

## Real-time risk assessment of road vehicles based on inverse perspective mapping

Qin Shi<sup>a,\*</sup>, Yan Chen<sup>a</sup>, Haoxiang Liang<sup>b</sup>

<sup>a</sup> School of Information Engineering Chang'an University, Xi'an, ShaanXi Province China

<sup>b</sup> School of Electronic and Control Engineering, Chang'an University, Xi'an, ShaanXi Province China

### ARTICLE INFO

**Keywords:**

Camera calibration  
Vehicle speed  
Vanishing point  
Marking line  
Traffic monitoring  
Real-time road risk assessment

### ABSTRACT

Pan/Tilt/Zoom (PTZ) cameras play an important role in traffic scenes due to their wide monitoring fields and high flexibility. However, since the focal length and angle of PTZ cameras change irregularly with the monitoring needs, it is difficult to obtain accurate physical information about the real world from the image information of PTZ cameras. Aiming to address the need for real-time risk assessment of road vehicles in traffic monitoring scenarios, a vehicle position and velocity measurement scheme based on camera inverse perspective transformation is proposed, along with a method for real-time risk assessment based on the position and velocity. Specifically, Firstly, the vehicle target in the video is detected and tracked by deep learning YOLO detection algorithm and optical flow tracking algorithm. According to the obtained trajectory set, the vanishing points in the road direction are calculated by Cascade Hough Transform and the road marking lines are detected. Then, according to the vanishing point and marking line, the camera calibration task is accomplished via exploratory focal length. After camera calibration, the camera-to-road inverse perspective transformation is applied to project the image plane onto the road surface and obtain, the actual position information of vehicles. Finally, the vehicle speed measurement and real-time road risk assessment are achieved by calculating the average of instantaneous velocities across multiple frames. Simulation experiment results in a traffic monitoring scenario demonstrate that this perspective-based method for real-time road vehicle risk assessment achieves good stability and practicality, which meets the requirements for vehicle speed measurement and real-time road risk assessment.

### 1. Introduction

Intelligent traffic monitoring systems have been widely used in traffic scenarios such as highways, urban roads, tunnels and bridges. And their widespread applications are benefit from the integration and development of various technologies inevitably, such as pattern recognition, video image processing and network communication [1,2]. Meanwhile, intelligent traffic monitoring systems also suffer from many significant challenges in extremely complex traffic monitoring scenarios as well as special weather conditions.

The most challenging issue is how to obtain real-time risk assessment of vehicles. Currently, road traffic safety analysis primarily relies on the regression analysis for predicting road traffic accidents. Regression analysis can effectively reflect the causal relationship between road accidents and various influencing factors. However, regression analysis requires a large sample size, minimal data fluctuations, and strong

regularities, which has great effects on prediction accuracy. Furthermore, the analysis methods based on road traffic accidents are retrospective and cannot respond promptly to the unforeseen road incidents.

Therefore, this study proposes a real-time road vehicle risk analysis method based on camera inverse perspective transformation that considers complex traffic scenarios. Firstly, camera calibration is performed and then the camera-to-road inverse perspective transformation is utilized to obtain the actual position information of vehicles by projecting the image plane onto the road surface. Finally, vehicle speed measurement and real-time risk assessment are achieved by calculating the average of instantaneous velocities across multiple frames. This method determines the vehicle risk value by analyzing the relative speeds and distances between vehicles. By aggregating the vehicle risk values, real-time road risk assessment can be obtained.

Recently, there exist lots of work to address this issue. Specifically, the speed measurement system can be divided into intrusive and non-

\* Corresponding author.

E-mail address: [qinshi@chd.edu.cn](mailto:qinshi@chd.edu.cn) (Q. Shi).

intrusive [3]. The intrusive speed measurement is based on the speed measurement method with the induction loop detector, which is widely used in many countries, but more difficult to install and maintain and can even accelerate problems such as pavement deterioration. The non-intrusive speed measurement can avoid the above-mentioned problems, such as radar speed measurement, laser speed measurement [4]. However, the price of this method is relatively high. In recent years, with the camera technology and video image processing technology continuing to mature and develop [5], video-based speed measurement method becomes a part of non-intrusive speed detection, reducing system maintenance costs, realizing easy and simple installation and better detection accuracy, and meeting the actual demand. Besides, the existing highways have been installed a large number of monitoring devices, thus video vehicle detection technology will be the trend of real-time traffic information acquisition and processing. An important part of the video-based velocity measurement approach is camera calibration [6,7], an important process for evaluating and acquiring camera parameters, which enables the mapping of pixel information from camera images to real-world coordinates. A wide range of camera calibrations are also used in 3D reconstruction, robotic navigation and scene measurement. In recent years, PTZ cameras have been widely used in intelligent surveillance systems to compensate for the narrow field of view and low flexibility of fixed camera monitoring. The calibration method based on PTZ cameras has been widely researched. Camera calibration in real traffic scenarios is more difficult than traditional camera calibration, for the fact that it is impossible to acquire manual information about the calibration scene. Therefore, many traditional calibration methods are difficult to be applied in the actual scenario. Zhang et al. [8] achieved camera calibration by acquiring different directions of vanishing points based on a variety of road signs. Schoepflin et al. [9] dynamically calibrated PTZ cameras using lane lines in active areas of the road. Song et al. [10] used edge detection to extract lane lines from a static background image and then estimate the vanishing point assuming camera height and lane width are known.

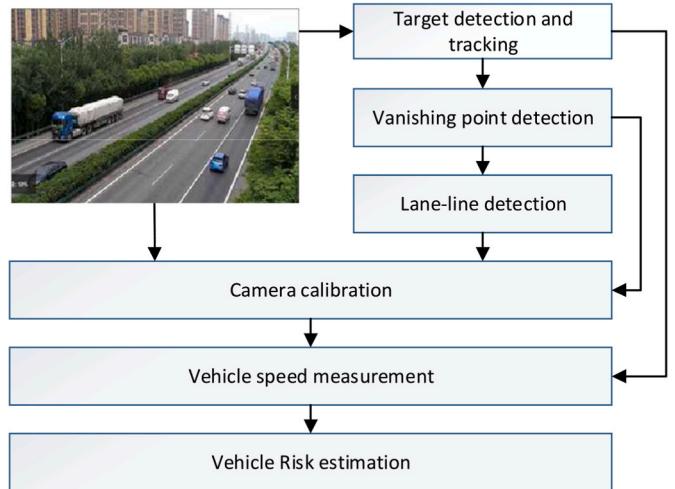
Though the above methods can deal with the issue well in some extent, there are some major limitations in the practical use of the above methods. For instance, some methods need human intervention, while others need to get the intrinsic parameters or the installation height of cameras in advance. However, in the actual highway scene, the installation height of cameras are relatively large, which makes it difficult to obtain the horizontal edges of the vehicle. In addition, even if the horizontal edges of the vehicle could be obtained, the horizontal vanishing point is unstable due to its short length. Therefore, the application of the above method in the highway scene is limited.

