Research Article

An Obstacle Detection and Distance Measurement Method for Sloped Roads Based on VIDAR

Guoxin Jiang (ID), **Yi Xu** (ID), **Xiaotong Gong** (ID), **Shanshang Gao** (ID), **Xiaoqing Sang** (ID), **Ruoyu Zhu** (ID), **Liming Wang** (ID), **and Yuqiong Wang** (ID)

School of Transportation and Vehicle Engineering, Shandong University of Technology, Zibo 255000, China

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

Environmental perception systems can provide information on the environment around a vehicle, which is key to active vehicle safety systems. However, these systems underperform in cases of sloped roads. Real-time obstacle detection using monocular vision is a challenging problem in this situation. In this study, an obstacle detection and distance measurement method for sloped roads based on Vision-IMU based detection and range method (VIDAR) is proposed. First, the road images are collected and processed. Then, the road distance and slope information provided by a digital map is input into the VIDAR to detect and eliminate false obstacles (i.e., those for which no height can be calculated). The movement state of the obstacle is determined by tracking its lowest point. Finally, experimental analysis is carried out through simulation and real-vehicle experiments. The results show that the proposed method has higher detection accuracy than YOLO v5s in a sloped road environment and is not susceptible to interference from false obstacles. The most prominent contribution of this research work is to describe a sloped road obstacle detection method, which is capable of detecting all types of obstacles without prior knowledge to meet the needs of real-time and accurate detection of slope road obstacles.

1. Introduction

With increasing public attention to the field of traffic safety, the automobile industry is developing in the direction of intelligence, with many studies on autonomous driving by engineers and scientific researchers. Autonomous driving does not refer to a single technological field, but it is a product of the development and integration of automotive electronics, intelligent control, and breakthroughs related to the Internet of Things [1, 2]. The principle is that autonomous driving systems obtain information on the vehicle and the surrounding environment through an environmental perception system. Then, the information is analyzed and processed by the processor, and the obstacle information in front of the vehicle is detected. Combining with the vehicle dynamics model, the obstacle avoidance path planning and lateral control of the vehicle are realized [3–7].

Environmental perception systems, which need to perform functions such as object classification, detection, segmentation, and distance estimation, have become a key component of autonomous vehicles. These systems can not only provide important traffic parameters for autonomous driving but also perceive surrounding obstacles, such as stationary or moving objects, including roadblocks, pedestrians, and other elements [8]. During the vehicle's movement, radar (laser, millimeter wave), infrared and vision sensors are used to collect environmental information to determine whether a target is in a safe area [9–11]. However, the price of infrared sensors and radars is relatively high, and most of them are limited to advanced vehicles [12]. Compared with other sensor systems, monocular vision requires only one camera to capture images and analyze scenes, thereby reducing the cost of detection solutions. Moreover, the camera can work at a high frame rate and provide rich information from long distances under good lighting and favorable weather conditions [13]; therefore, detection methods based on machine vision are being more and more widely adopted.

Machine learning can be used to achieve object classification for vision-based obstacle detection [14, 15]. However, traditional machine learning methods can only detect known types of obstacles (see Figure 1). If the vehicle cannot detect an unknown type of obstacles accurately, it is very likely that a traffic accident will occur. This situation is not conducive to the safe driving of the vehicle; therefore, in this study, we propose an unsupervised learning-based obstacle detection method, which allows the detection of both known- and unknown-type obstacles in complex environments.

Traditional obstacle detection methods, such as motion compensation [16–18] and optical flow methods [19–22], allow the detection of obstacles of different shapes and at various speeds. However, these methods require the extraction and matching of a large number of object points, which increases the computational load. Therefore, in this study, we adopt a Vision-IMU (inertial measurement unit)-based detection and ranging method, abbreviated as VIDAR, which can realize fast matching and feature point processing of the detection area and improve the obstacle detection speed and detection effectiveness.

VIDAR is an obstacle detection method developed for horizontal roads. When obstacles and test vehicles are located on different slopes, there will be imaging parallax, which will lead to the detection of false obstacles as real ones, resulting in a large measurement error, thereby affecting the detection accuracy. To cope with the impact of slope changes, in this study, we take the slope of road into account during the model establishment, and analyze the specific situation according to the position relationship between the detected vehicle and the obstacle. We thus propose an obstacle detection and distance measurement method for sloped roads based on VIDAR. In the proposed method, slope and distance information are provided by digital maps [23–26].

The rest of this study is structured as follows: in Section 2, we review the research on obstacle detection and visual ranging. In Section 3, the conversion process from world coordinates to camera coordinates and the ranging principle of VIDAR are introduced. In Section 4 the detection process of real obstacles on sloped roads is outlined and the ranging and speed measurement models are established. Simulated and real experiments are presented in Section 5 and the experimental results are compared with the detection results of YOLO v5s to demonstrate the detection accuracy of the proposed method. In Section 6, the proposed method and our findings are summarized, and the study is concluded.

2. Related Work

Obstacle detection still forms one of the most significant research foci in the development of intelligent vehicles. With the improvement and optimization of monocular vision, obstacle detection based on monocular vision has attracted the attention of researchers. Most of the research on the detection of obstacles using monocular vision is based on the optimization of machine vision and digital image processing to improve the accuracy and speed of detection. S. Wang

proposed a novel image classification framework that integrates a convolutional neural network (CNN) and a kernel extreme learning machine to distinguish the categories of extracted features, thus improving the performance of image classification [27]. Nguyen proposed an improved framework based on fast response neural network (Fast R-CNN). The basic convolution layer of Fast R-CNN was formed using the MobileNet architecture, and the classifier was formed using the deep separable convolution structure of the MobileNet architecture, which improved the accuracy of vehicle detection [28]. Yi proposed the improved YOLO v3 neural network model, which introduced the concept of Faster R-CNN's anchor box, and used a multiscale strategy, thus greatly improving the robustness of the network in small object detection [29]. Wang K.W. proposed an efficient fully convolutional neural network, which could predict the occluded part of the road by analyzing foreground objects and the existing road layout, thereby improving the performance of the neural network [30]. Although the above methods improved the accuracy of obstacle detection, they require a large number of sample data for network training and the range of samples must cover all obstacle types; otherwise, the obstacles cannot be detected.

Monocular ranging pertains to the use of a single camera to capture images and perform distance calculations. Zhang et al. used a stereo camera system to compute a disparity map and use it for obstacle detection. They applied different computer vision methods to filter the disparity map and remove noise in detected obstacles, and a monocular camera in combination with the histogram of oriented gradients and support vector machine algorithms to detect pedestrians and vehicles [31]. Tkocz studied the ranging and positioning of a robot in motion, considering the scale ambiguity of monocular cameras. However, only experimental research has been done on the speed and accuracy of measurement [32]. Meng C designed a distance measurement system based on a fitting method, where a linear relationship between the pixel value and the real distance is established according to the pixel position of the vehicle in the imaging plane coordinate, thus realizing adaptive vehicle distance measurement under monocular vision [33]. Zhe proposed a method for detecting vehicles ahead, which combined machine learning and prior knowledge to detect vehicles based on the horizontal edge of the candidate area [34]. These methods were only used for the measurement of distance to other vehicles and are not applicable to other types of obstacles.

Rosero proposed a method for sensor calibration and obstacle detection in an urban environment. The data from a radar, 3D LIDAR, and stereo camera sensors were fused together to detect obstacles and determine their shape [35]. Garnett used a radar to determine the approximate location of obstacles, and then used bounding box regression to achieve accurate positioning and identification [36]. Caltagirone proposed a novel LIDAR-camera fusion fully convolutional network and achieved the most advanced performance on the KITTI road benchmark [37]. Although sensor fusion methods reduce the processing load and achieve improved detection accuracy, these methods are based on flat roads and are not suitable for complex slope road environments.

FIGURE 1: Fast R-CNN. Normal cars are detected, but the overturned car and the box are not detected.

To solve the above problems, we propose an obstacle detection and distance measurement method for sloped roads based on VIDAR. This method does not require a priori knowledge of the scene and uses the road slope information provided by a digital map and the vehicle driving state provided by an IMU to construct distance measurement and speed measurement models, which allow the detection of obstacles in real time, as well as the distance and movement state of the obstacles.

3. Methodology

The obstacle detection model of VIDAR is based on pinhole camera model, which can accurately calculate the distance between vehicles and obstacles.

3.1. Coordinate Transformation.
The camera can map the coordinate points of the three-dimensional world to the two-dimensional imaging plane. This imaging principle is consistent with the pinhole model principle, so camera imaging can be described by pinhole model.

If we want to determine the correspondence between the object point and the image point, we must establish the coordinate system needed by vision system, including world coordinate system, camera coordinate system, imaging plane coordinate system, and pixel coordinate system. The transformation process from the world coordinate system to the pixel coordinate system is shown in Figure 2.

Pixel coordinate (u, v) and image plane coordinate (x, y) are on the same plane, and the X and Y axes are parallel. The corresponding position of the original point in the image plane coordinate system is (u_0, v_0). Both the world and the camera coordinate systems are 3D coordinates, which are associated through the camera. According to the principle of keyhole imaging, the camera coordinate system can be obtained through a transformation of the coordinate axes of the world coordinate system, so the conversion relation between the two coordinate systems must be deduced. The conversion

equation from the world to the pixel coordinate system is shown in

$$Z_C = \begin{bmatrix} \mu \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} 1/d_x & 0 & \mu_0 \\ 0 & 1/d_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} + T \end{bmatrix}$$

$$= \begin{bmatrix} a_x & 0 & \mu_0 \\ 0 & a_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R \begin{matrix} X_w \\ Y_w \\ Z_w \end{matrix} \end{bmatrix} + T,$$

(1)

where R and T are the external parameters. The internal and external parameters can be obtained through camera calibration.

3.2. Obstacle Ranging Method.
The obstacle ranging principle is also based on the pinhole model principle. For the convenience of expression, we installed the camera on a test vehicle and a vehicle on a sloped road was regarded as the obstacle. The feature points of the obstacle were detected, and the lowest point was taken as the intersection point between the obstacle and the road surface (see Figure 3). In the case of normal detection by the system, the camera collects image information, and by processing the image information, feature points in the image can be extracted. By measuring the distance of the feature point, it can be determined whether the obstacle where the feature point is located has a height. For real obstacles, tracking the feature point at the lowest position can calculate the moving speed of the obstacle, judge the motion state of the obstacle, and provide data support for the safe driving of the vehicle. As long as the camera can capture images normally, all obstacles in the captured scene can be detected. The number of detected obstacles is related to the number of extracted feature points.

