

## Research Article

# Application of Image Processing and Three-Dimensional Data Reconstruction Algorithm Based on Traffic Video in Vehicle Component Detection

Gang Li,<sup>1</sup> Huansheng Song,<sup>1</sup> Shuyu Wang,<sup>1</sup> and Jinliang Kong<sup>2</sup>

<sup>1</sup>School of Information Engineering, Chang'an University, Xi'an, Shaanxi 710064, China

<sup>2</sup>China Highway Engineering Consulting Corporation, Beijing 100097, China

Correspondence should be addressed to Gang Li; [gangli@chd.edu.cn](mailto:gangli@chd.edu.cn)

Received 28 February 2017; Revised 12 August 2017; Accepted 23 October 2017; Published 15 November 2017

Academic Editor: Simone Bianco

Copyright © 2017 Gang Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Vehicle detection is one of the important technologies in intelligent video surveillance systems. Owing to the perspective projection imaging principle of cameras, traditional two-dimensional (2D) images usually distort the size and shape of vehicles. In order to solve these problems, the traffic scene calibration and inverse projection construction methods are used to project the three-dimensional (3D) information onto the 2D images. In addition, a vehicle target can be characterized by several components, and thus vehicle detection can be fulfilled based on the combination of these components. The key characteristics of vehicle targets are distinct during a single day; for example, the headlight brightness is more significant at night, while the vehicle taillight and license plate color are much more prominent in the daytime. In this paper, by using the background subtraction method and Gaussian mixture model, we can realize the accurate detection of target lights at night. In the daytime, however, the detection of the license plate and taillight of a vehicle can be fulfilled by exploiting the background subtraction method and the Markov random field, based on the spatial geometry relation between the corresponding components. Further, by utilizing Kalman filters to follow the vehicle tracks, detection accuracy can be further improved. Finally, experiment results demonstrate the effectiveness of the proposed methods.

## 1. Introduction

With the rapid development of intelligent traffic control, computer vision has attracted much attention from designers of intelligent transportation systems (ITSs) [1], due to its importance in information collection in the real-time environment. Monitoring systems based on computer vision technology have become very important in the development of ITSs, and a detailed introduction to vehicle monitoring methods and video monitoring system frameworks is available in [2].

At present, video-based vehicle monitoring systems can be divided into two categories according to two different kinds of vehicle features: vehicle appearance and vehicle moving character. In the method based on vehicle appearance [2, 3], a vehicle target is detected by means of geometric structures, color information, and texture features of the

entire or part of the vehicle, such as the symmetry of the vehicle structure, the outline of the vehicle, the local car lights, and license plates. There are a variety of feature representation operators to describe vehicle appearance, such as high-order Godunov (HOG) features, Haar-like operators, scale-invariant feature transform (SIFT) features, and speeded-up robust features (SURFs). These feature descriptors are combined with classification algorithms for vehicle monitoring systems in ITSs, such as artificial neural networks, support vector machines, AdaBoost, and sparse representation classification. In the method based on moving features [4, 5], the entire or part of the vehicle is tracked, and the corresponding tracking trajectories used to detect vehicle targets and analyze vehicle behaviors are obtained.

The method based on a 3D vehicle model [6] comprises building a 3D model of the vehicle target and then detecting the vehicle target using the vehicle identification method.

The difficulty of this method is in how to build a better vehicle model. There are different vehicle sizes and shapes among different vehicle models and types. It is not easy to put forward a unified 3D model to include all kinds of vehicles in real life. When the traffic situation is complex, the detection of a 3D model is usually ineffective. Although 3D information of the vehicle is useful, occlusion [7], adhesion, and other issues still cannot be resolved.

The method based on features [8] is different from the above methods. All types of vehicle detection methods view vehicles as the smallest unit of target detection, but regarding the vehicles themselves, they have many variable features [9, 10], such as headlights, taillights, license plates, and vehicle symmetries. The organic combination of these features can often represent vehicle targets. We can detect a vehicle's local characteristics instead of detecting vehicle targets. How to detect a vehicle's local characteristics is a key and challenging aspect of this method. The images obtained by video cameras distort the shape and size of the vehicle due to the perspective relation, which makes it difficult to extract the local features of the vehicle from the 2D image.

In this paper, we aim to solve how to (1) detect vehicle local features or components, (2) avoid complex scenes caused by occlusion and adhesion problems, and (3) establish the relationship between a 2D image and a 3D scene, using the local features or inherent properties of size and shape components while avoiding image distortion caused by the camera.

Under the conditions of normal light illumination and the basic rules of the road, it is not a problem for a human being to locate the target vehicle's headlights and its color information quickly. However, if a computer vision method is utilized to design a robust algorithm to fulfill these tasks, it becomes extremely challenging to detect vehicle headlights, license plates, and taillights accurately.

## 2. Contributions of the Paper

In this paper, we mainly study the target recognition algorithm in traffic monitoring systems. Based on the probabilistic model of the spatial relation, the detection of the target vehicle's components is exploited instead of detection of the entire vehicle.

There are two main innovations in our paper: on one hand, the shape and size of the vehicle components are distorted due to camera projection transformation. In this paper, therefore, an inverse projection algorithm is proposed to construct an inverse projection map. The inherent shape and size of the vehicle can be obtained on the inverse projection map, and the saliency components of the vehicle can be detected using these inherent attributes. Details of this public detection of vehicles during both day and night are presented in Section 3.1. On the other hand, the target vehicle can be replaced by detecting organic combinations of vehicle parts. At night, headlights can be selected as detection object and are first detected according to their geometric characteristics. A Gaussian mixture model is established using the distance

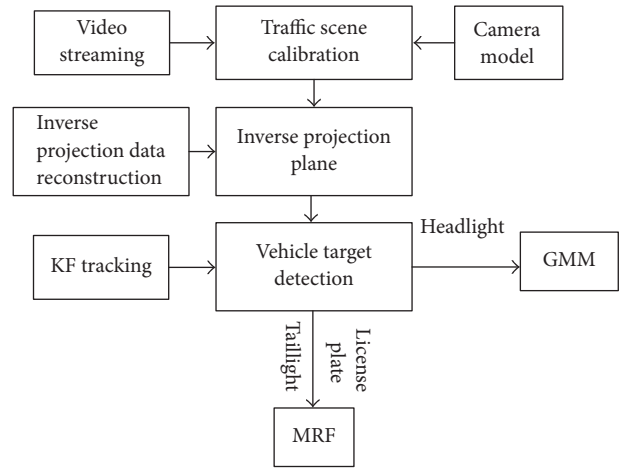


FIGURE 1: Main flow of proposed algorithm.

between the headlights and their height, and the resulting probability model is used to achieve accurate detection. Details of this process appear in Section 3.2. In the daytime, the taillights and license plate are selected as the objects of detection through use of a color model to detect them. Then, by determining the geometric relationship between the parts, a Markov random field is established to complete vehicle detection. Details of this process appear in Section 3.3.

*2.1. Summary of the Proposed Nighttime Algorithm.* In the evening, since sunlight is limited, only the headlight information of the vehicles can be used. The target vehicle components at night are detected by the following steps:

- (1) Search the real data according to the car design and manufacturing standard, and then establish a Gaussian mixture model (GMM) model based on the real data [11].
- (2) Detect the dominant information of the vehicle components by using the inverse projection map.
- (3) Use the GMM model to train the dominant information, and obtain the GMM stochastic value.
- (4) Judge the stochastic value to achieve final vehicle detection.

*2.2. Summary of the Proposed Daytime Algorithm.* During the daytime, solar illumination is adequate, and thus the color information of the vehicle's parts can be utilized to detect the vehicle license plate and taillights. In the daytime, the detection of the vehicle's license plate and taillight can be realized by using background subtraction and a Markov random field (MRF) along with the spatial geometry relation between the corresponding components.

The main flow of the proposed algorithm is shown in Figure 1. In this paper, we introduce discrete algorithms for day and night.



FIGURE 2: Schematic of inverse projection plane.

### 3. Vehicle Detection Algorithm Based on Components

#### 3.1. Inverse Projection Plane and Inverse Projection Map.

In this paper, an inverse projection plane is preliminarily determined in the traffic scene that has been calibrated [12]. Once the position of the space inversion projection plane is determined, the corresponding relationship between the point in the reverse projection plane and the position of the pixel in the projection image is also determined. Then, the data of the projected image can be projected onto the inverse projection plane to obtain an inverse projection map, which copies the spatial information of the inverse projection plane. The inverse projection [5] process consists of two parts: the design of the inverse projection plane and the construction of the inverse projection map.

The proposed method relies on the space information with three dimensions; therefore, the calibration procedure for traffic scenes is necessary, and there are many calibration methods for traffic scenes. In this paper, the direct linear transformation (DLT) as proposed by Abdel-Aziz and Karara [12] is used for scene calibration.

**3.1.1. Design of Inverse Projection Plane.** The inverse projection plane [13] is designed according to the characteristics and spatial position of the target to be detected. Depending on the specific circumstances, it can be set to be parallel of the road, perpendicular to the road, or at a certain angle to the road. The number can be designed as 1 or more.

The part of the surface of the vehicle can be approximated as a plane with some geometric features. If the vehicle in three-dimensional (3D) space is regarded as a polyhedron, when the characteristics of different faces of the vehicle body are selected as the detection objects, the reverse projection plane is attached to the corresponding plane of the vehicle body, making the data after the construction of the inverse projection able to effectively show some of the apparent characteristics of the body (see Figure 2).

