RGB images and generate images with only red pixels, because taillights are mainly red.

$$R_{i,j}(i \in [1:720], j \in [1:480]). \tag{8}$$

The red pixels of $G_{i,j}$ vary in gradient from 0 to 255, but they are all 255 on $R_{i,j}$. So, if you subtract the gray level in the red channel, you can see pixels that are not absolutely white (p_{ij} <255) and absolutely black ($p_{ij} = 0$).

$$S_{i,j}(i \in [1:720], j \in [1:480]) = R_{i,j} - G_{i,j}. \tag{9}$$

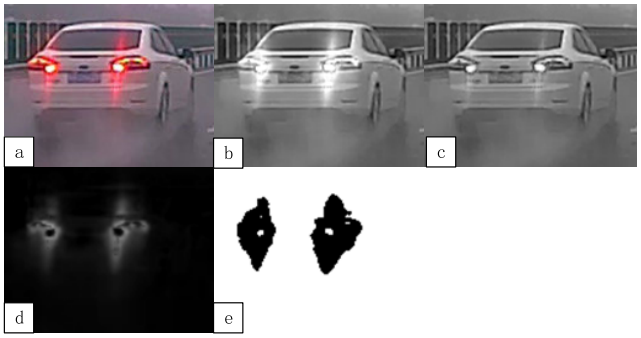

FIGURE 4. Example of the taillight segmentation algorithm steps and results. (a) Original image. (b) Grey image. (c) Red channel of original image. (d) Subtraction of grey image from red channel. (e) Pixels turned to white over threshold.

Since the gray-scale image of the taillight also contains noise, it is necessary to threshold the image $S_{i,j}$. Through the analysis of the image, it is determined that the threshold be set to 45. When the pixel is greater than the threshold, it becomes 255, and when the pixel is lower than the threshold, it becomes 0.

$$I_{i,j}(i \in [1:720], j \in [1:480]) = S_{i,j}(p_{i,j} = 255), \text{ if } p_{i,j} > 45)|_{i=1}^{720}\Big|_{j=1}^{480}. \quad (10)$$

where $p_{i,j}$ is a pixel on i^{th} column and j^{th} row and $|_{i=1}^{720}|_{j=1}^{480}$ represents the nested loop. In other words, loop through I and J twice, and if the pixel value is greater than 45, round it to 255. An example of the algorithm steps and results is shown in Fig.4.

2) TAILLIGHT MATCHING

After the vehicle taillight is extracted, because there may be multiple vehicles in the image, a certain method must be adopted to associate the left and right taillights belonging to the same vehicle. At the same time, taillight matching is helpful to eliminate the interference of residual noise in the vehicle taillight area, so as to detect vehicle position more accurately. By comparing a large number of vehicle samples, it is found that the width and height of vehicle taillights are within a certain range, and the space and position of two taillights have certain rules. According to these geometric rules, the following restrictions are set to match the extracted taillight area of the vehicle to determine whether the two taillights belong to the same vehicle.

1) The size difference between the two areas to be determined is within a certain range:

$$0 < S_1 < i, 0 < S_2 < j$$
$$|S_1 - S_2| < K \times \min(S_1, S_2). \quad (11)$$

2) The vertical distance between the center of two areas to be determined is within a certain range:

$$|y_{c1} - y_{c2}| < G \times \min(y_{c1}, y_{c2}). \quad (12)$$

3) The horizontal distance between the center of two areas is within a certain range:

$$M \times \min(x_{c1}, x_{c2}) < |x_{c1} - x_{c2}| < N \times \min(x_{c1}, x_{c2}). \quad (13)$$

FIGURE 5. The final effect of identifying the vehicle through taillight segmentation.

In formula (5-7), S_1,S_2 are the area of two taillights to be paired in pixels; x_{c1},x_{c2},y_{c1},y_{c2} are the distance of pixels in the x- and y-axis at the center of the area; i and j are the maximum area; K is the area proportional coefficient; and G, M, and N are the proportional coefficients of distance. The coefficients in the formula need to be set according to prior knowledge in practice. The prior values used in this paper are i = 1000, j = 1000, K = 2, G = 0.35, M = 1, and N = 7. The final effect of identifying the vehicle through taillight segmentation is shown in Fig.5.

