

Article

An Obstacle Detection Method Based on Longitudinal Active Vision

Shuyue Shi ¹, Juan Ni ¹, Xiangcun Kong ¹, Huajian Zhu ¹, Jiaze Zhan ¹, Qintao Sun ¹ and Yi Xu ^{1,2,*}

¹ School of Transportation and Vehicle Engineering, Shandong University of Technology, Zibo 255000, China; 22502060007@stumail.sdut.edu.cn (S.S.); 22502060003@stumail.sdut.edu.cn (J.N.); 22502060006@stumail.sdut.edu.cn (X.K.); 23402030186@stumail.sdut.edu.cn (H.Z.); 23502060271@stumail.sdut.edu.cn (J.Z.); 23402030185@stumail.sdut.edu.cn (Q.S.)

² Qingte Group Co., Ltd., Qingdao 266106, China

* Correspondence: xuyi@sdut.edu.cn

Abstract: The types of obstacles encountered in the road environment are complex and diverse, and accurate and reliable detection of obstacles is the key to improving traffic safety. Traditional obstacle detection methods are limited by the type of samples and therefore cannot detect others comprehensively. Therefore, this paper proposes an obstacle detection method based on longitudinal active vision. The obstacles are recognized according to the height difference characteristics between the obstacle imaging points and the ground points in the image, and the obstacle detection in the target area is realized without accurately distinguishing the obstacle categories, which reduces the spatial and temporal complexity of the road environment perception. The method of this paper is compared and analyzed with the obstacle detection methods based on VIDAR (vision-IMU based detection and range method), VIDAR + MSER, and YOLOv8s. The experimental results show that the method in this paper has high detection accuracy and verifies the feasibility of obstacle detection in road environments where unknown obstacles exist.

Keywords: longitudinal active vision; image processing; distance estimation; camera rotation strategy



Citation: Shi, S.; Ni, J.; Kong, X.; Zhu, H.; Zhan, J.; Sun, Q.; Xu, Y. An Obstacle Detection Method Based on Longitudinal Active Vision. *Sensors* **2024**, *24*, 4407. <https://doi.org/10.3390/s24134407>

Academic Editor: Sergio A. Velastin

Received: 19 April 2024

Revised: 24 June 2024

Accepted: 3 July 2024

Published: 7 July 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With the improvement in people's living standards and the progress of science and technology, automobiles have become the most common means of transportation in modern travel. With the increase in the number of automobiles in countries all over the world, the numbers of related automobile safety problems have also risen. According to the World Health Organization, the number of road traffic deaths has decreased by 5% per year since 2010, down to 1.19 million per year [1]. However, road traffic crashes remain a persistent global health crisis, with pedestrians, cyclists and other vulnerable road users facing a serious and rising risk of death. In order to improve road traffic safety, research in the field of automotive safety has evolved towards smarter Advanced Driving Assistance Systems (ADASs) [2]. An ADAS is the use of environmental sensing technology to collect dynamic data from the vehicle, driver and surrounding environment and analyze and process them, in order to achieve safe and comfortable driving and minimize traffic accidents by alerting the driver or actuators to intervene in the vehicle's maneuvering. In addition, it also provides reliable safety support for FSD (full self-driving). Therefore, obstacle detection is an important element in the structure of intelligent-vehicle environment-sensing technology, and it is also a prerequisite for realizing advanced assisted-driving functions.

The complex road environment brings great difficulties for the intelligent vehicle's environment perception; the reliable and accurate detection of obstacles is one of the core problems that need to be solved to realize automobile assisted driving [3]. In addition to the innovation of algorithms, developments in technology have made great progress in obstacle detection, and sensor types, target objects, and detection methods have become the main

research content. The sensor aspect includes cameras [4], radar [5,6], etc. Camera detection technology is widely used in obstacle detection due to its relative maturity and low cost. In this paper, a novel obstacle detection method is proposed, which uses a monocular camera as a visual sensor and servo motor to form a longitudinal active camera detection system to detect various types of obstacles in front of the vehicle in the road environment, including vehicles, pedestrians, and conical obstacles, spherical obstacles, rod obstacles, cartons and so on. The principle of the detection method in this paper is to use the two frames of images acquired before and after the camera rotates along the longitudinal plane as the detection object, use the height difference characteristics of the obstacle point and the ground point to detect the obstacle, obtain the distance information between the obstacle and the vehicle, and realize the detection of the obstacle in front of the vehicle.

The rest of the paper is organized as follows: Section 2 describes the existing research methods. Section 3 presents the principles of an obstacle detection method based on longitudinal active vision, including the detection process, obstacle extraction algorithms based on fast image region matching in stable extreme regions (MSER), monocular ranging models, and static and dynamic obstacle detection models. Section 4 designs the obstacle detection experiments for the theoretical method proposed in this paper, compares the obstacle detection methods based on YOLOv8s and VIDAR, and verifies the superiority of the method in this paper. Section 5 summarizes the content of this paper.

2. Related Work

Obstacle detection is a core component of automatic driving technology, which is crucial for promoting the development of automatic driving technology and has become a key area of research in recent years. Monocular vision sensors collect rich information, with strong applicability and a low price; many scholars have conducted relevant research on this, mainly divided into two kinds of obstacle detection algorithms, either based on image features or based on deep learning. With the rapid development of computer technology, deep learning is increasingly used in target detection. Target detection algorithms based on deep learning use convolutional neural networks (CNN) to learn features; this feature learning method is mainly through input training set to automatically learn the feature information of the target, and the neural network to generate more advanced, more abstract and more reliable feature information. Han et al. [7] proposed CNN-based real-time semantic segmentation of class-aware edge information based on channel-aware attention and constructed a new EdgeNet network with significant improvement in MIOU compared to other networks. Nguyen [8] proposed an improved framework based on a fast, reactive neural network using MobileNet architecture to build the base convolutional layer of fast R-CNN, which utilizes the deeply divisible convolutional structure in the mobile network architecture to build the classifier and improve the accuracy of vehicle detection. Zaghari et al. [9] used the YOLO non-maximum suppression (NMS) algorithm for obstacle detection, which has higher detection accuracy and faster detection speed compared to other algorithms. Yasmin et al. [10] proposed an approach based on transfer learning to detect small obstacles under strict lighting and illumination conditions using UNet++, PSPNet, PANNet, LinkNet and DeepLabV3+ semantic segmentation models. He et al. [11] proposed a flexible and efficient multi-scale unipolar target detector FE-YOLO for image obstacle detection. Sun et al. [12] proposed a real-time fused semantic segmentation network RFNet, which can effectively utilize depth complementary features, multi-dataset training and depth streaming in the architecture makes the network very effective in detecting unexpected small targets. However, the RFNet network performs poorly in detecting features such as sidewalks, walls, and motorcycles.

Detecting obstacles using deep learning methods requires prior knowledge about the obstacle; therefore, obstacle detection methods based on deep learning have some limitations. In obstacle detection based on conventional image features, GÜNGÖR et al. [13] proposed a new non-artificial intelligence method for finding ground lines and detecting obstacles on the road using v-parallax data. Yuan et al. [14] used U-V parallax images to

detect the drivable area of an unmanned vehicle using U–V parallax images as well as the geometric relationship between the size of an obstacle and its parallax to detect obstacles in the drivable area. Xu et al. [15] proposed a generalized obstacle detection method based on VIDAR combined with a fast image region-matching method based on MSER with high detection accuracy. Kumar et al. [16] accomplished the identification of defective regions on the surface of an obstacle by means of a multilevel color threshold segmentation method. Wang et al. [17] used a single wide-angle camera for real-time obstacle detection, based on the background difference method for detecting static and moving objects. Xue et al. [18] proposed a method for detecting obstacles on a reflective ground, using a pre-calibrated ground-based detection scheme that characterizes the difference between obstacles and non-obstacles, and locates obstacles by means of appearance–geometry fusion models. Optical flow methods can detect moving obstacle targets based on changes in the image optical flow field caused by object motion, with good robustness and without the need to know a priori information about the obstacle. Therefore, optical flow methods have been widely studied and applied [19–21].