**3.1.2. Construction of Inverse Projection Map.** According to the above-described method, an inverse projection plane that can be fitted with a certain local surface of the target is arranged in the space and divided into a grid with a certain resolution (such as  $1\text{ cm} \times 1\text{ cm}$ ). The camera's perspective relation is that the information contained in the grid is projected onto each pixel of the corresponding projection area on the image, where the inverse projection relationship from the image projection area to the reverse projection plane

is determined, that is, a small inverse of the inverse projection plane in the space. The grid corresponds to a pixel on the image.

The inverse projection map is a pixel representation of the inverse projection plane, which means that a small grid in the inverse projection plane is represented by a pixel in the inverse projection map. The process of building inverse projection map data is as follows: (1) the inverse projection plane of each small grid information is copied to the inverse projection map and (2) the inverse projection map in each pixel then represents the information of the inverse projection plane on each square of the grid information.

Suppose that  $m$  represents a small grid of the inverse projection plane,  $p$  the image's pixel of  $m$  projected onto the map, and  $mp$  the grid  $m$  corresponding to the inverse projection map of the pixel; the reverse projection process is then mapping the image pixel  $m$  to the inverse projection map pixel  $mp$ . The inverse projection map construction principle is illustrated in Figure 3.

It can be seen from Figure 3 that an inverse projection plane is provided on the spatial plane of the target surface. The inverse projection after restoration of the data is the copy of the target surface, which not only eliminates the perspective of the camera (some of the shape features of the target surface in the captured image are geometrically deformed) but well reflects the true dimensions of the local features of the target surface.

In this paper, the road traffic scene calibration during the course of the experiment is designed as two inverse projection planes, respectively, perpendicular to the road, as shown in Figure 4. The former is placed in the world coordinate system  $Y = 2\text{ m}$  plane of size  $2 \times 3.75\text{ m}$  and denoted inverse projection plane 1; the latter is placed in the  $X = 6\text{ m}$  plane of size  $2 \times 5\text{ m}$  and denoted inverse projection plane 2. The red border in the figure is the projection area of the two inverse projection planes. After a target vehicle and the tail plane and the side are completely fitted to the two inverse projection planes, using the inverse perspective transform method, the projection area data needed to construct the inverse projection map are obtained. As the inverse projection map shows, when just the vehicle tail plane is fitted to the inverse projection plane, the characteristics of the vehicle tail plane are recovered (blue boxes) and are not fitted to the still deformed tail plane. When the inverse projection plane is fitted to the vehicle side (blue box), the positive view of the real space on the side of the vehicle is also constructed. In the inverse projection map experiments, 1 pixel represents the world coordinate system,  $1 \times 1\text{ cm}$  square.

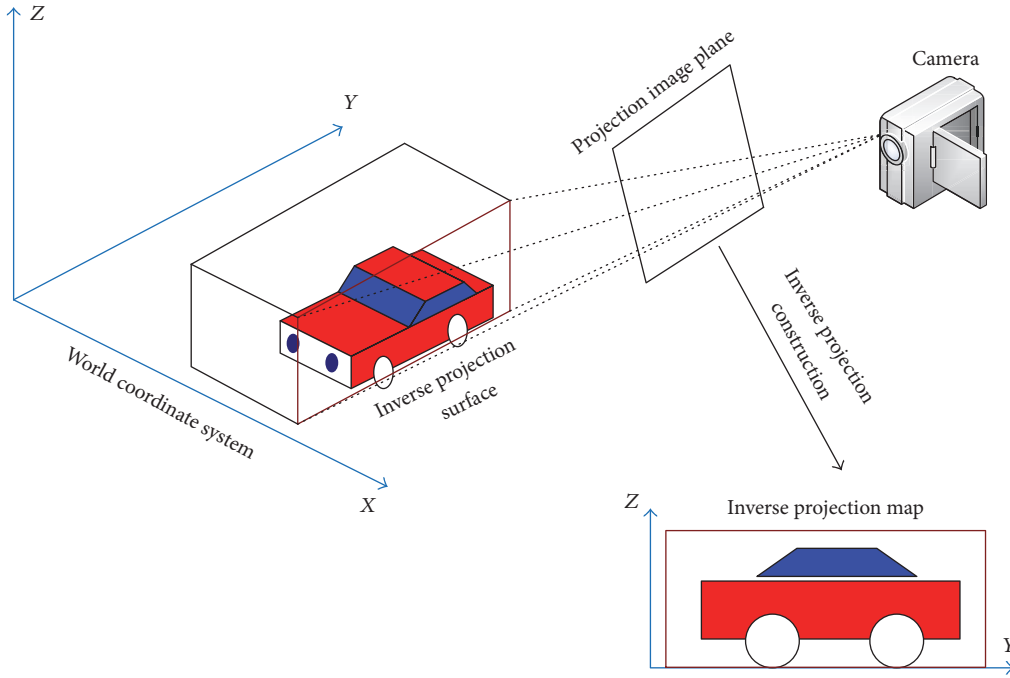


FIGURE 3: Imaging model of cameras and inverse projection transformation.

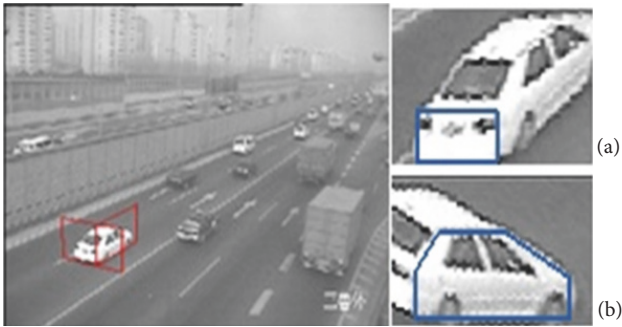


FIGURE 4: Original image: (a) inverse projection map 1 (tail) and (b) inverse projection map 2 (side).

**3.2. Algorithm for Nighttime Vehicle Detection.** Nighttime vehicle detection can use the center surround extreme as in [10]. In this paper, we detect headlights on the inverse projection map. Vehicle headlights as a significant feature of a vehicle at night have an apparent brightness, and on the inverse projection map targets have real structure characteristics and are an approximation of the real vehicle lights. The height of shadow on the projection plane is zero, and the headlights have some characteristics of height information, which can be used to remove the interference from headlight shadows. The flow of the main detection algorithm flow is as shown in Figure 5. In the evening, our method utilizes the shape characteristics of the headlights, rather than the color information, and thus grayscale video can be adopted.

**3.2.1. Segmentation of the Headlights.** The background difference method [14] is a commonly used foreground object extraction method in static camera imaging. Its function is based on the principle of using a background extraction algorithm to obtain the video background, because the pixels' gray values of foreground moving objects and those of the background exhibit a certain difference; thus, we can perform a differential operation between video pixel values and the background pixel value of the same position. If the difference is greater than the threshold, we can consider it the foreground target.

The mathematical representation of the background subtraction method is as follows: the hypothetical image size is  $H \times W$ , the point in the current frame's  $P(x, y)$  gray value is  $F(x, y)$ , and the gray value of the corresponding pixel in the background image is  $B(x, y)$ . The difference and binarization after determining the foreground pixel gray value is  $D(x, y)$ :

$$D(x, y) = \begin{cases} 0 & \text{if } |F(x, y) - B(x, y)| < T \\ 255 & \text{if } |F(x, y) - B(x, y)| \geq T, \end{cases} \quad (1)$$

where  $T$  is the image of the preset threshold for binarization processing.

The background extraction environment chosen for this paper is nighttime traffic scenes, so the impact of weather and shadows is relatively small, and it is used to reveal a significant difference in the brightness of nighttime vehicle headlights upon block segmentation of the background, so the background extraction requirements are not too difficult to meet. As shown in Figure 6, in this step the proposed algorithm utilizes the gray information of the headlights



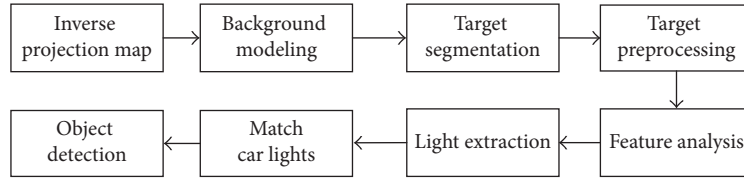


FIGURE 5: Flowchart of nighttime vehicle target detection algorithm.

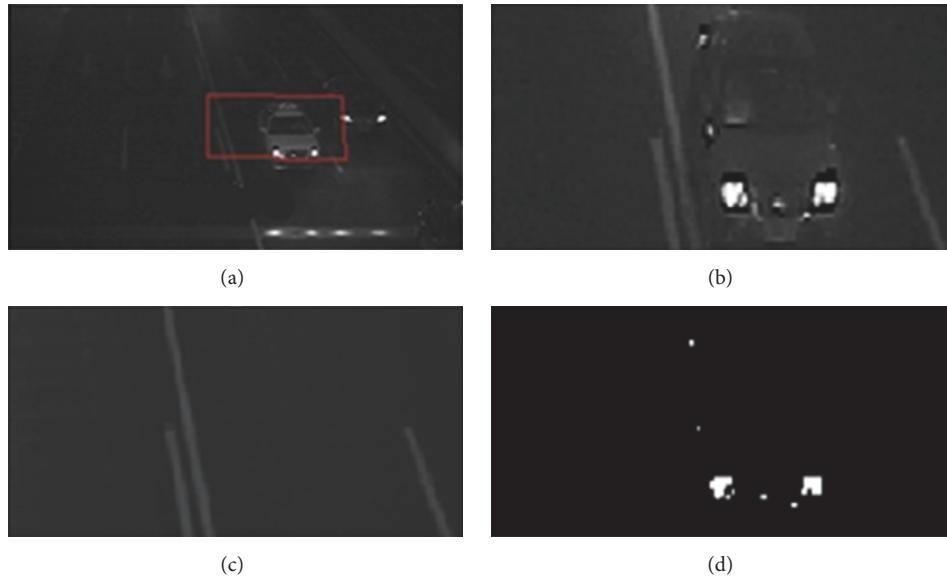


FIGURE 6: (a) Set inverse projection plane, (b) inverse projection map, (c) background extraction, and (d) foreground segmentation and binarization.

rather than the color information, and thus either gray or color video can be adopted.

