Research Article

Conflict Judgment and Safety Assessment at Unsignalized Intersections Based on Machine Vision

Yuqiong Wang,1 Liming Wang,1 Ruoyu Zhu,1 Yi Xu,1,2 Guoxin Jiang,1 and Xiaotian Ma3

1 School of Transportation and Vehicle Engineering, Shandong University of Technology, Zibo 255000, China
2 Collaborative Innovation Center of New Energy Automotive, Shandong University of Technology, Zibo 255000, China
3 Ou Ling Auto, Geely Commercial Vehicle, Zibo 255000, China

Correspondence should be addressed to Yi Xu; xuyisdut@163.com

Received 1 November 2022; Revised 18 January 2023; Accepted 30 January 2023; Published 17 February 2023

1. Introduction

Traffic conflict can be expressed as the result of interaction between one traffic participant and other traffic participants or traffic facilities during the course of movement in space [1]. In addition to rear-ending collisions, cross-collisions are a common type of collision. In a cross-collision between two vehicles, there is a certain angle between the motion directions of the two vehicles, so the collision is a kind of side collision. Most of the cross-collisions occur at the front right or left front of each vehicle. For the convenience of expression, all collisions involving a collision between the left and right fronts of the involved vehicles are collectively referred to as cross-collisions in this article, while the traffic conflicts before the cross-collisions are referred to as cross-conflicts.

Commonly used sensors for obstacle detection include millimeter wave radar, laser radar, and camera [2]. Millimeter wave radar has excellent real-time performance [3, 4], but its detection accuracy is low. Laser radar has high detection accuracy, but its price is quite high. Compared to other sensors, the camera has the advantages of high detection speed, good real-time performance, and robust adaptability [5, 6]. Obstacle detection based on machine vision uses methods such as machine learning and feature matching to identify the obstacles in a target region based on the region of interest (ROI) partition and obtains the deep information of the obstacles using a vision ranging model [7–10]. Liu et al. [11] proposed the SSD (single shot multibox detection) model based on the precise positioning method based on the Faster R-CNN candidate area proposal and the fast detection method based on YOLO, improving the end-to-end target detection method. Redon et al. [12] proposed the detection and recognition framework of YOLOv2, optimized the detection accuracy, detection speed, and detection category of the network model, and improved the running speed of the feature extraction network and the convergence speed of the model. Zhao et al. [13] proposed an integrated microscopic traffic flow model to describe...
human-driven vehicle maneuvers under interactions and verified that the model can accurately describe vehicle passing orders of interacting maneuvers, paths, and speeds against empirical data. Zhao et al. [14] proposed a novel method to model the trajectories of vehicles in two-dimensional space and speed, and the descriptive power, plausibility, and accuracy of the proposed model are investigated by comparing the calculated results under several cases, which can be solved from symmetry or analytically. This article proposes a detection method based on machine vision. Three monocular cameras are used to build a monocular-binocular vision switching system based on an artificial potential field for detecting the obstacle vehicles of the ego vehicle. The Kalman filter algorithm is used to locate, track, and predict the positions of the target vehicles. Monocular and binocular ranging models are established to calculate the distance, speed, direction, and other parameters of obstacle vehicles.

An accurate real-time prediction of vehicle trajectories at unsignalized intersections is a crucial part of traffic collision detection and early warning [15–18]. Formosa et al. proposed a centralized digital architecture that predicts traffic conflicts using deep learning. This model detects lane markings and tracks vehicles according to the images captured by a single front-end camera. The data are combined with traffic variables and calculated surrogate safety measures (SSMs) through the centralized digital architecture to develop a series of deep neural network (DNN) models for predicting traffic conflicts [19]. Huang et al. developed an extended driving risk estimation model considering the conflict between vehicles and pedestrians based on an artificial potential field, which projects the vehicles at the intersection into virtual lanes and measures the conflict risk in dynamic risks using the minimum safety distance between conflicting vehicles in the virtual lane [20]. Rocha et al. developed a framework [21] that can utilize the self-organizing mechanism to group traffic conflicts based on the similar characteristics and contributing factors of traffic accidents. The causes and characteristics of different types of conflict events can be revealed based on the grouping results, allowing the traffic management department to formulate strategies to reduce the risk of collision. Qadri et al. captured road traffic information using a handheld digital camera and extracted vehicle information using a video playback and flow recording device. On this basis, they modeled cross-collision using a multiregression model. They analyzed the types of traffic conflicts at the unsignalized intersections of six intersecting road segments from the perspective of time, accident, and travel speed [22]. Zheng et al. [23] validate the Bayesian hierarchical extreme value model that is developed for estimating crashes from traffic conflicts. The study shows that Bayesian hierarchical extreme value models significantly outperform the on-site models in terms of crash estimation accuracy and precision. This article focuses on unsignalized intersections and only considers the obstacle vehicles traveling in the right front and left front directions of the ego vehicle. A model based on trajectory prediction, collision zone, collision time, and safety distance is established based on machine vision detection to enable vehicle detection and safety assessment at the unsignalized intersection. Thus, an early warning can be given to the driver before the accident actually occurs, allowing the driver to take braking measures in case of an emergency to reduce the occurrence of cross-collision accidents and improve the safety of driving.