However, today’s research techniques still contain deficiencies. The obstacle detection method based on VIDAR, although highly accurate, is limited in usage scenarios, and their applicability decreases when vehicles are stationary or at low speeds. Deep learning is limited by the sample type and cannot detect obstacles of unknown shape. Therefore, this paper proposes an obstacle detection method based on longitudinal active vision, which does not require a priori knowledge of the scene to achieve more accurate obstacle recognition and distance detection.

3. Methods

In this paper, we combine a fast image region-matching method based on MSER with an obstacle detection method based on longitudinal active vision to simplify the matching process, improve the matching speed, and use fewer feature points for obstacle detection. The process of the road obstacle detection method based on longitudinal active vision studied in this paper is as follows and is shown in Figure 1.

(1) Calibration of initial camera parameters

Calibrate the initial internal and external parameters of the camera mounted on the vehicle, obtain the camera’s focal length f by camera calibration, the pixel size p of the photosensitive chip, obtain the camera mounting height h by measurement calculation, and the camera rotation radius K .

(2) Image region matching and obstacle detection

- ① Acquire the initial frame image. At the moment $t = 0$, the camera obtains the first frame image I_{ft} . The feature points in I_{ft} are extracted, and the lowest point of the extracted feature points is considered as the intersection point P_i of the obstacle and the road plane. The distance d_i from point P_i to the camera is then calculated, resulting in the required rotation angle θ_i for the camera to obtain the next frame image.
- ② Acquire the second frame image. The second frame image I_{bt} is acquired after the camera is rotated by an angle θ_i . The fast image region-matching method based on MSER is performed on I_{ft} and I_{bt} to find the center of mass of the matched region as feature points.
- ③ Calculate the horizontal distance from the camera to the center of rotation. Calculate the horizontal distance $K \cos \theta$ from the camera to the center of rotation based on the camera rotation angle θ_i and camera rotation radius K obtained in ①.
- ④ Horizontal distance calculation. Assuming that the feature point is on the horizontal plane, the horizontal distance d_1, d_2 from the feature point to the camera is calculated based on the monocular ranging model at the before and after moments, respectively.

- ⑤ Obstacle judgment. According to the internal and external parameters of the camera as well as the acquired valid information, compare Δl ($\Delta l = d_2 + K \cos \theta - d_1 - K$) and k (k is a set threshold, $k > 0$). The feature point is not on the horizontal plane and the region is an obstacle if $\Delta l > k$. The feature point is on the horizontal plane and the region is not an obstacle if $\Delta l \leq k$.

(3) Camera reset

After completing the two-frame image detection, the camera resets and repeats step (2) to proceed to the next stage of active obstacle detection.

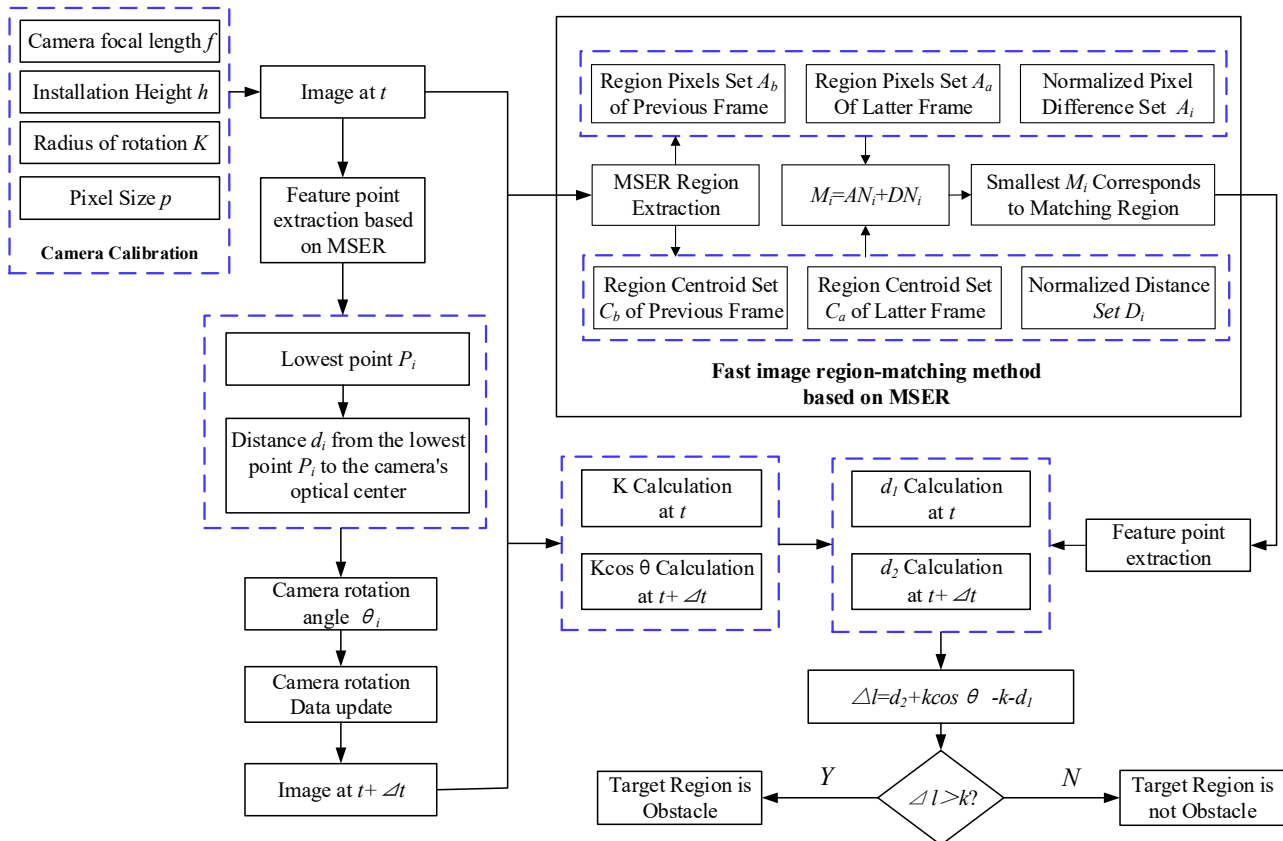


Figure 1. An obstacle detection based on longitudinal active vision.

3.1. Fast Image Region-Matching Method Based on MSER

The MSER algorithm is a region-based feature-extraction algorithm that is widely used in image matching and target tracking [22–25]. Since the detection object of this method is the region features of the image rather than the local features, the detected feature points are more stable, and the obstacles in the image can be detected quickly and the non-obstacles can be eliminated. The fast image region-matching method based on MSER ignores the position and shape differences of the MSER algorithm between two images, and the specific matching method process is as follows:

- (1) Extract the region of maximum stable extremes using the MSER algorithm.
- (2) Perform region range difference AN_i calculation. For the two frames captured in the experiment, it is assumed that the MSER region sets of the two frames before and after are $Af = \{Af_1, Af_2, \dots, Af_n\}$ and $As = \{As_1, As_2, \dots, As_n\}$, respectively. A_i is the set of the difference between the f th MSER region range in the previous frame and

the unmatched region in the next frame. The set A_i is normalized and the effect of normalization is denoted by AN_i , where

$$AN_i = \frac{A_i - \min(A_i)}{\max(A_i) - \min(A_i)} \quad (1)$$

- (3) Perform region set spacing DN_j calculation. It is assumed that the MSER region center-of-mass sets in the front and back images are $Df = \{Df_1, Df_2, \dots, Df_m\}$ and $Ds = \{Ds_1, Ds_2, \dots, Ds_m\}$, respectively. D_j is the set of distances between the f th MSER region range in the previous image and the unmatched region in the latter image. The set D_j is normalized and the result is denoted by DN_j , where

$$DN_j = \frac{D_j - \min(D_j)}{\max(D_j) - \min(D_j)} \quad (2)$$

- (4) Extract the matching region M_l . Let M_l be the set of matching values of the s th MSER, and extract the MSER corresponding to the smallest M_l as a matching region, denoted as $M_l = AN_i + DN_j$.

