



Article

Research on a Vehicle Recognition Method Based on Radar and Camera Information Fusion

Fang Ding *, Bo Wang, Qianbin Zhang and Aiguo Wang

College of Automobile and Rail, Anhui Technical College of Mechanical and Electrical Engineering, Wuhu 241000, China

* Correspondence: 0125000283@ahcme.edu.cn

Abstract: To improve the accuracy and real-time performance of vehicle recognition in an advanced driving-assistance system (ADAS), a vehicle recognition method based on radar and camera information fusion is proposed. Firstly, the millimeter-wave radar and camera are calibrated jointly, the radar recognition information is mapped on the camera image, and the region of interest is established. Then, based on operator edge detection, global threshold binarization is performed on the image of the region of interest (ROI) to obtain the contour information of the vehicle in front, and Hough transform is used to fit the vehicle contour edge straight line. Finally, a sliding window is established according to the symmetry characteristics of the fitting line, which can find the vehicle region with the highest symmetry and complete the identification of the vehicle. The experimental results show that compared to the original recognition region of the radar, the mean square error of this algorithm is reduced by 13.4 and the single frame detection time is reduced to 28 ms. It is proven that the algorithm has better accuracy and a faster detection rate, and it can solve the problem of an inaccurate recognition region caused by radar error.

Keywords: vehicle detection; millimeter-wave radar; machine vision; information fusion



Citation: Ding, F.; Wang, B.; Zhang, Q.; Wang, A. Research on a Vehicle Recognition Method Based on Radar and Camera Information Fusion. *Technologies* **2022**, *10*, 97. <https://doi.org/10.3390/technologies10040097>

Academic Editor: Vijayakumar Varadarajan

Received: 24 July 2022

Accepted: 19 August 2022

Published: 22 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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

Environmental perception is a key technology in the research field of advanced driving-assistance systems (ADAS). The sensors currently used in environmental perception mainly include a millimeter-wave radar, lidar, and camera. A millimeter-wave radar can detect the distance, relative velocity, and azimuth information of the target in front, but cannot acquire category and size information of the target. Lidar can create the point cloud of the measured target with depth information, but due to the expensive price and complex algorithm, it has not been popularized on a large scale. A camera can collect images with rich environmental information, but distance information is missing in the detection targets of images. Through information fusion, more accurate and less redundant information can be obtained than with a single sensor, which can improve the safety of vehicle driving. Therefore, the use of multi-sensor information fusion is an inevitable development trend.

For the research on multi-sensor sensing systems, Teoh and Bräunl [1] proposed a fast vehicle-detection method based on vehicle edge symmetry and horizontal shadows, which searched the symmetrical area overall based on the characteristics of the vehicle. However, the algorithm only used a single visual sensor and did not consider the identification information of a millimeter-wave radar. During the vehicle driving process, the environmental information will add noise to the image, making it impossible to accurately identify areas of vehicle symmetry. Therefore, Satzoda and Trivedi [2] proposed a symmetry-detection algorithm based on closed contour corner information. The symmetry error detected by the algorithm became smaller and solved some problems, but it still needs to be improved to meet real-time requirements.

In the work of Mingchi et al. [3], a camera and millimeter-wave radar were fused, where the millimeter-wave radar was used to provide the camera with a region of interest.

However, when symmetry analysis was performed for the region, the symmetry axis was only determined by calculating the midpoint of the shadow at the rear of the vehicle. Affected by the visual ranging error, it could not guarantee high accuracy. Parvin et al. [4] first acquired the approximate position of the vehicle in the image by extracting the overall features of the image, then projecting the millimeter-wave radar detection area into the image to match the vehicle features for information fusion, but this method had a slow processing speed, and the acquisition of the overall features required a complicated operation.

To enhance the image feature-extraction ability, Sun et al. [5] used the vehicle recognition algorithm based on the AdaBoost classifier, which achieved some beneficial results, but this algorithm only explored the two-dimensional detection performance, failing to analyze its fusion effect and ranging accuracy. On this basis, Hu et al. [6] calculated the intersection ratio between the visual detection frame and the millimeter-wave radar detection frame, and they used this as the basis for determining the front target, but the generation of a visual detection frame in this algorithm was affected by the training data. In the case of multi-vehicle occlusion, leak detection of the algorithm would occur.

Wang et al. [7] proposed a vehicle target-detection algorithm based on the fusion of the millimeter-wave radar and a monocular camera using rectangular boundary constraints and active contour detection. The algorithm used the active contour method to detect the vehicles within the boundary. Yet, in the detection process, the active contour method was seriously affected by the occlusion of light and shadow, so it was difficult to adapt to a traffic scene with multiple vehicles. Han et al. [8] proposed a fusion detection framework of the millimeter-wave radar and a camera based on probabilistic reasoning, which completely inputted the recognition results of the classifier and the detection results of the millimeter-wave radar into the probabilistic reasoning module to estimate the location and category of the target. However, the algorithm involved a large amount of computation and time delay.

Jiang et al. [9] proposed a target-detection algorithm based on millimeter-wave radar and camera fusion in foggy weather. In the fusion stage, the weighted method was used to simply combine the neural network detection results with the radar target estimation results, but the accuracy of the detection results was not high enough.

In view of the shortcomings of the millimeter-wave radar and camera fusion in environmental perception, the method proposed here first uses coordinate transformation to achieve the spatial fusion of the millimeter-wave radar and camera, then uses projected target points to generate a region of interest on the image to reduce the cost of pixel searching. In this area, image preprocessing, edge detection, and roof line fitting are performed. Furthermore, the prior information on the vehicle size is used to establish a sliding window, and a symmetry function is proposed to guide the translation of the window. Finally, the identification area with the highest linear symmetry on the roof is found to accurately complete the vehicle identification. This paper qualitatively and quantitatively proves the effectiveness of the proposed fusion method for correcting the inaccurate regions detected by radar through real-vehicle experiments, and explores the determination of the sliding window size in scenes under different lighting.