2. Vehicle Detection and Tracking in Left and Right front Direction Based on the Monocular-Binocular Vision Switching System

The cross-collision detection for vehicles is to determine whether there is a threat in the left and right front directions of the ego vehicle. There are many detection methods based on machine learning at present. The YOLOv5 model launched by Ultralytics in 2020 is implemented in the ecologically mature PyTorch, which has the advantages of small volume, high precision, and fast speed, and the effect of YOLOv5 has been recognized [24–26]. In this article, the detection of the obstacle vehicles in the left and right front directions of the ego vehicle is based on YOLOv5, and a monocular-binocular vision switching system is developed using an improved artificial potential field.

2.1. Vehicle Detection in Left and Right Front Direction Based on the Monocular-Binocular Vision Switching System

A monocular vision ranging model and a binocular vision ranging model are established to identify the vehicles in the left and right front directions of the ego vehicle and calculate the distance $d$, travel speed $v$, and driving direction $\alpha$.

Image information acquisition is the process of converting a three-dimensional object into a two-dimensional plane, which can be simplified as a pinhole camera model (see Figure 1). Assuming that the effective focal length of the camera is $f$, the installation height of the camera is $h$, the pitch angle of the camera is $\delta$, and the origin of the plane coordinate system $(x_0, y_0)$ (the intersection of the image plane and the optical axis of the camera) is usually set to $(0, 0)$. The intersection of the obstacle ahead and the road plane is $P$, and the coordinates of point $P$ in the image plane coordinate system are $(x, y)$. The horizontal distance $d$ between point $P$ and the camera is obtained through monocular camera ranging:

$$d = \frac{h}{\tan(\delta + \arctan((y_0 - y)/f))}$$ (1)

Binocular stereo vision entails using two cameras to capture two images of the same scene from two perspectives, thus obtaining a disparity image containing three-dimensional information about the scene. By processing the image with the stereo correction technique, the ordinate values of the same point on the two imaging planes can be equalized. Then, the distance information of the object can be calculated after the disparity is obtained. It is assumed that the left imaging plane and the right imaging plane are coplanar, the main optical axes are strictly parallel, $O_L$ and $O_R$ are the left and right focal points of the two cameras, respectively, and
the focal lengths \( f \) of the two cameras are equal. Assuming that point \( P(X, Y, Z) \) is an arbitrary feature point in the three-dimensional scene, then, its imaging points in the camera coordinate systems on the left and right imaging planes are \( P_l(x_l, y_l) \) and \( P_r(x_r, y_r) \), respectively (the ordinate values of the left and right imaging points \( y_l \) and \( y_r \) are equal). Taking the left camera as the reference, we can translate \( P_r(x_r, y_r) \) (the imaging point of point \( P(X, Y, Z) \) on the imaging plane of the right camera) to the imaging plane of the left camera to obtain \( P'_l(x'_l, y'_l) \); thus, \( |x_l - x'_l| \) is the vision disparity \( D \) between the two cameras. The depth information \( Z \) can be derived according to the similar triangle principle, as shown in equation (2). The principle of binocular stereo vision ranging is shown in Figure 2, where \( B \) is the baseline length of the binocular camera:

\[
\frac{|x_l - x'_l|}{B} = \frac{f}{Z} \Rightarrow Z = \frac{fB}{D}.
\]  

(2)

The obstacle vehicles are identified using the images captured by vision sensors, and their speeds are calculated using the interframe difference method. The frame rate of a general camera is greater than 25frames/s; the calculated speed can be closer to the real speed. Since the ego vehicle is moving, the relative speed and absolute speed of the obstacle vehicle in the left and right front directions can be calculated based on the travel speed of the ego vehicle. Assuming that the moment when the obstacle vehicle first appears in the camera’s field of view (FOV) is \( t_1 \), the distance between the two vehicles at \( t_1 \) is \( d_1 \), the next moment is \( t_2 \), the distance between the two vehicles at \( t_2 \) is \( d_2 \), and the interframe interval of the camera is \( t_3 \). Then, the time difference between the two moments and the distance difference between the two vehicles can be calculated using the following formulas:

\[
\Delta t = t_2 - t_1,
\]

\[
\Delta d = d_2 - d_1.
\]

(3)

The relative speed of the two vehicles is

\[
\Delta v = \frac{\Delta d}{\Delta t}.
\]

(4)

The value of \( \Delta t \) should be determined on the premise of ensuring driving safety and high efficiency, and 10\( t_1 \) is taken as a time interval. Plugging it into formula (4), we can calculate the relative speed of the two vehicles. The travel speed \( v_0 \) of the ego vehicle is obtained by the speed sensor. Thus, the travel speed of the obstacle vehicle is

\[
v = v_0 - \Delta v.
\]

(5)

The driving direction angle \( \alpha \) of the obstacle vehicle relative to the ego vehicle is calculated based on the \( x \) and \( y \) coordinates:

\[
\alpha = \arctan\left(\frac{|y|}{|x|}\right).
\]

(6)

2.2. Monocular-Binocular Vision Switching Strategy Based on Improved Artificial Potential Field. Considering the differences between monocular vision and binocular vision in terms of perception accuracy and range accuracy at different distances, we developed a monocular-binocular vision switching system based on the improved artificial potential field. The monocular-binocular switching strategy is formulated in real time to enable rapid and accurate calculation of obstacle vehicles, as well as effective tracking of obstacle vehicles. When cross-conflicts can be detected effectively, driving safety at unsignalized intersections under various driving conditions can be ensured.