Using the stabilized feature points extracted using the algorithm based on MSER for obstacle detection can shorten the post-processing image time and improve the speed of obstacle detection.

3.2. Static Obstacle Detection Model

Measurement methods based on monocular vision need to obtain position information in three-dimensional space from two-dimensional image information, and the obstacle ranging principle can be described by the pinhole model principle, which can accurately calculate the distance between the vehicle and the obstacle. As shown in Figure 2, a camera is loaded on the auto-vehicle, the object in front of the auto-vehicle is regarded as an obstacle, and the lowest point of the obstacle is regarded as the intersection of the obstacle and the road surface.

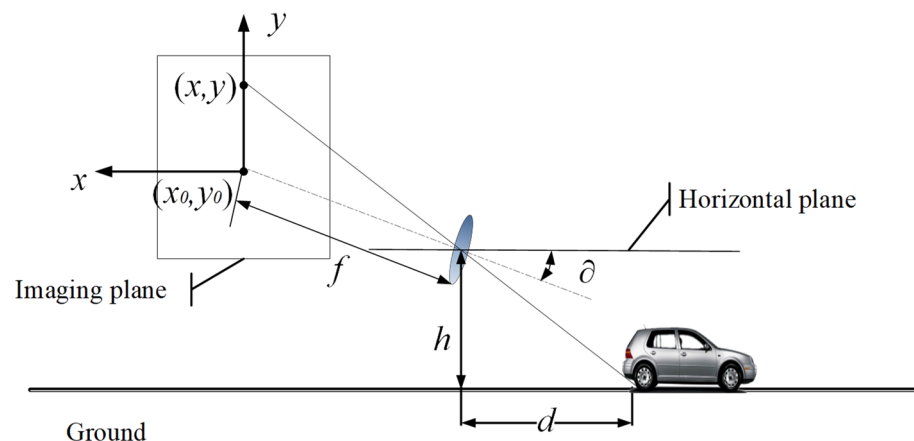


Figure 2. Obstacle ranging model.

Here, the focal length of the camera is f and the pixel is μ . The pitch angle of the camera installation is ϑ and the height of the installation is h . Let (x_0, y_0) be the coordinate origin in the image plane and (x, y) be the coordinates of the intersection of the obstacle and the road plane in the imaging plane. Then the horizontal distance between the camera and the obstacle can be found by Equation (3).

$$d = \frac{h}{\tan(\vartheta + \arctan[(y_0 - y)\mu / f])} \quad (3)$$

As shown in Figure 3, the first imaging point of the vertex of the obstacle at the initial moment is A . At the next moment, the camera rotates and translates from the image plane from y_1 to y_2 due to the camera making a longitudinal rotation in the plane with the point M as the center of rotation, and the second imaging point of the vertex of the obstacle is B . The horizontal distance from the camera optical center to the point M is K in the previous frame, and the horizontal distance from the optical center to the point M is $K \cos \theta$ after the camera is rotated by an angle θ . A' and B' are points on the road plane projected from the vertex of the obstacle, and Δl is the horizontal distance between the two projected points. The horizontal distance from the camera optical center to A' is d_1 , and the horizontal distance to B' is d_2 . d_1 and d_2 can be calculated by Equation (3). d_1 and d_2 are related by the equation $d_1 + K + \Delta l = d_2 + K \cos \theta$. If the feature point is a point on the ground, the equation for d_1 and d_2 should be $d_1 + K = d_2 + K \cos \theta$. If $d_1 + K \neq d_2 + K \cos \theta$, it means that the target point is a point with height and therefore the obstacle can be recognized by Δl . Figure 3 shows the theoretical model for detecting static obstacles when the vehicle is stationary, and the displacement in the horizontal direction due to camera rotation is $K - K \cos \theta$. And when the vehicle is in motion, the actual displacement of the camera in the horizontal direction is Δd . Δd can be calculated by Equation (5). Δl can be calculated by Equation (6).

$$t_s = \frac{\theta}{v_s} \quad (4)$$

$$\Delta d = t_s \times v_c - K + K \cos \theta \quad (5)$$

$$\begin{aligned} \Delta l &= d_2 + \Delta d - d_1 \\ &= d_2 + \frac{\theta \times v_c}{v_s} - K + K \cos \theta - d_1 \end{aligned} \quad (6)$$

where t_s is the time to acquire two frames of image camera rotation, v_s is the speed of the servo motor, and v_c is the vehicle speed.

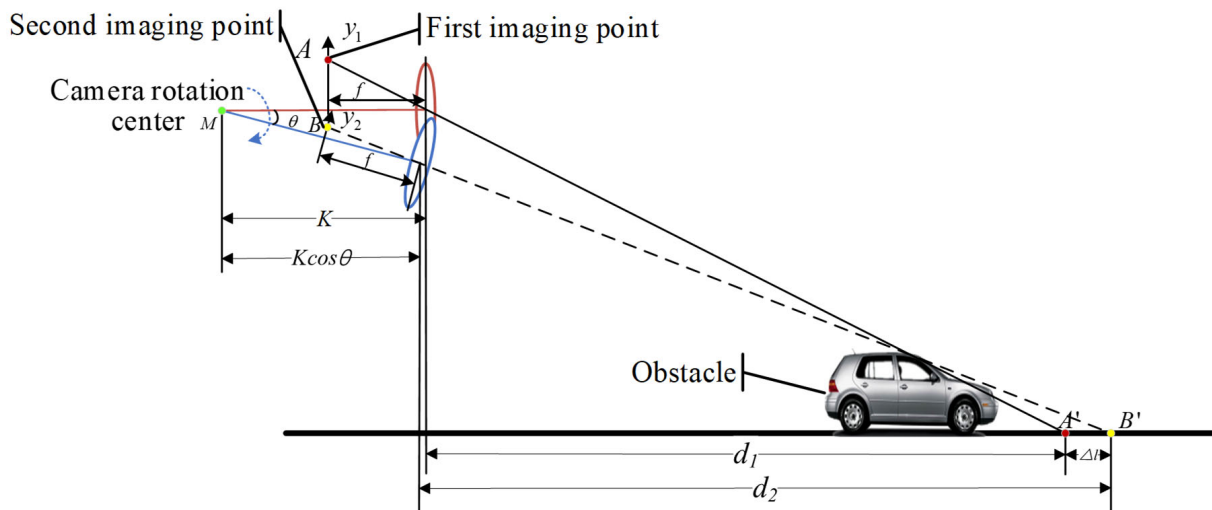


Figure 3. Schematic diagram of the static obstacle imaging.

3.3. Dynamic Obstacle Detection Model

When the obstacle in front of the vehicle moves in the horizontal direction, the imaging of the moving obstacle is shown in Figure 4. The camera pitch angle at the initial moment is 0. Therefore, the distance from the camera rotation center point M to the ground is equal to the distance from the camera optical center to the ground, which is h_1 . The distance between the optical center and the ground after the camera is rotated by an angle θ is h_2 . The horizontal distance between the optical center of the camera and the vertex of the obstacle at the initial moment is s_1 . The horizontal distance between the optical center and the vertex of the obstacle after the camera is rotated by an angle θ is s_2 . The distance

moved by the obstacle is s . When the target obstacle is not moving straight, s is the distance that the actual moving distance of the obstacle projects onto the direction of the vehicle's movement. Then the relationship between d_1, d_2, s_1, s_2 and s is:

$$\begin{cases} K \cos \theta + d_2 = K + d_1 + \Delta l \\ K - K \cos \theta + s_1 + s = s_2 \end{cases} \quad (7)$$