2. Radar and Camera Information Fusion

Radar and cameras are used to perceive the road environment and collect information about vehicles ahead. In this paper, the 77 GHz ARS404 mm wave radar from Continental Germany and the MV-SUA134GC monocular industrial camera from MindVision are used as sensors. The installation positions of the millimeter-wave radar and the camera are shown in Figure 1. The millimeter-wave radar is installed in the center of the front bumper of the vehicle and fixed with bolts. The camera is fixed on the front windshield of the vehicle with a suction cup, which is specifically installed under the rearview mirror of the vehicle. Taking the horizontal road surface as the reference, the radar axis and the camera optical axis are on a vertical plane perpendicular to the road surface.

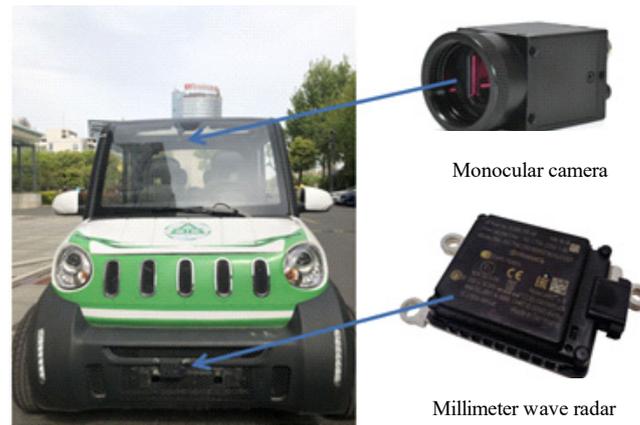


Figure 1. Radar and camera installation locations.

2.1. Coordinate Fusion of Sensors

The camera imaging model follows the ideal pinhole imaging principle. The principle is that the light in the three-dimensional space needs to be mapped to the two-dimensional plane of the camera. Therefore, to establish the three-dimensional world coordinate system and pixel coordinate system, the camera needs to be calibrated to measure its internal and external parameters. Figure 2 shows the camera imaging model, which represents the process of the conversion of a point from the world coordinate system to the image pixel coordinate system.

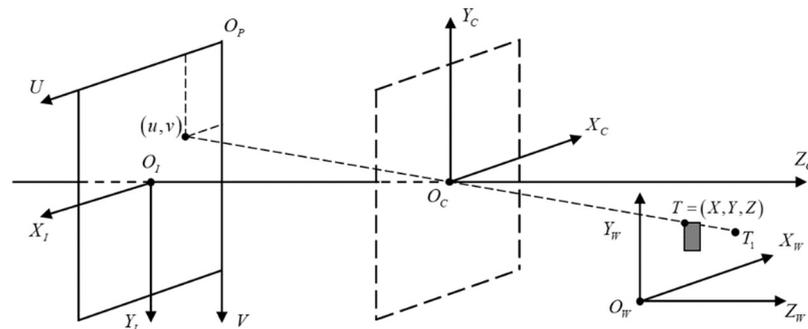


Figure 2. Camera imaging model.

In Figure 2, $O_w - X_w Y_w Z_w$ is the world coordinate system, $O_c - X_c Y_c Z_c$ is the camera coordinate system, $O_i - X_i Y_i$ is the image coordinate system, and $O_p - UV$ is the pixel coordinate system. According to the mapping principle, the transformation relationship between the world coordinate system and the pixel coordinate system is shown in Formula (1):

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = M_2 M_1 \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (1)$$

where (u, v) is the pixel coordinate, (u_0, v_0) is the pixel coordinate of the optical center, f_x and f_y are the normalized focal lengths in the x and y axes directions of the camera, respectively, where the unit is pixels, R represents the three-dimensional rotation matrix, T represents the three-dimensional translation vector, M_2 is the camera's internal parameter matrix, and M_1 is the camera's external parameter matrix.

As shown in Figure 3, $O_c - X_c Y_c Z_c$ is the camera coordinate system, and $O_R - X_R Y_R Z_R$ is the millimeter-wave radar coordinate system. The millimeter-wave radar has a distance of Z_{RC} from the camera in the Z axial direction, and a distance of Y_{RC} from the camera in

the Y axial direction. The transformation relationship between the millimeter-wave radar coordinate system and the camera coordinate system is shown in Formula (2).

$$\begin{bmatrix} X_C \\ Y_C \\ Z_C \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} X_R \\ Z_R \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ -Y_{RC} \\ Z_R \\ 1 \end{bmatrix} \quad (2)$$

By combining Formulas (1) and (2), the conversion from the radar coordinate system to the pixel coordinate system can be completed. The position information of a point detected by the radar in the three-dimensional world can be mapped to the pixel coordinate system, thus achieving coordinate fusion between sensors.

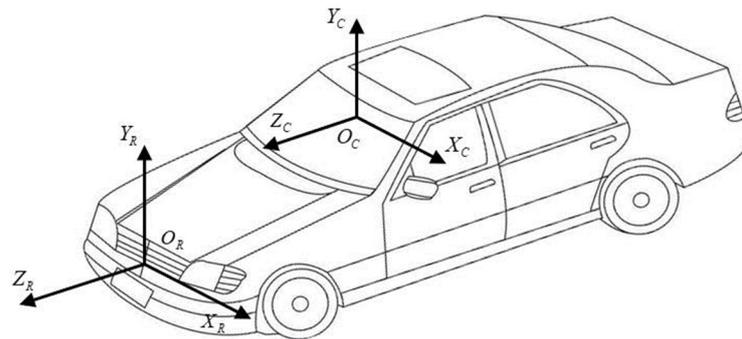


Figure 3. Schematic diagram of radar coordinate system and camera coordinate system.