**3.2.2. Headlight Pairing.** As the most obvious characteristic of nighttime vehicle detection [15], headlights have some geometrical features of their own, such as geometric features of the area, circularity, and the similarity of the headlights, as shown in Figure 7.

Calculating the foreground object block of these characteristics can exclude the nonheadlights block. The mathematical definitions and expressions of these geometric features, as well as the threshold values, are set as follows, which expresses the computing method for area  $A$ :

$$A = \sum_{m=a}^{a+w} \sum_{n=b}^{b+h} f(m, n), \quad (2)$$

where  $f(m, n)$  is the gray value of the pixel point  $(m, n)$  located in the two-value image and  $A$  is the area of the foreground object. The formula statistics, the number of pixels of the foreground object within  $f(m, n) = 1$ , is also the number of white pixels in the rectangle frame that is connected to the foreground object.  $(a, b)$  are the lower left vertex coordinates of the outer rectangle,  $h$  is the height of the outer rectangle frame, and  $w$  is the width of the outer

rectangle frame. According to the vehicle manufacturer's production specifications, the headlights are of a size that is in a certain range, because if the foreground object block is too large or too small, it is not a headlight target. The upper  $TA2$  and the lower  $TA1$  thresholds are determined by the conversion ratio of the pixel to the size when the inverse projection map is constructed. When the foreground object blocks meet the condition  $TA1 < A < TA2$ , the foreground object block is marked as a kind of vehicle headlight block.

The method of computing the degree of circularity  $C$  is

$$C = \frac{L^2}{A}, \quad (3)$$

where  $A$  is the area of the foreground object and  $L$  is the contour perimeter of the foreground object. According to the calculation, in a variety of geometries, the degree of circularity of a circular object is the minimum, the degree for square circularity is 16, the degree for rectangular circularity is greater than 16, and as the ratio of length to width increases, the degree of circularity also increases. A convex polygon's degree of circularity is smaller, but not smaller than that of a circle's circularity; that is, a concave polygon has a greater degree of circularity. According to a priori knowledge of headlight shape, usually headlight shape is circular or rectangular, but the degree of rectangularity will not be



FIGURE 7: Headlights of the foreground segmentation region and its geometric features.

too large. Owing to the light scattering effect, the shape of headlight bright blocks will to some extent be close to circular, indicating that the headlight circularity is not too large. In this paper, the degree of circularity  $C$  threshold value is  $TC = 18$ . When the foreground target block's degree of circularity is  $C > 18$ , it is not a headlight target block; otherwise, the prospect target block will be marked as a headlight block.

The method of computing the geometric similarity of headlights is

$$\begin{aligned} A\_Ratio &= \frac{A_i}{A_k}, \\ C\_Ratio &= \frac{C_i}{C_k}, \end{aligned} \quad (4)$$

where  $i$  and  $k$  are the foreground object block labels,  $A_i$  and  $A_k$  represent the areas of the  $i$  and  $k$  object blocks, and  $C_i$  and  $C_k$  represent the degree of circularity of the  $i$  and  $k$  foreground object blocks. As shown, from the geometric features, the vehicle is described in the inverse projection map. The two headlights of a vehicle have obvious symmetry, which means that the two headlights have the same area and geometry, and, coupled with the proposed inverse projection diagram, we can effectively restore the image of the target shape in order to apply the area ratio and the circularity ratio of headlights tending to 1 as the prior knowledge required to identify possible headlights. In addition, this allows us to satisfy the matching conditions of the foreground object block. In this paper, the threshold settings are as follows:  $T_{A,R} = T_{C,R} = 0.65$  and  $T_{A,R'} = T_{C,R'} = 1.6$ , when the area ratio and the circularity ratio of the  $i$  and  $k$  object blocks are both satisfied with  $T_{A,R} < A\_Ratio < T_{A,R'}$  and  $T_{C,R} < C\_Ratio < T_{C,R'}$ , and  $i$  and  $k$  are possibly a pair of headlights.

Using the above judgment conditions to detect headlights, the rough matching results are shown in Figure 8. In this step, our algorithm utilizes the shape characteristics of the headlights rather than the color information, and thus gray or color video can be adopted.

**3.2.3. Modeling the Spatial Relation of the Headlights.** At this stage of the proposed method, rough matching of the vehicle headlights has been completed. Now, using the headlight parts and the spatial relation of the headlights for recognition and positioning and for alternative vehicle identification and localization is a very important part of the process. According to the actual life of the spatial relationship features of the



FIGURE 8: Frame 567: left, inverse projection plane set; right top, building the inverse projection map; right bottom, rough matching.

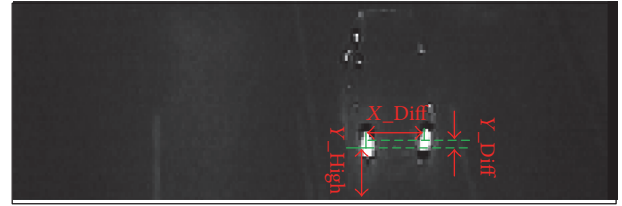


FIGURE 9: Schematic of the spatial relationship between parts of vehicle headlights.

vehicle components, one can extend the spatial relationship of the headlight features: it belongs to the center of gravity of the vehicle if the same pair of target headlights is presented in the same horizontal line and at the same vertical distance with parameter dimensions for practical production of vehicles. The spatial relationship is shown in Figure 9.

The mathematical expression of the spatial relationship between the headlights is as follows:

$$\begin{aligned} X\_Diff &= |X_i - X_j|, \\ Y\_Diff &= |Y_i - Y_j|, \\ Y\_High &= \frac{1}{2}(Y_i - Y_j), \end{aligned} \quad (5)$$

where  $X\_Diff$  and  $Y\_Diff$  denote two headlights in the  $X$  direction and  $Y$  direction distance difference and  $Y\_High$  denotes the height of the headlights above the ground.

TABLE 1: Values of variables of the spatial relationship of headlight components.

X-Diff (cm)	126	148	136	141	128	139	110	133	132
Y-Diff (cm)	1	2	0	2	1	0	2	1	0
Y-High (cm)	51	73	58	64	70	80	79	62	52
X-Diff (cm)	133	149	212	198	147	146	157	145	215
Y-Diff (cm)	0	1	3	0	2	1	3	1	2
Y-High (cm)	79	80	89	97	51	52	70	67	91
X-Diff (cm)	160	142	134	154	135	144	133	146	223
Y-Diff (cm)	3	2	2	1	0	3	2	1	4
Y-High (cm)	59	86	52	73	69	83	57	81	96

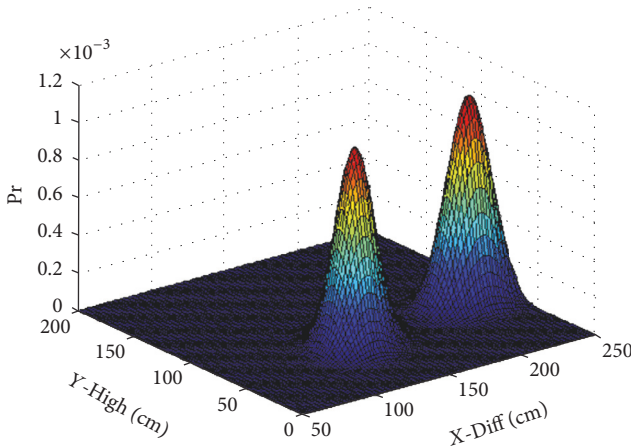


FIGURE 10: Spatial relationship model of headlight components for different types of vehicles.

In this paper, vehicle headlight distances of samples of two headlights were statistically calculated in the  $X$  direction and the  $Y$  direction differences, as were the ground height values of headlights. The values of variables of the spatial relationship of headlight components for a portion of the total sample set of 500 vehicles are shown in Table 1.

From the table, we can see that the difference of headlights on samples of the  $Y$  direction distance (i.e.,  $Y\_Diff$ ) is close to zero. Therefore, in this paper, the modeling of these samples can be considered as  $Y\_Diff \approx 0$ . As a result, we submit that the other two values (i.e.,  $X\_Diff$  and  $Y\_High$ ) constitute the probability distribution map and ignore the  $Y$  direction data. It is thus easy to find a spatial relationship model between the headlights. The results of using the GMM with these samples are shown in Figure 10. A 2D GMM is established via two attributes: the height of the headlights and the distance between them.

**3.3. Algorithm for Daytime Vehicle Detection.** Using the blue component of vehicle license plates and taillights, respectively, we studied color detection of the target vehicle components. In RGB color mode, the blue component of the license plate background is larger than the red and

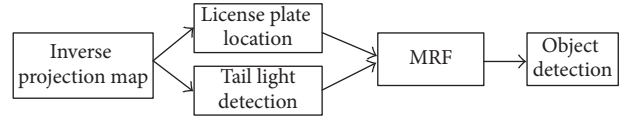


FIGURE 11: Flowchart of daytime vehicle target detection.

green components, and the red and green components are very small. The red weight of the rear lights is greater than the blue and green component weights, and the blue and green component weights are also small. Based on these features, after processing, the target video sequence color space conversion can locate the license plate and the taillights. The detection flowchart is shown in Figure 11. In the daytime, since our method depends on the color information of the blue license plate and the red taillight, color video is utilized.