In this paper, an automatic calibration method for PTZ cameras is proposed based on the constraints of vanishing point and lane line model. And the camera calibration matrix is obtained by calculating the longitudinal vanishing point, lane marking line and camera height. The camera calibration in road scenarios is accomplished with the vanishing point calibration method according to the actual requirements of traffic scenarios, which lays the foundation for speed measurement.

Based on the camera calibration results, a speed measurement algorithm is proposed in this paper. The algorithm framework is shown in Fig. 1, which can detect and track vehicle targets, extract vanishing points, lane lines, achieve automatic camera calibration, automatic measurement of vehicle speed without human assistance.

To sum up, the contributions are listed as follows.

- A lane line detection method via pixel variance in combination with vanishing points is proposed. And the optimal PTZ camera calibration parameters are automatically obtained by matching the lane line model with the detected lane lines.
- A real-time vehicle speed measurement algorithm based on camera inverse perspective mapping is proposed.
- Utilizing vehicles' real-time speed and position relationship, a method for evaluating the level of driving risks is introduced, which



**Fig. 1.** Processing of the algorithm frame.

enables quantitative analysis of road traffic safety factors and prevention of traffic accidents.

## 2. Camera calibration

### 2.1. Vehicle detection and track extraction

With the maturing of AI technology, object detection using deep learning convolutional neural networks is gradually becoming mainstream. And convolutional neural network methods ranging from two-stage representative methods RCNN [11] and Fast R-CNN [12] to Faster R-CNN [13], to one-stage representative methods YOLO [14–17] and SSD [18] have been gradually implemented to the task of real-time target detection. Additionally, it also includes image-based 2D and 3D object detection algorithms [19,20]. In this paper, the YOLO algorithm that can achieve real-time target detection is selected to detect the vehicle targets in highway scenes. The detection frame is output for each frame of the vehicle target in the video stream. At the same time, the center coordinate point, pixel length and width, and vehicle classification of each vehicle target are output. The YOLO algorithm treats target detection as a single regression, which includes dividing the grid, predicting the target window and confidence level in each grid, then removing the target window with lower confidence level based on the set threshold, and finally eliminating duplicate detection frames by non-maximal suppression (NMS) merging. The method can achieve a detection speed of 45 frames per second on the TitanX series GPUs, and 155 frames per second on the lightweight version, which is still a great improvement in accuracy compared to R-CNN. The results of vehicle object detection using the YOLO detection algorithm are shown in Fig. 2 (a).

(a) Vehicle detection (b) Vehicle optical flow tracking

To extract feature points in the vehicle object detection frame, this paper selects Harris feature points to extract the vehicle feature information within the detection frame, as the starting point of the optical flow tracking. Harris corner information detection is a very classical corner point extraction algorithm, which can be very good for the vehicle target corner information extraction. Then target tracking is performed for the extracted feature points using optical flow tracking algorithm [21]. In the optical flow tracking process, all the vehicle trajectories are collected, forming a collection of trajectories  $T_0$ , in which optical flow tracking is performed for the highway vehicles. The results are shown in Fig. 2 (b).



**Fig. 2.** Vehicle detection and optical flow tracking.

## 2.2. Vanishing point detection

The accuracy of vanishing point detection is an important task for vanishing point-based camera calibration. In this paper, we use vehicle trajectory to determine the vanishing point  $vp_0(u_0, v_0)$  along the road direction. The analysis of the set of vehicle linear trajectories  $T_0$  reveals that most of the vehicles will travel along the lanes, and thus most of the linear sets  $T_0$  are parallel to each other in 3D space. In other words, most of the linear trajectories intersect at one point in the image plane. In this paper, a cascading Hoff transform polling method [22] is used to detect vanishing points along a road, which can present each line  $(a, b, c)$  in image space in a finite diamond space, i.e., from a line to a point and then from a point to a polyline segment through two Cartesian to parallel line coordinate system transformations, where the polyline segment is represented by four sets of coordinate points. The change is shown in equation (1). Finally, the traditional Hoff transform voting method is used to obtain the vanishing point. The detection results of the vanishing point along the road are shown in Fig. 3, where the maximum value of the voting result in the diamond space is the vanishing point  $vp_0(u_0, v_0)$ .

$$(a, b, c) \rightarrow \left[ \frac{\alpha a}{c + \gamma a}, \frac{-\alpha c}{c + \gamma a} \right], \left[ \frac{b}{c + \beta b}, 0 \right], \left[ 0, \frac{b}{a + \alpha b} \right], \left[ \frac{-\alpha a}{c + \gamma a}, \frac{\alpha c}{c + \gamma a} \right] \quad (1)$$

Where  $\alpha = \text{sgn}(ab), \beta = \text{sgn}(bc), \gamma = \text{sgn}(ac)$ .

Various marking lines such as broken and long solid lines are usually distributed on highways [23]. In order to complete the camera calibration in Section 2.3, the broken and solid lines on the road need to be detected accurately. The solid lines are white and the broken lines are blue or gray (i.e., the color of the road surface). With this distribution law and the detected vanishing point  $vp_0(u_0, v_0)$ , the broken and solid lines can be detected.

Determine a horizontal line  $L$  over the vanishing point  $vp_0(u_0, v_0)$ . Set the color of the line to a gray similar to the road surface. Take the part

below the horizontal line of the image as the area to be detected. Then rotate the line  $L$  over the vanishing point clockwise at a certain angle, from  $0^\circ$  to  $180^\circ$ . Record the trajectory of each rotation and calculate the variance between the grayscale information of all the pixels on the trajectory and the grayscale pixel information of the image where the rotated line is located. For the rotation line through the long solid line or road grayscale plane, the variance of all its pixels usually shows less change, but stable state. For the rotation line through the broken line, due to the alternate distribution of the interval of the broken line, the variance of its pixels will also show some alternate distribution of the state. For the rotation line through the non-road plane, its pixels the passage of variance will show an irregular distribution state due to the road sides. The straight line trajectories showing alternating distribution states are filtered out and the image at the junction of the broken line on each line is identified as  $p = \{(x_{if}, y_{if}), (x_{ib}, y_{ib}), i = 1 \dots n\}$ .

## 2.3. Camera modeling and parameter solving

The camera calibration is a mapping between the actual road plane and the image plane based on the camera pinhole model [24], which consists of an in-camera and an out-camera reference. The in-camera reference describes the transformation between the image coordinate system and the camera coordinate system, including parameters such as focal length, tilt angle, and aspect ratio. And the out-camera reference describes the spatial transformation relationship between the camera coordinate system and the world coordinate system, including a rotation matrix and a translation vector. The overall mapping from the world space coordinate system  $P(x, y, z)$  to the image pixel coordinate system  $p(u, v)$  can be represented by equation (2).

$$sp = HP = KRTP \quad (2)$$

Where  $K$  is the camera's internal parameter,  $R$  is the rotation matrix,  $T$  is



**Fig. 3.** Vanishing point along the road direction.

the translation vector, and  $s$  is the scale factor. The KRT can be represented by the  $H_{3 \times 4}$ , where  $H_{3 \times 4}$  is the mapping matrix.