Let f be the effective focal length of the camera, ∂ be the pitch angle, μ be the pixel size, h be the mounting height of the camera and the camera center be the optical center of the lens. Let (x_0, y_0) be the coordinate origin of imaging plane
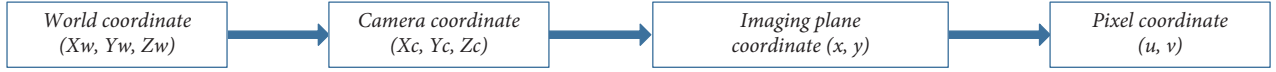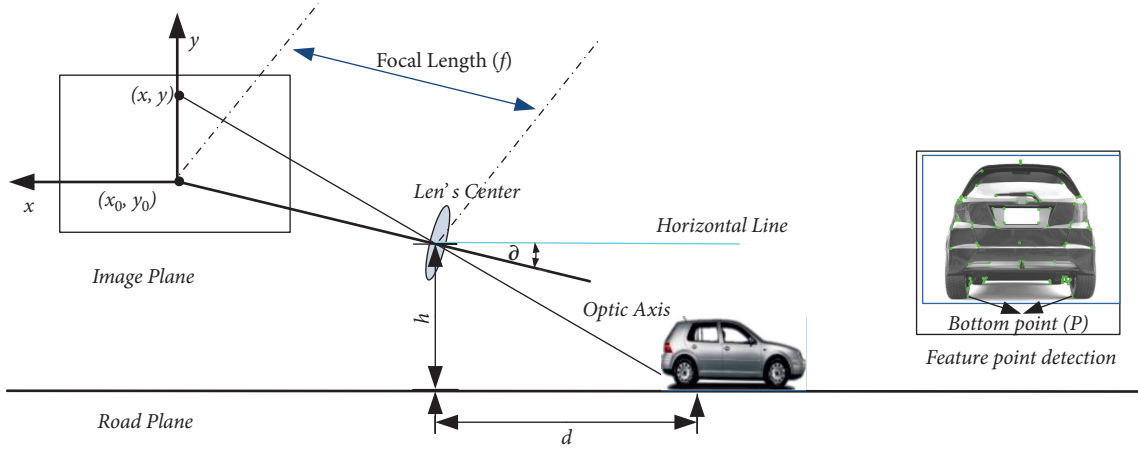
FIGURE 2: Transformation between coordinate systems.

FIGURE 3: Schematic diagram of obstacle ranging model (in order to visualize the detection principle, the nonreal proportional relationship is shown in the figure).

coordinate system, and (x, y) be the intersection coordinate of the obstacle and the road plane in the image plane coordinate system. The horizontal distance between the camera and the obstacle can be obtained using

$$d = \frac{h}{\tan\left(\vartheta + \arctan\left[(y_0 - y)\mu/f\right]\right)}. \tag{2}$$

4. Research Approach

In the traditional VIDAR model, it is assumed that the test vehicle and obstacles are on the same plane. However, when the test vehicle and the obstacles are on roads with different slopes, this will cause a deviation of the distance measurement. In order to enhance the visual detection accuracy and expand the visual ranging application scenarios, in this study, we take the slope into account and establish an obstacle detection model for the sloped road.

4.1. Establishment of the Distance Measurement Model. The sloped road mentioned in this study refers to a road where the test vehicle and the obstacles are not on the same slope. When measuring distance, the above situation can be simplified into two models.

The distance model between the camera and obstacles on a sloped road with obstacles in front of the test vehicle are shown in Figure 4. Let the light blue line be the auxiliary line, and the red dot on the obstacle be any detected object point.

Let C' be a point on the road's surface, C be the image point of C' on the sloped road's surface, C'' be the intersection point where CC' extends to the imaginary horizontal plane, and S' be the distance from the camera to the beginning of road slope change. Let d_i be the horizontal

distance between the camera and C'', and d_{ii} be the horizontal distance between the camera and C'.

Using triangle similarity, equation (3) can be obtained through the geometric relationships shown in Figure 4:

$$\begin{cases} \dfrac{h_1}{d_{ii} - S'} = \tan\alpha, \\[3mm] \dfrac{h + kh_1}{d_{ii}} = \tan\theta_i, \\[3mm] \theta_i = \partial + \dfrac{\arctan(y_0 - y_i)\mu}{f}. \end{cases} \tag{3}$$

The expression of d_{ii} is further derived by

$$d_{ii} = \frac{S' * \tan\alpha + kh}{\tan\alpha + k\tan\theta_i}. \tag{4}$$

When the slope of the road where the obstacle is located is larger than that of the test vehicle, $k = -1$. In the opposite case, $k = 1$.

4.2. Determination of the Real Obstacle on Sloped Roads. In the process of the test vehicle's movement, road images were collected twice. The imaging diagram of the stationary obstacles is shown in Figure 5. Let A' and B' be points on the road surface, and A and B be the corresponding image points. The first point of the obstacle on the image plane is A. As the camera moves with the test vehicle, and the y axis on the image plane moves from axis y_1 to the axis y_2, we obtain the point B of the obstacle on the image plane. A'' is the intersection, where AA' extends to the imaginary horizontal plane, and accordingly for B''. Δd is the movement distance
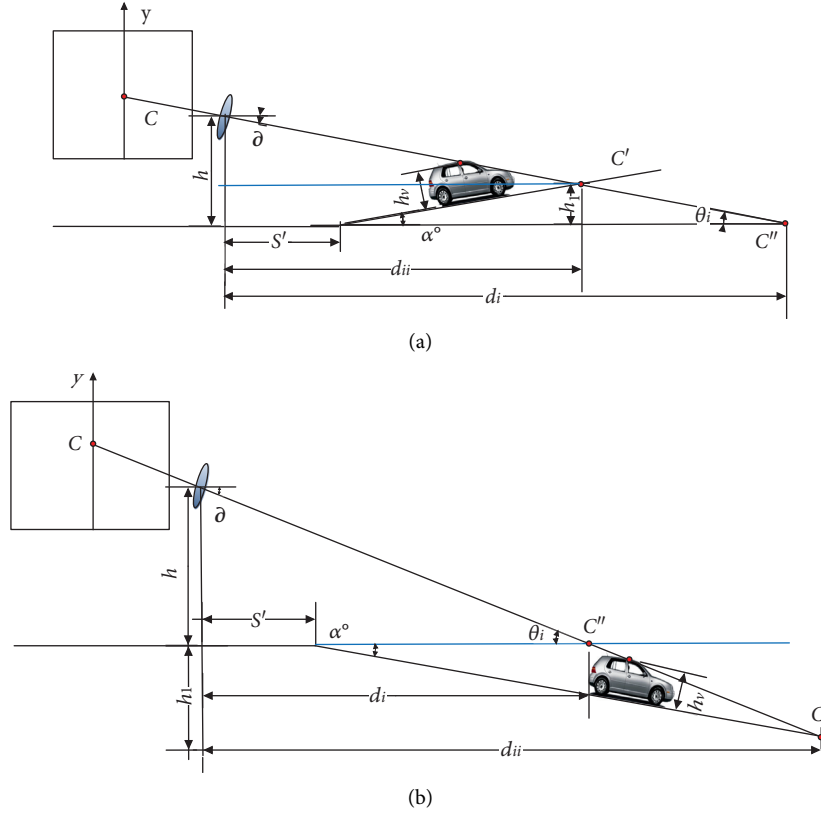
(a)

(b)

FIGURE 4: Diagram of distance models. (a) Situation 1: the test vehicle on a flat road and obstacles on an uphill road. (b) Situation 2: the test vehicle on a flat road and obstacles on a downhill road.

of the camera (i.e., the test vehicle), d_1 is the horizontal distance from the camera to A'', d_2 is the same, d_{11} is the horizontal distance from the camera to A', and accordingly for d_{22}.

d_{11} and d_{22} can be calculated using equation (4). The relationship between d_{11} and d_{22} can be approximated as $d_{11} = d_{22} + \Delta d$, but the real relationship is $d_{11} = d_{22} + \Delta d + \Delta l$. If $d_{11} \neq d_{22} + \Delta d$, the object points are not on the road surface. Using this method, it can be determined whether the obstacle has a height (i.e., it is a real obstacle).

4.3. Special Case of Obstacle Detection. A special case should be excluded during obstacle detection. When the test vehicle and the obstacles are moving at the same time, the imaging point of the camera light on the road surface through an object point of the obstacle coincides with each other. VIDAR is unable to detect obstacles in this case.

The diagrams of obstacle detection in complex environments are shown in Figure 6. Let L_A be the distance (along the road where the obstacle is located) between the highest point of the obstacle and the object point of the road surface when the test vehicle is moving for the first time. Similarly, L_B is the distance when the vehicle moves for the second time. The letters in Figure 6 have the same meaning as the letters above.

Let the speeds of the detecting vehicles and obstacles be v and v', respectively. When the imaging point of the road's intersection point and the obstacle's object point passes through the camera, the relationships between h_v, v, v', L_A and L_B are as follows:

$$
\begin{cases}
\dfrac{h_v}{\tan(\alpha + b\theta_1)} = L_A, \\[3mm]
\dfrac{h_v}{\tan(\alpha + b\theta_2)} = L_B, \\[3mm]
L_A - L_B = v't, \\[3mm]
\Delta d = vt.
\end{cases}
\tag{5}
$$

When the slope of the road where the obstacle is located is larger than that of the test vehicle, $b = 1$, while $b = -1$ in the opposite case.