2.2. Acquisition of Millimeter-Wave Radar Region of Interest

The acquisition of the region of interest includes the determination of the location and size of the region. As shown in Figure 4, the millimeter-wave radar consists of one transmitter (TX) and four receivers (RX), and the three-dimensional coordinates and speed of the front target are obtained according to the reflected waves of different phases. The three-dimensional coordinates (X_R, Y_R, Z_R) of the target are used as the center point of the region of interest. Specifically, Formula (2) is used to complete the conversion from the millimeter-wave radar coordinate system to the camera coordinate system and combined with Formula (1), the center point is projected onto the image (the red point in Figure 5), and the center position of the region of interest is completed.

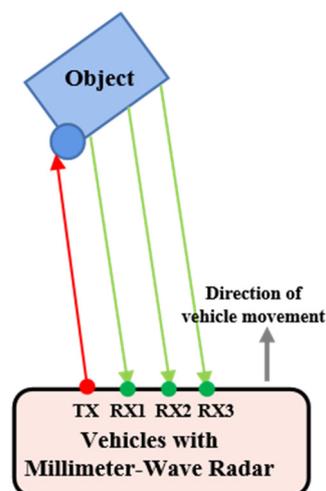


Figure 4. Millimeter-wave radar antenna principle.



Figure 5. Radar detection target is projected on the image.

Considering that the vehicle outline is roughly rectangular, a rectangular region of interest is established with the projection point of the radar detection target as the center. After acquiring the mapping point of the millimeter-wave radar data in the image, the point is taken as the center of the rectangular frame. The size of the rectangular frame is determined according to the perspective transformation of the camera. Figure 6 is a schematic diagram of a perspective transformation where L is the distance from the camera to the vehicle, H is the vehicle height, W is the vehicle width, H_c is the height of the field of view, W_c is the width of the field of view, β is the vertical angle of view, α is the horizontal angle of view, X_b is the width of the captured image, Y_b is the height of the captured image, w is the width of the vehicle in the image, and h is the height of the vehicle in the image.

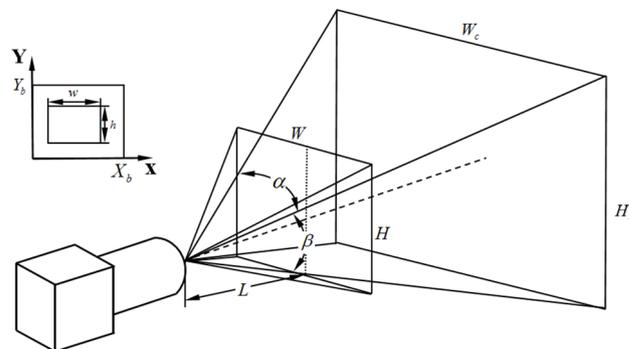


Figure 6. Schematic diagram of perspective transformation.

Since the camera imaging model follows the ideal pinhole imaging principle, the similar triangle principle can be used, and the rectangular frame size of the vehicle in the image can be calculated according to Formula (3):

$$\begin{aligned} \frac{H}{h} &= \frac{2H_c}{Y_b} & H_c &\approx L \tan \frac{\beta}{2} & H &= \frac{2L \tan \frac{\beta}{2}}{Y_b} h \\ \frac{W}{w} &= \frac{2W_c}{X_b} & W_c &\approx L \tan \frac{\alpha}{2} & W &= \frac{2L \tan \frac{\alpha}{2}}{X_b} w \end{aligned} \quad (3)$$

According to the relevant laws and regulations of road traffic and the national standard of China GB 1589–2004 “Road Vehicle Outline Dimensions, Axle Loads and Mass Limits” regarding vehicle dimensions, the vehicle height is 1.6 times that of the width. To ensure that the region of interest can completely reflect vehicle information, a rectangular frame with a width of 2.6 m and a height of 4.2 m is used as the region of interest of the radar, and the region of interest is mapped onto the image using perspective transformation based on Formula (3) to complete the size determination of the region of interest.

According to the radar data map shown in Figure 7, combined with the position and size of the region of interest, the initial radar region of interest, that is, the red box region in Figure 8, can be determined.



Figure 7. The projected point of the radar's target detection point in the image.



Figure 8. Initial radar region of interest.

In the actual detection of the millimeter-wave radar, the target signal reflected by its beam is not necessarily at the center of the vehicle, and the vehicle is affected by road conditions and working conditions during driving, resulting in the deviation of the radar target. Hence, the initial region of interest will be corrected below to obtain a more precise location.

3. Feature-Based Vehicle Recognition

3.1. Image Preprocessing

After the initial vehicle region of interest is determined, the features in the region need to be analyzed to determine the best identification region, and before that, the image should be preprocessed. As shown in Figure 9, the original image is converted into a grayscale image. Since the grayscale image has less information than the color image, it can increase the running speed without losing the image texture information.

The noise in the image can be removed by using 3×3 Gaussian filtering for the smoothing of the grayscale image; the edge part of the image is more prominent after smoothing, meaning it is ready for the following edge detection.



Figure 9. Image preprocessing. (a) Original image. (b) Grayscale image after smoothing.