In order to complete the calibration of the camera in the traffic surveillance scenario, the road calibration model is reasonably simplified in this paper, assuming that the main point of the camera is coincident with the image center, and the calibration process is not affected by the camera distortion. And the camera's spin angle can be kept to zero by rotating the image. The road calibration model is shown in Fig. 4. The camera is set up on a pillar beside the road, at the height of  $h$ . The camera coordinate system  $O_c$  is set up with the camera center, and the world coordinate system  $O_w$  is set up directly below the camera. The angle between the camera coordinate system  $Z_c$  and the road plane is  $\varphi$ , and the angle between the plane  $Z_w O_w Z_c$  and the  $Y_w$  axis is  $\theta$ . The camera coordinate system and the world coordinate system are both right-handed coordinate systems.

Based on this road camera calibration model, the internal parameter matrix can be determined as equation (3).

$$K = \begin{bmatrix} f & 0 & C_x \\ 0 & f & C_y \\ 0 & 0 & 1 \end{bmatrix} \quad (3)$$

Where  $f$  is the focal length of the camera in pixels, and  $(C_x, C_y)$  is the principal point position which is a known parameter.

$$R = R_x R_y \quad (4)$$

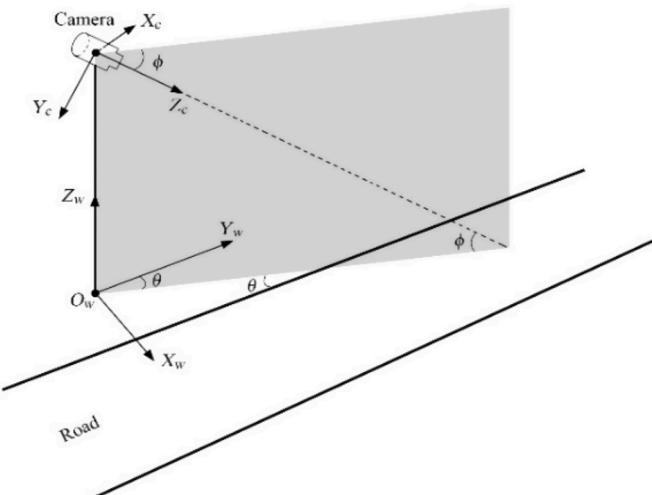
Where the rotation matrix  $R$  is the combination of two parts of the world coordinate system that rotate first clockwise  $\theta$  around the  $Z$  axis and then counterclockwise  $\varphi + \pi/2$  around the  $X$  axis. The camera height is known to be  $h$ , so the translation vector  $T$  is defined as the origin moving  $h$  distance along the  $Z$  axis of the world coordinate system to the camera origin, which can be expressed as.

$$T = \begin{bmatrix} 0 \\ 0 \\ -h \end{bmatrix} \quad (5)$$

$$\text{Thus, } H_{3 \times 4} = \begin{bmatrix} f \cos \theta & -f \sin \theta & 0 & 0 \\ f \sin \theta \sin \varphi & -f \sin \varphi \cos \theta & -f \cos \varphi & fh \cos \varphi \\ -\sin \theta \cos \varphi & \cos \theta \cos \varphi & -\sin \varphi & h \sin \varphi \end{bmatrix}.$$

In particular, the three-dimensional point  $(x, y, 0)$  on the road surface can be substituted into Equation (2) to obtain the corresponding image coordinates  $(u, v)$ :

$$u = \frac{su}{s} = \frac{X_w H_{11} + Y_w H_{12}}{X_w H_{31} + Y_w H_{32} + H_{34}} \quad (6)$$



**Fig. 4.** Road camera calibration model.

$$v = \frac{sv}{s} = \frac{-X_w H_{12} + Y_w H_{22} + H_{24}}{X_w H_{31} + Y_w H_{32} + H_{34}} \quad (7)$$

The calibration based on vanishing point in traffic scenes can be roughly divided into single, double and triple vanishing point calibration. The double vanishing point calibration [25] shows that the determination of the camera's internal reference and rotation matrix is directly related to the vanishing point, i.e., the focal length  $f$ , pitch angle  $\varphi$  and yaw angle  $\theta$  can be determined by two vanishing points, as shown in equation (8)-equation (10), where  $vp_0(u_0, v_0)$  refers to the vanishing point along the road direction and  $vp_1(u_1, v_1)$  refers to the vanishing point perpendicular to the road direction. Finally, camera height  $h$  can be added to complete the camera calibration.

$$f = \sqrt{-(v_0^2 + u_0 u_1)} \quad (8)$$

$$\varphi = \arctan\left(\frac{-v_0}{f}\right) \quad (9)$$

$$\theta = \arctan\left(\frac{-u_0 \cos \varphi}{f}\right) \quad (10)$$

In practical experiments, it is possible to determine the vanishing point  $vp_0(u_0, v_0)$  along the road direction with the method in Section 2.2. However, it is difficult to extract the vanishing point  $vp_1(u_1, v_1)$  in this direction due to the low amount of information in the vertical road direction. Observational equations (8)-(10) reveal that when the focal length  $f$  is determined, pitch angle  $\varphi$  and yaw angle  $\theta$  are only related to the vanishing point  $vp_0(u_0, v_0)$  but not to the vanishing point  $vp_1(u_1, v_1)$ . Therefore, the enumeration method based on the lane line model is used to obtain the focal length  $f$  so as to achieve the camera calibration. The idea is to assume an initial focal length  $f_0$  based on prior knowledge. Then, through this initial focal length  $f_0$  and vanishing point  $vp_0(u_0, v_0)$ , pitch angle  $\varphi$  and yaw angle  $\theta$  can be calculated. These parameters can determine the mapping matrix  $H_{3 \times 4}$ , through which the points in the world coordinate system are projected onto the image coordinate system, comparing the projection points with the actual detected points, and then the focal length  $f$  being adjusted so that the projection points coincide with the actual detected points. In this way the focal length is obtained.

The specific enumeration algorithm contains 4 steps.