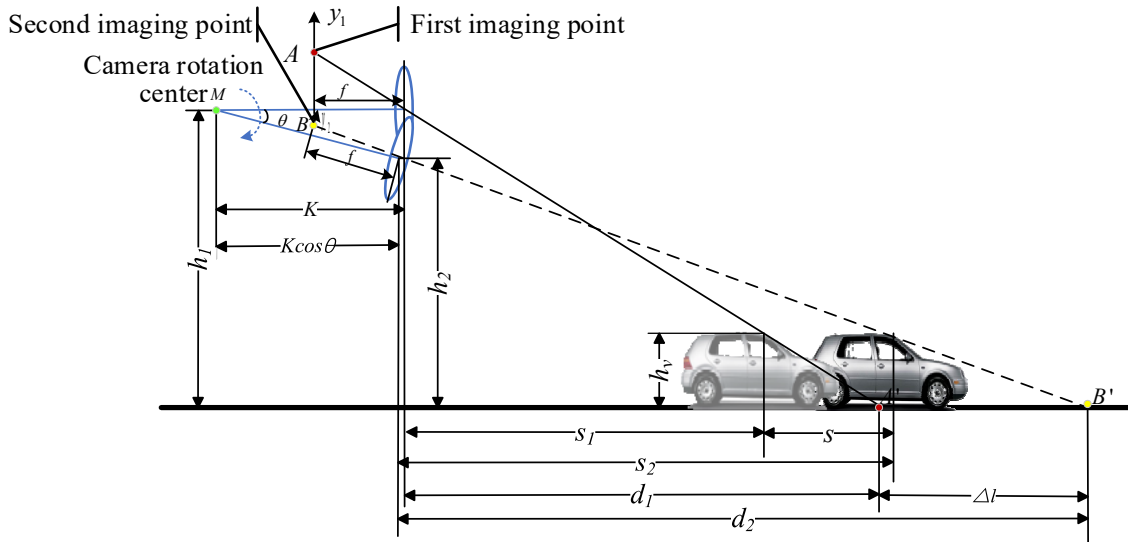


Figure 4. Schematic diagram of dynamic obstacle imaging.

The height of the obstacle is h_v . According to the principles of a right triangle, the relationship between $h_v, h_1, h_2, d_1, d_2, s_1$ and s_2 can be expressed as:

$$\begin{cases} \frac{h_v}{h_1} = \frac{d_1 - s_1}{d_1} \\ \frac{h_v}{h_2} = \frac{d_2 - s_2}{d_2} \end{cases} \quad (8)$$

From Equations (4) and (5):

$$\Delta l = \frac{(h_1 - h_v)(sh_2 - sh_v + s_2h_v) - s_1h_v(h_2 - h_v)}{(h_1 - h_v)(h_2 - h_v)} \quad (9)$$

Dynamic obstacles can be recognized when $(h_1 - h_v)(sh_2 - sh_v + s_2h_v) \neq s_1h_v(h_2 - h_v)$. Thus obstacles on the road can be recognized using camera rotation. The judgment process only requires tracking of the feature points, calculating their positions, and obtaining the height of the optical center above the ground after camera rotation, thus reducing the consumption of time and space for obstacle detection. From Equation (9), it can be seen that the dynamic obstacle detection model is only related to the parameters such as the height above the ground before and after the movement of the camera, the moving distance of the target obstacle, and has nothing to do with the state of the self-vehicle, the camera rotating speed and the camera rotating angle, so that the vehicle in both stationary and motion state can realize the accurate detection of obstacles.

3.4. Camera Rotation Strategy

The theoretical models of static and dynamic obstacle detection presented in Sections 3.2 and 3.3 are based on two frames of images captured before and after camera rotation, and the target always remains within the field of view before and after camera rotation. The rotation angle of the camera is calculated by pointing the optical axis to the lowest point of the obstacle based on the positional information of the obstacle. As shown in Figure 5, the obstacle is located in front of the vehicle, when the camera is located in the

initial position to acquire the first frame image. Based on the MSER algorithm to obtain a stable extreme value region, the lowest point of the extreme value region as a feature point, the lowest point of the extracted feature point is regarded as the intersection point P_i (i is the number of P -points) of the obstacle and the road plane, the optical axis points to the intersection point a at the next moment. The coordinates of point P_i in the imaging plane are (x_i, y_i) , the distance d_i from the camera to P_i can be calculated from Equation (3), and the camera's optical axis rotation angle θ_i can be found based on the trigonometric relationship:

$$\theta_i = \arctan \frac{h}{d_i + K} = \arctan \frac{h \tan(\arctan[(y_0 - y_i)\mu/f])}{h + K \tan(\arctan[(y_0 - y_i)\mu/f])} \quad (10)$$

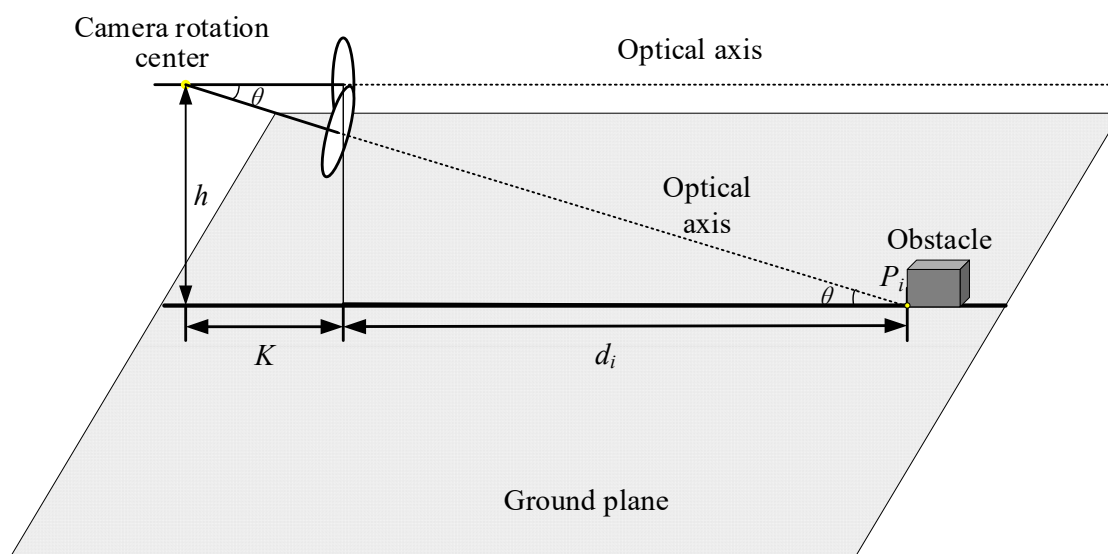


Figure 5. Schematic diagram of the camera rotation.

4. Experiments and Results

In order to verify the feasibility of the theoretical method in this paper, simulation experiments and real-vehicle experiments under controlled scenarios are designed. The experimental results show that the obstacle detection method based on longitudinal active vision proposed in this paper can effectively detect obstacles in the road environment. By analyzing and comparing with the traditional method, it proves that the method has high detection accuracy and detection speed.

4.1. Experimental Equipment

In order to realize the active rotation of the camera as well as the accurate control of the camera angle and direction of rotation, this paper selects the MS9015 V1 model three-phase permanent magnet synchronous motor (Shanghai Lingkong Technology Co., Ltd., Shanghai, China), U-shaped bracket, and magnetic suction cup for the design of the active camera gimbal. The camera model selected in this paper is the RER-USB8MP02G (RERVISION Technology Co., Ltd., Shenzhen, China), which has high resolution and dynamic range, is not easily affected by temperature changes, and meets the requirements of environment sensing for the autonomous vehicle. The STM32F103VET6 microcontroller is utilized as the controller to achieve the control of the motor through voltage space-vector pulse width modulation (SVPWM). CAN communication is used between the motor and the STM32F103VET6 microcontroller (Wildfire Technology Co., Ltd., Dongguan, China) to control the camera rotation. This paper designs the longitudinal active camera obstacle detection system, as shown in Figure 6.