All cross-conflicts are assumed to occur within 135° left and 135° right from the longitudinal axis of the ego vehicle, the detection radius of the camera is \( r \leq 55 \) m (\( r \) is the effective ranging depth of the camera), and the FOV = 150°. Three cameras are mounted on the ego vehicle to ensure the coverage area of the camera FOV and the effective switching between monocular and binocular cameras. In the FOV of each camera, some parts are only covered by one camera, while others are covered by two cameras (overlap region). In the part of FOV covered by a single camera, the environmental perception is based on monocular vision. In the FOV overlap region of two cameras, the advantages of both monocular vision and binocular vision can be combined for more effective detection [27, 28]. The monocular-binocular vision switching strategy is formulated based on the real-time traffic situation at the intersection. The FOVs of the three cameras are shown in Figure 3.

The artificial potential field method is a kind of virtual force method [29–31]. The premise of establishing a repulsion force field is that the ego vehicle will purposefully avoid obstacle vehicles. However, during the actual driving process, the movements of the ego vehicle and the obstacle vehicle are not linked. Therefore, we improved the artificial potential field method. The intersection area of the gravitational field \( G_0 \) between the ego vehicle and its target point and the gravitational field \( G_{n}(G_1, G_2, G_3, \ldots) \) between each obstacle vehicle and its target point is identified using the principle of an artificial gravitational field. Then, a real-time monocular-binocular vision switching strategy is developed based on the intersection area. Due to the difficulty of obtaining the target points of obstacle vehicles, the maximum distance that the vehicle can travel within 3 seconds in a straight line is used as the maximum length of the gravitational field, which is set to 50 m in this article according to the speed range of vehicles near an intersection.

Considering the requirements in the real driving environment, an overlap region \( W \) is defined according to the angle parameter \( \alpha \) (the driving direction angle of the obstacle

---

**Figure 1:** Model of monocular camera ranging.
vehicle relative to the ego vehicle) and the intersection area of the artificial gravitational fields $G_0$ and $G_n$, and the monocular-binocular vision switching strategy is formulated in real time according to the overlap region. In the single-camera area, the advantages of monocular vision, such as its small computing workload and high detection efficiency, can be exploited to detect obstacle vehicles, and the gravitational field $G_0$ between the ego vehicle and its target point is established in real time. For the obstacle vehicles that appear in the FOV overlap region of two cameras, the artificial gravitational field $G_n$ between each obstacle vehicle and its target point is established, and the overlap region $W$ is determined according to the intersection of $G_0$ and $G_n$. In the FOV overlap region, the distance and speed of the nearest vehicle are calculated, and whether to use binocular vision is decided according to the size of $W$. Gravity is linearly related to the distance $d$ between the two vehicles. The greater the distance $d$, the greater the gravity. When the overlap region $W > 0$, it indicates that the obstacle vehicle has begun to threaten the normal travel of the ego vehicle. At this time, binocular vision is used to identify the obstacle vehicle and calculate its distance and speed. The switching process is shown in Figure 4; the process of determining monocular-binocular vision switching strategy is shown in Figure 5.

The specific process is as follows:

1. The three cameras are mounted on the ego vehicle to acquire the environmental information in their own FOVs. The monocular vision technology based on YOLOv5 is used to identify the obstacle vehicles. We calculate the distance $d$ from the target vehicle to the camera by using similar triangles. $\beta$ is calculated according to the operating angle $\alpha$ and distance parameter $d$. The driving angle $\alpha$ of the obstacle vehicle relative to the ego vehicle is calculated according to the $x$ and $y$ coordinates of the obstacle vehicle relative to the ego vehicle:

$$d = \sqrt{x^2 + y^2},$$

$$\alpha = \arctan\left(\frac{y}{x}\right). \quad (7)$$

An orthogonal coordinate system is established, in which the moving direction of the ego vehicle is taken as the positive direction of the $X$-axis, and the positive direction of the $Y$-axis is obtained by rotating the $X$-axis clockwise by 90°. The acute angle $\beta$ between the moving direction of the centroid of the obstacle vehicle and the $Y$-axis in a time interval is determined according to the target point of the obstacle vehicle, and the gravitational field $G_n$ between the obstacle vehicle and its target point is determined accordingly. The scene is shown in Figure 6.

2. When $\alpha > 75°$, monocular vision technology is used to identify and track the obstacle vehicle, as well as calculate its distance and speed.

3. When $\alpha \leq 75°$, the gravitational field $G_0$ is established, and the gravitational field $G_n$ is established according to the driving direction $\beta$ of the obstacle vehicle. When $W < 0$, monocular vision technology is used to track the obstacle vehicle as well as calculate its speed and distance. When $W \geq 0$, binocular vision technology is used to track the obstacle vehicle as well as calculate its distance and speed.

4. When the angle parameter $\alpha$ or $W$ deviates out of the original range, real-time monocular-binocular vision switching is needed to enable high-precision tracking.
(5) For a new obstacle vehicle or an obstacle vehicle that reappears in the vision system, we repeat the above four steps.

2.3. Vehicle Motion State Estimation Based on Extended Kalman Filter. The position of the monitored vehicle at a certain moment in the future can be obtained based on a reasonable prediction of the trajectory of the vehicle. This provides a basis for tracking the obstacle vehicle [32–36]. Taking the centroid of the vehicle as the origin of coordinates, we can establish a three-dimensional coordinate system, in which the positive direction of the $X$-axis is the driving direction of the vehicle. The positive direction of the $Y$-axis is obtained by rotating the positive direction of the $X$-axis clockwise by $90^\circ$, and the positive direction of the $Z$-axis is the upward direction perpendicular to the XOY plane, as shown in Figure 7.