3.2. Operator-Based Edge Detection

In digital image processing, it is considered that pixels at the edge have relatively severe grayscale changes, so the Sobel operator [10] is used to calculate the grayscale gradient in the x and y directions between adjacent pixels, and the magnitude and direction of the gradient are used to indicate the grayscale change for image edge detection. The specific detection steps are as follows:

- (1) Define the Sobel convolution factor, as shown in Formula (4). The operator includes two groups of 3×3 matrices, representing the vertical and horizontal directions, respectively:

$$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \quad (4)$$

- (2) Use the Sobel convolution factor to perform the convolution of the image [11] so the gray gradient approximation G_X , G_Y in the two directions can be obtained. The pixel gradient size G is calculated by Formula (5):

$$G = \sqrt{G_X^2 + G_Y^2} \quad (5)$$

- (3) The gradient direction can be calculated according to Formula (6):

$$\Theta = \arctan\left(\frac{G_X}{G_Y}\right) \quad (6)$$

- (4) After calculating G and Θ , according to the structural characteristics of the driving scene, the maximum pixel gradient size G_{\max} is set to 0.5, the minimum pixel gradient size G_{\min} is set to 0 as the threshold, and the pixels located at the edge are extracted. Figure 10a shows the edge detection result using the Sobel operator. It can be seen that a significant vehicle contour curve can be plotted.

Canny edge detection [12] is used to suppress the magnitude of non-maximum values according to the gradient direction based on the Sobel operator, which uses the double threshold algorithm to connect the pixels between the thresholds and locate more accurate edge pixels, thus effectively suppressing false edges. Figure 10b shows the result of Canny edge detection. It can be seen that Canny edge detection can obtain more accurate vehicle edges than the Sobel operator.

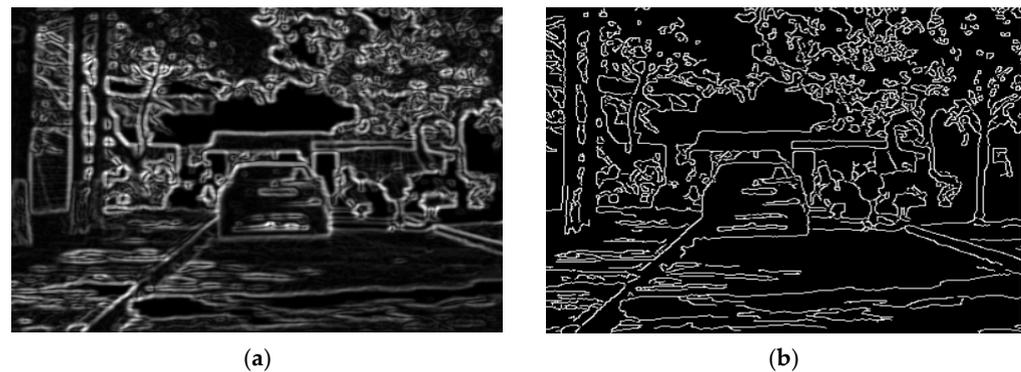


Figure 10. Operator edge-detection results. (a) Sobel edge detection. (b) Canny edge detection.

3.3. Vehicle Contour Line Fitting

When the vehicle is driving in a structured scene, the gray value of the surrounding contour of the vehicle is quite different from that of the environment, which is helpful to extract the vehicle contour. Firstly, the edge-enhanced image is obtained by Canny edge detection performed on the radar region of interest, then the global threshold of the image is obtained by using the maximum interclass variance of pixels [13], which is transformed into a binary image. Finally, the vehicle contour information is fitted by the probabilistic Hough transform.

In this paper, the probabilistic Hough transform [14] is used to detect the contour line of the vehicle shape by voting. The transformation process takes place in the parameter space. By calculating the local maximum value of the accumulated result, a set conforming to the specific shape is established as a result of the Hough transform. Figure 11b is determined after edge detection and the binarization of the radar identification area in the red frame in (a), then the Hough transform is performed to fit the vehicle contour line. By setting the accumulated plane threshold, the minimum straight-line length and the maximum interval of the line segment, as shown by the red line in (c), the vehicle contour with straight-line characteristics can finally be fitted, and the coordinates of each pixel in the fitted line can be determined.

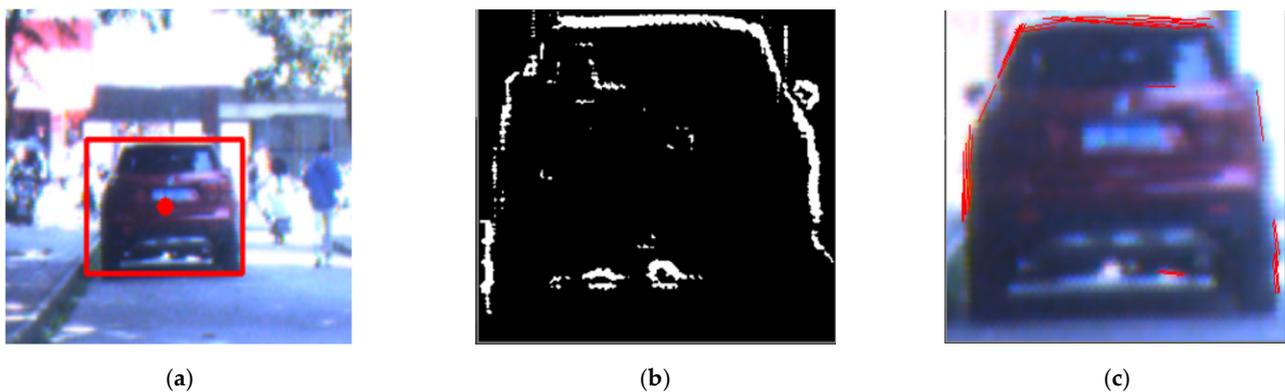


Figure 11. Vehicle contour fitting. (a) Radar recognition area. (b) Binary image of vehicle contour edge. (c) Hough transform to fit vehicle contour.