III. VEHICLE TRACKING

A. EXHAUSTIVE SCALE SPACE TRACKING

Target tracking algorithms can be roughly divided into two types: generative and discriminant. The main idea of generative algorithm is to model the target, and use the model to match with the search area of the next frame of image. The highest matching degree is target area. Common models are the Markov Model (MM) and the Gaussian Mixed Model (GMM). However, the simple idea of model building does not distinguish between target and background information, so the accuracy of the model is greatly affected by background clutter. Moreover, such algorithms are generally not as fast as discriminant algorithms. The main idea of the discriminant algorithm is to solve the target tracking problem as a classification problem [27]. By constructing affine transformation, cyclic displacement, window translation, and so on [28]–[32], the target negative sample is constructed, and then the classifier (tracker) is obtained by learning the positive and negative samples. The learning strategy of the classifier determines the stability of the tracker. Since the classification problem can be solved by machine learning, the discrimination-based tracking algorithm is generally more accurate than the generative method. Classical discriminant method Minimum Output Sum of Squared Error (MOSSE) was proposed by Bolme *et al.* [32]. Because of the idea of the proposed correlation operation, the learning process of classifiers could be quickly realized in the Fourier domain, thus achieving an operation speed of more than 300 frames. Later, Henriques *et al.* [28] proposed Kernel Correlation Filter (KCF) on the basis of MOSSE, and used cyclic matrix to replace the radiation transformation of MOSSE to achieve intensive sampling of training samples. The target feature also changes from single-channel grayscale feature to multi-channel HOG feature, which enables the tracker to have more stable tracking performance in complex

environments while maintaining the computational speed of over 100 frames. Therefore, the DSST is considered here for vehicle tracking. It is an improvement based on the classical discriminant methods MOSSE and KCF, which mainly introduce the multi-feature fusion mechanism and scale estimation [33] and its superior performance meets the requirements of accuracy and real-time performance. The DSST is mainly realized through the following steps. First, the HOG feature is used to translate the filter and connect it with the usual image intensity feature. Then, one-dimensional (1D), two-dimensional (2D), and three-dimensional (3D) filters are used to estimate the scale, translation, and exhaustive scale-space localization of the target, respectively.

We represent a signal (e.g., an image) in a D-dimensional feature map and extract the rectangular patch of a target from the feature map and set it as f. We denote the feature dimension number l∈{1,...,d} of f by J. An optimal correlation filter h is found by minimizing the cost function (8). The optimal filter exists on the per feature dimension h^l.

$$\varepsilon = \left\| \sum_{l=1}^{d} h^l * f^l - g \right\|^2 + \lambda \sum_{l=1}^{d} \left\| h^l \right\|^2. \tag{14}$$

where g refers to the expected correlation output associated with the training example f. The parameter $\lambda \geq 0$ is used to limit the regularization term. Equation (8) can be solved by the following formula:

$$H^l = \frac{\bar{G}F^l}{\sum_{k=1}^{d} \bar{F}^k F^k + \lambda}. \tag{15}$$

The spectrum of f is limited by regularization, as a result of which some unimportant or negligible features are reduced to 0 to reduce the complexity of the model and effectively avoid overfitting [32]. An optimal filter can be obtained by minimizing the output error over all training patches [34], [35]. However, this requires a complicated calculation process of d × d linear system per pixel. To solve this problem, a robust approximation of the update method of the numerator A_t^l and denominator B_t of the correlation filter H_t^l was designed in equation (9).

$$A_t^l = (1-\eta)A_{t-1}^l + \eta \bar{G}_t F_t^l. \tag{16}$$

$$B_t = (1-\eta)B_{t-1} + \eta \sum_{k=1}^{d} \overline{F_t^k} F_t^k. \tag{17}$$

where η is a learning rate parameter. The new target state is obtained by maximizing the score y. The correlation score y of the feature map at the rectangular region z is thus calculated using equation (12).

$$y = \mathcal{F}^{-1} \left\{ \frac{\sum_{l=1}^{d} \overline{A^l} Z^l}{B + \lambda} \right\}. \tag{18}$$

Here, a tracking method based on the joint translation scale of a learning 3D scale spatial correlation filter is adopted.

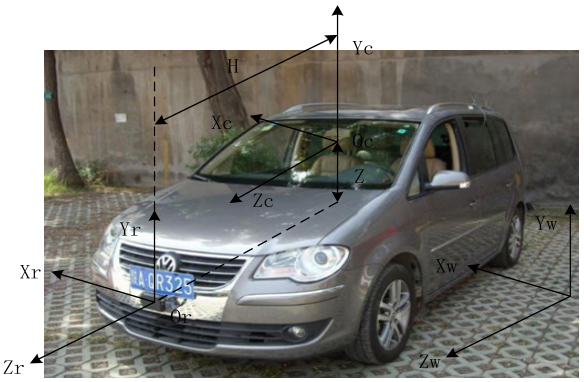

FIGURE 6. Vehicle coordinate system.

The size of the filter is fixed as M × N × S, where M, N, and S are the height and width of the filter and the number of scales, respectively. In order to update the filter, we first calculate a feature pyramid in the rectangular area around the target so that the estimated scale of the target is M × N. The training sample f is then set as a rectangular cuboid with a size of M × N × S in the feature pyramid, and the center is the estimated position and scale of the target. A 3D Gaussian function is then used as the corresponding expected correlation output g. Finally, equations (10) and (11) are used to update the scale space tracking filter. In order to locate the target in the new framework, we extract an M × N × S cuboid z from the feature pyramid, as shown above. The cuboid is centered on the predicted target location and scale. Equation (12) is then used to calculate the correlation score y and get the new target position and scale by finding the maximum score in y.