- 1 Based on the filtered trajectory set  $T$ , determine the vanishing point  $vp_0(u_0, v_0)$ , and determine the coordinates of the real and imaginary lane line junction by pixel variance method. The actual lane real and imaginary line junction coordinates are expressed as  $p = \{(x_{if}, y_{if}), (x_{ib}, y_{ib}), i = 1 \dots n\}$ . Details are shown in Section 2.2.
- 2 Set the initial focal length  $f_0$ , and then calculate the pitch angle  $\varphi$  and yaw angle  $\theta$  at this focal length by equations (9) and (10). Then the mapping matrix  $H_{3 \times 4}$  is determined.
- 3 Establish the lane line model  $p' = \{(x_{jf}, y_{jf}), (x_{jb}, y_{jb}), j = 1 \dots n\}$ . The broken lane lines on Chinese highways meet the layout specification of "6 solid, 9 blank", that is, the length of the solid line in the broken line is 6 m, and the length of the blank part is 9 m. Select the first real line segment starting point A in Step1, and calculate the coordinates of point A in the world coordinate system according to formula (6) and formula (7), that is, point A1 in Fig. 5. Combined with the actual size of the lane line in the world coordinate system, calculate the coordinates of the



**Fig. 5.** Diagram of the road model.

other points of the lane line. The results are shown in [Table 1](#). Then according to [formula \(6\)](#) and [\(7\)](#), the coordinates  $p' = \{(x_{jf}, y_{jf}), (x_{jb}, y_{jb}), j = 1 \dots n\}$  in the image coordinate system can be calculated.

4. By adjusting the focal length  $f_o$  and calculating the corresponding pitch angle  $\varphi$  and yaw angle  $\theta$ , it is possible to match the actual lane broken line junction coordinate  $p = \{(x_{if}, y_{if}), (x_{ib}, y_{ib}), i = 1 \dots n\}$  with the 2D lane line model  $p' = \{(x_{jf}, y_{jf}), (x_{jb}, y_{jb}), j = 1 \dots n\}$ . When the distance  $D$  in equation [\(11\)](#) reaches the minimum value, the matching of  $p$  and  $p'$  is completed, thus establishing the final camera focal length  $f$ , pitch angle  $\varphi$  and yaw angle  $\theta$ , and completing the calibration of the camera in the final traffic monitoring scenario.

$$D = \sum_{i,j=1}^n |p_i p'_j| \quad (11)$$

In equation [\(11\)](#),  $D$  is the sum of the actual detected lane line and the distance to the corresponding point of the lane line calculated from the parameters obtained from the test focal length. If the value of  $D$  decreases when the focal length is increased, the focal length continues to increase, otherwise it decreases; if the value of  $D$  decreases when the focal length is decreased, the focal length continues to decrease, otherwise it increases. Adjust the focal length value until the  $D$  value is minimum, then the camera calibration is completed.

### 3. Vehicle speed measurement

#### 3.1. Model assumption

All positions in the road monitoring image can be mapped to the world coordinate system satisfying the condition  $Z_w = 0$  by camera calibration. However, the accurate measurement of vehicle speed not only depends on the camera calibration but also has a great relationship with the trajectory of the vehicle. In order to better realize the vehicle speed measurement, the following assumptions on the speed model are made: (1) In the highway scenario, the road is relatively flat, that is, satisfying the condition  $Z_w = 0$ , and there is no big change of road surface undulation. (2) The vehicles in the surveillance area of the highway are traveling in a straight line, and there is no substantial arbitrary lane change. (3) The time interval between each frame of the video surveillance on the highway is the same. In this way, this paper can calculate the vehicle speed by the time interval between frames after obtaining the exact vehicle location [\[26–28\]](#).

#### 3.2. Model design and implementation

Based on the above assumptions of the speed model, the target vehicle speed detection scheme is established, as shown in [Fig. 6](#). We use  $s_i$  represents the true physical distance on the road plane,  $d_i$  represents the projection of this distance on the image plane. The detailed process of speed detection is now explained. First, the target detection and camera calibration are completed in Section 2. Using the YOLO target detection algorithm, the coordinates of the upper left corner of the image detection frame can be obtained. And the coordinates of the bottom edge center of the detection frame can be obtained through the length and width of the detection frame, which can more reliably ensure that the vehicle speed is close to the real speed. For each target vehicle in each frame of the video stream, a set of vector relations can be obtained, as in equation [\(12\)](#), where  $u_i(t)$  denotes the coordinates of the lower bottom edge center of the vehicle target detection frame in the current video frame,  $u_i(t - \Delta t)$  denotes the coordinates of the lower bottom edge center of the vehicle target detection frame in the previous frame,  $\Delta t$  is the time interval between the two frames, and  $i = \{1, 2, \dots, n\}$  denotes the track point of the track (see [Fig. 7](#)).

**Table 1**  
World coordinates of lane line point (in metres).

Point	x	y	z
A1	$xw$	$yw$	0
A2	$xw$	$yw + 6$	0
A3	$xw$	$yw + 15$	0
A 4	$xw$	$yw + 21$	0
A 5	$xw$	$yw + 30$	0
A 6	$xw$	$yw + 36$	0

$$\vec{d}_i = u_i(t) - u_i(t - \Delta t) \quad (12)$$

The matrix  $H_{3 \times 4}$  created by the camera calibration can be used to map  $u_i(t)$  and  $u_i(t - \Delta t)$  to the world coordinate system, to represent the real distance that the target vehicle moves from the previous frame to the current frame, as shown in [Eq. \(13\)](#).  $\|\vec{s}_i\|$ , in meters, is the Euclidean paradigm of  $\vec{s}_i$  and represents the physical distance that the target vehicle moves from moment  $t - \Delta t$  to moment  $t$  in world coordinates.  $\|\vec{s}_i\|$  can be accomplished by measuring the speed of the vehicle target through equation [\(14\)](#), where  $\Delta t$  is the time between two frames, measured in seconds, which is considered a constant, the reciprocal of the frame rate. Typically 1 s contains 25 frames in highway surveillance video, thus  $\Delta t = 1/25$ .

$$\vec{s}_i = H_{3 \times 4} \cdot u_i(t) - H_{3 \times 4} \cdot u_i(t - \Delta t) \quad (13)$$

$$v_i = \frac{\|\vec{s}_i\|}{\Delta t} = \frac{\|H_{3 \times 4} \cdot u_i(t) - H_{3 \times 4} \cdot u_i(t - \Delta t)\|}{\Delta t} \quad (14)$$

If a vehicle trajectory contains  $m$  frames of trajectory points, i.e., the target speed of the vehicle between every two adjacent images in the first  $m$  frames of the video is  $v_1, v_2, \dots, v_{m-1}$ , then based on equation [\(14\)](#) we can obtain:

$$\begin{aligned} v_1 &= \frac{\|\vec{s}_1\|}{\Delta t} = \frac{\|H_{3 \times 4} \cdot u_2(t) - H_{3 \times 4} \cdot u_1(t)\|}{\Delta t} \\ v_2 &= \frac{\|\vec{s}_2\|}{\Delta t} = \frac{\|H_{3 \times 4} \cdot u_3(t) - H_{3 \times 4} \cdot u_2(t)\|}{\Delta t} \\ &\vdots \\ v_{m-1} &= \frac{\|\vec{s}_{m-1}\|}{\Delta t} = \frac{\|H_{3 \times 4} \cdot u_m(t) - H_{3 \times 4} \cdot u_{m-1}(t)\|}{\Delta t} \end{aligned} \quad (15)$$