State and observation variables are determined using the $x$ coordinate, $y$ coordinate, longitudinal component $v_x$ of the travel speed, and yaw angle $\theta$ of the driving direction of the vehicle. Assuming that the motion state of the vehicle is $X = (x, y, v_x, \theta)$ and the observation state is $Z = (x, y, v_y, \theta)$, we can obtain the position of the obstacle vehicle by predicting the $x$ and $y$ coordinates of the obstacle vehicle. We can obtain the state equation (8) of vehicle motion and the observation equation (9) of the vehicle motion system by ignoring the system noise and inputting the longitudinal acceleration $a_x$ and yaw rate $\omega_v$ of the vehicle calculated by the environmental sensing system as specific quantities:

$$ X_{k+1} = \begin{bmatrix} x_{k+1} \\ y_{k+1} \\ v_{y(k+1)} \\ \theta_{(k+1)} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \Delta h \cos \theta_k & 0 \\ 0 & 1 & \Delta h \sin \theta_k & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_k \\ y_k \\ v_{y(k)} \\ \theta_{(k)} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ a_{x(k+1)} \\ \omega_v_{(k+1)} \end{bmatrix} \Delta h, \quad (8) $$

$$ Z_{k+1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} X_k + \begin{bmatrix} 0 \\ 0 \\ a_{x(k+1)} \\ \omega_v_{(k+1)} \end{bmatrix} \Delta h + R, \quad (9) $$

where $\Delta h$ represents the actual sampling time (i.e., the time interval between two samples) and $R$ is the observation noise, assuming that the position calculation error of the system is less than 1 m, the velocity calculation error is less than 2 m/s, and the acceleration calculation error is less than 1.3 m/s$^2$. 

![Figure 3: FOVs of three cameras.](image-url)
The matrix of the system can be expressed as
\[
\Phi_k = \begin{bmatrix}
1 & 0 & \Delta h \ast \cos \theta_{k+1,k} & -\Delta h \ast v_y \sin \theta_{k+1,k} \\
0 & 1 & \Delta h \ast \sin \theta_{k+1,k} & \Delta h \ast v_y \cos \theta_{k+1,k} \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix},
\]
(10)

During the actual driving process, the motion state of the obstacle vehicle may be moving straightforward, left turning, or right turning, and it is also possible that turning left or right is followed by moving straightforward. Under the traffic condition of moving straightforward, the motion state is characterized by the regular variation of the position parameter \(x\), slight variation of the position parameter \(y\), and slight variation of \(v_x\) and \(v_y\) in direction. Under the traffic condition of left and right turning, obvious variations of the position parameters \(x\), \(y\), and \(v_y\), as well as \(\theta\), can be observed.

The vehicle is considered to be a particle. After the vehicle motion state is predicted using the extended Kalman filter algorithm, its motion state equation is modified, and the time period length of each prediction is changed in order to obtain the predicted value after a period of \(t\) (\(t\) represents the time difference between the two predictions). The new equation for vehicle motion is

\[
X_{k+1} = \begin{bmatrix}
x_{k+1} \\
y_{k+1} \\
v_{y(k+1)} \\
\theta_{(k+1)}
\end{bmatrix} = \begin{bmatrix}
1 & 0 & t \cos \theta_k & 0 \\
0 & 1 & t \sin \theta_k & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix} \begin{bmatrix}
x_k \\
y_k \\
v_y(k) \\
\theta(k)
\end{bmatrix} + \begin{bmatrix}
0 \\
0 \\
0 \\
0
\end{bmatrix} + \begin{bmatrix}
0 \\
a_x(k) \\
0 \\
w_r(k)
\end{bmatrix} t.
\]
(11)

The vehicle motion state equation at the moment \(T\) is

\[
X(T_k + it) = \begin{bmatrix}
1 & 0 & t \cos \theta_{(T_k+iT_s)} & 0 \\
0 & 1 & t \sin \theta_{(T_k+iT_s)} & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix} \begin{bmatrix}
x_k \\
y_k \\
v_y(k) \\
\theta(k)
\end{bmatrix} + \begin{bmatrix}
0 \\
a_x(T_i) \\
0 \\
w_r(T_i)
\end{bmatrix} t.
\]
(12)
where \( i \) is given a value from 1 to \( n \), and \( n = T/t \), in which \( n \) is the iteration number of the state equation. We can change the value of \( n \) to predict the trajectory of the vehicle in a certain period of time.

The process of tracking the detected obstacle vehicle using the Kalman filter algorithm is as follows: First, the detected vehicle is marked by a rectangular frame, and the vehicle’s position in the image is obtained. The center of the rectangular frame is taken as the centroid of the vehicle for marking, and the marked position is the state variable in the tracking process. Then, the algorithm judges whether the vehicles in different images are the same vehicle through feature matching. Finally, the motion state equation is continuously updated in order to determine the predicted position of the vehicle at a certain moment in the future. The tracking result is shown in Figure 8. The vehicle position is indicated by a yellow frame, and the vehicle tracking result is marked with a red frame. The experimental results show that the vehicle tracking method based on the Kalman filter algorithm can achieve a position matching degree of 94%, which satisfies the requirement for tracking vehicles at intersections.