4. Vehicle Symmetry Detection

4.1. Sliding Window Detection

In vehicle symmetry detection, it is very important to judge whether the vehicle in the region is symmetrical and find the best recognition frame according to the region of interest identified by the radar. Therefore, this paper adopts sliding window technology [15], takes the center point of the region of interest of the initial millimeter-wave radar as the starting point, translates 20 pixels to the right and left, respectively, and sets the sliding step as four

pixels. As shown in Figure 12, the sliding window range covers the recognition frame with the highest vehicle symmetry.



Figure 12. Sliding window detection.

4.2. Vehicle Symmetry Analysis

In symmetry analysis, the concept of symmetry is introduced [16]. Firstly, the straight line of the vehicle contour fitted above is analyzed, as shown in Figure 13a. Since there is a good boundary between the roof and the environment, this paper firstly uses the roof fitting line as the detection standard, then calculates the horizontal distance U_{left} between the left endpoint of the roof line and the image origin. Following this, the horizontal distance U_{right} between the right endpoint and the image origin, and the horizontal distance between the middle point of the roof line and the image origin, can be calculated according to Formula (7).

$$U_{mid} = \frac{U_{right} - U_{left}}{2} + U_{left} \quad (7)$$

Finally, the difference between the center point of the roof line and the image center point Q in the horizontal direction of the image can be calculated, and this difference is used to measure the symmetry sym of the vehicle position in the detection frame, as shown in Formula (8):

$$sym = |U_Q - U_{mid}| \quad (8)$$

where U_Q is the horizontal distance between the image center point Q and the origin.

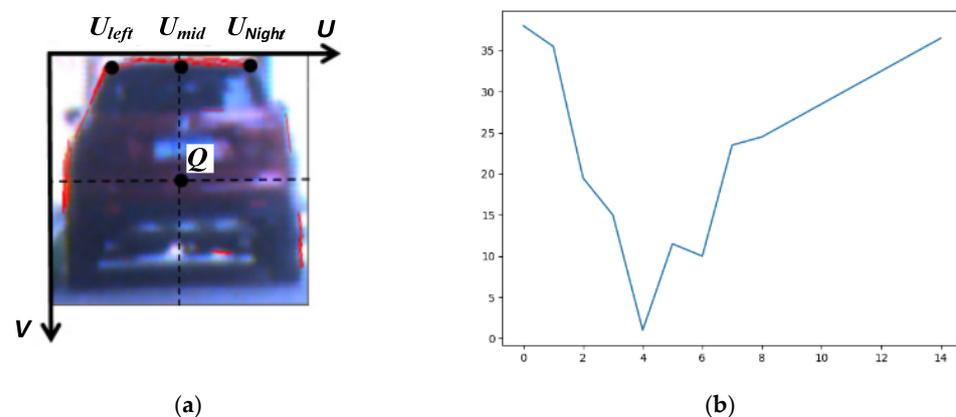


Figure 13. Vehicle symmetry analysis. (a) Schematic diagram of symmetry analysis. (b) Sliding window symmetry curve.

The symmetry problem can be converted into the solution of the minimum value sym_{min} of the symmetry degree by Formulas (7) and (8). As shown in Figure 13b, the symmetry curve formed by the sliding window at different positions exhibits a single peak

characteristic, and the detection frame with the minimum symmetry degree is the detection frame with the optimal vehicle contour symmetry.

5. Experimental Results

5.1. Internal and External Parameter Calibration

To verify the performance of the vehicle symmetry detection algorithm, the camera is first calibrated. As shown in Figure 14, this paper uses an 11×8 checkerboard diagram, in which the side length of the small square is 20 mm. The internal parameters are calibrated according to 20 multi-position and multi-angle chessboard diagrams taken by the fixed camera.

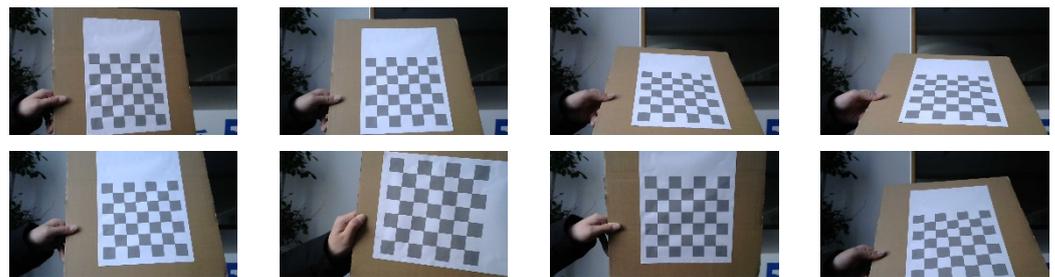


Figure 14. Calibration of the camera's internal parameters.

After the camera's internal parameter calibration, the external parameter matrix of the camera and the radar can be calculated by using a laser rangefinder to measure the vertical and horizontal distances between the radar and the camera, and at the same time, measure the height of the radar from the ground. Table 1 shows the external parameter calibration of the camera. The transformation of different spaces can be realized according to the camera's internal parameters, rotation matrix, and translation matrix.

Table 1. Calibration table of camera parameters.

Calibration Parameters	
camera's internal parameters	$K = \begin{bmatrix} 2364.5 & 0 & 901.5 & 0 \\ 0 & 2364.5 & 518.3 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$
rotation matrix	$R = \begin{bmatrix} -1 & -0.0011 & 0.0024 & -18.64 \\ 0.0009 & -0.8732 & -0.034 & 12.3652 \\ 0.0024 & -0.034 & 0.8732 & 3125.6 \\ 0 & 0 & 0 & 1 \end{bmatrix}$
translation matrix	$T = [241.6 \quad 489.3 \quad 1278.17]$

5.2. Real Vehicle Experiment and Visualization

After the coordinate conversion between the camera and the radar, this paper collects the camera and radar data on a campus road based on the built experimental platform. The algorithm is written in Python, with radar frequency band 77 GHz and camera resolution 1280×720 . Figure 15 shows the best symmetry recognition results at different distances. Figure 15a is the radar's initial region of interest. It can be seen that this region does not completely contain the entire vehicle. Figure 15b shows the optimal detection region corrected by the algorithm. It can be seen that the vehicle has good symmetry characteristics in this region, and analysis of the experimental results shows that the single-frame processing speed of the algorithm is as fast as 28 ms, which means that the algorithm has a faster processing speed.