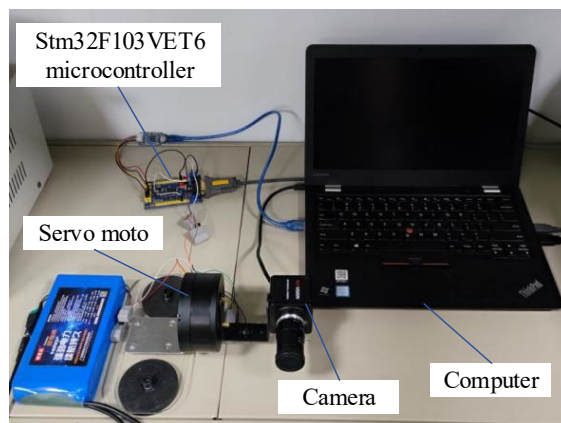


Figure 6. Architecture of the longitudinal active camera obstacle detection system.

4.2. Obstacle Detection Simulation Experiment

The longitudinal active camera is mounted on the experimental platform, a vehicle scale model is used to simulate obstacles, and common non-obstacles such as traffic signs and road patches are simulated with pieces of paper attached to a flat surface, and two other unknown types of obstacles are set up. One set of obstacle simulation experiments is processed as follows.

In the simulation experiment of obstacle detection, the camera pitch angle $\theta = 0^\circ$ at the initial moment, the effective focal length of the camera $f = 6.779$ mm, the installation height $h = 18$ cm. The camera rotation radius and steering angle affect the measurement accuracy of the target distance, so the corresponding camera steering angle is calculated by setting different K values, and the measured target distance is compared and analyzed with the real distance. As shown in Figures 7 and 8, the steering angle decreases with the increase in the rotation radius, and the calculated camera steering angle $\theta = 10.1^\circ$ at $K = 9$ cm, when the corresponding distance error is minimized. Therefore, the rotation radius of the camera was set to 9 cm. The camera acquires the first frame of the image at the initial moment, and detects the extreme value region in the image based on the MSER algorithm through the filtering and thresholding operation of the first frame of the image at multiple scales. According to the stability and distribution density of the region, the lowest point of the stable extreme value region is selected as the feature point, and the point with the largest value of the vertical coordinates among the various feature points under the imaging coordinate system is regarded as the intersection point of the obstacle and the road plane (Figure 9c), and the angle of rotation of the camera at the next moment $\theta = 9^\circ$ is calculated by Equation (7).

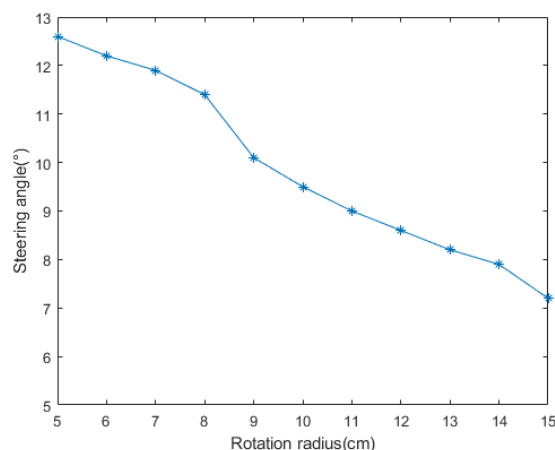


Figure 7. Steering angles corresponding to different radii of rotation.

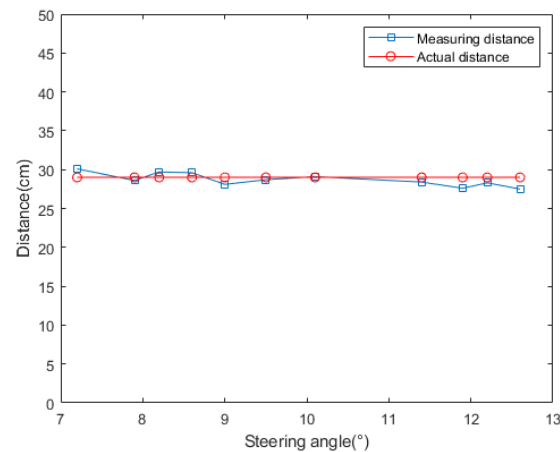


Figure 8. Distance measurements corresponding to the different steering angles.

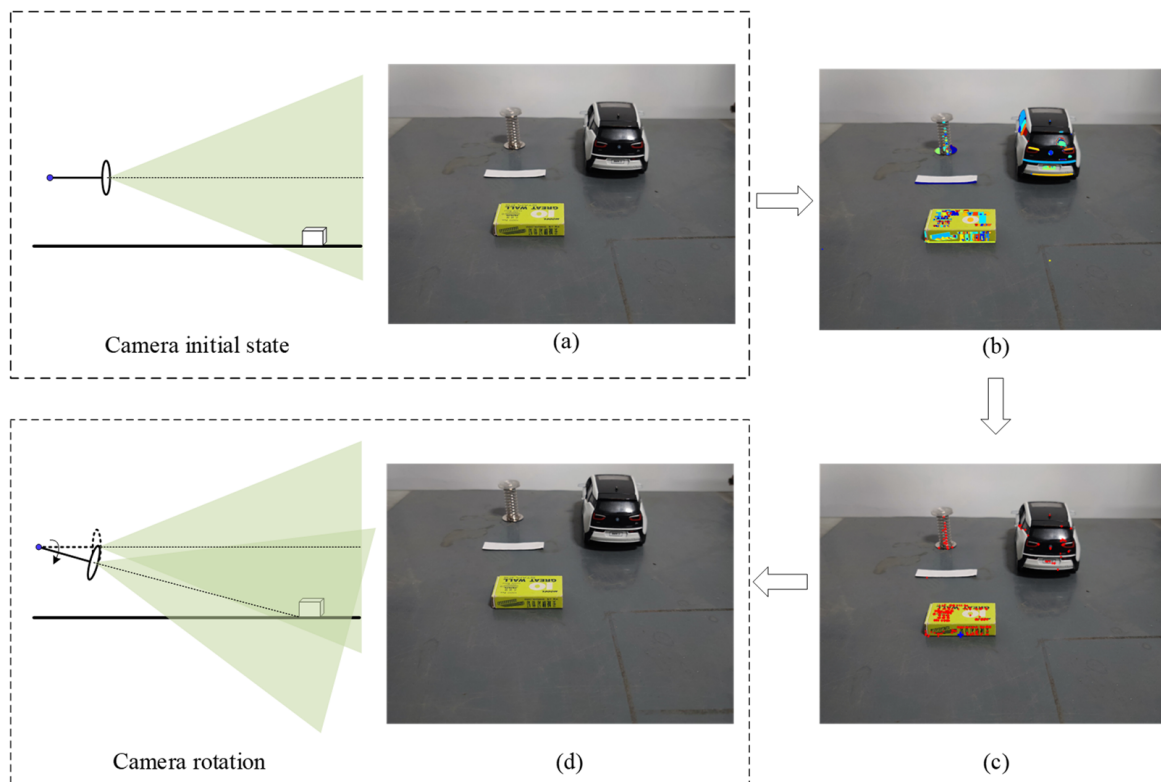


Figure 9. Two-frame image acquisition before and after camera rotation; (a) is the obstacle image at the initial moment, (b) is the feature region extraction based on MSER, and (c) is the feature point extraction, where red * is the lowest point of each extreme region and blue + is the intersection point of the obstacle and the road plane. (d) The second frame image acquired after camera rotation.

Based on the rotation angle calculated from the above experiments, the camera rotates to acquire the second frame image. The horizontal distances from the optical center to the center of rotation at the before and after moments were calculated by the monocular ranging model as 10 cm and 9.74 cm, respectively, and the experiments used the fast image region-matching method based on MSER to process the two frames acquired at the before and after moments. The centers of mass of the 14 extracted matching regions are used as feature points, as shown in Figure 10.