Therefore, VIDAR can be used in all cases except when $L_A - L_B = v't$, and $v' = v/\Delta d (h_v/\tan(\alpha + b\theta_1) - h_v/\tan(\alpha + b\theta_2))$. Therefore, the proposed method using a monocular camera to detect obstacles on sloped roads is convenient and feasible. The detection process only includes tracking and calculating the
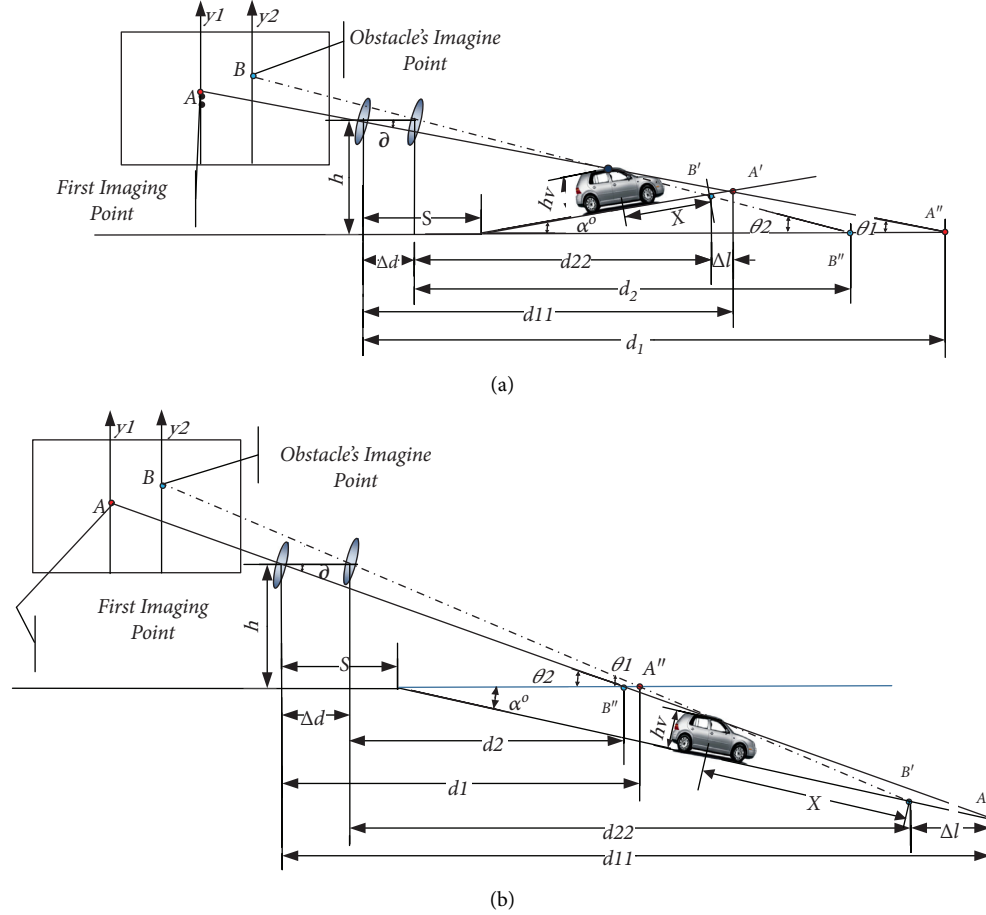
(a)

(b)

FIGURE 5: Schematic diagram of stationary obstacle imaging. (a) Situation 1: the test vehicle moving on a flat road and stationary obstacle on an uphill road. (b) Situation 2: the test vehicle moving on a flat road and stationary obstacle on a downhill road.

position of the object point, which can shorten the detection time and reduce computational resource consumption.

4.4. Speed Measuring Model of the Sloped Road Obstacle.
Obstacles can be imaged in the camera photosensitive element. By extracting and calculating the feature points of the collected obstacle images, we can calculate the feature points that are not on the road surface, that is, the feature points whose height is not zero. The object points with nonzero height are morphologically processed to obtain the obstacles' areas. The movement state of the obstacles can be determined through tracking and calculating the speed of the lowest point.

When the test vehicle is moving, the obstacles, camera, and the lowest point of the road will form images (see Figure 7). At this time, the horizontal distance between the lowest point of the obstacle and the camera can be expressed as d_{ii}.

Let A be the image plane point corresponding to the lowest point of the obstacle at time t and B corresponding point at $t + \Delta t$. The relationship between v' d_{11}, d_{22} and Δd is as follows:

$$\frac{|d_{11} - (d_{22} + \Delta d)|}{\cos \alpha} = v' \cdot \Delta t, \tag{6}$$

where $\Delta d = v * \Delta t$, with v being the speed of the test vehicle. When $d_{11} = d_{22} + v * \Delta t$, the obstacle is stationary; otherwise, it is moving with a speed of

$$v' = \frac{|d_{11} - d_{22} - v \cdot \Delta t|}{\Delta t * \cos \alpha}. \tag{7}$$

4.5. Obstacle Detection on Sloped Roads Using VIDAR.
In this study, an obstacle detection and distance measurement method for sloped roads based on VIDAR is proposed, which can quickly judge and eliminate false obstacles that without height, and at the same time identify real obstacles and judge their movement state. The detection process is as follows (see Figure 8).

Step 1. Update camera parameters using the IMU:

(1) Calibration of the camera's initial internal and external parameters: the camera's parameters, such as the focal length f, mounting height h, pixel μ, and pitch angle ∂ are obtained through calibration.

(2) Data acquisition: the camera is used to collect images and the IMU is used to collect inertial data. The
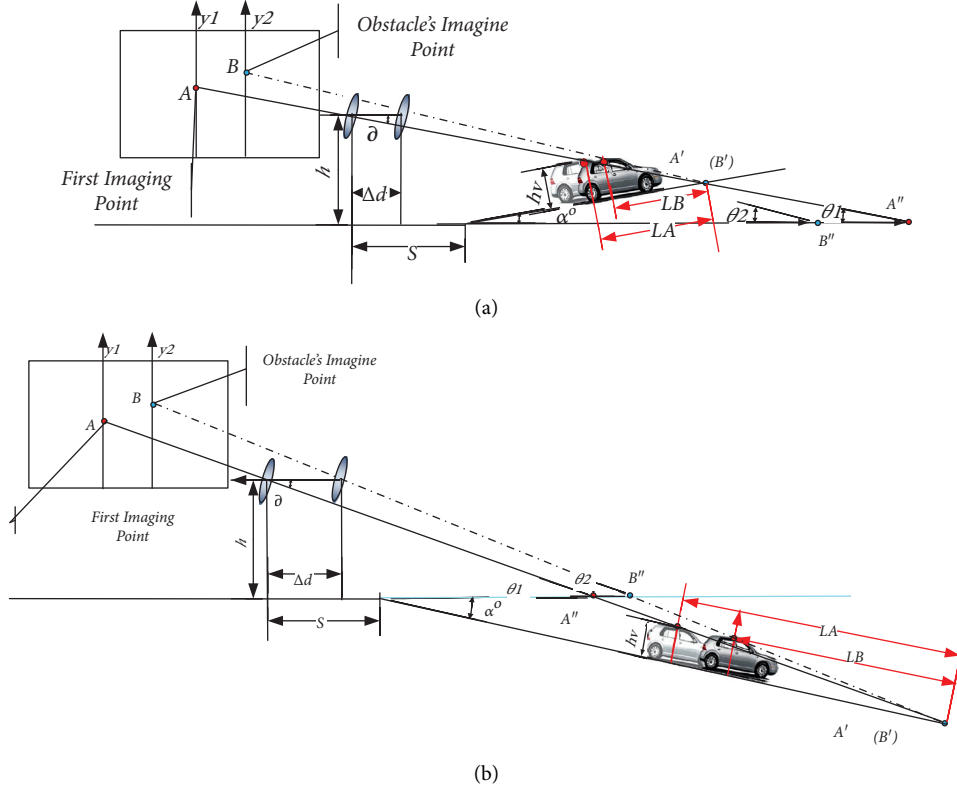
(a)

(b)

FIGURE 6: Schematic diagram of obstacles and camera imaging in complex environments, (a) Situation 1: Test vehicle driving on a flat road and dynamic obstacles on an uphill road, (b) Situation 2: Test vehicle driving on a flat road and dynamic obstacles on a downhill road.

acquisition frequency of the IMU is larger than that of the camera.

(3) Update of camera parameters: the frequency relationship between the IMU and camera is established and the camera parameters at time t are calculated periodically according to the inertial data.

Step 2. Obtain the road information.

Acquire the road slope α and the distance S from the test vehicle to the sloped road using the digital map.

Step 3. Regional background extraction:

(1) Two consecutive images are taken as the total obstacle detection area B_i during the running of the test vehicle (see Figures 9(a) and 9(b)).

(2) The lane line is detected and the image within the lane line is extracted as G_i.

(3) Machine learning is used to process the images, detect and classify specific types of obstacles. The area set F_i of known types of obstacles is obtained, where $F_i = \{f_1, f_2, \ldots, f_k\}$, and k is the number of known obstacles.

(4) The known obstacle area F_i in the total obstacle detection area G_i is eliminated and the background area N_i ($N_i = G_i - F_i$) is extracted as the VIDAR data to be detected.

Step 4. Image processing and obstacle detection:

(1) Object points are extracted from the background areas N_i and N_{i+1} of two consecutive images. With N_i as the background region template map and N_{i+1} as the background region real-time map, the matching regions M_i and M_{i+1} are obtained using a fast image region matching method based on region feature extraction, as shown in Figure 9(c).

(2) The object points set P_i of matching area M_{i+1} is extracted, as shown in Figure 9(d).

(3) The distance between the test vehicle and the object point is calculated. The horizontal distance between the camera and the imaged object point on the imaginary road is $d_i = h/\tan \theta_i$. The horizontal distance between the camera and the imaged object point on the real road is $d_{ii} = S' \cdot \tan \alpha + kh/\tan \alpha + k \tan \theta_i$. The calculation process of d_{ii} is shown in Figure 10. First, the pixel coordinates of object points are obtained through the transformation of the coordinate axes. Then the slope information is obtained through Step 2, and finally the distance is obtained through the ranging model.