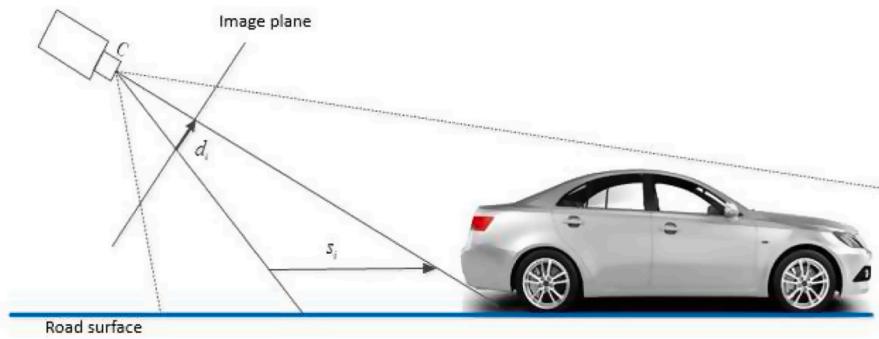
Then the average travel speed of the final target vehicle in the first  $m$  frames is as [Equation \(16\)](#), and the speed detection of the target vehicle is achieved by calculating the average value of the instantaneous speed between multiple frames.

$$v = \frac{\sum_{i=1}^{m-1} v_i}{m-1} \quad (16)$$

### 4. Vehicle risk estimation

#### 4.1. Vehicle Dynamics Model and risk assessment

The three fundamental elements of traffic are traffic flow rate ( $Q$ ), vehicle speed ( $v$ ), and traffic density ( $K$ ). Traffic flow rate is defined as the number of traffic entities passing through a specific location or section of a road within a designated time period. Vehicle speed is typically divided into spot speed and section speed. Spot speed refers to the instantaneous speed of vehicles passing through a specific location. Section speed refers to the ratio of the distance traveled by a vehicle to



**Fig. 6.** Vehicle speed measuring schematic.

the total time taken to travel that distance. Section speed is a comprehensive performance indicator used to evaluate the level of traffic congestion on a road and estimate travel delay. Traffic density refers to the number of vehicles per unit length of road at a specific moment, indicating the level of traffic density on a road.

There are safety risks when vehicles exceed the speed limit or maintain an insufficient following distance with the preceding vehicle during their operation. The three fundamental elements of traffic cannot directly reflect the road safety situation. Currently, road traffic safety analysis primarily relies on regression analysis to predict road traffic accidents. Regression analysis can effectively demonstrate the causal relationship between road accidents and various influencing factors. However, regression analysis requires a large sample size, low data variability, and strong regularity. Otherwise, its predictive accuracy may be compromised. Additionally, the regression analysis method based on road traffic accident prediction is a post event analysis method that cannot respond to road emergencies in a timely manner. Vehicle risk can be assessed based on the relative speed with preceding vehicles and whether the following distance is less than the safe distance. Therefore, this paper proposes a real-time vehicle risk analysis method that determines the magnitude of vehicle risk values based on the relative speed and distance between vehicles. This method is very important for timely prevention of traffic accidents and improvement of road safety. It can help traffic managers and policy makers understand the weakness of the transportation system and take corresponding measures to reduce accident risks and protect the safety of pedestrians and drivers.

To facilitate the calculation of vehicle risk values, this paper establishes a micro-level model for vehicle operation using a driver model to describe the behavior of individual drivers/vehicles. This system is a multi-agent system, meaning that each vehicle operates independently (see Fig. 7).

In the micro-level model, each vehicle is assigned a unique identifier, denoted as  $i$ . The vehicle  $(i - 1)$  is followed by the vehicle  $i$ . For the  $i$ th vehicle, we use  $x_i$  to represent its position along the road,  $v_i$  to represent its velocity, and  $l_i$  to represent its length. The distance from the rear bumper of the preceding vehicle to the front bumper of the following vehicle is denoted as  $s_i$ , while the velocity difference between the  $i$ th vehicle and the preceding  $(i - 1)$  vehicle is represented as  $\Delta v_i$ .

#### 4.2. Calculation of vehicle risk value

The braking distance of the preceding vehicle when encountering an emergency situation can be calculated based on the current vehicle speed  $v_{i-1}$  and the maximum deceleration rate  $b$ .

$$s_{i-1} = \frac{v_{i-1}^2}{2b} \quad (17)$$

The safe following distance between the following vehicle, with a speed of  $v_i$ , and the preceding vehicle, characterized by the speed difference  $\Delta v_i$ , is given by Equation (18):

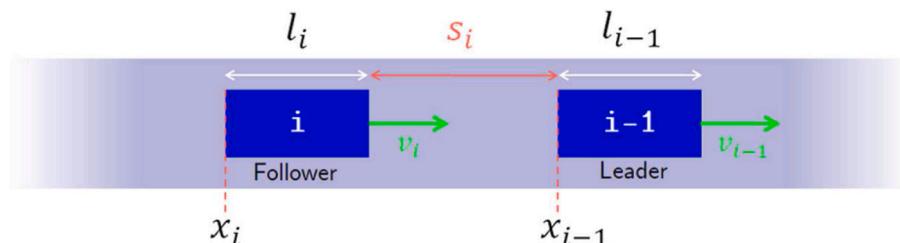
$$S_a = S_{min} + v_i T_i + \frac{v_i^2}{2b} - \frac{v_{i-1}^2}{2b} \quad (18)$$

In the above equation,  $b$  represents the maximum deceleration of the vehicle.  $T_i$  denotes the driver's reaction time.  $S_{min}$  represents the minimum safe distance when the vehicle is stationary. Assuming the actual distance between the two vehicles is  $S$ , the current risk value of the following vehicle can be calculated as the following equation:

$$r = \frac{S_a}{S} \quad (19)$$

When  $r \leq 1$ , the braking distance of the vehicle (also known as the safe following distance) is smaller than the traveled distance, indicating that the vehicle has no driving risk. When  $1 < r \leq 2$ , the braking distance exceeds the traveled distance, indicating a relatively low driving risk for the following vehicle, triggering a yellow warning in the system. When  $r > 2$ , the braking distance is significantly larger than the traveled distance, indicating a high driving risk for the following vehicle, triggering a red warning in the system.

The maximum deceleration of the vehicle "b" in the equation is closely related to the characteristics of the road and weather conditions. When the road is downhill, the maximum deceleration of the vehicle decreases, resulting in a longer braking distance. Consequently, the driving risk of the vehicle at the same speed significantly increases. Similarly, when the vehicle passes through a section affected by rainy or snowy weather, the maximum deceleration of the vehicle also decreases, leading to a higher driving risk at the same speed.