### 3.3.1. License Plate Detection

**(1) Color Conversion Model.** Considering the RGB color pattern, we mainly used the background color of the license plate to analyze the blue and white license plate components and then obtain each component's histogram. The results are shown in Figure 12, in which it is obvious that in the license plate background pixel value the blue component is much larger than the other two component values, and both component values are relatively small.

Therefore, by analyzing the apparent characteristics of the license plate, we convert the video to a special color space:

$$C_{x,y} = B_{x,y} - \min \{R_{x,y}, G_{x,y}\}, \quad (6)$$

where  $C_{x,y}$  is converted after the color pixels, and  $R_{x,y}$ ,  $G_{x,y}$ , and  $B_{x,y}$  are the red, green, and blue components of the pixel value, respectively. This transformation can enhance the license plate region, while suppressing the non-license-plate region. The conversion results are shown in Figure 13. From the transformed image sequence, we can see that the gray gradient of the license plate region is obviously enhanced, and the gray gradient of the non-license-plate region is significantly suppressed.

In this step, our algorithm utilizes the blue information of the license plate, and thus color video is adopted (gray-level video has no color information).

After transformation of the image sequence, a single channel image is obtained, and the pixels of the license-plate-region characteristics are quite obvious. The brightness is prominent, but brightness also exists around the target region. In order to exclude the interference of the surrounding pixel brightness, the histogram of the single channel image can be used, as shown in Figure 14(a). In the figure, one can see that the pixel value is greater than 50 and that the peak is in the vicinity of the license plate area. After threshold segmentation and further processing, the results shown in Figure 14(c) were obtained.

**(2) Gradient Extraction of License Plate Region.** After color conversion, the gray gradient [16] of the license plate region is

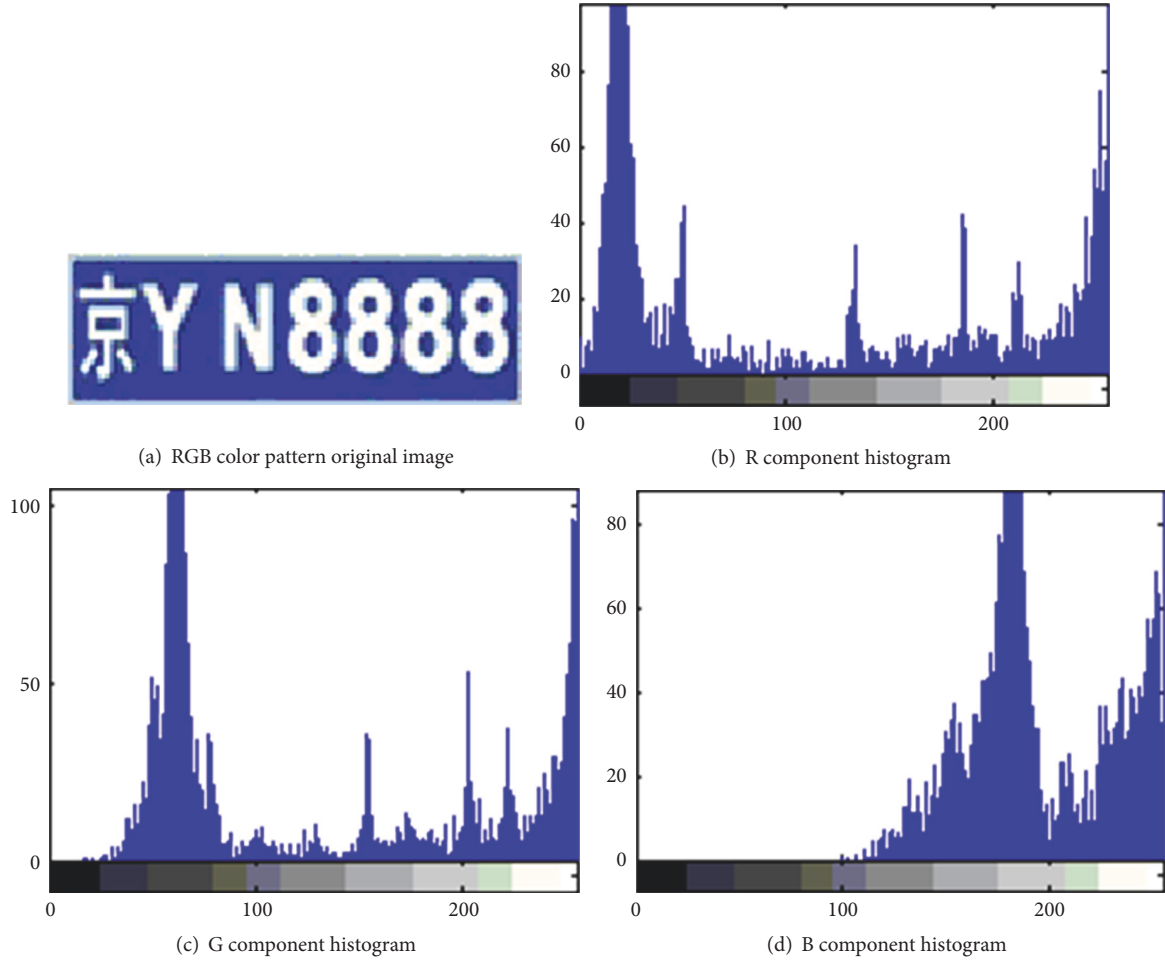


FIGURE 12: Blue license plate RGB component histogram statistics.

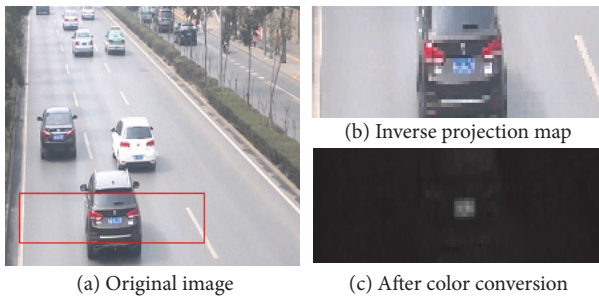


FIGURE 13: Color conversion results.

enhanced, so it can be expressed by the pixel-value difference of the pixels in the image sequence. The formula for the gray gradient calculation is

$$\text{Grad}_{x,y} = \max \{ \text{Nb}_{x,y} \} - \min \{ \text{Nb}_{x,y} \}, \quad (7)$$

where  $\text{Grad}_{x,y}$  is the gradient value of the pixel point  $(x, y)$  in the converted video image sequence and  $\text{Nb}_{x,y}$  is the neighborhood pixel value of the pixel  $(x, y)$ . The gradient extraction result is shown in Figure 15. A sliding window

is used to scan the gradient image, and then the average gray gradient of the window region is calculated. Because the image processing operation described in this paper is based on a 3D inverse projection map, the size of the license plate area and the actual size of license plate are in a certain proportion. When the target vehicle coincides with the inverse projection plane fitting, width and height of the license plate in the image can be calculated and used to set the width and height of the scanning window. The average gray-level calculation is expressed as

$$S_{x,y} = \frac{\sum_{x'-x-pw/2}^{x+pw/2} \sum_{y'-y-ph/2}^{y+ph/2} \text{Grad}_{x',y'}}{pw \times ph}, \quad (8)$$

where  $S_{x,y}$  is the average gray gradient of the center  $(x, y)$  of the license plate region and  $pw$  and  $ph$  are the width and height of the scanning window.

(3) *License Plate Location*. Using the average gradient, we used non-maxima suppression (NMS) [17] to find the local maximum value of the current frame image average gradient: if a local maximum value is larger than a preset threshold, the region is a candidate region for a license plate.



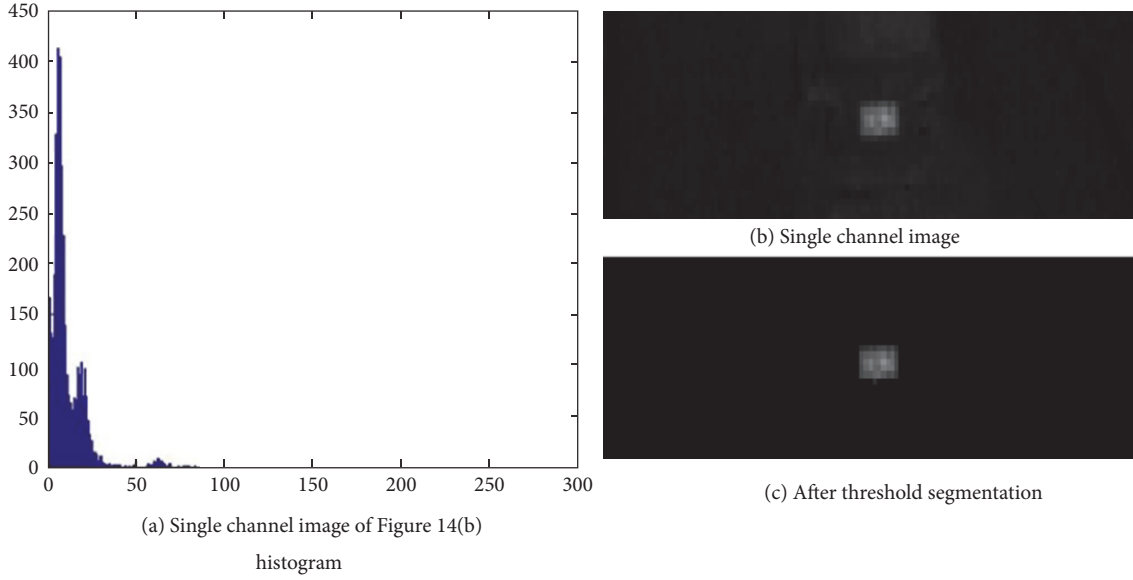


FIGURE 14: Threshold segmentation results.