(4) The object points with height in set P_i are extracted (see Figure 9(e)). Calculate d_{ii} and d_{i+1i+1} as the vehicle is moving continuously. If $d_{ii} = d_{i+1i+1} + \Delta d$, then the object points are on the road surface (without height), so the object points p_{ij} are
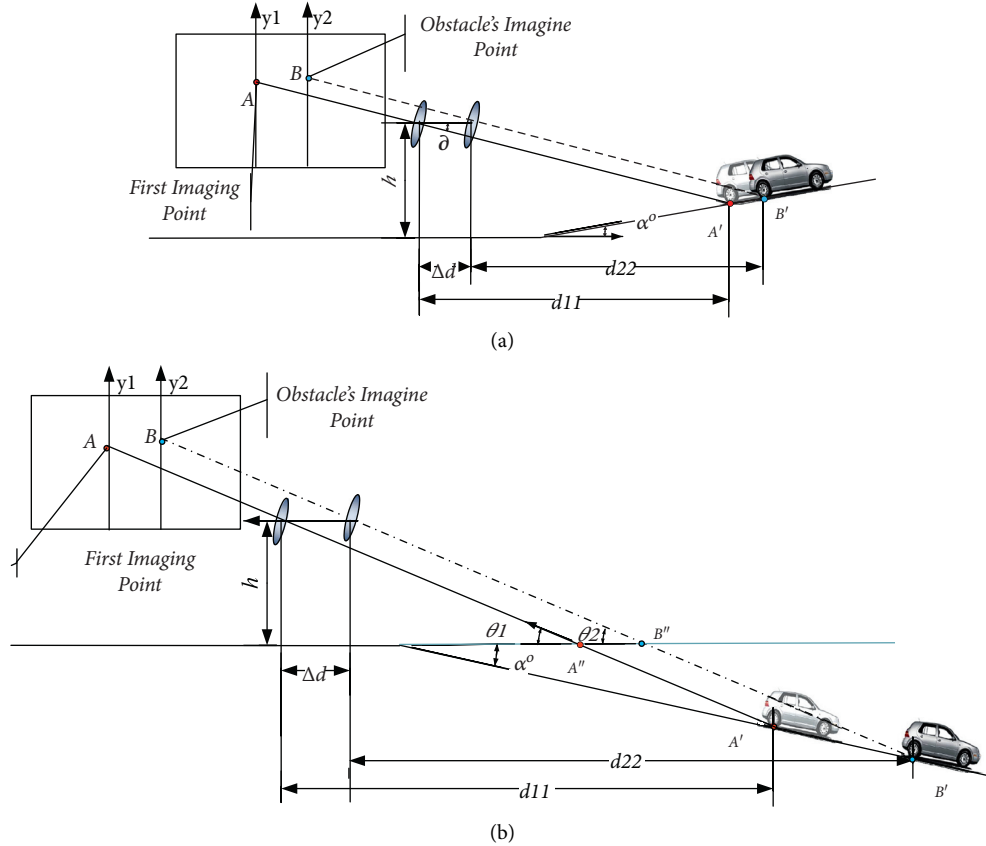
(a)

(b)

FIGURE 7: Schematic diagram of the camera and the lowest point on the road. (a) Situation 1: the test vehicle moving on a flat road and dynamic obstacles on an uphill road. (b) Situation 2: the test vehicle moving on a flat road and dynamic obstacles on a downhill road.

eliminated. If $d_{ii} \neq d_{i+1i+1} + \Delta d$, then the object points are not on the road surface (i.e., they have a nonzero height). The object points are extracted to obtain the object point set p_i'.

(5) Morphological processing is applied to the image of the object point set p_i' (Figure 9(f)). The target image is E_i, and the structural element is B, which is used to apply a closing operation on E_i and obtain C connected regions. The real obstacle region O_i is thus obtained, where $O_i = E_i \cdot B_i = (E_i \oplus B_i) \ominus B_i$, and $O_i = (o_{i,1}, o_{i,2}, \ldots, o_{i,c})$.

(6) Edge detection of real obstacles, shown in Figure 9(g).

(7) According to the detection result of (6), the lowest object point of each obstacle area is extracted, as shown in Figure 9(h). The lowest object point set P_i'' constitutes the obstacle area.

(8) Each object point in P_i'' is tracked during the movement of the test vehicle.

(9) Get the movement state of the obstacles is obtained. The movement speed of the obstacle where the object point is located can be obtained by tracking each object point in P_i''. If $|d_{i+1i+1} - d_{i+2i+2}| = v_{i+1} \cdot \Delta t$, the obstacle on which these object points are located is static. If $|d_{i+1i+1} - d_{i+2i+2}| \neq v_{i+1} \cdot \Delta t$, the obstacle is

moving with an instantaneous speed $v_{i+1}' = |d_{i+1i+1} - d_{i+2i+2} - v_{i+1} \cdot \Delta t| / \Delta t * \cos \alpha_i$.

The proposed obstacle detection method can be used to detect real obstacles in complex environments and determine their movement state, which is beneficial for vehicles to take timely measures and avoid accidents.

5. Experiment and Evaluation

The proposed method can be used for obstacle detection in complex environments with improved accuracy, as well as distance and speed measurement of obstacles. Obstacle detection and distance measurement were realized in Matlab, whereas all experiments were performed on a desktop PC with the Intel(R) Xeon(R) Silver 4210 CPU.

5.1. Simulation Experiment. In this study, experimental equipment was used to simulate a detection environment so as to verify the detection effect of obstacles on sloped roads based on VIDAR. The experimental equipment included: a test vehicle equipped with an OV5640 camera unit and a JY61p IMU (Figure 11(a)), vehicle scale models (Figure 11(b)), bottle caps and paper (Figure 11(c)), and simulated sloped road (Figure 11(d)). Among them, the test vehicle was used to analyze the road environment and detect its own driving state, scaled vehicle models were used to
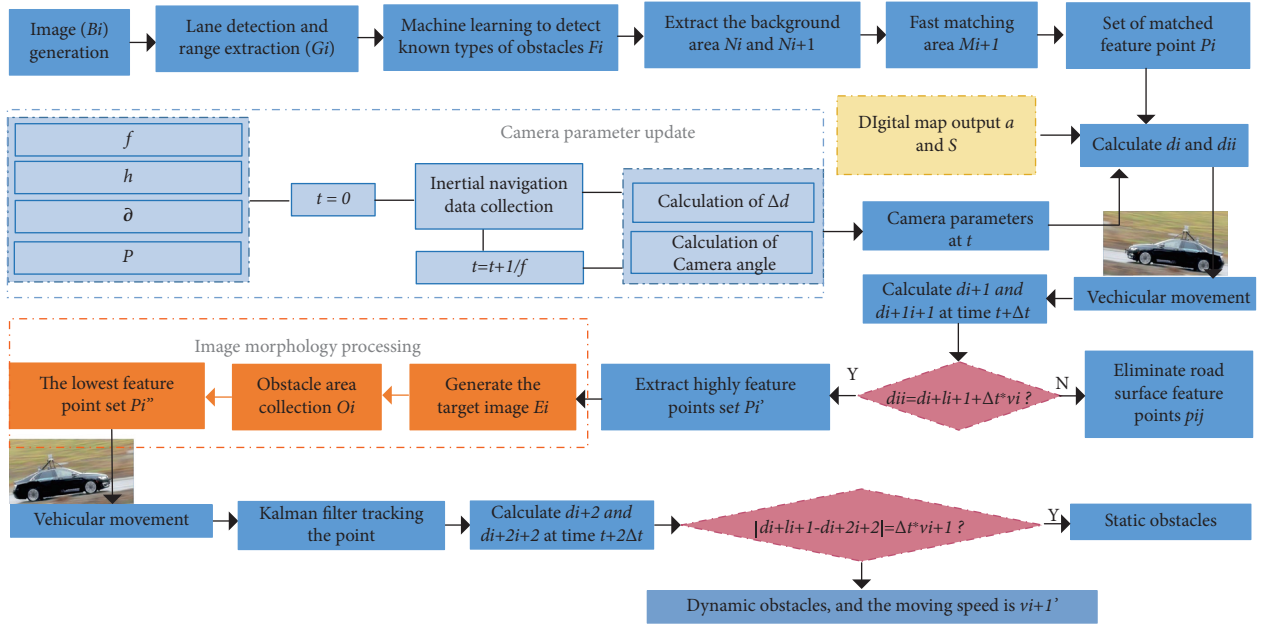
FIGURE 8: Flow chart of slope road obstacle detection method based on VIDAR (see the following for specific steps).
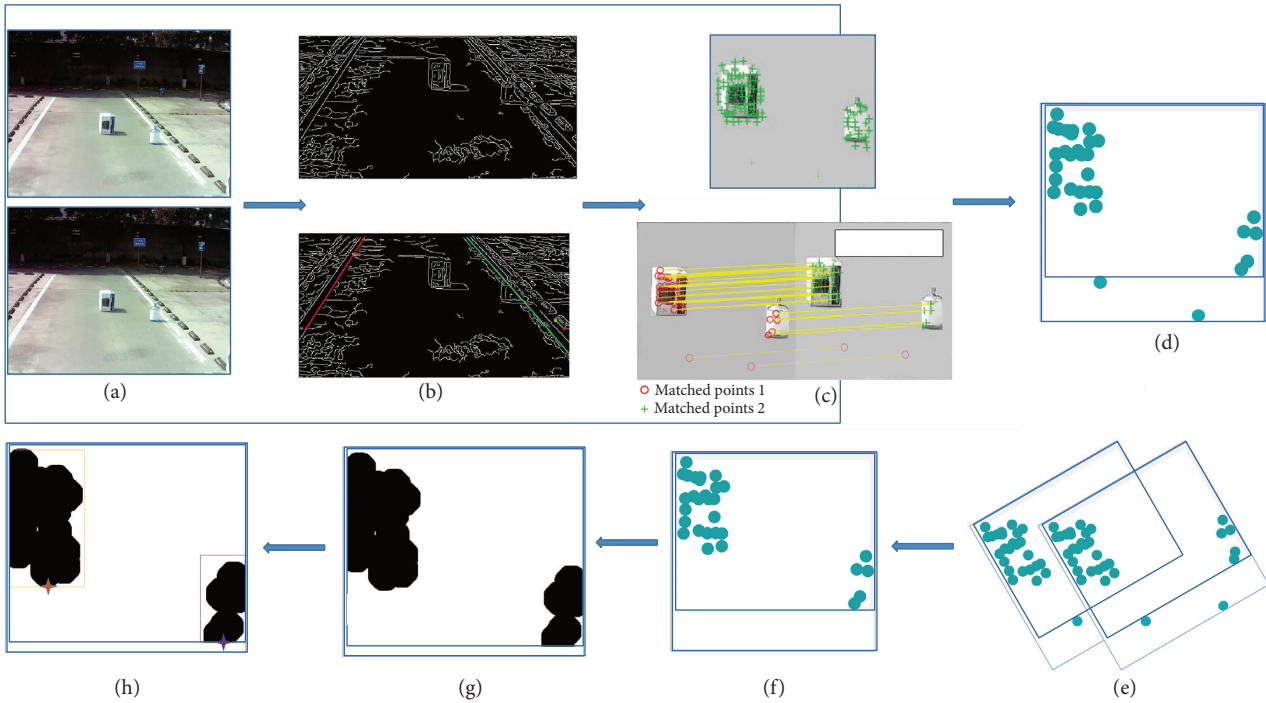

FIGURE 9: Obstacle detection based on VIDAR and digital map. (a) Two images collected during the movement of the vehicle. (b) Lane line detection on the image and extraction of the detection range. (c) Feature point detection and matching on the images within the range. (d) Extraction of feature points. (e) Determination of whether the extracted feature points have height. (f) Removal of feature points without height. (g) Morphological processing on feature points. (h) Tracking of the lowest point of obstacles.

simulate known obstacles, while the bottle caps and paper were used to simulate unknown obstacles. The road slope was set to 13°.