Figure 15. Comparison of vehicle recognition regions. (a) Initial radar identification region. (b) Identification region after improvement.

To verify the robustness of the proposed method, we select a night-time scene with poor lighting for qualitative analysis. As shown in Figure 16, an initially inaccurate sliding window can be generated in the image from the raw radar detection points. When we use the 225×255 window (blue box), commonly used during the day for correction, the right endpoint (yellow dot) of the roof is wrongly positioned on other targets on the road, and it cannot be accurately translated to the symmetrical area of the vehicle, resulting in detection failure. We further reduce the size of the sliding window and find that the window size of 180×180 (green box) does not greatly affect the performance of the Canny edge detection; more importantly, it can effectively overcome the interference of irrelevant factors. Therefore, selecting sliding windows of different sizes is a solution to adapt to scenes of different levels of complexity.

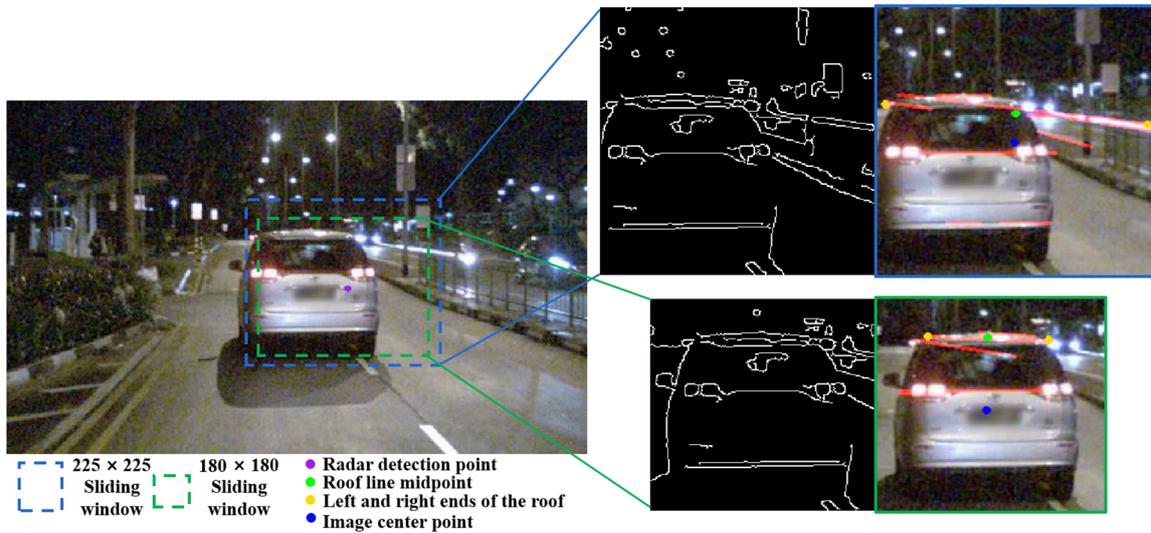


Figure 16. Detection of sliding windows of different sizes in night scenes.

5.3. Experimental Results' Analysis

To verify the recognition accuracy of the algorithm in this paper, the initial video of the camera and the raw data of the millimeter-wave radar are collected. Figure 17a shows the front target-detection results of the millimeter-wave radar at different times, representing the distance between the target and the x and y directions of the radar, and the azimuth angle between the target and the radar. Figure 17b shows the visual interface of radar detection, which helps us better understand the relative position between the target and the radar at the current moment. The interface includes the target information, target properties, radar status, number of targets, and number of filters.

1	ObjTimes	Obj_ID	Obj_Dist	Obj_Dist	Obj_Vrel	Obj_Vrel	Obj_Dyn	Obj_RCS	Obj_Dist	Obj_Dist	Obj_Vrel	Obj_Vrel
2	18:18:05	0	7	5.8	0.75	0	1	-5.5	20	20	19	20
3	18:18:05	1	1.4	1.4	0.75	0	1	-9	20	20	19	20
4	18:18:05	0	7	5.8	0.75	0	1	-5.5	20	20	21	20
5	18:18:05	1	1.4	1.2	0.5	0	1	-14	20	20	21	20
6	18:18:05	0	7	5.8	0.75	0	1	-5.5	20	20	22	20
7	18:18:05	1	1.6	1.2	0.5	0	1	-14	20	20	22	20
8	18:18:05	1	1.8	1.2	0.5	0	1	-15	20	19	21	20
9	18:18:05	7	3.8	1	0.25	0	1	-8	20	20	19	20
10	18:18:05	8	9.8	3.8	0.25	0	1	-5.5	20	20	19	20
11	18:18:05	11	49	0.8	0.25	0	1	9	20	21	19	20
12	18:18:05	1	1.8	1.2	0.25	0	1	-16	20	19	22	20
13	18:18:05	7	4	1	0.25	0	1	-8	20	20	21	20
14	18:18:05	8	9.4	4.2	0.25	0	1	-1	19	19	20	20
15	18:18:05	14	9.6	5.6	0.25	0	1	5.5	20	20	19	20
16	18:18:05	11	49	0.8	0.25	0	1	-0.5	20	21	21	20
17	18:18:05	1	1.8	1.2	0.25	0	1	-16	20	19	22	20
18	18:18:05	7	4	1	0.25	0	1	-10	20	20	22	20
19	18:18:05	8	9.4	4.2	0.25	0	1	-1	19	19	21	20
20	18:18:05	14	9.6	5.6	0.25	0	1	5.5	20	20	21	20
21	18:18:05	11	49	0.8	0.25	0	1	-0.5	20	21	22	20
22	18:18:05	1	2	1.2	0	0	1	-16	20	20	23	20
23	18:18:05	5	8.8	1.6	0	0	1	-5	20	20	19	20
24	18:18:05	8	9.4	4.2	0.25	0	1	-1	19	19	22	20