3. Safety Assessment Based on Trajectory Prediction, Collision Zone, and Collision Time

After determining the accurate motion state and trajectory of the obstacle vehicle, the next step is to determine the cross-conflict point between the ego vehicle and the obstacle vehicle. The time-to-collision method is used to detect cross-collisions and assess the obstacle vehicle’s safety. Calculating the time for the ego vehicle and the obstacle vehicle to enter and leave the collision zone, respectively, and judging whether the cross collision occurs according to the \( W \) value of the time period of the ego vehicle and the obstacle vehicle passing through the collision zone, the conflict level is determined based on the distance between the two vehicles when one of them reaches the cross-conflict point, and the safety of the obstacle vehicle is evaluated based on the conflict level. The specific process is shown in Figure 9.

3.1. Vehicle Trajectory Crossing Model Based on Motion State

When there is the possibility of a cross-collision between the ego vehicle and the obstacle vehicle, their trajectories will intersect. In cross-conflicts, the X-shaped trajectory intersection is the most common type. An X-shaped trajectory intersection indicates that the smaller included angle after the intersection of the extension lines of the driving directions of two vehicles is an acute angle, i.e., the two lines form an X shape. Perpendicular trajectory intersection is a special case where the included angle is 90°. Suppose \( \gamma \) is the trajectory intersection angle of the ego vehicle and the obstacle vehicle. In this article, we performed cross-collision detection for X-shaped trajectory intersection conditions and established a safety assessment model (Figure 10).

Suppose both the ego vehicle and the obstacle vehicle move along a straight line in a time period \( \Delta t \); the position of the ego vehicle at the moment \( t_1 \) is taken as the origin of coordinates; the driving direction of the ego vehicle is the positive direction of the X-axis, and the position direction of the Y-axis is obtained by rotating the positive direction of the X-axis clockwise by 90°. At the moment \( t_{1} \), the position of the ego vehicle is the origin of coordinates \( o \). The position of the ego vehicle is \((x_{11}, y_{11})\) and \((x_{12}, y_{12})\) at moments \( t_{1} - \Delta t \) and \( t_{1} \), respectively. The positions of the obstacle vehicle at the moments \( t_{1} - \Delta t \) and \( t_{1} \) are \((x_{21}, y_{21})\) and \((x_{22}, y_{22})\), respectively. The heading angle of the obstacle vehicle is the angle between the positive direction of the X-axis and the connecting line of the positions of the obstacle vehicle and ego vehicle at a certain moment. At the moments \( t_{1} - \Delta t \) and \( t_{1} \), the heading angles of the obstacle vehicle are \( \theta_1 \) and \( \theta_2 \), respectively. Moreover, the corresponding distances between the two vehicles are \( d_1 \) and \( d_2 \), respectively. The trajectory intersection point is denoted as \((x_0, y_0)\). Thus, the positions of the obstacle vehicle at the moments \( t_{1} - \Delta t \) and
vision on Vehicle1

The location of the vehicle at a certain moment

The position of the vehicle at a certain moment

Use binocular vision on Vehicle1

Use monocular vision on Vehicle2

Figure 6: Schematic diagram of intersection scene at a certain moment.

The location of the vehicle at a certain moment

The position of the target vehicle at a certain moment

Use binocular vision on Vehicle1

Use monocular vision on Vehicle2

Figure 7: Three-dimensional coordinate system of the vehicle.

3.2. Cross-Conflict Judgment Based on Collision Zone and Collision Time. We use an X-shaped trajectory intersection as an example to establish a model of collision zone and collision time for the scenario in which the ego vehicle encounters a single obstacle vehicle. Assuming that the position of the intersection point is \((x_0, y_0)\), we can substitute \(x = x_0\) and \(y = 0\) into equation (14) to obtain

\[
x_0 = \frac{x_{21}y_{22} - x_{22}y_{21}}{y_{22} - y_{21}}.
\] (15)

The collision zone is a circular area with a radius of \(\rho\), and its center is the intersection point \((x_0, y_0)\). A cross-conflict will occur if two vehicles appear in the collision zone simultaneously. \(T_{11}, T_{12}, T_{21}\), and \(T_{22}\) are the collision times, which are, respectively, the times it takes for the ego vehicle and the target vehicle to enter and exit the collision zone. Since the size of the vehicle cannot be ignored, the length and width of the vehicle should be taken into account when calculating the radius of the collision zone. The formula for calculating \(\rho\) is as follows:

\[
\rho = \frac{b/2}{\tan(\gamma/2)} + \frac{a}{2} + c,
\] (17)

where \(v_1\) represents the speed of the ego vehicle and \(a\) and \(b\) are the length and width of the vehicle, respectively. The time periods for the ego vehicle and the target vehicle to enter the collision zone are \((T_{11}, T_{12})\) and \((T_{21}, T_{22})\), respectively. The following inferences can be made:

1. When \((T_{11}, T_{12})\) \(\cap\) \((T_{21}, T_{22})\), that is, \(T_{11} < T_{21} < T_{12}\) or \(T_{21} < T_{12} < T_{22}\), it indicates that two vehicles appear in the collision zone simultaneously, resulting in cross-conflict. The ego vehicle must make emergency braking if the available reaction time is limited.

2. When \(T_{11} > T_{22}\), it indicates that the obstacle vehicle will cross the collision zone at first. Thus, cross-conflict will not occur.

3. If \(T_{21} > T_{12}\), it indicates that the ego vehicle will cross the collision zone at first. Thus, cross-conflict will not occur.