B. JOINT CALIBRATION OF LIDAR AND CAMERA

Calibration is the preparation condition for multi-sensor information fusion. Different sensors have independent coordinate systems and different acquisition frequencies, so data from different coordinate systems must be converted to the same coordinate system and time registration to achieve fusion. For the joint calibration of mm-W radar and camera, the purpose is to obtain the conversion relationship between mm-W radar and camera data, that is, to find the corresponding pixel points in the target azimuth data and image measured by radar at the same time. Due to the fact that the camera capture frequency is higher than the laser radar, the camera and laser radar image data point cloud data is a synchronized timestamp. Therefore, this paper adopts the method of time nearest neighbor matching in the time registration to find the image data with the smallest time interval with each frame of lidar data for processing,, so as to realize the time registration of lidar and camera.

Next, coordinate transformation is undertaken. The world coordinate system $O_0 - X_0Y_0Z_0$ is established as shown in Fig. 6, and the coordinates are expressed as (X,Y,Z). mm-W radar is installed at the O_0 point, where $X_0O_0Z_0$ is the detection plane and the detection target position is (r, α). There are three coordinate systems in the camera system: pixel coordinate system, image coordinate system,

and camera coordinate system. The coordinates of image coordinate system are expressed as (x,y), where the origin is the main point of the camera, that is, the intersection of the camera optical axis and the image plane, the x axis is parallel to the U axis, and the Y axis is parallel to the V axis. The camera coordinate system can describe the relative position of the object and the camera. The coordinates are expressed as (X_C, Y_C, Z_C), where the origin is the O point of the camera optical center, the XC axis is parallel to the X axis, the Y_C axis is parallel to the Y axis, and the Z_C axis is parallel to the camera optical axis and perpendicular to the image plane.

The coordinate transformation process can be generally divided into three steps:

1) From a millimeter wave radar coordinate system to a world coordinate system, Z_0 is the distance between plane $X_rO_rZ_r$ and plane XO_0 Y and H is the distance between plane XO_0 Y and plane $X_rO_rZ_r$.

$$\begin{cases} X = r \times \sin\alpha \\ Y = -H \\ Z = Z_0 + r \times \cos\alpha. \end{cases} \tag{19}$$

2) The transformation from the world coordinate system to the camera coordinate system can be represented by the rotation matrix R and the translation matrix T, where R is a matrix with a size of 3×3, representing the rotation of spatial coordinates, and T is a matrix of size 3×1, which represents a shift in space coordinates.

$$\begin{bmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{bmatrix} = \begin{bmatrix} R & T \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}. \tag{20}$$

3) The transformation from camera coordinate system to image coordinate system is a process of transformation from a 3D coordinate system to a 2D coordinate system, which belongs to a perspective projection relationship and satisfies the similarity theorem of triangles, where f is the focal length of the camera.

$$Z_C \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{bmatrix}. \tag{21}$$

4) The transformation from an image coordinate system to a pixel coordinate system mainly involves a flex and shift transformation.

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{dx} & 0 & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}. \tag{22}$$

According to the equation (13-16), the conversion relation between world coordinates and image pixel coordinates can be determined.

FIGURE 7. Experiment roadmap.

IV. EXPERIMENTAL DESIGN AND DATA ACQUISITION

A. EXPERIMENTAL REQUIREMENTS

This experiment requires data relevant for vehicle identification in conditions of rain and fog to be obtained and analyzed, so strict requirements were imposed on the experimental environment, which had to include poor visibility in heavy rain and fog. The visibility range had to be between 150 and 200 m, which is considered to pose a significant driving risk.

This experiment aimed to analyze the maximum distance at which a car in front of the observer could be recognized in conditions of poor visibility. This required that there be no shielding between the two vehicles, which required the traffic flow on the highway of the experiment to be small.

The experiment was completed using cameras and radar. To shoot clear objects in a low visibility environment, digital High Definition (HD) cameras were required, and the radar measurement distance was more than 200 m.

B. EXPERIMENTAL SITE

According to the requirements of this experiment, the locations of the experiment were the Xi'an Expressway and the Shanghai–Shaanxi Expressway. The city of Xi'an is located in the territory of Xi'an City, and the Xi'an Ring expressway – G3001 National expressway and Shanghai–Shaanxi expressway – G40 national expressway are two-way six lane highways, as shown in Fig.7. The two expressways are characterized by low traffic flow, high speeds, high construction standards, complete traffic facilities, and large traffic capacities and thus ensure fewer chances for the camera to be blocked in the shooting process and less interference in the radar measurement distance, which is convenient for the later data search and analysis.

C. INTRODUCTION OF EXPERIMENTAL EQUIPMENT

A TOURAN car, as shown in Fig.8, was selected for the experiment and equipped with a digital HD camera and Delphi Electronically Canning Radar (ESR). The camera was mounted on the roof of the test car. The scene ahead of the vehicle was filmed through the windshield, and video data was recorded and stored approximately every minute.

FIGURE 8. Experimental equipment. (a) Camera installation position. (b) mm-W radar. (c) mm-W radar measurement range. (d) mm-W radar installation position. (e) TOURAN car.