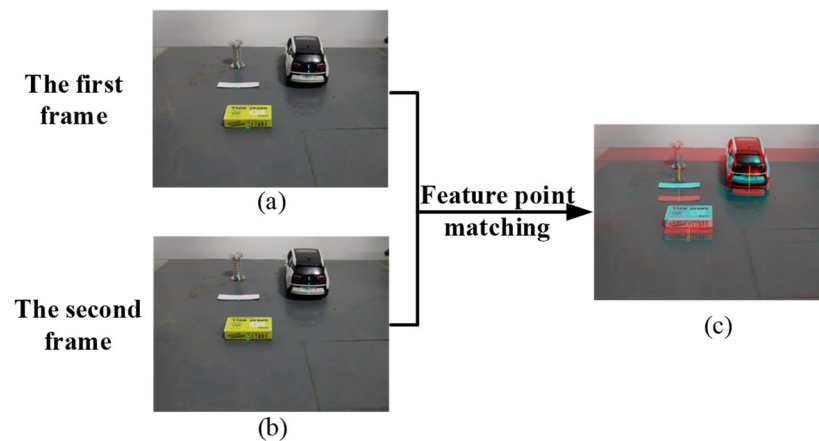


Figure 10. MSERs feature region extraction; (a) is the obstacle image in the initial moment, (b) the obstacle image in the next moment, and (c) is the region matching image in the two moments before and after, where the red region and + are the center of mass of MSERs and MSERs in the initial moment in the image, and the cyan region and o are the center of mass of MSERs and MSERs in the next moment in the image.

Feature point extraction is performed based on the front and rear frames acquired by the active camera (Figure 11), assuming that the feature points are on the horizontal plane, and detecting the horizontal distance d_1 and d_2 of each feature point from the camera in the front and rear frames through the monocular ranging model, and comparing Δl and k ($k = 2$ cm) to confirm the obstacles. d_1 , d_2 and Δl are calculated as shown in Table 1.

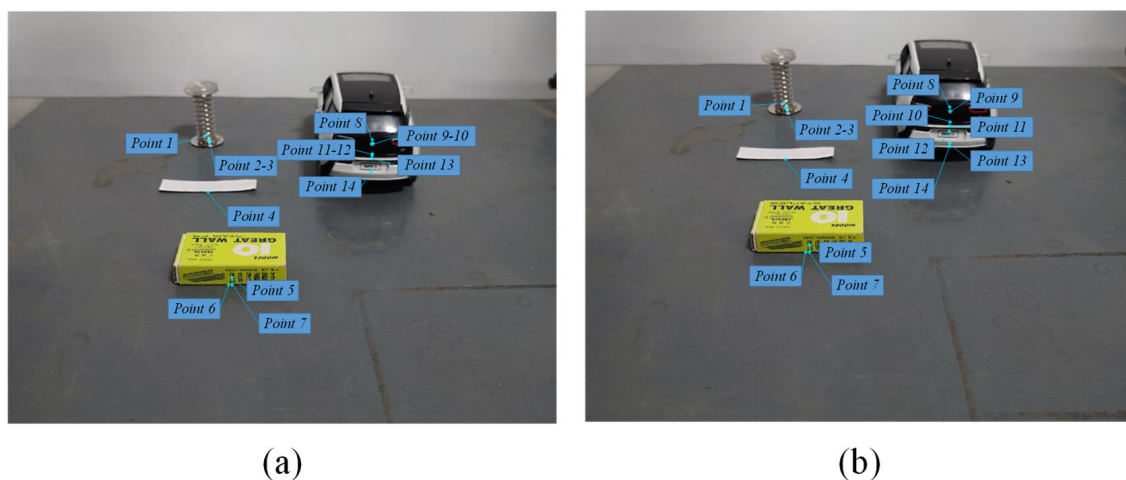


Figure 11. Feature point location; (a) shows the location of the feature point located in the image at the initial moment and (b) shows the location of the feature point located in the image at the next moment.

The calculation results show that the value of feature point four is 0.83 cm, which is smaller than the set threshold and is a feature point on the horizontal plane. The rest of the feature points except feature point four are not on the horizontal plane and are feature points on obstacles with height targets, and the corresponding MSERs all belong to the obstacle region. Non-obstacle points are excluded based on a set threshold, and the intersection of the bottom boundary of the obstacle region and the road plane is taken as the ranging point of the obstacle, and the distance from the obstacle to the camera is calculated using the monocular ranging model. Figure 12 shows the divided obstacle region, and the distances from the camera to the target obstacle are 42.09 cm, 23.78 cm, and 32.55 cm, respectively.

Table 1. Calculation results of d_1 , d_2 and Δl .

Feature Point	d_1/cm	d_2/cm	$\Delta l/cm$
1	39.88	42.27	2.21
2	39.74	42.09	2.17
3	39.78	42.12	2.16
4	29.12	30.14	0.83
5	21.04	23.54	2.32
6	21.22	23.78	2.37
7	21.19	23.75	2.37
8	29.77	32.97	3.22
9	29.43	32.76	3.15
10	29.56	32.86	3.12
11	29.40	32.75	3.17
12	29.37	32.75	3.20
13	29.17	32.55	3.20
14	29.25	32.58	3.15

**Figure 12.** Obstacle area division (where the yellow box is the detected obstacle area and the upper number is the distance from the obstacle to the camera).

4.3. Obstacle Detection Real Vehicle Experiment

In order to verify the reliability of the method in this paper, a real vehicle experiment was set up. A Honda Fit (FIT) vehicle was used as the experiment vehicle, and a longitudinal active camera was mounted at the front position of the roof to collect environmental information, and the camera was mounted at a height of 1.60 m. The equipment for the real vehicle experiment is shown in Figure 13. A computer with an NVIDIA GeForce RTX 2080 Ti graphics card was used. It is equipped with Intel(R) Xeon(R) Silver 4210 CPU running at 2.20 GHz with 32 GB of RAM. Figure 14 shows the route of the real vehicle experiment. The experimental vehicle traveled in the campus environment at a speed of 0–30 km/h (including a straight section and a turning section), and the environmental information was collected by the camera. The total number of obstacles in the target area was 4479. The obstacles in the road were classified into two categories: 1. Pseudo obstacles with height and no risk of obstructing the vehicle's travel. 2. Real obstacles with height and risk of obstructing the vehicle's travel.

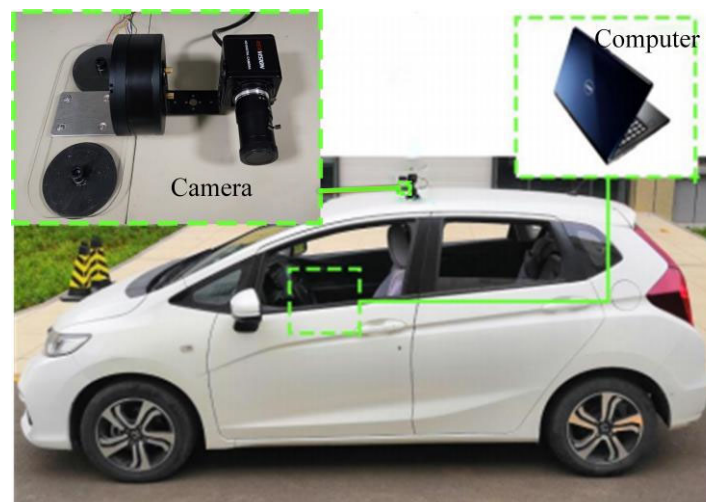


Figure 13. Experimental equipment for real vehicle.

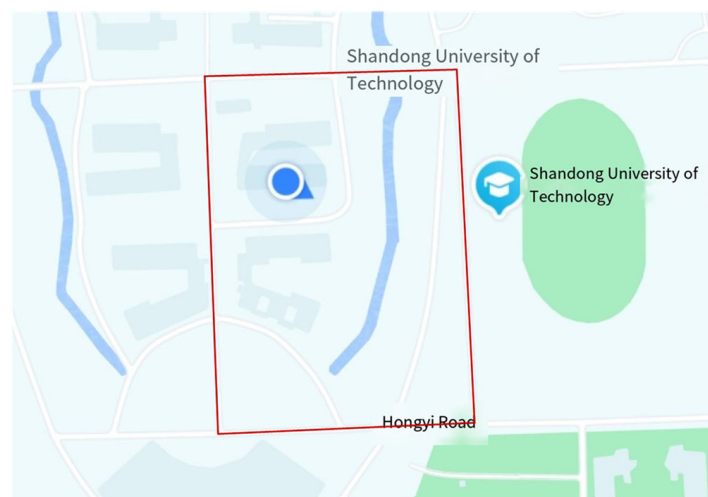


Figure 14. Real vehicle experiment route.