Figure 12 is a schematic diagram of the collision zone and time at an intersection.

3.3. Vehicle Safety Assessment Based on the Safety Distance. Assuming that the safety distance (the distance just long enough to avoid a collision if both drivers apply emergency braking) is \(D_s\), we can take \(D_s\) as the critical value of the highest conflict level to calculate the safety distance:

\[
\rho = \frac{b/2}{\tan(\gamma/2)} + \frac{a}{2} + c,
\] (16)
\[ D_5 = v(\tau_1 + \tau_2) + \frac{v^2}{2a_{\text{max}}}, \]  

where \( v \) represents the real-time speed of the vehicle arriving at the cross-conflict point, \( \tau_1 \) is the driver’s reaction time, \( \tau_2 \) is the delay time of the braking system, and \( a_{\text{max}} \) represents the maximum braking deceleration of the vehicle.

When \( (T_{11}, T_{12}) \cap (T_{21}, T_{22}) \), cross-conflict between the ego vehicle and the obstacle vehicle will occur. It is assumed that one vehicle arrives at the cross-conflict point first, and the distance between the other vehicle and the cross-conflict point is \( D_n \). In the case of cross-conflict, it is stipulated that the two thresholds are \( D_1 \) and \( D_2 \). When \( D_n \leq D_2 \), the conflict is a level-3 one; when \( D_2 \leq D_n < D_1 \), the conflict is a level-2 one; if \( D_1 \leq D_n \), the conflict is a level-1 one.

4. Real Vehicle Verification and Simulation

4.1. Reliability Analysis of Monocular-Binocular Vision Switching System. In this article, a HAVAL H7 vehicle serves as the experimental vehicle, a RERVISION camera was used as a sensor, and a Core (TM) i7-9700 Lenovo notebook was used as a processing unit. The camera’s optical center is 1.73 m above the ground. The procedure for verifying the reliability of the monocular-binocular vision switching strategy is as follows:

(1) Three RERVISION cameras are mounted on a HAVAL H7 vehicle, traveling at a certain unsignalized intersection and collecting traffic video, as shown in Figure 13.
The improved artificial potential field principle is programmed. The monocular-binocular vision switching system was implemented to process the captured images and calculate the motion states of the obstacle vehicles in real-time. The artificial gravitational field was established to determine the value of $W$, and monocular-binocular vision switching was implemented based on the values of $W$ and $a$. The selective monocular-binocular vision switching strategy was executed to process the captured images so as to identify the obstacle vehicle and calculate its distance, speed, driving angle, and other parameters.

A nonselective monocular-binocular vision switching system was set up. A program was used to process the captured images so as to identify the obstacle vehicle in the monocular vision area and calculate its distance, speed, driving angle, and other parameters. In the overlap zone, only binocular vision was used for vehicle identification and parameter calculation.

The experimental results obtained using the three methods were compared. The reliability of the monocular-binocular vision switching strategy proposed in this article was verified based on the performance evaluation of each method in detection accuracy, ranging accuracy, ranging speed, and camera detection area. A frame of camera picture was taken every 0.04 s to collect information about obstacle vehicles around the experimental vehicle (RERVISION camera frames are 25 frames per second). The proposed monocular-binocular vision switching strategy was used to identify each obstacle vehicle and calculate its distance, speed, driving angle, and other parameters. The calculated data, including the number, distance, speed, and driving angle of obstacle vehicles, were compared with the data measured by on-board Lidar so as to assess the performance in detection accuracy, ranging accuracy, ranging speed, and camera detection area, thus verifying the feasibility of the proposed monocular-binocular vision switching strategy. Figure 14 shows the construction of the real vehicle experimental platform, Figure 15 demonstrates the result of detecting a vehicle ahead of the ego vehicle, and Table 1 shows the distance calculation results obtained based on monocular and binocular vision.

The ranging accuracy values calculated in the monocular vision mode and the binocular vision mode were assessed using the ranging accuracy results of Lidar as a benchmark. Table 1 demonstrates that the distance calculation method based on monocular vision can reach an average accuracy of 96.18% within 60 m. The method based on binocular vision can achieve an average accuracy rate of 96.63% within 60 m, which can meet the accuracy level required in measuring the distances of moving targets in a real driving environment in real time.

The detection areas of cameras under different vision types were calculated using the formula for fan area calculation based on camera FOV. The detection accuracy and ranging accuracy of the monocular-binocular vision switching system are expressed by the average of the accuracy of monocular vision and the accuracy of binocular vision. Table 2 compares the experimental results of various vision switching strategies, assuming that the actual distance is $a$, the calculated distance is $b$, and the ranging accuracy is $1 - |a - b|/a$.

Table 2 demonstrates that the monocular-binocular vision switching system proposed in this article outperforms pure monocular vision in terms of detection accuracy and is superior to pure binocular vision in terms of the detection area. In the performance of distance calculation, the
Collision Zone. A PreScan simulation experiment was conducted on the safety assessment model based on the collision zone. The parameters of the ego vehicle and the driving environment were set according to the driving data and environmental conditions obtained in a real vehicle experiment, and cross-conflict classification and safety level division were carried out following the procedure below:

1. We set the driving scene as an unsignalized intersection.
2. We set the travel speed and direction of the ego vehicle.
3. We set the initial position, driving direction, travel speed, and initial azimuth $\beta_o$ of the obstacle vehicle.
4. We obtain the length and width data of the ego vehicle and obstacle vehicle according to their models.
5. We calculate the collision radius $\rho$ using formula (16) and calculate the values of $(T_{11}, T_{12})$ and $(T_{21}, T_{22})$.