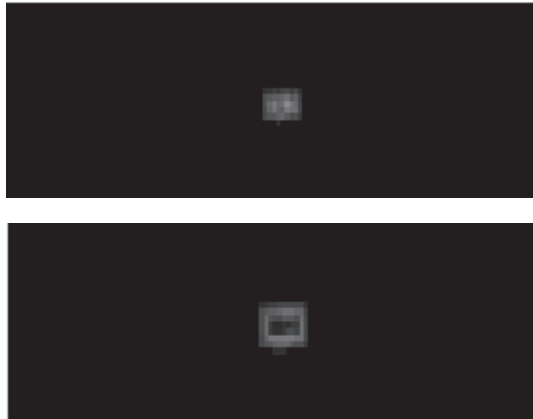


FIGURE 15: Results of gradient extraction.

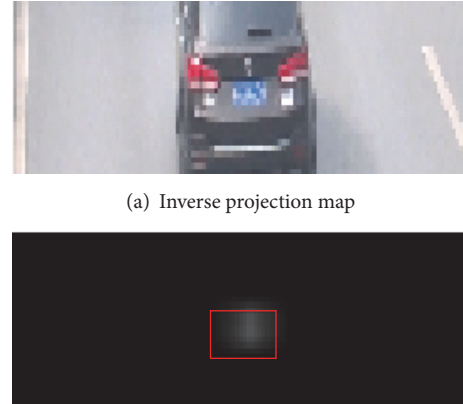


FIGURE 16: License plate location results.

In the video image sequence, except for the license plate area, there are still some similar license areas that also have a gray gradient, but that can also match a color feature, which may lead to a license plate detection error. Through observation, it is found that the license plate region has not only a significant gray gradient, but also a more uniform distribution. Therefore, we further improve the validity of the algorithm via the texture consistency of the vehicle license plate region [18], the expression of which is

$$U = \sum_{I=0}^{L-1} p^2(z_i), \quad (9)$$

where  $U$  represents the consistency of the image,  $L$  the ash step, and  $p(z_i)$  the gray-level histogram. In this experiment, the consistency of the license plate area is 0.7, and the license plate location results are shown in Figure 16.

**3.3.2. Taillight Detection.** Taillights comprise a significant feature of a vehicle, and note that it is obvious that the taillight color is red in our experiment. We therefore use the red color to detect taillight.

*(1) Color Conversion Model.* As with the license plate location, taillight detection is based on the RGB color model, and through the statistical histogram it can be found that the red component is significantly greater than the other two components, as shown in Figure 17.

After the calculation of the color conversion model, the vehicle taillight area is enhanced, and the nontaillight area is suppressed. As observed, in the RGB color mode the red component of the vehicle taillight is greater than the blue and green components, and the difference between the green and blue components is small. Upon analyzing this feature, we

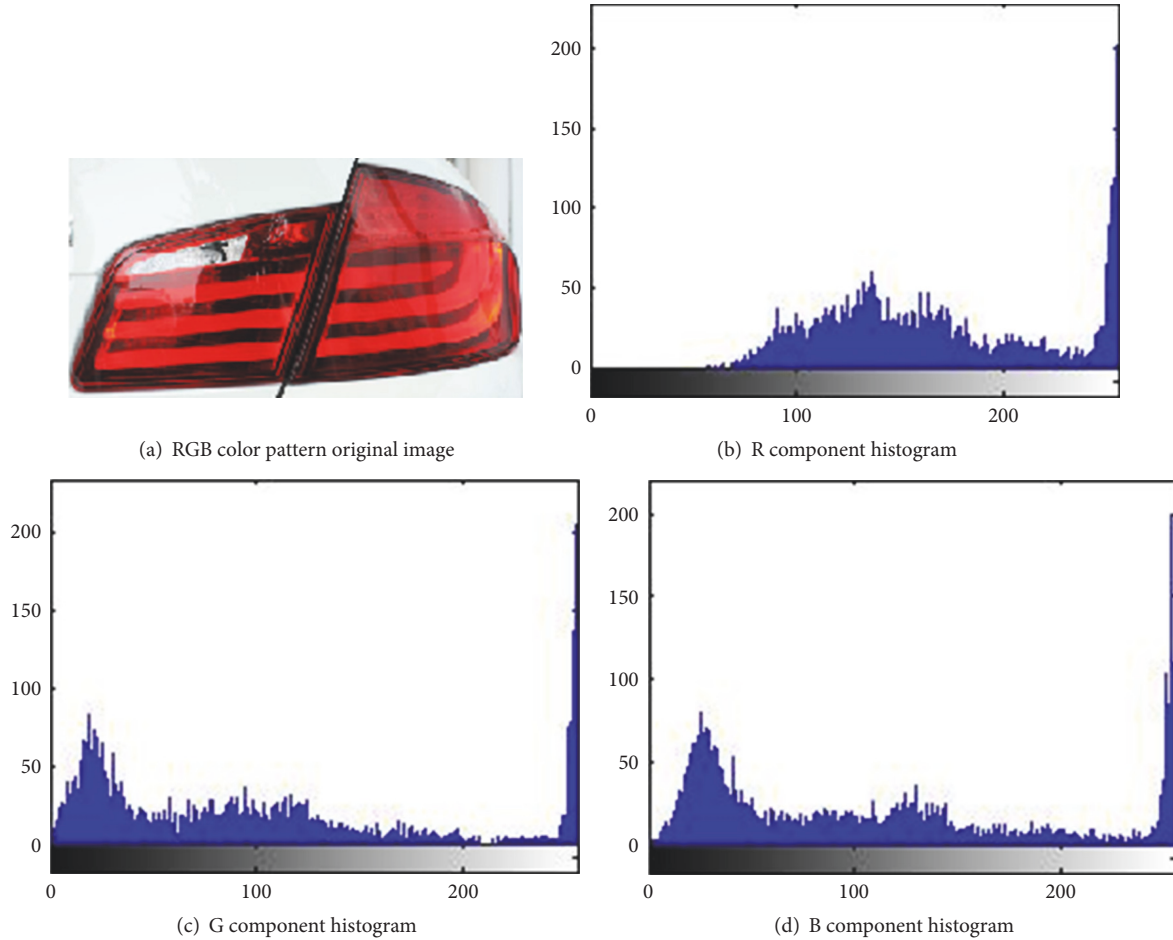


FIGURE 17: Taillight RGB component histogram statistics.

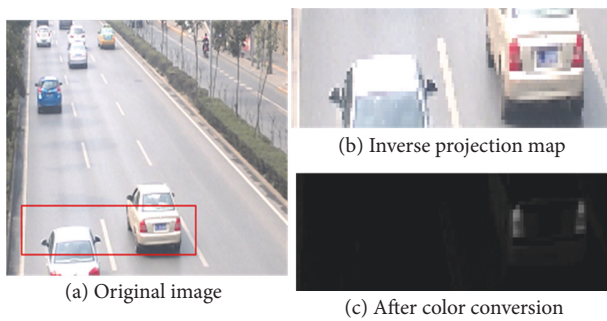


FIGURE 18: Color conversion results.

obtained the color conversion model, as shown in Figure 18. The conversion expression is

$$C_{x,y} = R_{x,y} - \max \{G_{x,y}, B_{x,y}\} - 2 \times \|G_{x,y} - B_{x,y}\|, \quad (10)$$

where  $C_{x,y}$  is converted after the color pixels, and  $R_{x,y}$ ,  $G_{x,y}$ , and  $B_{x,y}$  are the red, green, and blue components of the pixel value, respectively. In this step, our algorithm utilizes the red

information of the taillights, and thus color video is adopted (gray-level video has no color information).

(2) *Threshold Segmentation.* The advantages and disadvantages of target segmentation have a great influence on follow-up target detection. Holes still exist in the target after threshold segmentation. Considering the morphological operation mentioned above, the closed image can fill the target image hole and connect the adjacent target block. The boundary of the target can be smoothed and the results of the closed operation are shown in Figure 19.

3.3.3. *Vehicle Detection Based on MRF.* Considering the space geometry relationship between the license plate and headlight, the Markov random field (MRF) model was used to detect the target.

(1) *MRF Principle.* Markov random fields [19] comprise a class of stochastic processes originally based on the Markov chain proposed by Russian mathematician Markov. In 1907, A. A. Markov described the characteristics of the process: under a given current state, the change of its future status will not rely on the previous state. The Brownian motion of

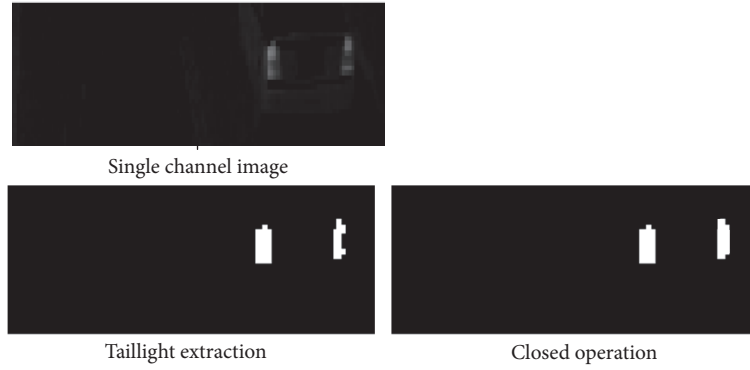


FIGURE 19: Extraction of taillight and closed operation.

particles in a liquid and a frog jumping in a lotus pond are Markov processes. According to the time and the state of being continuous or not, Markov processes can be divided into four types: (a) time and state are continuous, (b) time is continuous and state is discrete, (c) time is discrete and state is continuous, and (d) time and state are discrete. Here, the time and state are discrete Markov processes called a Markov chain.