(a)

(b)

Figure 17. Radar raw data and visualization. (a) Initial video of the camera and data of the radar. (b) Visual interface of radar detection.

The 520 frames of the collected video and the radar data at the corresponding moment are selected, and we project them onto the image (as shown by the red dots in Figure 18). Labellmg software is used to label the vehicle in the image (green box in Figure 18), ensuring that the labeled boxes completely and symmetrically surround the entire vehicle area, where the center point of the area is taken as the real center point. It is worth noting that for a fair evaluation of the performance of the proposed method, we do not label the false and missed targets of the radar (the yellow area in Figure 18). This is because the false detection object lacks the image features of the vehicle, and the missed detection object lacks the position information provided by the millimeter-wave radar. After labeling, the accuracy of the initial radar recognition area and the recognition area generated by the algorithm in this paper are verified, respectively, as shown in Formula (9). The mean square

error $Loss$ is used to measure the error between the recognition area and the real area in the horizontal direction of the image:

$$Loss = \frac{\sum (x_t - x_p)^2}{n} \quad (9)$$

where x_t is the x direction coordinate value of the real center point, x_p is the x direction coordinate value of the center point of the recognition area, and n is the number of detection samples. As shown in Figure 19, the blue point indicates the projection point error of the raw radar data, and the yellow point indicates the projection point error after improvement of the algorithm in this paper. It can be seen that using image information to correct the raw radar data can greatly reduce the shift of the original detection frame in the horizontal direction. Compared to the raw radar data $Loss = 32.9$, the $Loss$ in this paper is 19.5, with the error reduced by 13.4, which proves the effectiveness of the radar and the camera information's fusion in this paper.

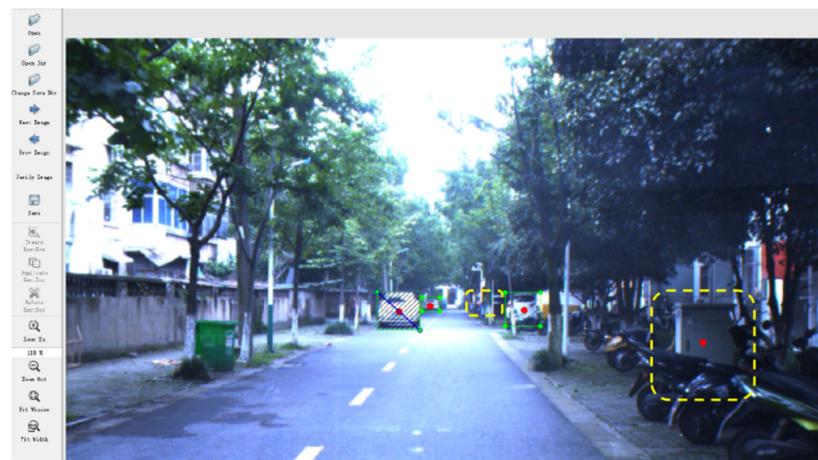


Figure 18. Annotation of ground-truth labels.

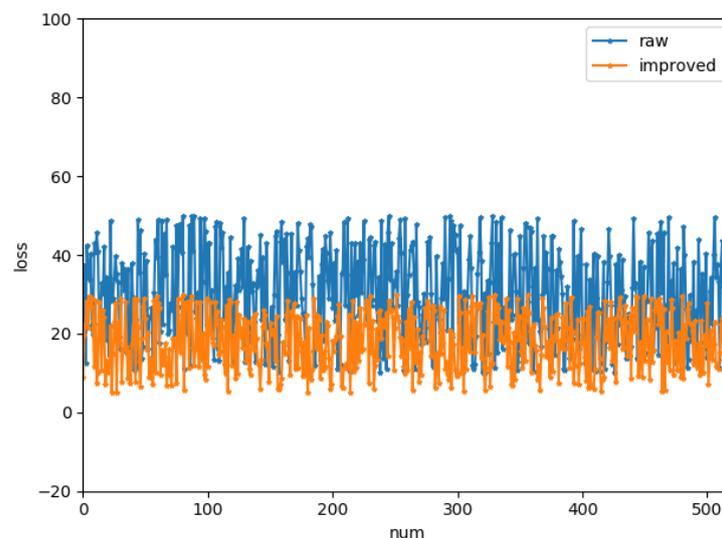


Figure 19. Comparison of the initial radar data and the algorithm error of this paper.

6. Conclusions

In this paper, a vehicle recognition method was proposed based on the fusion of radar and machine vision. The following conclusions can be drawn:

- (1) The algorithm first uses the camera's extrinsic and intrinsic parameters to realize the spatial fusion of the radar and the image, then projects the vehicle-recognition target point onto the camera image. Next, it further uses image smoothing, Canny edge detection, and probabilistic Hough transform to fit the roof contour line. Finally, a sliding window and symmetry function are established to detect the symmetry of vehicles in the region of interest, and in that way, we dynamically find the optimal recognition region;
- (2) The experimental results show that the algorithm has better accuracy and a faster processing speed; the mean square error of the algorithm is reduced by 13.4 and the single-frame detection time is reduced to 28 ms, which can meet real-time requirements in low-computing-power scenarios. In the end, we analyzed the influence of the size of the sliding window on the detection performance under different illuminations, and determined that the optimal, small 18×18 window can effectively reduce the interference of background pixels;
- (3) In this paper, only the fusion of a monocular industrial camera and millimeter-wave radar was attempted. In the future, we will consider using more types of sensors to extend this approach to the fusion of time-of-flight cameras and infrared cameras.