![Figure 12: Schematic diagram of collision zone and collision time at an intersection: (a) condition of one obstacle vehicle; (b) condition of two obstacle vehicles.](image)

![Figure 13: Collection of images.](image)
(6) We randomly generate any value for a speed between 10 km/h and 60 km/h of the ego vehicle.

(7) We set the traffic condition of one obstacle vehicle and the traffic condition of two obstacle vehicles. The initial positions and driving directions of the obstacle vehicles are set randomly, and the travel speed of each obstacle vehicle is set randomly between 10 km/h and 60 km/h.

(8) We record the motion states of the ego vehicle and the obstacle vehicle every 0.5 s and judge the safety situation.

In the PreScan software, a driving scene model of unsignalized intersection was developed to facilitate the simulation and analysis of the results of judging conflicts resulting from X-shaped trajectory intersection. The scene simulation is shown in Figure 16. There is only one obstacle vehicle in traffic conditions 1–3, and there are two obstacle vehicles in traffic condition 3. Among them, $D_1 = 16.7$ m and $D_2 = 5$ m. The simulation results are as follows.

### 4.2.1. Traffic Condition 1

The speed of the ego vehicle is $v_0 = 60$ km/h, the speed of the obstacle vehicle is $v_1 = 16$ km/h, the initial distance from the ego vehicle to the cross-conflict point is set to 120 m, the initial distance from the obstacle vehicle to the cross-conflict point is 52 m, and the initial heading angle of the obstacle vehicle is $\beta_1 = 15^\circ$. The simulation results are as follows: $T_{11} = 5.9$ s, $T_{12} = 8.5$ s, $T_{21} = 6.8$ s, and $T_{22} = 16.6$ s, and the distance variation between the two vehicles during driving is shown in Figure 17. When $T_{21} > T_{12}$, it indicates that the ego vehicle passes through the collision zone before the target vehicle and there is no danger of a collision. As shown in Figure 18, when the ego vehicle arrives at the cross-conflict point at 7.2 s,

---

<table>
<thead>
<tr>
<th>Lidar ranging range (cm)</th>
<th>Ranging mode</th>
<th>Actual distance (cm)</th>
<th>Calculated distance (cm)</th>
<th>Error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–300</td>
<td>Monocular</td>
<td>300</td>
<td>309.8</td>
<td>3.3</td>
</tr>
<tr>
<td></td>
<td>Binocular</td>
<td></td>
<td>291.3</td>
<td>2.9</td>
</tr>
<tr>
<td>300–600</td>
<td>Monocular</td>
<td>600</td>
<td>620.4</td>
<td>3.4</td>
</tr>
<tr>
<td></td>
<td>Binocular</td>
<td></td>
<td>617.3</td>
<td>2.9</td>
</tr>
<tr>
<td>600–900</td>
<td>Monocular</td>
<td>900</td>
<td>931.8</td>
<td>3.5</td>
</tr>
<tr>
<td></td>
<td>Binocular</td>
<td></td>
<td>928.1</td>
<td>3.1</td>
</tr>
<tr>
<td>900–2000</td>
<td>Monocular</td>
<td>2000</td>
<td>1928.2</td>
<td>3.6</td>
</tr>
<tr>
<td></td>
<td>Binocular</td>
<td></td>
<td>1935.4</td>
<td>3.3</td>
</tr>
<tr>
<td>2000–4000</td>
<td>Monocular</td>
<td>4000</td>
<td>4167.9</td>
<td>4.2</td>
</tr>
<tr>
<td></td>
<td>Binocular</td>
<td></td>
<td>4164.1</td>
<td>4.1</td>
</tr>
<tr>
<td>4000–6000</td>
<td>Monocular</td>
<td>6000</td>
<td>5704.6</td>
<td>4.9</td>
</tr>
<tr>
<td></td>
<td>Binocular</td>
<td></td>
<td>6292.7</td>
<td>4.9</td>
</tr>
</tbody>
</table>

---

**Figure 14:** The construction of the real vehicle experimental platform.

**Table 1:** Distance obtained by monocular and binocular vision.
the distance from the obstacle vehicle to the cross-conflict point is $D_n = 13\,\text{m}$ (which constitutes a second-class conflict). The ego vehicle is far from the cross-conflict point when the obstacle vehicle arrives at the cross-conflict point at 11.7 s. Thus, this scenario will not be discussed further.

### 4.2.2 Traffic Condition 2

The speed of the ego vehicle is $v_0 = 50\,\text{km/h}$, the speed of the obstacle vehicle is $v_1 = 36\,\text{km/h}$, the distance from the ego vehicle to the cross-conflict point is set to 65 m, the distance from the obstacle vehicle to the cross-conflict point is 60 m, and the initial heading angle of the obstacle vehicle is $\beta_1 = 67^\circ$. The simulation results are $T_{11} = 4.3\,\text{s}$, $T_{12} = 5.1\,\text{s}$, $T_{21} = 4.4\,\text{s}$, and $T_{22} = 5.5\,\text{s}$, and the radius of the collision zone is $\rho = 4.57\,\text{m}$. The distance variation between the two vehicles during driving is shown in Figure 18. If $T_{11} < T_{21} < T_{12}$, a cross-collision between the two vehicles will occur. As shown in Figure 18, when the ego vehicle reaches the cross-conflict point, the obstacle vehicle is 14 m away from the cross-conflict point. Thus, the conflict level of this obstacle vehicle is level 2.