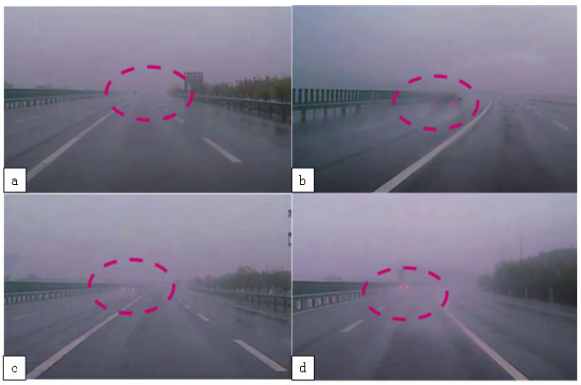

FIGURE 9. Different lighting effects on the car in front of the camera. (a) No taillights. (b) Clearance lamp. (c) Emergency flasher. (d) Fog lamp.

The Delphi ESR radar is installed in the front bumper of the experimental vehicle to ensure that there is no object blocking the front. The data acquisition system collects the distance and Angle data from the radar through the Controller Area Network (CAN) bus, which has a high measurement accuracy and good stability performance, and can carry out long-term continuous and stable work in a vehicle-mounted environment. At the same time, the radar has a high-level protection function and can operate normally under various weather conditions and road conditions.

D. DATA ACQUISITION AND PROCESSING

In the experiment, the vehicle-mounted camera simulated shooting and recording the front scene, and the vehicle-mounted radar measured and recorded the distance, relative speed, and angle from the obstacles and vehicles in front to the experimental vehicle. The time of the data recorded by the camera and the data recorded by the radar was synchronous. The test car is running normally on the middle or right lane of the expressway. When the vehicle on the left overtakes the test car and continues to move away from the test car, there will be a critical state where the vehicle can be recognized and cannot be recognized. The driver in the car reported this moment, and the co-pilot staff recorded this moment. Taking the time recorded by the staff as the combining point, the video was watched to find the scene when the vehicle in front was in the critical state that could be recognized or not recognized under four states: no lights, taillights, double flash, and fog lights, and the corresponding data was found in the radar data.

In the above experimental method, the visual acuity of the driver in the experimental vehicle is an important factor affecting the experimental results, so it is required that the indicators of the driver's near vision, moving vision and light and dark adaptation be good. In addition, the subjective cognitive effect of the driver's eyes in the experimental car is different from the picture recorded by the car camera, so the picture read from the camera is only used to represent the relevant state, and the difference between the subjective perception of the driver and that of the driver is allowed. The figure shows the different lighting effects of the car in front as recorded by the camera.

TABLE 1. Experimental data summary.

Serial Number	Distance/m	State
1	110.4	No Taillights
2	108.7	Clearance Lamp
3	116.8	No Taillights
4	121.8	No Taillights
5	130.0	No Taillights
6	92.4	Clearance Lamp
⋮	⋮	⋮
85	131.4	No Taillights
86	103.3	Clearance Lamp
87	134.2	No Taillights
88	145.8	Clearance Lamp
89	167.4	Fog Lamp
90	124.7	No Taillights

E. SUMMARY AND GROUP ANALYSIS OF EXPERIMENTAL DATA

By sorting and screening the data obtained by the camera and radar in the experiment, the distance between the front car and the rear car in the fog can be recognized by the rear car and will disappear when the vehicle is not equipped with any light, taillight, double flash, or fog light, as shown in figure 9. The situation in which the car in front does not turn on any light is recognized by the car behind when it is about to disappear into the fog is divided into a group, known as the state of not turning on any light fixture, referred to as not turning on. The situation in which the front car indicator light can just be recognized in the fog is divided into the second group, the state in which the indicator light can just be recognized, abbreviated as clearance lamp. The groups double flash and fog light are created according to the same principle. The final effective experimental data are shown in TABLE I.