The bottle cap was taken as the real obstacle of unknown type, and the paper pasted on the simulated road was taken as the pseudoobstacle of unknown type. The angular

acceleration and acceleration data of the vehicle were obtained by the IMU installed by the vehicle. Quaternion method is used to solve the camera attitude, and the pitching angle of the camera is updated. The velocity data are used to calculate the horizontal distance between the vehicle and the obstacle. The height of the obstacle is calculated by the
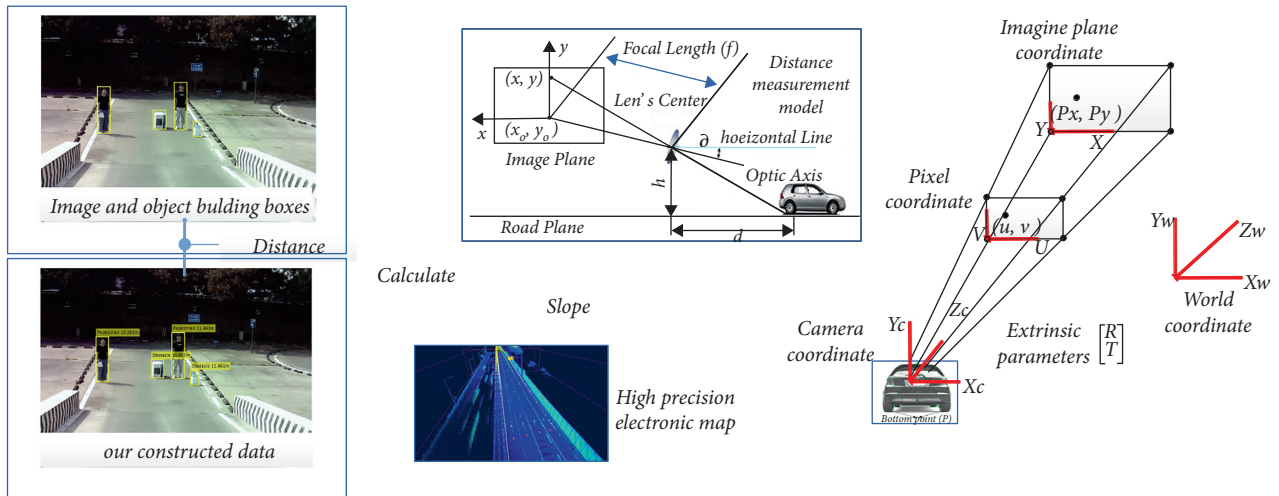
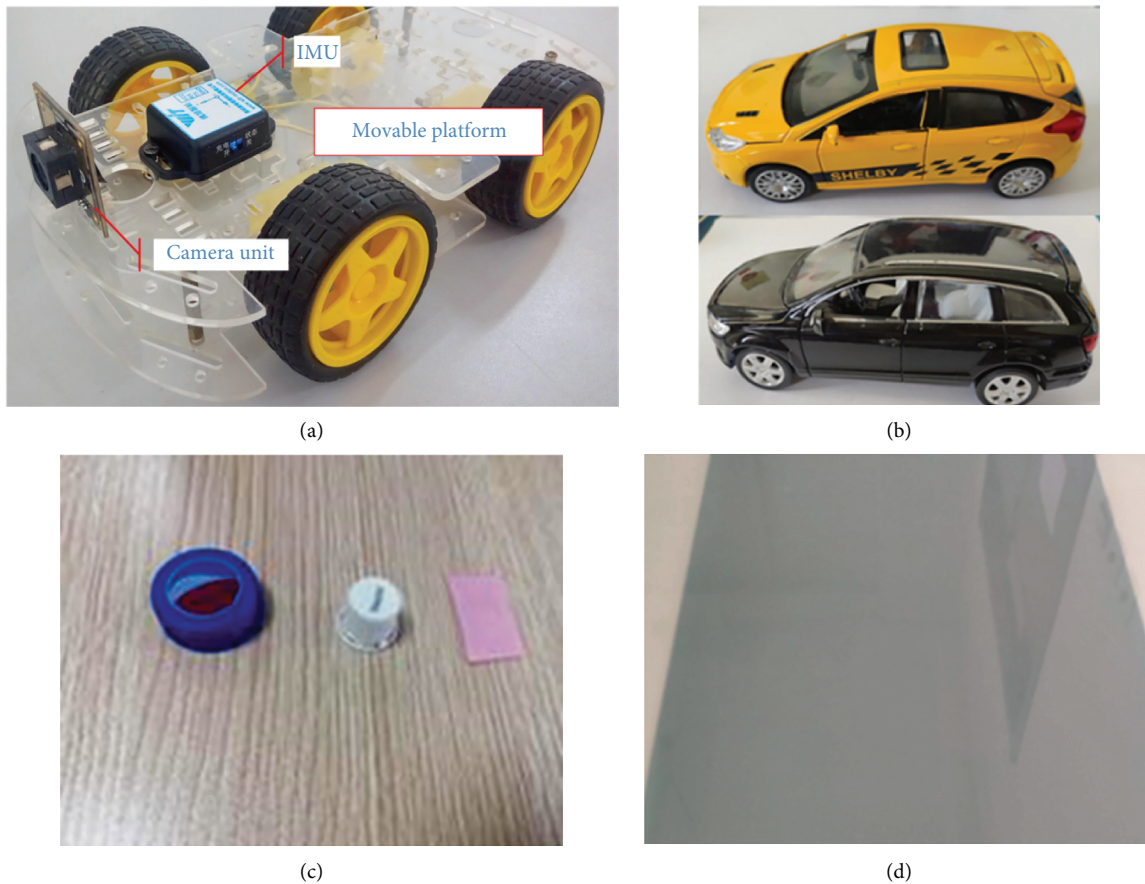FIGURE 10: Flow chart of horizontal distance calculation.

FIGURE 11: Equipment for simulating experiment. (a) Test vehicle. (b) Vehicle scale models. (c) Bottle caps and paper. (d) Simulated sloping road.

change of the distance before and after the movement, so as to determine whether the detected obstacle is a real obstacle. The video collected using the OV5640 camera comprised an image sequence at 12FPS, which was used for obstacle detection. The results obtained using the original VIDAR and VIDAR on sloped roads are shown in Figure 12, while the test results of the simulation experiment are summarized in Table 1.

The test results of the simulation experiment is shown in Table 1.

It can be seen from Figure 12 that the original VIDAR can detect unknown types of obstacles such as bottle caps, but it will detect false obstacles as real obstacles, resulting in low accuracy of obstacle detection. However, the obstacle detection method for sloping roads based on VIDAR can

(a)   (b)
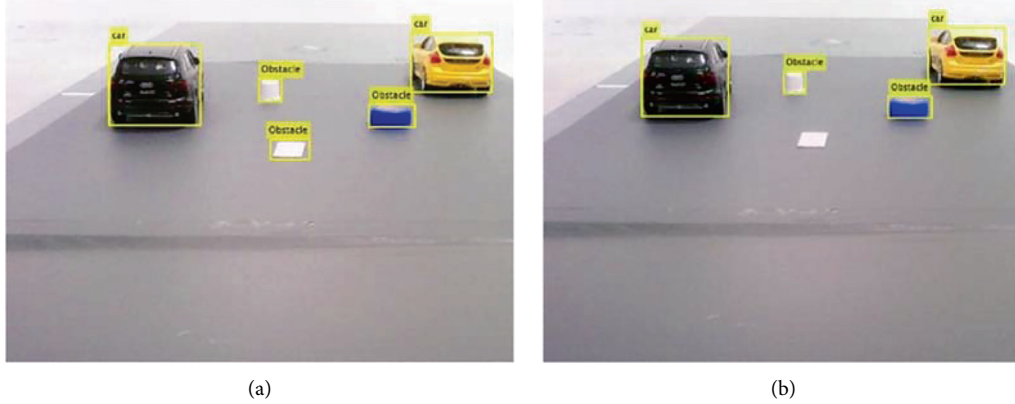
FIGURE 12: Comparison of obstacle detection effect in simulated environment. (a) Original VIDAR. (b) VIDAR on sloped roads.

TABLE 1: Test results of the simulation experiment.

Obstacle	The i th movement of the test vehicle	Detection distance (cm)	The test vehicle moving distance	Whether it has height	Whether to exclude
1	1	16.5	3	No	Yes
	2	19.5			
2	1	18.5	3	Yes	No
	2	21.2			
3	1	24.5	3	Yes	No
	2	27.0			

eliminate false obstacles, which makes up for the wrong detection of unknown types of obstacles on sloping roads; therefore, compared with the original VIDAR, our proposed method can detect obstacles more accurately.

5.2. Real Environment Experiment.

In the real environment, purely electric vehicles were used as test vehicles (see Figure 13). As a sensor, the camera can adapt to complex environments and collect environmental information in real time (we only used the left camera). The camera was installed at a height of 1.60 m. The IMU used for locating the test vehicle and reading of its movement state in real time was installed at the bottom of the test vehicle. GPS was used for accurate location positioning. Through the combination of GPS and IMU, the real-time position information of the test vehicles and obstacles can be obtained, and then the trajectory information of vehicles and obstacles can be obtained. A digital map was used to obtain accurate road information such as distance and slope. A calculation unit was used to process the data in real time.

Accurate calibration of camera parameters was a prerequisite for the whole experiment and is a very important task for obstacle detection methods. In this paper, Zhang Zhengyou's camera calibration method was adopted to calibrate the DaYing camera. First, the camera was fixed to capture images of a checkerboard at different positions and angles. Then, the key points of the checkerboard were selected and used to establish a relationship equation. Finally, the internal parameter calibration was realized. The camera calibration result is shown in Figure 14.

Camera distortion includes radial distortion, thin lens distortion, and centrifugal distortion. The superposition of the three kinds of distortion results in a nonlinear distortion, the model of which can be expressed in the image coordinate system as follows:

$$\begin{cases} \delta_x(x, y) = s_1 x(x^2 + y^2) + 2p_1 xy + p_2 y^3 + k_1 x(x^2 + y^2) \\ \delta_y(x, y) = s_2 y(x^2 + y^2) + 2p_2 xy + p_1 y^3 + k_1 x(x^2 + y^2) \end{cases},$$

(8)

where s_1 and s_2 are centrifugal distortion coefficients; k_1 and k_2 are radial distortion coefficients, and p_1 and p_2 are the distortion coefficients of thin lenses.