A MRF model mainly includes a Markov property and a random field:

- (1) A Markov property is a given sequence of random variables arranged with time; the state of  $N + 1$  time of the sequence only has a relationship with the state of  $N$ .
- (2) In a random field, a value is assigned according to a certain distribution of each phase location. It mainly contains two elements: the location and phase space.

A 1D Markov stochastic process is a sequence of random variables  $X_1, X_2, \dots, X_n$ , where  $X_n$  represents the variables' state at the time  $n$ , and the value of all variables is called the "state space," which can be expressed using a probability distribution function:

$$P(V_{n+1} = x \mid X_0, X_1, X_2, \dots, X_n) = P(X_{n+1} = x \mid X_n), \tag{11}$$

where  $x$  represents a state in the stochastic process; note that (11) is the most basic property of the MRF model.

(2) *Relationship between MRF Model and an Image.* An image is a collection of points on a 2D plane and can also be regarded as a 2D Markov random field [20]. We set  $S = \{(i, j) \mid 1 \leq i \leq M, 1 \leq j \leq N\}$  to show the location of  $MN$ , that is, the location of the airport.  $\Lambda = \{1, 2, \dots, L\}$  represents the state space, that is, the phase space of the random field.  $X = \{x_s \mid s \in S\}$  represents a random field defined in  $\forall s \in S$ ,  $X_s$  is expressed in the random field  $X$ , and the state space is a random variable of  $\Lambda$ .

In the application of image processing, for the convenience of modeling, the MRF model introduces the concept of a system and group in the field, which is used to define the

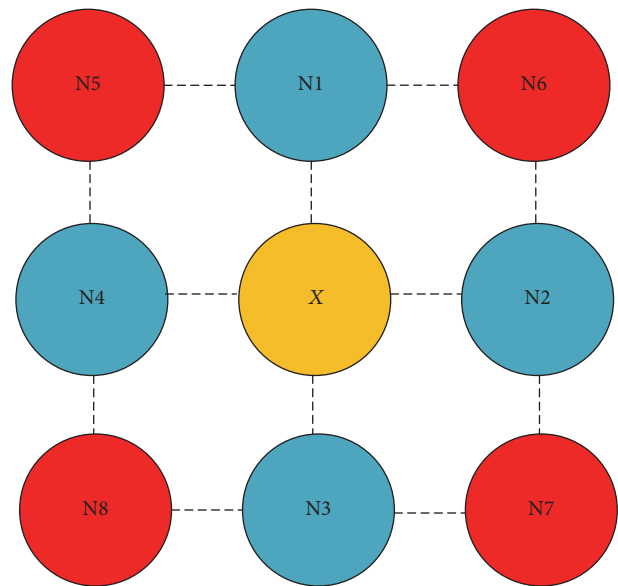


FIGURE 20: Field system.

relationship between a pixel and the surrounding pixels in the image. Common areas of the system include the first-order field system and the second-order system. The first-order system is the current pixel of the upper and lower and left and right four positions of pixels. In a second-order system, except for a first-order system in four positions of pixels and diagonal four pixels, as shown in Figure 20, a group is a subset of the realm.

(3) *Model Representation.* In using the MRF model, the concept of a tag is introduced. If  $L$  represents the entire tag set,  $P$  represents a frame of the image,  $p$  represents the image of an element, and  $L$  and  $P$  are mapping relations:  $f : P \rightarrow L$ ; that is, use every pixel  $p$  in the image to find the corresponding  $L$  values. If there are  $n$  pixels in  $P$  and  $M$  kinds of values in  $L$ , then  $f$  has  $Mn$  kinds of values.

We define the image of  $P$  of a set of random variables  $F = \{F_1, F_2, \dots, F_n\}$ , where  $F_i$  represents a value of  $L$  in the label set, and  $F_i = f_i$  denotes  $F_i$  values for  $f_i$  events. Events can be

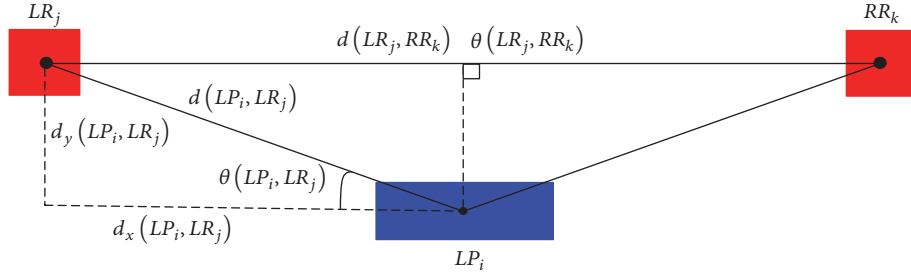


FIGURE 21: Relationship among graph nodes.

expressed as  $(F_1 = f_1, F_2 = f_2, \dots, F_n = f_n)$ , and supposing that  $f = \{f_1, f_2, \dots, f_n\}$ , the joint event can be abbreviated as  $F = f$ , the joint probability for  $P(f)$ .

Figure 20 shows a data structure that consists of a nonempty vertex set and multiple pairs of multiple relations. In this paper, we use the concept of graphs to construct the MRF model. We represent the vehicle license plate and taillights as a graph node, and the relationship between the graph nodes is shown in Figure 21. First, the license plate in the current frame is used as a node of the graph, and then the adjacent vehicle taillights are added as a candidate node.

$G = \{V, E\}$  represents a vehicle component model,  $V = \{v_1, v_2, \dots, v_n\}$  represents a vehicle license plate and taillight, and  $E = \{e_1, e_2, \dots, e_n\}$  represents the component relationship between the license plate and the taillight.  $G$  is a complete graph, and each pair of nodes has a side. Each node in the graph corresponds to a random variable  $F_i$ , which makes the tag set  $f = \{f_1, f_2, \dots, f_n\}$ , the value of  $f_i$  is  $D = \{0, 1, 2, 3\}$ , and there are four types of nodes in the MRF diagram presented in this paper:

- $d = 1$ , vehicle license plate node.
- $d = 2$ , left taillight node.
- $d = 3$ , right taillight node.
- $d = 0$ , detecting the wrong node.

In this paper, we use the probability distribution expression of the MRF model:

$$P(F) = \frac{1}{Z} \prod_{c \in C} \varphi_c(f_c) = \frac{1}{Z} \prod_{i \in V} \omega(f_i) \prod \psi(f_i, f_j), \quad (12)$$

$$Z = \sum_F \prod_{i \in V} \omega(f_i) \prod_{(i,j) \in E} \psi(f_i, f_j),$$

where  $Z$  is the normalized function,  $\omega(f_i)$  represents the possibility of nodes, representing the detection confidence of each node, and  $\psi(f_i, f_j)$  is the edge of the possibility of representing the relationship between the two nodes.

**3.4. Vehicle Tracking and Detection.** Vehicle tracking can be used to predict the position of the vehicle and to match the video image of the vehicle, and the Kalman filter [21] (KF) is suitable for vehicle tracking. For the daytime scene, we selected the license plate center and the vehicle velocity as the

state vector; for the nighttime scene, we selected the midpoint of the headlights and the vehicle velocity as the state vector:

$$x = [p_x, p_y, s_x, s_y]^t, \quad (13)$$

where  $p_x$  and  $p_y$  are the vehicle coordinates in the  $X$  and  $Y$  directions, respectively, and  $s_x$  and  $s_y$  are the speeds of the vehicle in the  $X$  and  $Y$  directions, respectively. Kalman filtering can be divided into two steps:

(1) *Prediction.* Here, we predict the state vector  $x$  and the state covariance matrix  $P$  of the current  $k$  time, as follows:

$$x_k = Fx_{k-1} \quad (14)$$

$$P_k = FP_{k-1}F^t + Q,$$

where  $x_{k-1}$  is the state vector at time  $k-1$  time,  $F$  is the state transition matrix,  $x_k$  is the state vector of the current time  $k$ ,  $P_k$  and  $P_{k-1}$  are the  $k$  moments and  $k-1$  moments of the covariance matrix, and  $Q$  is the process noise matrix.

(2) *Update.* Here, we select the nearest vehicle in the forecast position. If the predicted position and the current position of the vehicle are less than the set threshold values, the observed value is considered to be  $y_k$ . If the observed value is not accurate, then the update phase is skipped. The update phase can be described by the Kalman gain  $K_k$ :

$$K_k = P_k H^t (HP_k H^t + R)^{-1}, \quad (15)$$

where  $H$  is the measurement matrix and  $R$  is the measurement of the noise covariance matrix. We can then update the state vector and covariance matrix, as follows:

$$x_k = x_k + K_k (y_k - Hx_k) \quad (16)$$

$$P_k = (I - K_k H) P_k. \quad (17)$$

## 4. Experiment

**4.1. Nighttime Vehicle Detection Algorithm.** In Section 3.2.3, according to real-world vehicle specifications, in order to complete the two light components of the spatial relationship GMM modeling, we make use of the GMM in the detection process to join the space position and 3D size information.