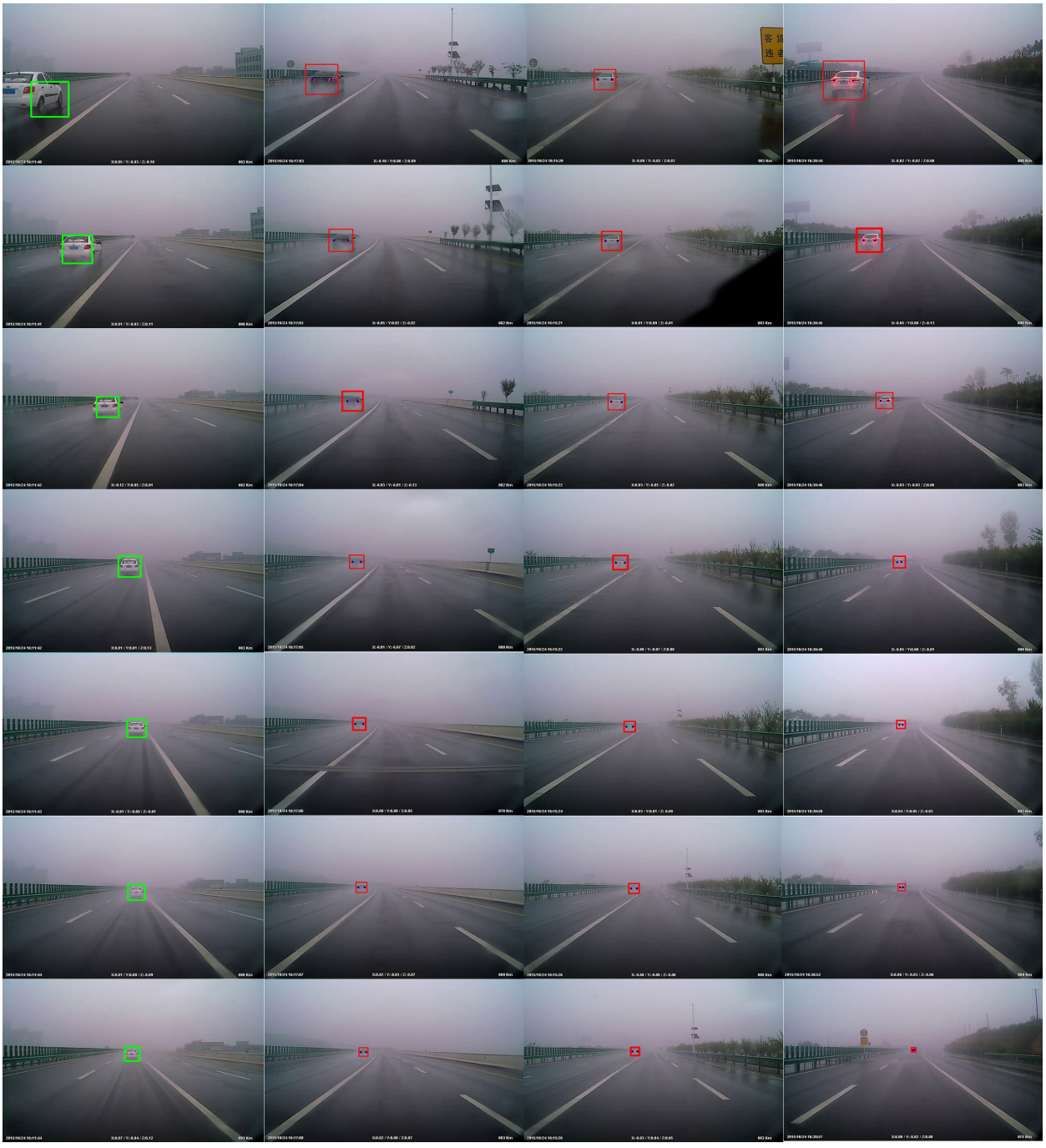

FIGURE 10. Visibility measurement effects of different taillights. (a) No taillights. (b) Clearance lamp. (c) Emergency flasher. (d) Fog lamp.

A total of 90 valid data were collected, including the longest distance that can be recognized by the experimental vehicle under four states of running without any light, clearance lamp, emergency flasher, and fog lamp and the color of the vehicle itself.

V. EXPERIMENTAL RESULTS AND ANALYSIS

A. VALIDATION OF VEHICLE CLASSIFIER

Although the cascade classifier selected for Haar feature training has a satisfactory accuracy in ordinary scenes, considering that it is rarely used in foggy conditions and that the recognition effect of the classifier directly affects the accuracy of the visibility, the performance of the trained classifier is verified. A random frame image is taken from the experimental video as a test sample. Since visibility measurement requires little real-time performance, precision and recall are only selected as evaluation indexes for this verification. The near and far samples were tested separately.

In all the test samples, there were 348 vehicles in total, of which 325 were correctly identified as vehicle targets

and 23 were not identified. The number of non-vehicles that were identified as vehicle targets is five. The accuracy of vehicle detection is 93.39%, and the missing rate is 6.61%. The results indicate that the classifier trained in this paper achieves satisfactory results in vehicle identification.

B. VERIFICATION OF ANTHROPOMORPHIC VISIBILITY

According to the different opening conditions of taillights, the 90 groups of effective data collected in the experiment were divided into 47 groups with taillights not turned on, 23 groups with enlightening width lights, 11 groups with double flashing, and 8 groups with fog lights. The tracking effect of various taillights is shown in Figure 10, the last picture of which is the critical moment when the tracking target is about to be lost. The distance measured by radar at that time is denoted as measured visibility (MV).

In order to quantitatively analyze the accuracy of the MV, the MV accuracy(acc_{MV}) was calculated by the visibility accuracy evaluation formula

$$acc_{\mathrm{i}} = \frac{d_{veri}}{d_{tr}} \tag{23}$$

Moreover, the visibility calculated by Koschmieder's law (KV) was selected for comparison, and its accuracy was expressed as acc_{KV}. In the equation (23), i represents the method used, d_{veri} is the visibility obtained by the method to be verified, and d_{tr} is the ground truth measured in TABLE I for different states.