**Fig. 7.** Vehicle dynamics model.

## 5. Experimental results and analysis

In this paper, the videos of urban trunk road (South Second Ring Road, Xi'an) and highway (G60) traffic surveillance scenes are selected for experiments. The urban trunk road and highway scenes are calibrated and the experimental vehicle speed of urban trunk road is measured. The experimental results are analyzed and explained. The video pixel of the city main road scene is 1920\*1080 and the video pixel of the highway scene is 1280\*720.

### 5.1. Camera calibration

With the method in Section 2.2, the road direction vanishing point is obtained and the lane line is detected. The camera calibration is performed using the camera calibration method in Section 2.3, with the final focal length being determined by matching the lane line model with the actual lane line. Fig. 8 shows the matching results for city road and highway scenes respectively (the green line in the figure shows the matching result).

By analyzing the lane line matching results of different scenes, it is found that the lane line model is a good match with the actual lane line model after many matches from the wrong position. And when the final match is reached, the camera focal length  $f$ , pitch angle  $\varphi$  and yaw angle  $\theta$  can be determined in the scene to complete the camera calibration.

Through the camera calibration results in two scenarios, this paper selects the wheelbase of eight types of vehicles that appear more frequently in urban main roads and compares them with the wheelbase after calibration testing. Meanwhile, this paper selects the wheelbase of eight types of vehicles that appear more frequently in expressways (except large trucks) and compares them with the wheelbase after calibration testing. The comparison results of vehicle wheelbase in different scenarios are shown in Figs. 9 and 10. In order to ensure the accuracy of the calibration results, 15 test calibrations and measurements are carried out repeatedly on 8 different types of vehicles with different axle lengths in the two test scenarios. And the average calibration error is 0.77 % in the urban main road scenario and 1.57 % in the highway scenario, indicating that the calibration error in the highway scenario is greater than that in the urban scenario. There are two possible reasons for the analysis: first, the camera is lower in urban main road monitoring, so the vehicle target information is clearer in monitoring. So, the proposed calibration parameters are more accurate. In contrast, the camera is higher on highways, so the vehicle target clarity is poorer than that in urban scenes, which leads to a larger error. Secondly, due to the use of existing traffic scenes, the camera pixels on urban main roads (1920\*1080) are larger than those on highways (1280\*720), which will also have an impact on the validation of calibration results. However, the overall analysis shows that the calibration error is within the allowable range and can provide protection for vehicle speed.

### 5.2. Vehicle speed measurement

The vehicle speed is measured by calculating the average value of the instantaneous speed between multiple frames, as described in Section 3. At the same time, in order to verify the accuracy of the speed measurement, the experimental vehicle passes through three different lanes of the urban main road monitoring scene three times, all driving at 60 km/h. In total, nine experiments are conducted to get the average value of the vehicle in different lanes and compare it with the vehicle driving speed.

With different scenes of vehicle speed measurement and experiments, results statistics (see Table 2) show that the average error between the speed measured by the calibration and speed measurement method proposed in this paper and the driving speed is less than 6 %. The error measured by this scheme is smaller than the actual electronic violation of the monitoring of the radar speed error. The main reason for the error is that, for the trajectory in the scene there is a certain jump, so that the instantaneous error measured between different frames sometimes for a sudden jump situation, resulting in an increase in the final average speed error.

### 5.3. Vehicle risk estimation

The risk value of vehicles is calculated based on the real-time speed and position relationship of vehicles on the road. The specific calculation method is referred to Section 4.1. To validate this calculation method, a traffic flow simulation is conducted to simulate real-world traffic scenarios. The traffic flow simulation in this paper utilizes an intelligent driver model, which assumes that all vehicles travel along a straight line without collisions, while considering the speed difference and spacing between adjacent vehicles. The discretized acceleration equation for the intelligent driver model is as follows:

$$\frac{dv_i}{dt} = a_{\text{freeroad}} + a_{\text{interaction}} \quad (20)$$

The vehicle's following acceleration is composed of the free-road acceleration  $a_{\text{freeroad}}$  and the interaction acceleration on non-free-road segments  $a_{\text{interaction}}$ .

$$\begin{cases} a_{\text{freeroad}} = a_{\max}^{(i)} \left( 1 - \left( \frac{v_i}{v_{\max}^{(i)}} \right)^{\delta} \right) \\ a_{\text{interaction}} = -a_{\max}^{(i)} \left( \frac{s^*}{s_i} \right)^2 \end{cases} \quad (21)$$

where  $a_{\max}^{(i)}$  represents the maximum acceleration of the vehicle  $i$ ,  $v_{\max}^{(i)}$  represents the maximum desired velocity of the vehicle  $i$ , and " $\delta$ " is the acceleration exponent that controls the "smoothness" of the acceleration,  $s^*$  is the desired distance between the vehicle  $i$  and the preceding vehicle  $i - 1$ , and  $s_i$  represents the actual distance.

Using the established traffic model, a simulated scenario of emergency braking is conducted to simulate traffic behavior. By employing the vehicle risk calculation method, the risk values for each vehicle are



**Fig. 8.** Calibration parameters match in the urban road and the highway.

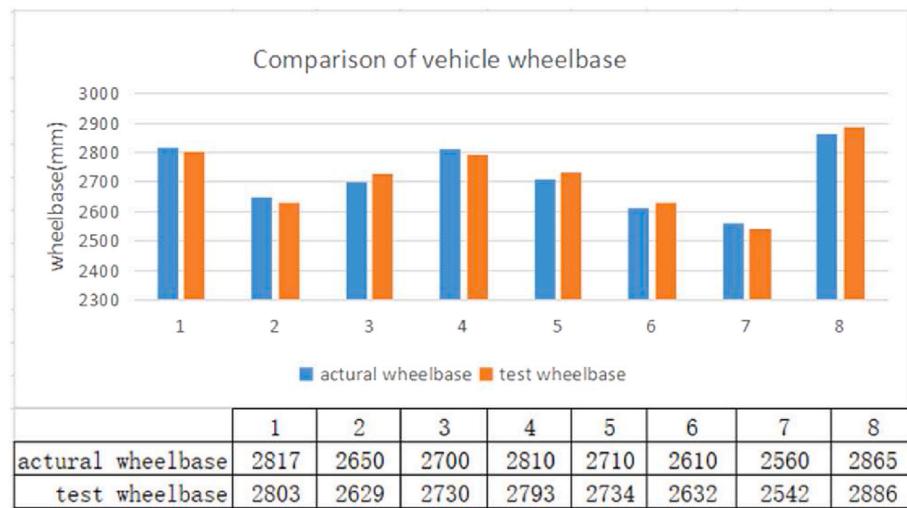


Fig. 9. Vehicle wheelbase comparison results on urban roads.