Because the centrifugal distortion of the camera is not considered in this study, the internal reference matrix of the camera can be expressed as shown in

$$M = \begin{bmatrix} 5.9774e + 03 & 0 & 949.8843 \\ 0 & 5.9880e + 03 & 357.0539 \\ 0 & 0 & 1 \end{bmatrix}.$$

(9)

The calibration of the camera's external parameters can be calculated by taking the edge object points of lane lines. The calibration results are shown in Table 2.

Since the images in the public data set were all data collected by other cameras, different camera parameters will affect the accuracy of ranging, so we used the VIDAR-Slope database (Figure 15), the images in which where collected using a DaYing camera. The collection frequency was 77 frames/min, and there are 2270 images in total. The experiment and image collection took place in Shandong

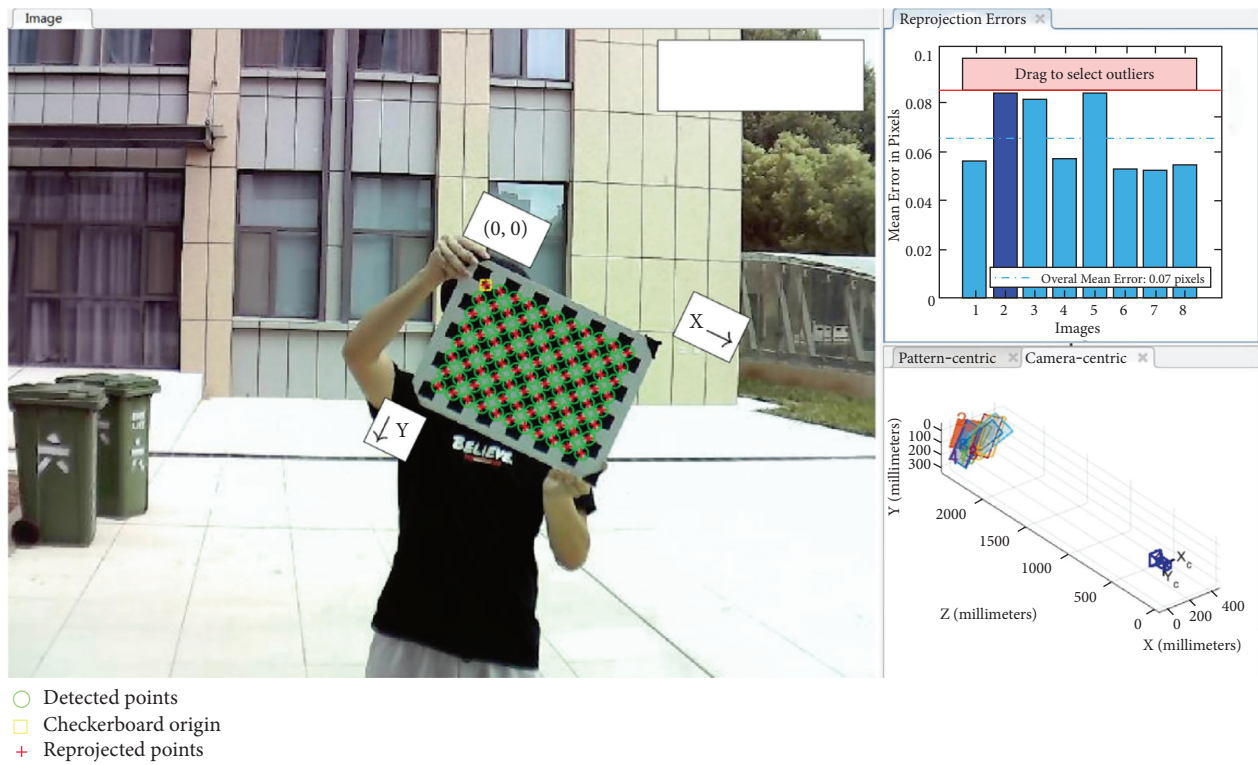FIGURE 13: Schematic diagram of the test vehicle.

○ Detected points
□ Checkerboard origin
+ Reprojected points

FIGURE 14: Schematic diagram of camera calibration results.

TABLE 2: Calibration results of camera external parameters.

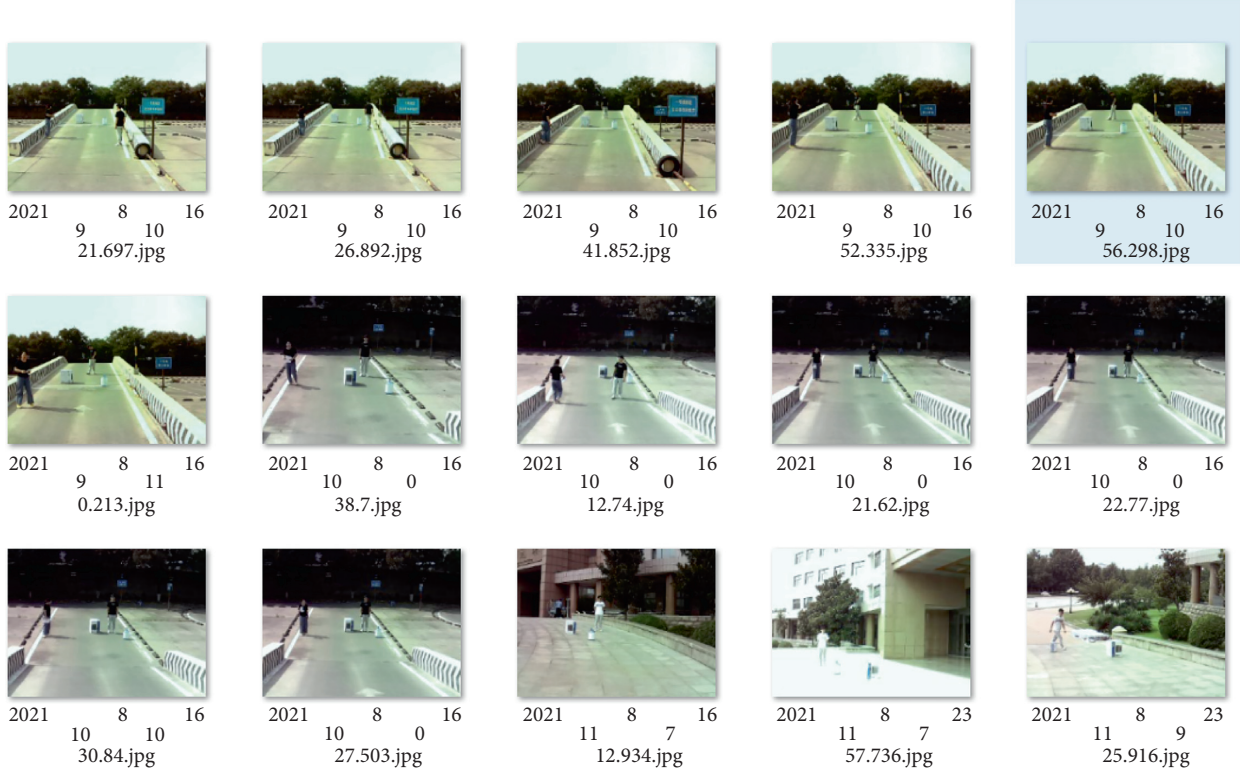| External parameter type	Parameter size
Pitch angle	1.20
Yaw angle	3.85
Rotation angle	2.37

FIGURE 15: The VIDAR-Slope database (part of data).

University of Technology's driving school and experimental building. We selected the downhill section of the parking lot for the experiment. In the process of obstacle detection, the test vehicle moves at a constant speed of 25 km/h.

The detection results of YOLO v5s and the method proposed in this study are shown in Figure 16. The accuracy of obstacle detection was measured through the number of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). Let a be an obstacle that is correctly classified as a positive example, b be an obstacle that is wrongly classified as a positive example, c be an obstacle that is correctly classified as a negative example, and d be an obstacle that is incorrectly identified as a negative example. Then, $TP = \sum_{i=1}^{n} a_i$, $FP = \sum_{i=1}^{n} b_i$, $TN = \sum_{i=1}^{n} c_i$, $FN = \sum_{i=1}^{n} d_i$.

The YOLO series is a representative target detection framework based on deep learning. There are four versions of the target detection network: namely YOLO v5s, YOLO v5m, YOLO v5l, and YOLO v5x. Among them, YOLO v5s is the smallest and has the fastest speed, so we choose it for comparative experiments.

Comparing the two methods, it can be seen that the stability of the proposed method is higher than that of YOLO v5s.

YOLO v5s lacks training in unknown types of obstacles, and will consequently offer reduced safety when used in realistic vehicle situations. However, the proposed obstacle detection method does not require training and can detect all types of obstacles, thus ensuring its effectiveness of obstacle detection results on sloped roads. The total number of obstacles in the target area in the VIDAR-Slope database was 9526. The results of YOLO v5s and proposed method are shown in Table 3.

In the results' analysis, Accuracy (A), Recall (R), and Precision (P) were used as evaluation indices for the two obstacle detection methods, calculated through the following equations:

$$A = \frac{TP + TN}{TP + TN + FP + FN}, \tag{10}$$

$$R = \frac{TP}{TP + FN}, \tag{11}$$

$$P = \frac{TP}{TP + FP}. \tag{12}$$

The Accuracy, Recall, and Precision of YOLO v5s and the method proposed in this study are shown in Table 4.
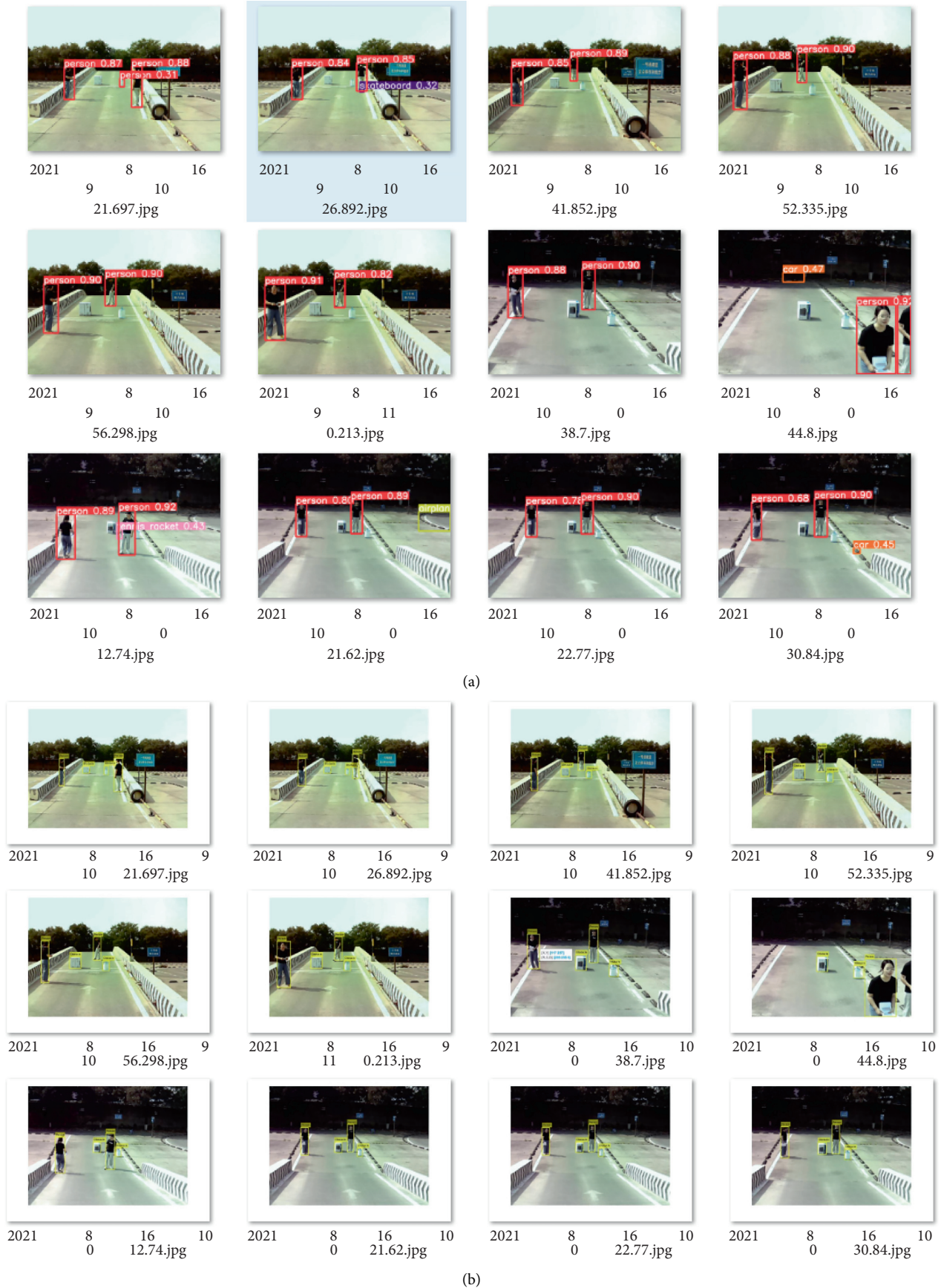
(a)

(b)

FIGURE 16: Comparison of partial test results of YOLO v5s and proposed method. (a) YOLO v5s can only detect known type obstacles. (b) The obstacle method proposed in this article can detect pedestrians (known types) and boxes (unknown types).

TABLE 3: Obstacle detection results in sloped roads of YOLO v5s and the proposed method.

Detection method	Input value	TP	FP	TN	FN
YOLO v5s	9526	6716	2808	915	1214
Proposed method	9526	9124	402	530	485

TABLE 4: Evaluation indices of YOLO v5s and the proposed method.

Detection method	A (%)	R (%)	P (%)
YOLO v5s	81.73	84.68	70.50
Proposed method	90.36	94.94	95.77

TABLE 5: Evaluation indices of YOLO v5s and the proposed method.

Detection method	A (%)
Fast R-CNN	76.53
SSD	71.28
Fast YOLO	78.95
SSD 300	73.96
YOLO v5s	81.73
Proposed method	90.36

TABLE 6: Evaluation indices of detections of YOLO v5s and the proposed method.

Detection method	Detection time (s)
YOLO v5s	0.164
Proposed method	0.201

The experimental results in Tables 3 and 4 show that due to vehicle fluctuations and other factors, misjudgment or misdetection may occur during vehicle movement. Compared with YOLO v5s, the accuracy of the obstacle detection method proposed in this study is increased by 8% and its precision is increased by 26.4%, which demonstrates its improved obstacle detection capability on sloped roads.

In terms of detection accuracy, we also compared our method with other commonly used target detection methods. The detection results are shown in Table 5. It is evident that the proposed obstacle detection method achieves an accuracy higher than state of the art methods.

The real-time nature of obstacle detection refers to the ability to process every image frame collected in time. In terms of detection speed, YOLO v5s and the proposed method were used to process 2270 images and the respective average obstacle detection times were calculated. The results are shown in Table 6.

Since the average detection time of the method proposed by us is 0.201s, in order to ensure the detection of obstacles under normal driving conditions, the speed of the detected vehicle must be less than or equal to the ratio of detection distance to the average detection time.

Compared with YOLO v5s, the method proposed in this study saves the training step of data set. The modified method firstly uses machine learning to detect obstacles of known types, but it needs to process feature points of obstacles of unknown types, so the final detection time is longer than that of YOLO v5s. But it can still meet the demand of real-time detection.

In order to verify the reliability of the distance measurement method proposed in this study and the feasibility of practical application, we have done a set of obstacle detection experiments. We first use a fixed camera to take pictures of the real road environment ahead and record the process. The result of IMU data processing is shown in Figure 17. Then we select a few frames of images during the progress of the obstacle for processing. Finally, the distance between the camera and the obstacle in front is calculated, and the detection result is shown in Figure 18.

The comparison results are shown in Table 7.

Analyzing the difference between the actual and measured distance results, it was found that the difference lied mostly between 0.013 and 0.191. This phenomenon is caused by the slight change in the posture of the vehicle.

This study is based on the obstacle detection method of VIDAR and the use of a digital map for distance measurement experiments. The experimental results show that the error of this method is less than 2% at short distances (<20 m), and the distance measurement effect is better than that reported by Guo Lei's. Moreover, existing vision-based ranging requirements call for a measurement error of less than 5% [36]. Therefore, from the distance measurement results, the vision-based ranging algorithm proposed in this article meets the requirements in measurement accuracy and can achieve accurate distance measurement to obstacles.
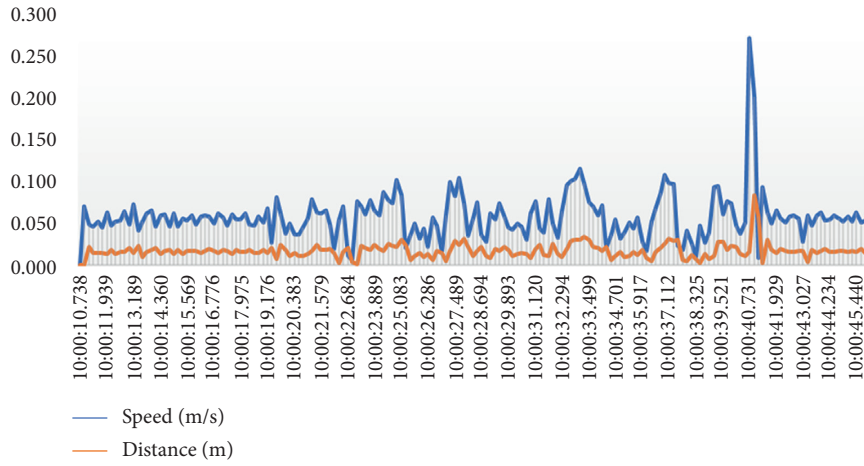
FIGURE 17: IMU data processing result (speed and distance of the test vehicle).
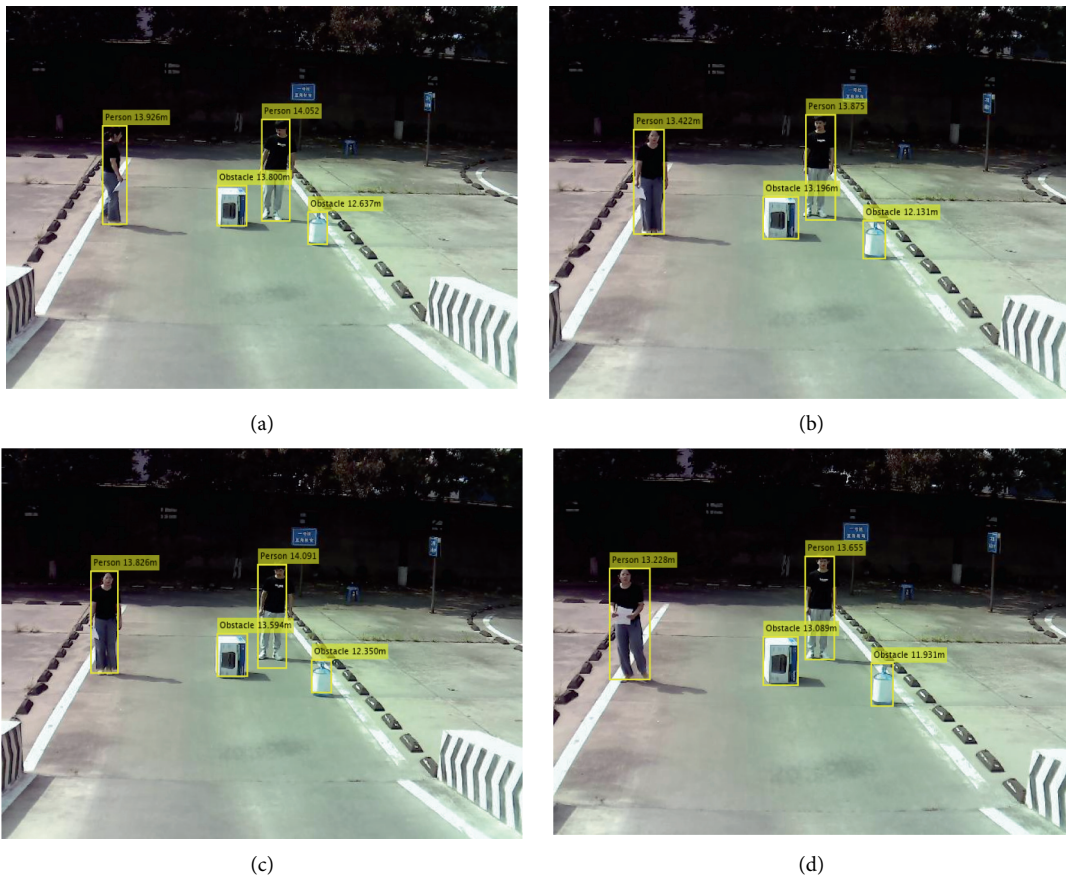

(a)

(b)

(c)

(d)

FIGURE 18: Obstacle distance detection. (a) 2021 8 16 10 0 32.825. (b) 2021 8 16 10 0 34.327. (c) 2021 8 16 10 0 35.766. (d) 2021 8 16 10 0 38.630.