It can be seen from Fig.11(a) that when the taillight is not turned on, the MV is lower than the ground truth, and the average acc_{MV} is about 0.88. The overall trend of KV is close to the ground truth, and its average acc_{KV} is slightly lower than AV, which is about 0.91. This is because the vehicle classifier is still not able to reach the visibility of human eyes, but can essentially reach the accuracy of traditional visibility measurement methods.

It can be seen from Fig.11(b) that after the clearance lamp was turned on, the acc_{MV} and acc_{KV} were both lower than the visibility of human eyes, but the acc_{MV} was generally higher than that of KV. From the average accuracy, the acc_{MV} showed an upward trend, reaching 0.91, while the acc_{KV} dropped to 0.85. This is because the traditional visibility calculation method is to calculate the visibility of the overall environment, and the taillight segmentation identification method can be used to identify vehicles with taillight on at a greater distance. This difference is also evident in the double-flash and fog lights. As shown in Fig.11(c∼d), acc_{MV} is about 0.90, while acc_{KV} is only 0.83 when the double flash is turned on. When the fog lamp is turned on, acc_{MV} is about 0.95, while acc_{KV} is only 0.81. However, since there is a period of time for the vehicle to turn off its lights when the dual flashing is on, it is easy to lose the tracking target in advance during tracking, so acc_{MV} must fluctuate.

Overall, acc_{MV} in doesn't open taillights flash and fog lamps, open revelation width modulation, double shows an ascendant trend. This is because the vehicles with taillights on in foggy environments than don't open the rear lights is more penetrating, and fog lamps in the strongest penetrability in several kinds of lights, this is also consistent with human visual characteristic, so as to achieve the effect of personification. However, acc_{KV} tends to decrease in these cases. This is because the ground truth is based on the visibility measured by the rear light of the car in front determined by the human eye, while the acc_{KV} is calculated as the visibility of the overall environment, which is inconsistent with the visual characteristics of the human eye, so there is thus a deviation.

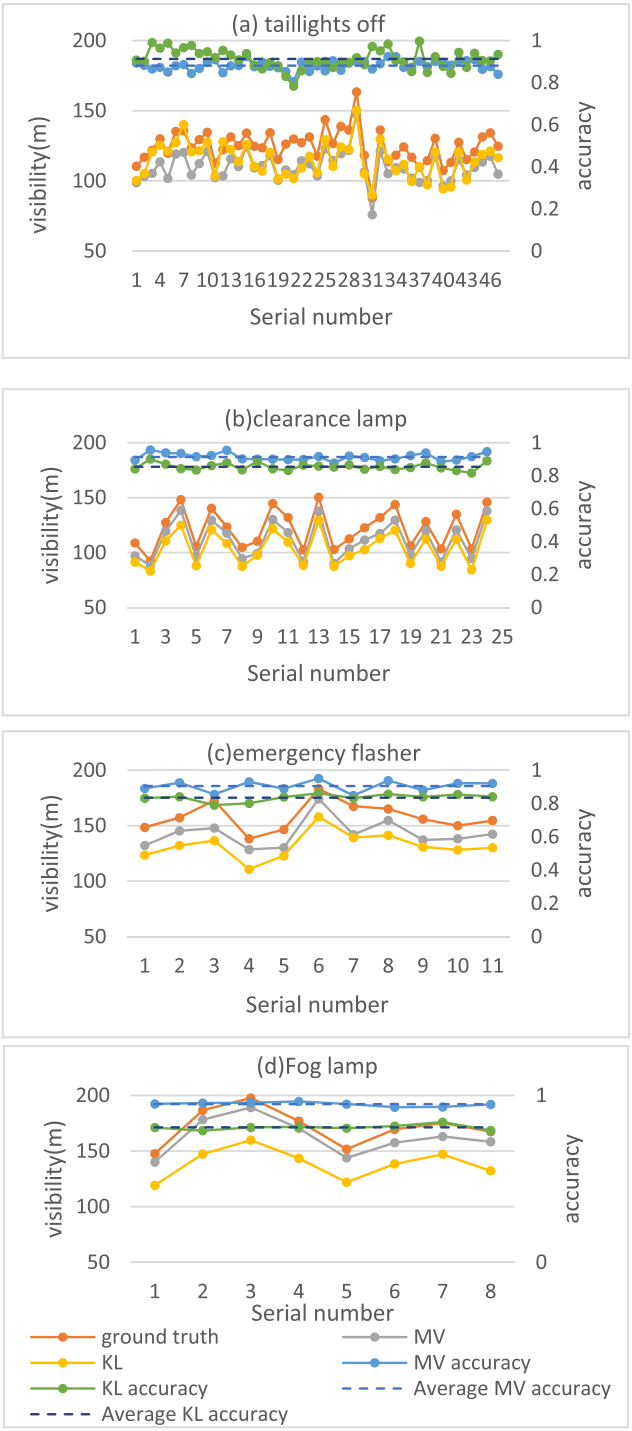

FIGURE 11. Anthropomorphic visibility comparison.