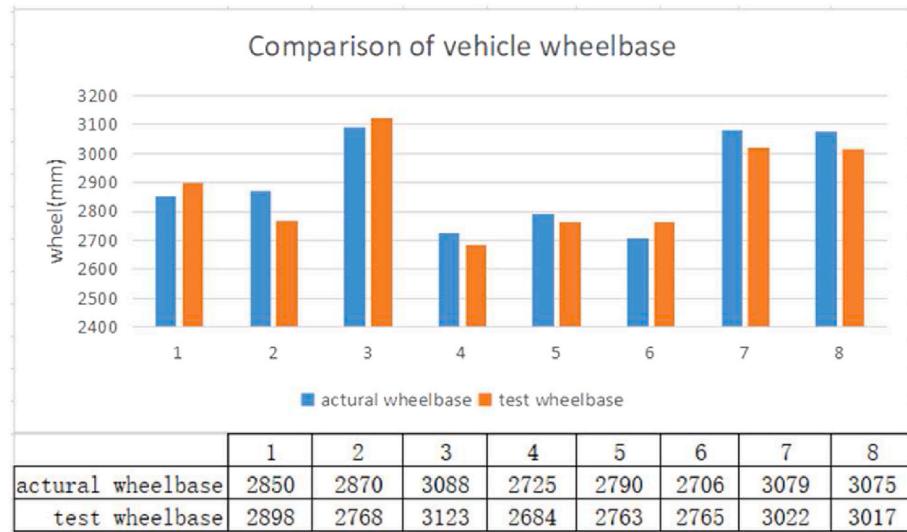


Fig. 10. Vehicle wheelbase comparison results on highways.

**Table 2**  
Vehicle speed experiment results statistics.

Lane	Vehicle speed (km/h)		Error/%
	Measurement	True speed	
Lane 1	63.2	60	+5.3
Lane 2	62.6	60	+4.3
Lane 3	57.3	60	-4.5

calculated. And the results are presented in the graph Fig. 11, where the horizontal axis of the graph represents the distance along the road, with a total distance of 2 km. The vertical axis represents time, displaying the spatio-temporal trajectory of the vehicles. The lower graph shows the resultant risk values for each vehicle. When the risk value "r" is less than or equal to 1, it indicates that the braking distance (i.e., safe following distance) of the vehicle is smaller than its traveled distance, which means a lower risk of the vehicle. When  $1 < r \leq 2$ , the braking distance exceeds the traveled distance, indicating a certain level of driving risk for the following vehicle, triggering a yellow warning in the system. When  $r > 2$ , the braking distance is significantly larger than the traveled distance, indicating a high level of driving risk for the following vehicle,

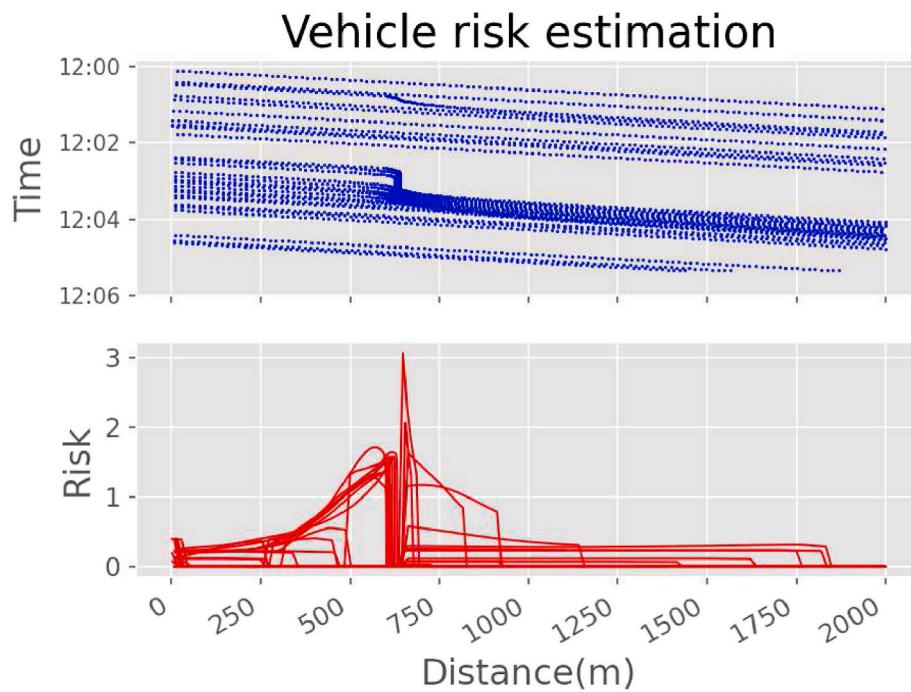
triggering a red warning in the system.

The road segments are divided into small sections with intervals of 100 m. The vehicle risk values for each interval are accumulated, with the results shown in Fig. 12.

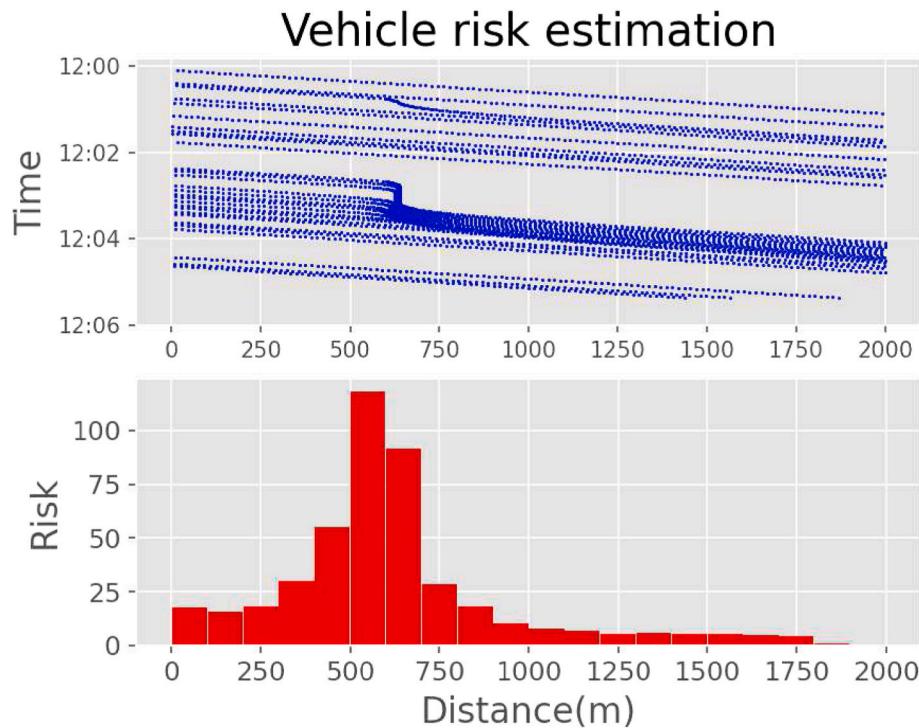
From the diagram, it can be observed that by aggregating the vehicle risk values based on road segments, we can obtain a risk assessment result for each segment, which reflects the road safety situation. For example, when there is an abandoned object on the road, drivers usually take measures to avoid potential risks and safety hazards. One common practice is to lower the vehicle speed through emergency braking, among other methods, to avoid collisions with the abandoned objects. This is reflected in the spatio-temporal trajectory of vehicles and the statistical graph of vehicle risks. Near the segment with abandoned objects, the road risk will significantly increase. Real-time monitoring of changes in road risk conditions and predicting future traffic situations can be achieved.