TABLE 7: Distance measurement results based on monocular vision.

Time	Slope (°)	Obstacle	Actual distance (m)	Measuring distance (m)	Distance difference (m)	Error (%)
2021 8 16 10 0 32.825	7.35	1	13.793	13.926	0.133	0.964
		2	13.609	13.800	0.191	1.403
		3	12.651	12.637	0.014	0.111
		4	14.198	14.052	0.146	1.028
2021 8 16 10 0 34.327	7.21	1	13.719	13.826	0.107	0.780
		2	13.499	13.594	0.095	0.704
		3	12.265	12.350	0.085	0.693
		4	14.270	14.091	0.179	1.254
2021 8 16 10 0 35.766	7.10	1	13.532	13.422	0.110	0.813
		2	13.358	13.196	0.162	1.213
		3	12.309	12.131	0.178	1.446
		4	13.998	13.875	0.123	0.879
2021 8 16 10 0 38.630	6.88	1	13.321	13.228	0.093	0.698
		2	13.102	13.089	0.013	0.099
		3	12.037	11.931	0.106	0.881
		4	13.713	13.655	−0.058	0.423

6. Conclusion

In this study, an obstacle detection method based on VIDAR is applied to complex environments, avoiding the drawbacks of machine learning methods that can only detect known obstacles. Moreover, by integrating slope information into the VIDAR detection method, real obstacles can be detected on sloped roads, and distance and speed measurement of obstacles can be realized, which has important research value for autonomous vehicles and active safety systems. It can be seen from the results that the proposed method is effective in improving the accuracy and speed of obstacles detection and can meet the requirements of obstacle detection in complex environments. Obstacle detection in complex road environment is the basis for safe driving of vehicles. Therefore, obstacle avoidance path planning and speed control based on obstacle detection are our future research directions.

Data Availability

Data are available on request to the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] S. Grigorescu, B. Trasnea, T. Cocias, and G. Macesanu, "A survey of deep learning techniques for autonomous driving," *Journal of Field Robotics*, vol. 37, no. 3, pp. 362–386, 2020.

[2] H. Fujiyoshi, T. Hirakawa, and T. Yamashita, "Deep learning-based image recognition for autonomous driving," *IATSS Research*, vol. 43, no. 4, pp. 244–252, 2019.

[3] I.-S. Weon and S.-G. Lee, "Environment recognition based on multi-sensor fusion for autonomous driving vehicles," *Journal of Institute of Control, Robotics and Systems*, vol. 25, no. 2, pp. 125–131, 2019.

[4] Z. Sun, Q. Zhao, N. Zhang, B. Zhu, and Wang, "Intelligent vehicle multi-objective tracking based on particle swarm algorithm," *Forest Engineering*, vol. 36, no. 4, pp. 70–75, 2020.

[5] M. Bucolo, A. Buscarino, C. Famoso, L. Fortuna, and M. Frasca, "Control of imperfect dynamical systems," *Nonlinear Dynamics*, vol. 98, no. 4, pp. 2989–2999, 2019.

[6] L. Xiong, X. Xia, Y. Lu et al., "IMU-based automated vehicle slip angle and attitude estimation aided by vehicle dynamics," *Sensors*, vol. 19, no. 8, p. 1930, 2019.

[7] M. Shang, B. Rosenblad, and R. Stern, "A novel asymmetric car following model for driver-assist enabled vehicle dynamics," *IEEE Transactions on Intelligent Transportation Systems*, 2022.

[8] X. Zhao, P. Sun, Z. Xu, H. Min, and H. Yu, "Fusion of 3D LIDAR and camera data for object detection in autonomous vehicle applications," *IEEE Sensors Journal*, vol. 20, no. 9, pp. 4901–4913, 2020.

[9] W. Huang, Z. Zhang, W. Li, and J. Tian, "Moving object tracking based on millimeter-wave radar and vision sensor," *Journal of Applied Science and Engineering*, vol. 21, no. 4, pp. 609–614, 2018.

[10] C.-C. Lin, W.-L. Mao, T.-W. Chang, C.-Y. Chang, and S. S. S. Abdullah, "Fast obstacle detection using 3D-to-2D LiDAR point cloud segmentation for collision-free path planning," *Sensors and Materials*, vol. 32, no. 7, pp. 2365–2374, 2020.

[11] A. Dairi, F. Harrou, M. Senouci, and Y. Sun, "Unsupervised obstacle detection in driving environments using deep-learning-based stereovision," *Robotics and Autonomous Systems*, vol. 100, pp. 287–301, 2018.

[12] H. Zhu, K.-V. Yuen, L. Mihaylova, and H. Leung, "Overview of environment perception for intelligent vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 10, pp. 2584–2601, 2017.

[13] W. Song, Y. Yang, M. Fu, Y. Li, and M. Wang, "Lane detection and classification for forward collision warning system based on stereo vision," *IEEE Sensors Journal*, vol. 18, no. 12, pp. 5151–5163, 2018.

[14] L. Sun, K. Yang, X. Hu, W. Hu, and K. Wang, "Real-time fusion network for RGB-D semantic segmentation incorporating unexpected obstacle detection for road-driving images," *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 5558–5565, 2020.

[15] Y. Qian, W. Zhou, J. Yan, W. Li, and L. Han, "Comparing machine learning classifiers for object-based land cover classification using very high resolution imagery," *Remote Sensing*, vol. 7, no. 1, pp. 153–168, 2015.

[16] Y. Yu, L. Kurnianggoro, and K.-H. Jo, "Comaring machine learning classifiers for object-based land cover classification using very high resolution imagery," *International Journal of Control, Automation and Systems*, vol. 17, no. 7, pp. 1866–1874, 2019.

[17] K. Dokka, P. R. MacNeilage, G. C. DeAngelis, and D. E. Angelaki, "Multisensory self-motion compensation during object trajectory judgments," *Cerebral Cortex*, vol. 25, no. 3, pp. 619–630, 2015.

[18] Z. Li, S. E. Dosso, and D. Sun, "Motion-compensated acoustic localization for underwater vehicles," *IEEE Journal of Oceanic Engineering*, vol. 41, no. 4, pp. 840–851, 2016.

[19] P. Agrawal, R. Kaur, V. Madaan, and M. Babu, "Moving object detection and recognition using optical flow and eigen face using low resolution video," *Recent Patents on Computer Science*, vol. 11, no. 2, pp. 1–10, 2018.

[20] J. Cho, Y. Jung, D.-S. Kim, S. Lee, and Y. Jung, "Moving object detection based on optical flow estimation and a Gaussian mixture model for advanced driver assistance systems," *Sensors*, vol. 19, no. 14, p. 3217, 2019.

[21] X. Zhao, F. Pu, Z. Wang, H. Chen, and Z. Xu, "Detection, tracking, and geolocation of moving vehicle from uav using monocular camera," *IEEE Access*, vol. 7, pp. 101160–101170, 2019.

[22] L. Chengmei, B. A. I. Hongyang, G. Hongwei, and L. Huaju, "Moving object detection and tracking based on improved optical flow method," *Chinese Journal of Scientific Instrumen*, vol. 39, pp. 249–256, 2018.

[23] I. Yang and W. H. Jeon, "Development of lane-level location data exchange framework based on high-precision digital map," *Journal of Digital Contents Society*, vol. 19, no. 8, pp. 1617–1623, 2018.

[24] M. ElMikaty and T. Stathaki, "Detection of cars in high-resolution aerial images of complex urban environments," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 10, pp. 5913–5924, 2017.

[25] J. Li, H. Cheng, H. Guo, and S. Qiu, "Survey on artificial intelligence for vehicles," *Automotive Innovation*, vol. 1, no. 1, pp. 2–14, 2018.

[26] K. Choi, J. K. Suhr, and H. G. Jung, "Map-matching-based cascade landmark detection and vehicle localization," *IEEE Access*, vol. 7, pp. 127874–127894, 2019.

[27] S. H. Wang and X. X. Li, "A real-time monocular vision-based obstacle detection," in *Proceedings of the Sept. 2020 6th International Conference on Control, Automation and Robotics (ICCAR)*, pp. 695–699, Singapore, April 2020.

[28] H. Nguyen, "Improving YOLO v5 framework for fast vehicle detection," *Mathematical Problems in Engineering*, vol. 2019, Article ID 3808064, 11 pages, 2019.

[29] Z. Yi, S. Yongliang, and Z. Jun, "An improved tiny-YOLO v3 pedestrian detection algorithm," *Optik*, vol. 183, pp. 17–23, 2019.

[30] K. Wang, F. Yan, B. Zou, L. Tang, Q. Yuan, and C. Lv, "Occlusion-free road segmentation leveraging semantics for autonomous vehicles," *Sensors*, vol. 19, no. 21, p. 4711, 2019.

[31] M. Kristan, V. Sulic Kenk, S. Kovacic, and J. Pers, "Fast image-based obstacle detection from unmanned surface vehicles," *IEEE Transactions on Cybernetics*, vol. 46, no. 3, pp. 641–654, 2016.

[32] M. Tkocz and K. Janschek, "Vehicle speed measurement based on binocular stereovision system," *Journal of Intelligent & Robotic Systems*, vol. 80, no. 3, pp. 475–489, 2015.

[33] C. Meng, H. Bao, Y. Ma, X. Xu, and Y. Li, "Visual Meterstick: preceding vehicle ranging using monocular vision based on the fitting method," *Symmetry*, vol. 11, no. 9, p. 1081, 2019.

[34] T. Zhe, L. Huang, Q. Wu, J. Zhang, C. Pei, and L. Li, "Inter-Vehicle distance estimation method based on monocular vision using 3D detection," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 5, pp. 4907–4919, 2020.

[35] L. A. Rosero and F. S. Osório, "Calibration and multi-sensor fusion for on-road obstacle detection," in *Proceedings of the 2017 Latin American Robotics Symposium (LARS) and 2017 Brazilian Symposium on Robotics (SBR)*, pp. 1–6, Curitiba, Brazil, November 2017.

[36] N. Garnett, S. Silberstein, and S. Oron, "Real-time category-based and general obstacle detection for autonomous driving," in *Proceedings of the IEEE International Conference on Computer Vision Workshops*, pp. 198–205, Venice, Italy, October 2017.

[37] L. Caltagirone, M. Bellone, L. Svensson, and M. Wahde, "LIDAR–camera fusion for road detection using fully convolutional neural networks," *Robotics and Autonomous Systems*, vol. 111, pp. 125–131, 2019.