### 4.2.3 Traffic Condition 3

The speed of the ego vehicle is $v_0 = 40\,\text{km/h}$, the speed of obstacle vehicle 1 is $v_1 = 45\,\text{km/h}$, and the speed of obstacle vehicle 2 is $v_2 = 47\,\text{km/h}$. The distance from the ego vehicle to trajectory intersection point 1 is 50 m, the distance from obstacle vehicle 1 to trajectory intersection point 1 is 80 m, and the initial heading angle of obstacle vehicle 1 is $62^\circ$. The distance from the ego vehicle to trajectory intersection point 2 is 60 m, the distance from obstacle vehicle 2 to trajectory intersection point 2 is 50 m, and the initial heading angle of obstacle vehicle 2 is $43^\circ$. The simulation results of obstacle vehicle 1 are $T_{11} = 2.5\,\text{s}$, $T_{12} = 6.5\,\text{s}$, $T_{21} = 6.26\,\text{s}$, and $T_{22} = 9.74\,\text{s}$. At time $T = 0.24\,\text{s}$, indicating that cross-conflict will occur. It can be seen from Figure 19(a) that when the ego vehicle reaches the intersection point, the distance from obstacle vehicle 1 to the intersection point is $D_n = 36\,\text{m}$, which constitutes a level 1 conflict. The simulation results of obstacle vehicle 2 are as follows: $T_{11} = 3.5\,\text{s}$, $T_{12} = 7.4\,\text{s}$, $T_{21} = 2.0\,\text{s}$, and $T_{22} = 5.5\,\text{s}$. At time $T = 2.0\,\text{s}$, indicating that cross-conflict will occur, and the collision radius is $\rho$.

### Table 2: Comparison of experimental results obtained under different vision switching strategies.

<table>
<thead>
<tr>
<th>Vision type</th>
<th>Detection accuracy (%)</th>
<th>Ranging accuracy (%)</th>
<th>Ranging speed (t/s)</th>
<th>Detection area (m²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monocular-binocular switching</td>
<td>95.7</td>
<td>96.4</td>
<td>1.1</td>
<td>7915.4</td>
</tr>
<tr>
<td>Monocular</td>
<td>94.3</td>
<td>96.1</td>
<td>0.9</td>
<td>7915.4</td>
</tr>
<tr>
<td>Binocular</td>
<td>95.7</td>
<td>96.6</td>
<td>1.3</td>
<td>3957.7</td>
</tr>
</tbody>
</table>

![Figure 15: Result of detecting a vehicle ahead of the ego vehicle.](image1)

![Figure 16: X-shaped cross-track scene simulation.](image2)
Figure 19(b) demonstrates that when obstacle vehicle 2 reaches the intersection point, the distance from the ego vehicle to the intersection point is 18 m, which constitutes a level 1 conflict. Figure 19 demonstrates the variations of the distances between the vehicles during the process of obstacle vehicles 1 and 2 passing the intersection points.

The simulation results of the three traffic conditions were compared with the theoretical calculation results. The comparison demonstrated that the cross-conflict judgment and safety assessment model based on trajectory prediction, collision zone, and collision time can effectively identify cross-conflicts and rate the severity of each conflict under the corresponding threshold. PreScan simulation results show that the effective recognition rate of cross conflict based on collision zone and collision time is more than 97%, and the recognition accuracy of conflict level is more than 95%. It shows that the intersection conflict judgment
method and safety evaluation model proposed in this article are suitable for low-speed driving of vehicles at unsignalized intersections.

5. Conclusion

Aiming to solve the rarely studied problem of how to ensure traffic safety at unsignalized intersections, we developed a method to detect cross-conflicts between the ego vehicle and vehicles traveling in the left and right direction based on machine vision and a safety assessment model. First, a monocular-binocular vision switching strategy based on an improved artificial potential field method was developed, and a monocular-binocular vision switching system was set up. Then, a single-binocular vision conversion strategy based on the basic principle of artificial potential field is formulated, and a single-binocular vision environment perception system is built to realize the effective detection of obstacle vehicles in front of the vehicle side. In addition, a vehicle motion state estimation model and a vehicle trajectory prediction model based on the extended Kalman filter algorithm were constructed. On this basis, a cross-conflict detection and safety assessment model based on vehicle trajectory, collision zone, and collision time was developed. With a reasonable setting of conflict level thresholds, the method proposed in this article can enable effective safety assessment for the vehicles traveling at low speeds at intersections, provide a reasonable process for judging the conflict level in the traffic environment at intersections, and provide a basis for the takeover of the driving assistance system and for human-computer interaction of autonomous vehicles.

Data Availability

No data were used to support this study.

Conflicts of Interest

There are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant 51905320, China Postdoctoral Science Foundation under Grant 2018M632696, Natural Science Foundation of Shandong Province under Grants ZR2022MF230 and ZR2021QF039, Shandong Province Major Science and Technology Innovation Project under Grant 2019JZZY010911, Shandong Key R&D Plan Project under Grant 2019GGX104066, Experimental Technology Upgrading Project under Grant 2022003, and Shandong Provincial Program of Introducing and Cultivating Talents of Discipline to Universities: Research and Innovation Team of Intelligent Connected Vehicle Technology.

References