## 6. Conclusion

In this paper, a camera calibration method and a vehicle speed measurement algorithm are proposed for surveillance videos in urban



**Fig. 11.** Results of vehicle risk values in emergency braking scenario.



**Fig. 12.** Results of vehicle risk values in emergency braking scenario.

trunk road and highway scenes. The vehicle target is detected and tracked by deep learning YOLO detection algorithm and optical flow tracking algorithm. Based on the set of tracked trajectories, a set of vanishing points in the road direction is obtained to accurately obtain the sign lines on the road. Then based on the detected vanishing points and the sign lines on the road, the camera calibration in traffic scenes is accomplished via exploratory focal length. Finally, the vehicle speed is measured by calculating the average instantaneous speed between multiple frames. The calibration error is 1.17 % and the speed

measurement error is 6 % in different monitoring scenes (city main roads and highways), which can be well applied to actual traffic scenes to achieve vehicle speed measurement in video images and lay the foundation for traffic control and event monitoring. A vehicle risk assessment calculation method is proposed based on vehicle detection and speed measurement. Through simulations, it has been verified that this method can reflect the real-time road safety situation. The real-time monitoring and prediction capability can assist traffic management authorities in formulating more scientific and effective traffic

management strategies, thereby reducing the occurrence rate of road accidents. However, during the experiment, it is found that the effective field of view of a single camera is generally within 200 m. In order to achieve complete risk monitoring coverage of a section of road, it is necessary to connect the vehicle trajectories across cameras to achieve safety risk assessment of the entire road.

### CRediT authorship contribution statement

**Qin Shi:** Writing – original draft, Validation, Formal analysis, Visualization, Software, Methodology. **Yan Chen:** Writing – review & editing. **Haoxiang Liang:** Resources.

### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Data availability

Data will be made available on request.

### References

- [1] Wan Y, Huang Y, Buckles B. Camera calibration and vehicle tracking: highway traffic video analytics[J]. Transport Res C Emerg Technol 2014;44:202–13.
- [2] Karon G, Mikulski J. Selected problems of transport modelling with ITS services impact on travel behavior of users[C]. Int. Confer. its Telecommunic. 2017:1–7.
- [3] Luvizon DC, Nassu BT, Minetto R. A video-based system for vehicle speed measurement in urban roadways[J]. IEEE Trans Intell Transport Syst 2016;(99):1–12.
- [4] Ju Zhiyong, Wang Chaonan, He Xiaolei. Vehicle speed detection algorithm research of real time video image[J]. Appl Res Comput 2017;(9):268–270+299.
- [5] He XC, C Yang NH. A novel algorithm for estimating vehicle speed from two consecutive images[J]. WACV: IEEE Workshop on Applications of Computer Vision; 2007.
- [6] Trajković M. Interactive calibration of a PTZ camera for surveillance applications [C]. Proc. Asian Conf. Comput. Vis. 2002:1–8.
- [7] Al-Hadrusi MS, Sarhan NJ, Davani SG. A clustering approach for controlling PTZ cameras in automated video surveillance[C]. IEEE international symposium on multimedia (ISM), vol. 2016. IEEE; 2016. p. 333–6.
- [8] Zhang Runchu, Du Qianyun, Yu Zhuliang, et al. A calibration method for road monitoring cameras exploiting reference images and roadway information[J]. J Highw Transp Res Dev 2014;31(11):137–41.
- [9] Schoepflin TN, Dailey DJ. Dynamic camera calibration of roadside traffic management cameras for vehicle speed estimation[J]. IEEE Trans Intell Transport Syst 2003;4(2):90–8.
- [10] Song KT, Tai JC. Dynamic calibration of Pan-Tilt-Zoom cameras for traffic monitoring[M]. IEEE Press; 2006.
- [11] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2014. p. 580–7.
- [12] Girshick R. Fast r-cnn[C]. Proceedings of the IEEE international conference on computer vision 2015:1440–8.
- [13] Ren S, He K, Girshick R, et al. Faster r-cnn: towards real-time object detection with region proposal networks[C]. Adv Neural Inf Process Syst 2015:91–9.
- [14] Redmon J., Divvala S., Girshick R., et al. You only look once: unified time object detection[C]. Proc of IEEE CVPR 2016:779–788.
- [15] Redmon J, Farhadi A. Yolov3: an incremental improvement[J]. 2018. arXiv preprint arXiv:1804.02767.
- [16] Zhu X, Lyu S, Wang X, et al. TPH-YOLOv5: improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios[C]. Proceedings of the IEEE/CVF Int. confer. Computer vision 2021:2778–88.
- [17] Wang CY, Bochkovskiy A, Liao HYM. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition 2023:7464–75.
- [18] Liu W, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector[C]. In: European conference on computer vision. Cham: Springer; 2016.
- [19] Gupta A, Anpalagan A, Guan L, et al. Deep learning for object detection and scene perception in self-driving cars: survey, challenges, and open issues[J]. Array 2021; 10:100057.
- [20] Chen W, Li Y, Tian Z, et al. 2D and 3D object detection algorithms from images. A Survey[J]. Array 2023:100305.
- [21] Jiang Zhijun, Yi Huarong. An image pyramid-based feature detection and tracking algorithm. Geomatics Inf Sci Wuhan Univ 2007;32(8):680–3.
- [22] Dubská Markéta. Real projective plane mapping for detection of orthogonal vanishing points. 2013 [M].
- [23] Yan Hongping, Wang Lingfeng, Pan Chunhong. Automatic self-calibration of expressway surveillance camera under dynamic condition[J]. J Comput Aided Des Comput Graph 2013;25(7).
- [24] Bhardwaj R, Tummala GK, Ramalingam G, et al. Autocalib: automatic traffic camera calibration at scale[C]. In: The 4th ACM international conference. ACM; 2017.
- [25] Kanhere NK, Birchfield ST. A taxonomy and analysis of camera calibration methods for traffic monitoring applications[J]. IEEE Trans Intell Transport Syst 2010;11(2): 441–52.
- [26] Cathey FW, Dailey DJ. A novel technique to dynamically measure vehicle speed using uncalibrated roadway cameras[C]//IEEE Intelligent Vehicles Symposium. IEEE; 2005.
- [27] Xu Yanxia, Yang Ming. Video speed measurement method based on ground reference system[J]. Laster Journal 2017;(11):101–5.
- [28] Zaarane A, Slimani I, Al Okaishi W, et al. Distance measurement system for autonomous vehicles using stereo camera[J]. Array 2020;5:100016.