Author Contributions: Conceptualization, F.D. and B.W.; data curation, F.D.; funding acquisition, B.W. and Q.Z.; methodology, F.D. and B.W.; project administration, A.W.; supervision, A.W.; visualization, F.D.; writing—original draft preparation, F.D.; writing—review and editing, F.D. and B.W. All authors will be informed about each step of manuscript processing including submission, revision, revision reminder, etc. All authors have read and agreed to the published version of the manuscript.

Funding: The authors would like to thank the financial supports for the key natural science research projects of colleges and universities in the Anhui Province, Nos. KJ2020A1101 and KJ2020A1116, and the academic support project for top talents in disciplines (majors) of the colleges and universities in the Anhui Province, No. GXBJZD20106.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Teoh, S.S.; Bräunl, T. Symmetry-based monocular vehicle detection system. *Mach. Vis. Appl.* **2012**, *23*, 831–842. [[CrossRef](#)]
2. Satzoda, R.K.; Trivedi, M.M. Multipart Vehicle Detection Using Symmetry-Derived Analysis and Active Learning. *IEEE Trans. Intell. Transp. Syst.* **2016**, *17*, 926–937. [[CrossRef](#)]
3. Mingchi, F.; Xiaoqian, G.; Huizong, F.; Bowang, S. Research on the Fusion Method for Vehicle Shape-position Based on Binocular Camera and Lidar. In Proceedings of the 2021 6th International Symposium on Computer and Information Processing Technology, ISCIPT 2021, Changsha, China, 11–13 June 2021; pp. 419–423.
4. Parvin, S.; Islam, M.E.; Rozario, L.J. Nighttime Vehicle Detection Methods Based on Headlight Feature: A Review. *IAENG Int. J. Comput. Sci.* **2022**, *49*, 79–93.
5. Sun, B.; Li, W.; Liu, H.; Yan, J.; Gao, S.; Feng, P. Obstacle detection of intelligent vehicle based on fusion of lidar and machine vision. *Eng. Lett.* **2021**, *29*, 722–730.
6. Hu, Y.; Liu, F.; Wei, Z.; Zhao, L. Vehicle Tracking of Information Fusion for Millimeter-wave Radar and Vision Sensor. *Zhongguo Jixie Gongcheng/China Mech. Eng.* **2021**, *32*, 2181–2188. (In Chinese)
7. Wang, X.; Xu, L.; Sun, H.; Xin, J.; Zheng, N. On-Road Vehicle Detection and Tracking Using MMW Radar and Monovision Fusion. *IEEE Trans. Intell. Transp. Syst.* **2016**, *17*, 2075–2084. [[CrossRef](#)]
8. Han, S.; Wang, X.; Xu, L.; Sun, H.; Zheng, N. Frontal object perception for Intelligent Vehicles based on radar and camera fusion. In Proceedings of the 35th Chinese Control Conference (CCC), Chengdu, China, 27–29 July 2016; Volume 2016, pp. 4003–4008. [[CrossRef](#)]

9. Jiang, Q.; Zhang, L.; Meng, D. Target detection algorithm based on MMW radar and camera fusion. In Proceedings of the 2019 IEEE Intelligent Transportation Systems Conference, ITSC 2019, Auckland, New Zealand, 27–30 October 2019; pp. 1–6.
10. Li, L. An improved image processing algorithms of sobel edge detection and functional verification. *Bol. Tec./Tech. Bull.* **2017**, *55*, 359–366.
11. Arslan, E.; Orman, Z.; Arik, S. Feature extraction of objects in moving images and implementation of the purification algorithm on the CNN Universal Machine. *Int. J. Circuits Syst. Signal Process.* **2011**, *5*, 488–495.
12. Rahman, M.A.; Amin, M.F.I.; Hamada, M. Edge detection technique by histogram processing with canny edge detector. In Proceedings of the 3rd IEEE International Conference on Knowledge Innovation and Invention 2020, ICKII 2020, Kaohsiung, Taiwan, 21–23 August 2020; pp. 128–131.
13. Zou, Y.; Zhang, J.; Upadhyay, M.; Sun, S.; Jiang, T. Automatic Image Thresholding Based on Shannon Entropy Difference and Dynamic Synergic Entropy. *IEEE Access* **2020**, *8*, 171218–171239. [[CrossRef](#)]
14. Manoharan, K.; Daniel, P. Image processing-based framework for continuous lane recognition in mountainous roads for driver assistance system. *J. Electron. Imaging* **2017**, *26*, 063011. [[CrossRef](#)]
15. Zhou, S.; Zhi, X.; Lin, F.; Ning, H.; Jiang, L.; Wu, N. Research on Vehicle Detection and Tracking Algorithm Based on Onboard Video Images. *Tongji Daxue Xuebao/J. Tongji Univ.* **2019**, *47*, 191–198. (In Chinese)
16. Zhao, W.; Li, B.; Shan, Y.; Xu, H. Vehicle Detection and Tracking Based on Fusion of Millimeter Wave Radar and Monocular Vision. *Wuhan Daxue Xuebao (Xinxi Kexue Ban)/Geomat. Inf. Sci. Wuhan Univ.* **2019**, *44*, 1832–1840. (In Chinese)