YOLOv8 is divided into different versions such as YOLOv8n, YOLOv8s, YOLOv8m, YOLOv8l and YOLOv8x according to the depth and width of the network. In this paper, YOLOv8s was selected for comparison experiments by taking the model size into account. The statistical data of the experimental results of four obstacle detection methods, this paper's method, VIDAR [15], VIDAR + MSER [26], and YOLOv8s were used to measure the accuracy of the obstacle detection methods with TP, FP, TN, and FN values [27–29], where TP denotes that the model correctly predicts real obstacles as real obstacles, FN denotes that the model incorrectly predicts real obstacles as pseudo-obstacles, FP denotes that the model incorrectly predicts pseudo-obstacles as real obstacles, and TN denotes that the model correctly predicts pseudo-obstacles as pseudo-obstacles. The TP, FP, TN, and FN values for each method are shown in Table 2.

Table 2. TP, FP, TN, FN values of the four methods.

Experimental Method	TP	FP	TN	FN
VIDAR	3710	397	118	254
VIDAR + MSER	3856	356	41	226
YOLOv8s	3527	362	289	301
Proposed method	4033	146	168	132

In the result analysis, accuracy, recall, precision, mAP were used as the evaluation index for obstacle detection methods, where mAP is the average of average precision (average precision, AP) of all categories. Accuracy indicates the proportion of correctly judged data to the total data. Recall rate for all positive cases in the data, indicating the proportion of correctly judged positive cases to all positive cases in the total data. The accuracy rate is also called the check rate, and a high accuracy rate indicates that most of the obstacles detected by the model are indeed obstacles. Since the methods in this paper, VIDAR and VIDAR + MSER are designed to categorize road obstacles into two categories, obstacles and non-obstacles and YOLOv8s contains multiple categories of obstacles, mAP is therefore used as the final evaluation metric for the four detection methods. The comparison of the detection precision of each method is shown in Table 3. Accuracy, recall and precision can be expressed, respectively, as follows:

$$Accuracy = \frac{FP + FN}{TP + TN + FP + FN} \quad (11)$$

$$Recall = \frac{FP}{TP + FP} \quad (12)$$

$$Precision = \frac{TP}{TP + FP} \quad (13)$$

Table 3. Results of the four methods.

Experimental Method	mAP/%	Recall/%	Accuracy/%	Precision/%	Time/s
VIDAR	89.3	93.6	85.5	90.3	0.324
VIDAR + MSER	92.7	94.4	87.6	91.5	0.343
YOLOv8s	88.1	92.1	85.2	90.7	0.205
Proposed method	96.7	96.8	93.8	96.5	0.317

The results show that the mAP of the method in this paper was improved by 7.4%, 4%, and 8.6% compared to the obstacle detection methods based on VIDAR, VIDAR+MSER, and YOLOv8s, respectively. The recall improved by 3.2%, 2.4% and 4.7%. The accuracy improved by 8.3%, 6.2% and 8.6%. Compared with YOLOv8s, this paper's method had no obvious advantage in detection time, but compared with the other two methods, this paper's method needed to deal with fewer feature points, and the final detection time was relatively short.

Figure 15 shows the obstacle detection results (partially) when the experiment vehicle is in motion and stationary. From the figure, it can be seen that the method in this paper can realize the detection of common obstacles in the road, such as vehicles, bicycles, pedestrians and so on. In addition, it can also achieve the detection of unknown types of obstacles, such as irregularly shaped tires due to breakage, road cones, barricades and so on. Poor detection is due to YOLOv8s lack of training on unknown types of obstacles. Moreover, in scenes with complex backgrounds, false detection occurs, which leads to lower detection accuracy. The second group in Figure 15 shows the images captured by the camera when the vehicle is completely stationary, and VIDAR and VIDAR + MSER cannot recognize the obstacles when the vehicle is stationary. The detection method proposed in this paper uses camera rotation to obtain image information and process the obstacle feature points. For a class of obstacles with similar color to the ground, the detection effect is poor, but the detection effect of this paper's method is better overall than the other methods. In addition, the method used in this paper calculates the distance between the camera and the obstacle in front of it. The distance detection results of obstacles in the second set of images are shown in Table 4. By analyzing the difference between the actual distance and the measured distance, the results show that the distance error is mainly between 0.09 and 0.23. The method has an error of less than 0.15 m over short distances (<10 m); the error increases with the increase in distance, and the distance error is relatively large in the presence of

obstacles that are partially obscured. From the overall results of distance measurement, the vision-based ranging algorithm proposed in this paper meets the requirements of measurement accuracy and can realize accurate distance measurement of obstacles within a short distance.



Figure 15. Detection results.

Table 4. Distance measurement results.

Obstacle	Measuring Distance (m)	Actual Distance (m)	Error (m)
1	4.79	4.88	0.09
2	5.13	5.24	0.11
3	7.60	7.73	0.13
4	8.33	8.44	0.11
5	11.09	11.24	0.15
6	10.10	10.27	0.17
7	12.78	12.97	0.19
8	15.72	15.92	0.20
9	15.90	16.08	0.18
10	18.68	18.89	0.21
11	22.53	22.76	0.23
12	12.19	12.34	0.15
13	13.50	13.64	0.14
14	14.58	14.66	0.18
15	9.50	9.65	0.15
16	12.98	12.12	0.14

5. Conclusions and Future Work

In this paper, we propose an obstacle detection method based on longitudinal active vision, which realizes the detection of all types of obstacles in the road, does not depend on the accurate classification of obstacles, and reduces the spatiotemporal complexity of road environment perception. Combining the fast image region-matching method based on MSER with the obstacle detection method based on longitudinal active vision improves the speed of image matching as well as the accuracy of obstacle recognition and realizes the distance measurement of obstacles. The experimental results show that this paper's method can effectively recognize obstacles, and the mAP of this paper's method is improved by 7.4%, 4%, and 8.6% compared with the obstacle detection methods of VIDAR, VIDAR+MSER and YOLOv8s. The obstacle ranging error over a short distance (<10 m) is less than 0.15 m. Compared with other obstacle detection methods, the method in this paper is more applicable and can meet the detection requirements in complex environments.

In this paper, we produced a methodological innovation in obstacle detection, using the obstacle detection method based on longitudinal active vision to realize the accurate detection of road obstacles, and there are still some shortcomings. Compared with other traditional obstacle detection methods, the detection speed of this paper's method is low and has no obvious advantages. The next step is to consider using more complex feature descriptors and matching algorithms to improve the detection speed. In addition, machine learning is being considered for combining with this paper's model in subsequent research to realize more comprehensive and efficient obstacle detection on the roads.