VI. CONCLUSION

In this paper, a sensor fusion–based anthropomorphic visibility measurement method was proposed for low visibility in foggy expressway. This method takes into account the fact that the driver mainly uses the taillight to identify the vehicle in front in cases of low visibility, so the AdaBoost algorithm and the taillight segmentation method were adopted to identify the vehicle with taillights on and off in the image collected by the camera in real time respectively. Then, the DSST algorithm was used to continuously track identified vehicles until they lose tracking. The distance of the vehicle measured by mm-W radar at the moment of losing tracking is recorded as anthropic visibility. The anthropomorphic visibility measured by the proposed method is basically the same as that of normal human eyes. Specifically, the visibility measured when the vehicle is not turned on or with different taillights is consistent with the difference observed by the driver, and the visibility error measured under various taillights and the visibility error of the normal vision driver is kept within 12%, with a minimum of 5%.

This paper presents a visibility measurement method which is more suitable for drivers' needs. As car networking technology matures and with the growing popularity of mm-W radar and camera use, moving vehicles in a highway network can have a method of using real-time measured personification of the current location on the local visibility and the visibility via snatched technology unified collection of the entire road network database of personification visibility. Similarly, the network technology can also transmit the visibility information of the location of vehicles on the road network ahead to the drivers of vehicles, providing an early-warning effect to reduce the risk of accidents.

Future research should focus on improving the recognition accuracy of vehicles without turning on the taillights while improving the overall accuracy of the vehicle recognition algorithm and improving the loss tracking of vehicles with double-flash on when lights are off. In addition, we will consider applying this method to the measurement of night visibility.

REFERENCES

[1] R. G. Hallowell, M. P. Matthews, and P. A. Pisano, "Automated extraction of weather variables from camera imagery," in *Proc. MCTRS*, Ames, Iowa, Aug. 2005, pp. 1–13.

[2] K. Mori, T. Kato, T. Takahashi, I. Ide, H. Murase, T. Miyahara, and Y. Tamatsu, "Visibility estimation in foggy conditions by in-vehicle camera and radar," in *Proc. 1st Int. Conf. Innov. Comput., Inf. Control (ICICIC)*, vol. 1. Beijing, China, Aug./Sep. 2006, pp. 548–551.

[3] M. Gabb, S. Krebs, O. Lohlein, and M. Fritzsche, "Probabilistic inference of visibility conditions by means of sensor fusion," in *Proc. IEEE Intell. Vehicles Symp.*, Dearborn, MI, USA, Jun. 2014, pp. 1211–1216.

[4] H. Koschmieder, "Theorie der horizontalen Sichtweite," in *Beitrage zur Physik der freien Atmosphare*. Frankfurt, Germany: Akademische Verlagsgesellschaft, 1924, pp. 33–53.

[5] C. Zhang, M. Wu, J. Chen, K. Chen, C. Zhang, C. Xie, B. Huang, and Z. He, "Weather visibility prediction based on multimodal fusion," *IEEE Access*, vol. 7, pp. 74776–74786, 2019.

[6] K. W. Kim, "The comparison of visibility measurement between image-based visual range, human eye-based visual range, and meteorological optical range," *Atmos. Environ.*, vol. 190, pp. 74–86, Oct. 2018.

[7] N. Hautiere, R. Babari, and E. Dumont, "Estimating meteorological visibility using cameras: A probabilistic model-driven approach," in *Proc. ACCV*, Berlin, Germany, 2010, pp. 243–254.

[8] F. Guo, H. Peng, J. Tang, B. Zou, and C. Tang, "Visibility detection approach to road scene foggy images," *KSII Trans. Internet Inf. Syst.*, vol. 10, no. 9, pp. 4419–4441, Sep. 2016.

[9] X. Cheng, B. Yang, G. Liu, T. Olofsson, and H. Li, "A variational approach to atmospheric visibility estimation in the weather of fog and haze," *Sustain. Cities Soc.*, vol. 39, pp. 215–224, May 2018.

[10] X. Cheng, B. Yang, G. Liu, T. Olofsson, and H. Li, "A total bounded variation approach to low visibility estimation on expressways," *Sensors*, vol. 18, no. 2, p. 392, Jan. 2018.

[11] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Miami, FL, USA, Jun. 2009, pp. 1956–1963.

[12] Q. Li, Y. Li, and B. Xie, "Single image-based scene visibility estimation," *IEEE Access*, vol. 7, pp. 24430–24439, 2019.

[13] J. Liu, "Visibility distance estimation in foggy situations and single image dehazing based on transmission computation model," *IET Image Process.*, vol. 12, no. 7, pp. 1237–1244, Jul. 2018.

[14] J. Ye, Y. Jin, and N. Wang, "A novel real-time highway visibility measurement system based on dark channel prior," in *Proc. Int. Conf. Image Graph. Process. (ICIGP)*, 2018, pp. 171–175.

[15] H. Qin and H. Qin, "Image-based dedicated methods of night traffic visibility estimation," *Appl. Sci.*, vol. 10, no. 2, p. 440, Jan. 2020.