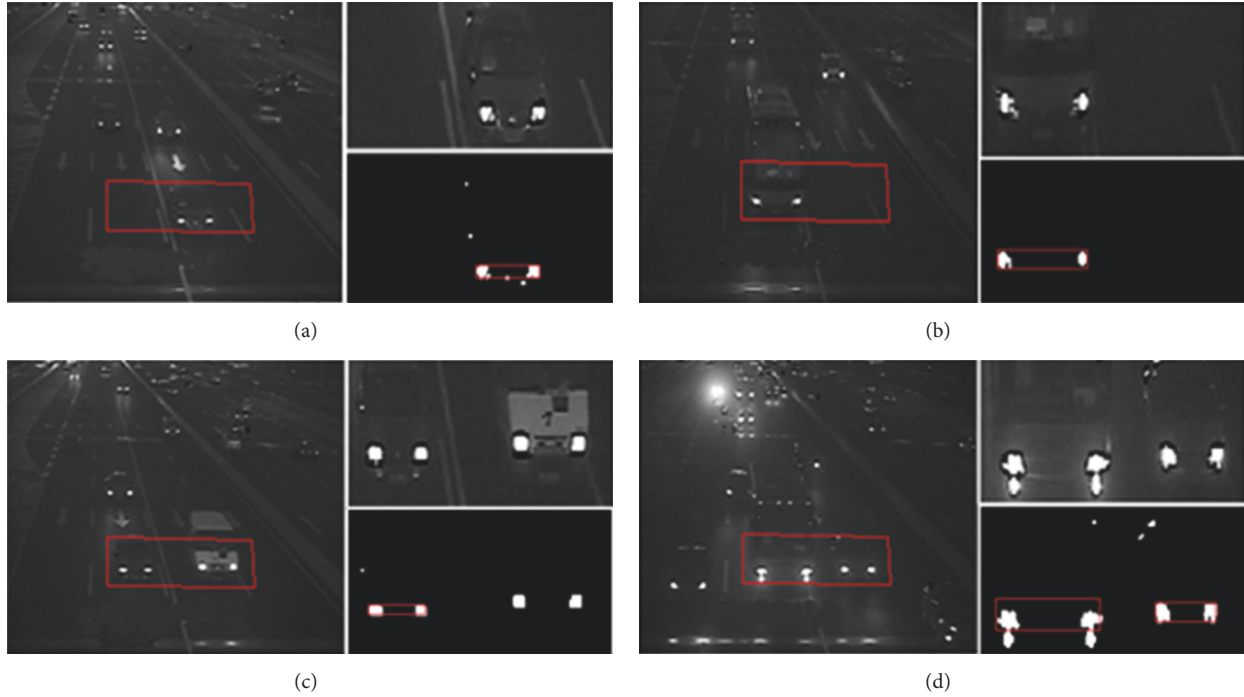


FIGURE 22: Nighttime vehicle detection results.

Adding the distance between the two headlights, our method can eliminate incorrect matching of the adjacent headlights on the same horizontal line. Adding the headlight heights to the space position information, our method can judge whether the vehicle front plane fits with inverse projection plane, to avoid the same vehicle multirecognition problem. The GMM of the headlights' spatial relationship contains the horizontal position, vertical position, and height information of the headlights; therefore, in the GMM further testing will obtain correct matching results for the headlights, thus allowing us to determine the representative headlights among the overall vehicle recognition results.

Assuming that the spatial probability model of the headlights is  $\Omega$ , the spatial variables relationship of the candidate headlight component is  $D = \{d\_X\_Diff, d\_Y\_Diff, d\_Y\_High\}$ , where  $P(D | \Omega)$  indicates that the detected headlight meets the probability of the established mode. The vehicle detection formula is

$$P(D | \Omega) > Tp, \quad (18)$$

where  $Tp$  is a preset threshold with atypical value of 0.2.

When the probability of the candidate headlight component is greater than a certain threshold in a given space probability model, then we can take these candidate headlight parts to be the target vehicle. The nighttime vehicle detection results are shown in Figure 22.

**4.2. Daytime Vehicle Detection Algorithm.** After building the model as described in Section 3.3.3, the probability distribution function must be solved according to the probability that

the confidence level [22] of the node and the edge can further detect the target.

**4.2.1. Energy Function of Node.** The energy function of the node is expressed as

$$\omega(f_i = p) = \begin{cases} \frac{1}{1 + \exp\{A_p \times S_i + B_p\}}, & \text{node is detected,} \\ 1 - \frac{1}{1 + \exp\{A_p \times S_i + B_p\}}, & p = 0, \\ \lambda, & \text{otherwise,} \end{cases} \quad (19)$$

where  $p$  is the probability of the current node, the parameters  $A_p$  and  $B_p$  are obtained by learning the  $S$  function ( $A_p = -1$  and  $B_p = 0$ ),  $\lambda = 0.01$ , and the taillight component detection factor  $p$  is 0.8.

**4.2.2. Energy Function of Edge.** The energy function of the edge is expressed as

$$\psi(f_i = p, f_i = q) = \begin{cases} \sum_{k=1}^M \alpha_k P_k(x | \theta_k), & p \neq 0, q \neq 0, p \neq q, \\ \frac{1}{2} \left( 1 - \sum_{k=1}^M \alpha_k P_k(x | \theta_k) \right), & p = 0 \text{ or } q = 0, p \neq q, \\ \lambda, & p = q. \end{cases} \quad (20)$$

Using the GMM method to determine the angle between the taillight and the license plate, as well as the horizontal

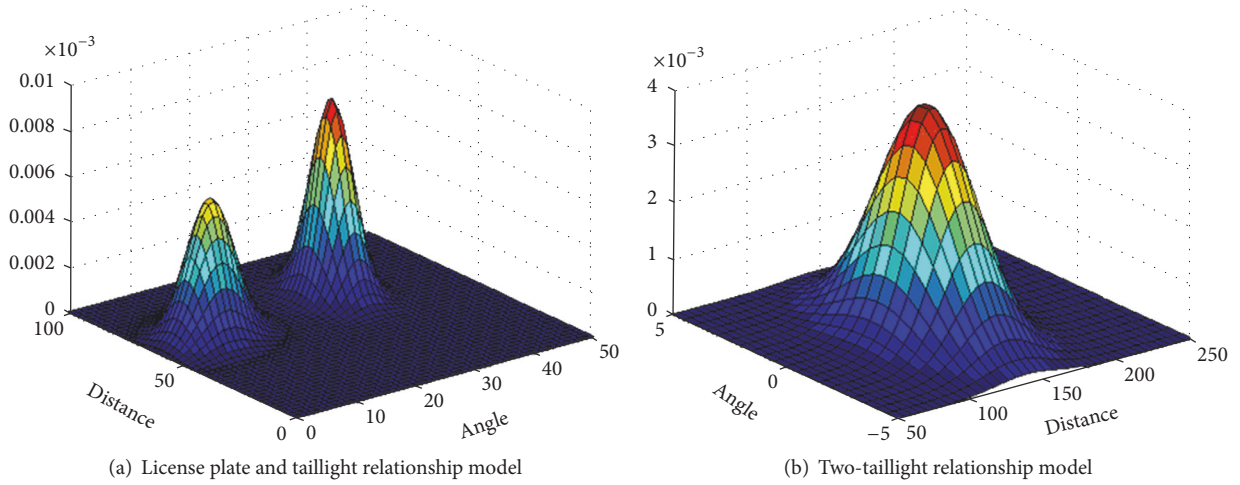


FIGURE 23

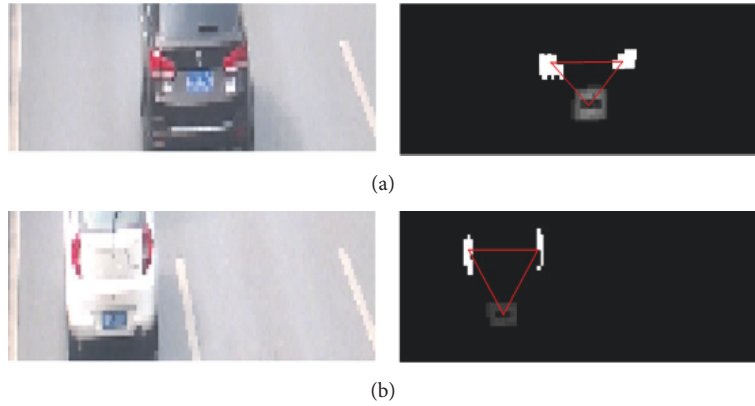


FIGURE 24: Daytime vehicle detection results.

angle between the two taillights, the distribution of the taillight and license plate is obtained and shown in Figure 23. In the expression parameter training process, the Expectation Maximization algorithm [23] is used to estimate the GMM parameters.

The confidence of nodes and edges determines the MRF. In this paper, the maximum a posteriori (MAP) can be calculated by all labels  $f_i \in F$  that can meet the MAP:

$$\begin{aligned} Q^* &= \arg \max P(F) \\ &= \arg \max \prod_{i \in V} \omega(f_i) \prod_{(i,j) \in E} \psi(f_i, f_j). \end{aligned} \quad (21)$$

Pearl [24] proposed a probabilistic inference method to solve this problem. For an acyclic graph, this method can obtain a precise solution; for a cyclic graph, it can obtain an approximate solution. After calculation, each node in the model has the best label and the target vehicle therefore can be detected, as shown in Figure 24.

If it is a blue car, detection is handled as follows: through analysis, after color model transformation, the gray gradient

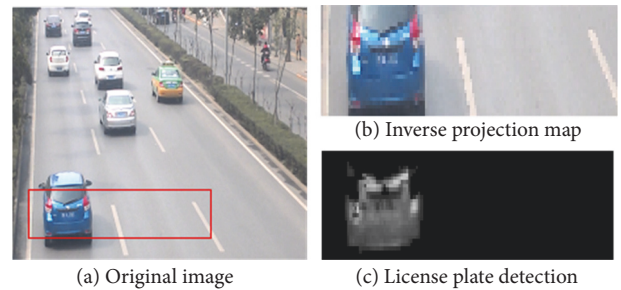


FIGURE 25: Processing method for blue car.

of the blue body is significantly enhanced, and there is a large target area on the transformation image.