Author Contributions: Conceptualization, S.S. and Y.X.; methodology, S.S., J.N. and X.K.; software, X.K., H.Z. and J.Z.; validation, Y.X. and Q.S.; investigation, J.Z., Q.S. and H.Z.; writing—original draft preparation, S.S. and J.N.; writing—review and editing, Y.X.; supervision, Y.X. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Shandong Province Youth Innovation Team Plan for Universities (Grant No. 2023KJ153), Shandong Province Major Science and Technology Innovation Project (Grant No. 2023CXGC010111), Small and Medium-sized Enterprise Innovation Capability Improvement Projects (Grant No. 2022TSGC2277), Experimental Technology Upgrading Project (Grant No. 2022003), Intelligent Automotive Sensor Technology Graduate Project (Grant No. 4053223015) and Natural Science Foundation of Shandong Province (Grant No. ZR2022MF230).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: Author Yi Xu was employed by the company Qingte Group Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. World Health Organization. *Global Status Report on Road Safety 2023*; World Health Organization: Geneva, Switzerland, 2023.
2. Ci, W.; Xu, T.; Lin, R.; Lu, S.; Wu, X.; Xuan, J. A Novel Method for Obstacle Detection in Front of Vehicles Based on the Local Spatial Features of Point Cloud. *Remote Sens.* **2023**, *15*, 1044. [[CrossRef](#)]
3. Badrloo, S.; Varshosaz, M.; Pirasteh, S.; Li, J. Image-Based Obstacle Detection Methods for the Safe Navigation of Unmanned Vehicles: A Review. *Remote Sens.* **2022**, *14*, 3824. [[CrossRef](#)]
4. Dodge, D.; Yilmaz, M. Convex vision-based negative obstacle detection framework for autonomous vehicles. *IEEE Trans. Intell. Veh.* **2022**, *8*, 778–789. [[CrossRef](#)]
5. Ding, N.; Mokhtarzadeh, A.A.; Geng, J. A Multi-lidar Based Dynamic Obstacle Detection and Prediction Method. In Proceedings of the 2023 International Conference on the Cognitive Computing and Complex Data (ICCD), IEEE2023, Huaian, China, 21–22 October 2023; pp. 162–167.
6. Popov, A.; Gebhardt, P.; Chen, K.; Oldja, R.; Lee, H.; Murray, S.; Bhargava, R.; Smolyanskiy, N. Nvradarnet: Real-time radar obstacle and free space detection for autonomous driving. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA/IEEE), IEEE2023, London, UK, 29 May–2 June 2023; pp. 6958–6964.

7. Han, H.; Chen, Y.; Hsiao, P.; Fu, L. Using Channel-Wise Attention for Deep CNN Based Real-Time Semantic Segmentation with Class-Aware Edge Information. *IEEE Trans. Intell. Transp. Syst.* **2021**, *22*, 1041–1051. [[CrossRef](#)]
8. Nguyen, H. Improving Faster R-CNN Framework for Fast Vehicle Detection. *Math. Probl. Eng.* **2019**, *2019*, 1–11. [[CrossRef](#)]
9. Zaghari, N.; Fathy, M.; Jameii, S.M.; Shahverdy, M. The improvement in obstacle detection in autonomous vehicles using YOLO non-maximum suppression fuzzy algorithm. *J. Supercomput.* **2021**, *77*, 13421–13446. [[CrossRef](#)]
10. Yasmin, S.; Durrani, M.Y.; Gillani, S.; Bukhari, M.; Maqsood, M.; Zghaibeh, M. Small obstacles detection on roads scenes using semantic segmentation for the safe navigation of autonomous vehicles. *J. Electron. Imaging* **2022**, *31*, 061806. [[CrossRef](#)]
11. He, D.; Ren, R.; Li, K.; Zou, Z.; Ma, R.; Qin, Y.; Yang, W. Urban rail transit obstacle detection based on Improved R-CNN. *Measurement* **2022**, *196*, 111277. [[CrossRef](#)]
12. Sun, L.; Yang, K.; Hu, X.; Hu, W.; Wang, K. Real-Time Fusion Network for RGB-D Semantic Segmentation Incorporating Unexpected Obstacle Detection for Road-Driving Images. *IEEE Robot. Autom. Lett.* **2020**, *5*, 5558–5565. [[CrossRef](#)]
13. Güngör, E.; Özmen, A. Stereo-image-based ground-line prediction and obstacle detection. *Turk. J. Electr. Eng. Comput. Sci.* **2024**, *32*, 465–482. [[CrossRef](#)]
14. Yuan, J.; Jiang, T.; He, X.; Wu, S.; Liu, J.; Guo, D. Dynamic obstacle detection method based on U-V disparity and residual optical flow for autonomous driving. *Sci. Rep.* **2023**, *13*, 7630. [[CrossRef](#)]
15. Xu, Y.; Gao, S.; Li, S.; Tan, D.; Guo, D.; Wang, Y.; Chen, Q. Vision-IMU based obstacle detection method, Green Intelligent Transportation Systems. In Proceedings of the 8th International Conference on Green Intelligent Transportation Systems and Safety, Changchun, China, 1–2 July 2017; Springer: Berlin/Heidelberg, Germany, 2019; pp. 475–487.
16. Kumar, M.P.; Ashok, D. A multi-level colour thresholding based segmentation approach for improved identification of the defective region in leather surfaces. *Eng. J.* **2020**, *24*, 101–108. [[CrossRef](#)]
17. Wang, S.; Li, X. A real-time monocular vision-based obstacle detection. In Proceedings of the 2020 6th International Conference on Control, Automation and Robotics (ICCAR), IEEE2020, Singapore, 20–23 April 2020; pp. 695–699.
18. Xue, F.; Chang, Y.; Wang, T.; Zhou, Y.; Ming, A. Indoor Obstacle Discovery on Reflective Ground via Monocular Camera. *Int. J. Comput. Vis.* **2024**, *132*, 987–1007. [[CrossRef](#)]
19. Capito, L.; Ozguner, U.; Redmill, K. Optical Flow based Visual Potential Field for Autonomous Driving. In Proceedings of the 2020 IEEE Intelligent Vehicles Symposium (IV), Las Vegas, NV, USA, 19 October–13 November 2020; pp. 885–891.
20. Lin, H.; Peng, X. Autonomous Quadrotor Navigation With Vision Based Obstacle Avoidance and Path Planning. *IEEE Access* **2021**, *9*, 102450–102459. [[CrossRef](#)]
21. Xu, H.; Li, S.; Ji, Y.; Cao, R.; Zhang, M. Dynamic obstacle detection based on panoramic vision in the moving state of agricultural machineries. *Comput. Electron. Agric.* **2021**, *184*, 106104. [[CrossRef](#)]
22. Ordóñez, A.; Acción, A.; Argüello, F.; Heras, D.B. HSI-MSER: Hyperspectral Image Registration Algorithm Based on MSER and SIFT. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 12061–12072. [[CrossRef](#)]
23. Tong, G.; Dong, M.; Sun, X.; Song, Y. Natural scene text detection and recognition based on saturation-incorporated multi-channel MSER. *Knowl. Based Syst.* **2022**, *250*, 109040. [[CrossRef](#)]
24. Kosala, G.; Harjoko, A.; Hartati, S. MSER-Vertical Sobel for Vehicle Logo Detection. *J. RESTI (Rekayasa Syst. Dan Teknol. Inf.)* **2023**, *7*, 1239–1245. [[CrossRef](#)]
25. Li, Y.; Li, Z.; Shen, Y.; Yang, J. Infrared small target detection using reinforced MSER-induced saliency measure. *Infrared Phys. Technol.* **2023**, *133*, 104829. [[CrossRef](#)]
26. Yi, X.; Song, G.; Derong, T.; Dong, G.; Liang, S.; Yuqiong, W. Fast road obstacle detection method based on maximally stable extremal regions. *Int. J. Adv. Robot. Syst.* **2018**, *151*, 1729881418759118. [[CrossRef](#)]
27. Liu, Z.; Zhao, S.; Wang, X. Research on Driving Obstacle Detection Technology in Foggy Weather Based on GCANet and Feature Fusion Training. *Sensors* **2023**, *23*, 2822. [[CrossRef](#)] [[PubMed](#)]
28. Dairi, A.; Harrou, F.; Senouci, M.; Sun, Y. Unsupervised obstacle detection in driving environments using deep-learning-based stereovision. *Robot. Auton. Syst.* **2018**, *100*, 287–301. [[CrossRef](#)]
29. Lambert, R.; Chavez-Galaviz, J.; Li, J.; Mahmoudian, N. ROSEBUD: A Deep Fluvial Segmentation Dataset for Monocular Vision-Based River Navigation and Obstacle Avoidance. *Sensors* **2022**, *22*, 4681. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.