[16] L. Yang, R. Muresan, A. Al-Dweik, and L. J. Hadjileontiadis, "Image-based visibility estimation algorithm for intelligent transportation systems," *IEEE Access*, vol. 6, pp. 76728–76740, 2018.

[17] T. M. Kwon, *Atmospheric Visibility Measurements Using Video Cameras: Relative Visibility*. Duluth, MN, USA: Univ. Minnesota Duluth, 2004.

[18] A. Palvanov and Y. Cho, "VisNet: Deep convolutional neural networks for forecasting atmospheric visibility," *Sensors*, vol. 19, no. 6, p. 1343, Mar. 2019.

[19] H. Chaabani, F. Kamoun, H. Bargaoui, F. Outay, and A.-U.-H. Yasar, "A Neural network approach to visibility range estimation under foggy weather conditions," *Procedia Comput. Sci.*, vol. 113, pp. 466–471, 2017.

[20] Y. You, C. Lu, W. Wang, and C.-K. Tang, "Relative CNN-RNN: Learning relative atmospheric visibility from images," *IEEE Trans. Image Process.*, vol. 28, no. 1, pp. 45–55, Jan. 2019.

[21] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, San Diego, CA, USA, Jun. 2005, pp. 886–893.

[22] T. Ojala, M. Pietikainen, and D. Harwood, "Performance evaluation of texture measures with classification based on kullback discrimination of distributions," in *Proc. 12th Int. Conf. Pattern Recognit.*, San Diego, CA, USA, Oct. 1994, pp. 582–585.

[23] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proc. CVPR*, Kauai, HI, USA, Dec. 2001, pp. I–I.

[24] R. Lienhart, A. Kuranov, and V. Pisarevsky, *Empirical Analysis of Detection Cascades of Boosted Classifiers for Rapid Object Detection* (Lecture Notes in Computer Science), vol. 2781. Berlin, Germany: Springer, 2003, pp. 297–304.

[25] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *J. Comput. Syst. Sci.*, vol. 55, no. 1, pp. 119–139, Aug. 1997.

[26] O. Alpar, "Corona segmentation for nighttime brake light detection," *IET Intell. Transp. Syst.*, vol. 10, no. 2, pp. 97–105, Mar. 2016.

[27] C. Liu, "On-line selection of discriminative tracking features," in *Proc. 9th IEEE Int. Conf. Comput. Vis.*, Nice, France, Oct. 2003, pp. 346–352.

[28] J. F. Henriques, R. Caseiro, P. Martins, and J. Batista, "High-speed tracking with kernelized correlation filters," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 3, pp. 583–596, Mar. 2015.

[29] J. Henriques, R. Caseiro, P. Martins, and J. Batista, "Exploiting the circulant structure of tracking-by-detection with kernels," in *Proc. ECCV*, Florence, Italy, 2012, pp. 702–715.

[30] M. Danelljan, G. Hager, F. S. Khan, and M. Felsberg, "Learning spatially regularized correlation filters for visual tracking," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Santiago, Chile, Dec. 2015, pp. 4310–4318.

[31] H. K. Galoogahi, A. Fagg, and S. Lucey, "Learning background-aware correlation filters for visual tracking," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Venice, Italy, Oct. 2017, pp. 1144–1152.

[32] D. Bolme, J. R. Beveridge, B. A. Draper, and Y. M. Lui, "Visual object tracking using adaptive correlation filters," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, San Francisco, CA, USA, Jun. 2010, pp. 2544–2550.
[33] M. Danelljan, G. Häger, F. Shahbaz Khan, and M. Felsberg, "Accurate scale estimation for robust visual tracking," in *Proc. Brit. Mach. Vis. Conf.*, Nottingham, U.K., 2014, pp. 1–11.
[34] V. N. Boddeti, T. Kanade, and B. V. K. V. Kumar, "Correlation filters for object alignment," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 2291–2298.
[35] H. K. Galoogahi, T. Sim, and S. Lucey, "Multi-channel correlation filters," in *Proc. IEEE Int. Conf. Comput. Vis.*, Sydney, NSW, Australia, Dec. 2013, pp. 3072–3079.

YANJUN GUO received the M.S. degree from Dalian Maritime University, Liaoning, China, in 2009, and the Ph.D. degree from Chang'an University, Xi'an, China, in 2016. Her research interests include vehicle safety technology and car sickness.

QINGJIN XU received the B.S. degree in vehicle engineering from the Henan University of Science and Technology, Luoyang, China, in 2017. He is currently pursuing the Ph.D. degree in vehicle engineering with the School of Automobile, Chang'an University. His research interests include driver behavior and intelligent vehicle safety auxiliary technology.

YANQI SU received the B.S. degree in transportation safety engineering from Chang'an University in 2017, where he is currently pursuing the Ph.D. degree in vehicle engineering with the School of Automobile. His research interests include driving behavior and vehicle active safety.

SIYANG JIANG received the B.Tech. degree in transportation engineering from Chang'an University, Xi'an, China, in 2019, where she is currently pursuing the M.S. degree in vehicle engineering with the School of Automobile. Her research interests include driving behavior and vehicle active safety.