According to the results of several experiments, if the total area of the target is greater than the total area of inverse projection map, which is 1/5, we can view the target as the rear of the vehicle, as shown in Figure 25.

Furthermore, the vehicle taillight is red, and the threshold segmentation results are shown in Figure 26. Note that the

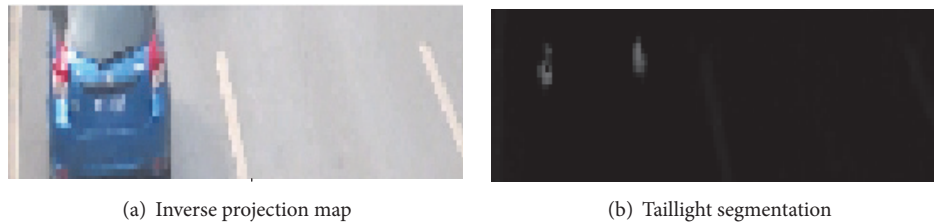


FIGURE 26: Taillights of blue car.

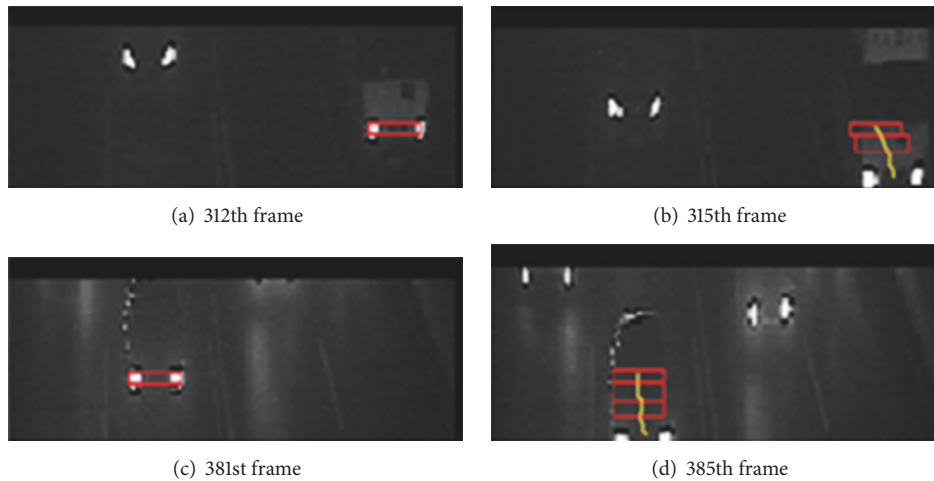


FIGURE 27: Tracking results.

taillight area brightness is more obvious. After extraction of the target area and analysis of the two regions, because the body and the taillights belong to a vehicle target, if the two regions intersect it is judged to be a vehicle target.

**4.3. Tracking and Testing Experiment.** In the nighttime scene, we selected the midpoint of the headlights' connection and vehicle speed as the state vectors, and the experimental results are shown in Figure 27. Figures 27(a) and 27(c) depict detection of the vehicle starting; Figures 27(b) and 27(d) show the yellow line for trajectory tracking.

## 5. Conclusions

This paper mainly introduces a method for realizing the recognition of a target component by using the spatial relationship model. The recognition result of the target was replaced by the result of the recognition of some parts of the target. In the real world, some inherent characteristics of the spatial relationship among the components of the target exist that would enhance the ability to describe and identify the content of the target. Based on the components of the vehicle detection algorithm, and by using multiple local components, a better detection effect was realized, but some missed and false detection phenomena still existed. Therefore, further improvement of the proposed method is required, mainly in the following aspects:

- (1) In the evening, detecting the headlight targets using the GMM, if only one headlight is turned on, or none are turned on, a detection error would be caused. Since the model is established based on the amount of statistical samples, more samples are needed in order to adapt the method to more models.
- (2) In the daytime, the vehicle is detected by the taillights and license plate color feature of the target. If the license plates are the same color as the car body, the detection results would be affected. Similarly, during taillight detection, a red-colored vehicle body would affect the test results. Therefore, the detection algorithm is still relatively imprecise and must be improved.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was funded by the Project of Shaanxi Provincial Science and Technology Program (Grant nos. 2014JM8351, 2015JZ018, and 2016JQ6011), the Fundamental Research Funds for the Central Universities (Grant nos. 2013G1241109 and 310824173603), and the National Natural Science Foundation of China (Grant nos. 61501058 and 61572083).

## References

- [1] S. Ozawa, "Image processing for intelligent transport systems," *IEICE Transaction on Information and Systems*, vol. 82, no. 3, pp. 629–636, 1999.
- [2] B. Tian, B. T. Morris, M. Tang et al., "Hierarchical and networked vehicle surveillance in ITS: a survey," *IEEE Intelligent Transportation Systems Magazine*, vol. 16, no. 2, pp. 557–580, 2015.
- [3] N. Buch, S. A. Velastin, and J. Orwell, "A review of computer vision techniques for the analysis of urban traffic," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 3, pp. 920–939, 2011.
- [4] S. Sivaraman and M. M. Trivedi, "Looking at vehicles on the road: a survey of vision-based vehicle detection, tracking, and behavior analysis," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 4, pp. 1773–1795, 2013.
- [5] H.-S. Song, S.-N. Lu, X. Ma, Y. Yang, X.-Q. Liu, and P. Zhang, "Vehicle behavior analysis using target motion trajectories," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 8, pp. 3580–3591, 2014.
- [6] S. Wender and K. C. J. Dietmayer, "3D vehicle detection using a laser scanner and a video camera," *IET Intelligent Transport Systems*, vol. 2, no. 2, pp. 105–112, 2008.
- [7] H. Yue, J. Wu, Y. Cao, and Z. Cui, "Research on moving vehicle detection in the presence of occlusion," in *Proceedings of the 9th International Symposium on Distributed Computing and Applications to Business, Engineering and Science (DCABES '10)*, pp. 514–517, Hong Kong, China, August 2010.
- [8] T.-Y. Lee, J.-S. Oh, and J.-H. Kim, "On-road vehicle detection based on appearance features for autonomous vehicles," in *Proceedings of the 15th International Conference on Control, Automation and Systems (ICCAS '15)*, pp. 1720–1723, Busan, South Korea, October 2015.
- [9] J. Kim, J. Baek, Y. Park, and E. Kim, "New vehicle detection method with aspect ratio estimation for hypothesized windows," *Sensors*, vol. 15, no. 12, pp. 30927–30941, 2015.
- [10] N. Kosaka, R. Ogura, and G. Ohashi, "Offline vehicle detection at night using center surround extremas," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E98A, no. 8, pp. 1727–1734, 2015.
- [11] C. R. Wang and J. J. Lien, "Automatic vehicle detection using local features—a statistical approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 9, no. 1, pp. 83–96, 2008.
- [12] Y. I. Abdel-Aziz and H. M. Karara, "Direct linear transformation into object space coordinates in close-range photogrammetry," in *Proceedings of the Symposium on Close-Range Photogrammetry*, pp. 1–18, 1971.
- [13] H. Li, M. Feng, and X. Wang, "Inverse perspective mapping based Urban road markings detection," in *Proceedings of the 2nd IEEE International Conference on Cloud Computing and Intelligence Systems (CCIS '12)*, pp. 1178–1182, Hangzhou, China, November 2012.
- [14] K. Dai, G. Li, and et al., "Prospects and current studies on background subtraction techniques for moving objects detection from surveillance video," *Journal of Image and Graphics*, vol. 11, no. 7, pp. 919–927, 2006.
- [15] W. Wang, C. Shen, J. Zhang, and S. Paisitkriangkrai, "A two-layer night-time vehicle detector," in *Proceedings of the Digital Image Computing: Techniques and Applications (DICTA '09)*, pp. 162–167, Melbourne, Australia, December 2009.
- [16] G. D. Finlayson and E. Trezzi, "Shades of gray and colour constancy," in *Proceedings of the Twelfth Color Imaging Conference: Color Science and Engineering Systems, Technologies, and Applications (CIC '04)*, pp. 37–41, Scottsdale, Ariz, USA, November 2004.
- [17] A. Neubeck and L. Van Gool, "Efficient non-maximum suppression," in *Proceedings of the 18th International Conference on Pattern Recognition (ICPR '06)*, pp. 850–855, Hong Kong, China, August 2006.
- [18] W. Li, D. Liang, X. Wang et al., "Segmentation method for car emblem based on texture homogeneity measure," *Application Research of Computers*, vol. 21, no. 10, pp. 141–142, 2004.
- [19] X. Li, S. Zhu et al., "A survey of the markov random field method for image segmentation," *Journal of Image and Graphics*, pp. 789–798, 2007.
- [20] C. Jiannong and W. Zhanhong, "Review and comparison of MRF and DMN methods in image segmentation," *Computer Engineering and Applications*, vol. 42, no. 20, pp. 19–24, 2006.
- [21] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.
- [22] B. Tian, Y. Li, B. Li, and D. Wen, "Rear-view vehicle detection and tracking by combining multiple parts for complex Urban surveillance," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 2, pp. 597–606, 2014.
- [23] H. Hong and D. Schonfeld, "Maximum-entropy expectation-maximization algorithm for image reconstruction and sensor field estimation," *IEEE Transactions on Image Processing*, vol. 17, no. 6, pp. 897–907, 2008.
- [24] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann Publishers Inc, San Francisco, Calif, USA, 1988.





# Hindawi

Submit your manuscripts at  
<https://www.hindawi.com>

