Developing an On-Road Object Detection System Using Monovision and Radar Fusion [†]

Ya-Wen Hsu [1] , **Yi-Horng Lai [2],*** , **Kai-Quan Zhong [1]** , **Tang-Kai Yin [3]** and **Jau-Woei Perng [1]**

[1] Department of Mechanical and Electro-Mechanical Engineering, National Sun Yat-sen University, Kaohsiung 80424, Taiwan; d043020006@student.nsysu.edu.tw (Y.-W.H.); wdf4200@gmail.com (K.-Q.Z.); jwperng@faculty.nsysu.edu.tw (J.-W.P.)

[2] School of Mechanical and Electrical Engineering, Xiamen University Tan Kah Kee College, Zhangzhou 363105, China

[3] Department of Computer Science and Information Engineering, National University of Kaohsiung, Kaohsiung 81148, Taiwan; tkyin@nuk.edu.tw

* Correspondence: lai81.tom@gmail.com

† This paper is an extended version of our paper published in 2018 IEEE Image and Vision Computing New Zealand Conference, Auckland, New Zealand, 19–21 November 2018.

check for updates

Abstract: In this study, a millimeter-wave (MMW) radar and an onboard camera are used to develop a sensor fusion algorithm for a forward collision warning system. This study proposed integrating an MMW radar and camera to compensate for the deficiencies caused by relying on a single sensor and to improve frontal object detection rates. Density-based spatial clustering of applications with noise and particle filter algorithms are used in the radar-based object detection system to remove non-object noise and track the target object. Meanwhile, the two-stage vision recognition system can detect and recognize the objects in front of a vehicle. The detected objects include pedestrians, motorcycles, and cars. The spatial alignment uses a radial basis function neural network to learn the conversion relationship between the distance information of the MMW radar and the coordinate information in the image. Then a neural network is utilized for object matching. The sensor with a higher confidence index is selected as the system output. Finally, three kinds of scenario conditions (daytime, nighttime, and rainy-day) were designed to test the performance of the proposed method. The detection rates and the false alarm rates of proposed system were approximately 90.5% and 0.6%, respectively.

Keywords: particle filter; histogram of gradient; sensor fusion; neural network; support vector machine; object recognition

1. Introduction

In recent years, the development of advanced driving assist systems (ADAS) has attracted a large amount of research and funds from major car factories and universities. The key issues of ADAS include on road object detection, anti-collision technology, park assist system, etc. Three kinds of sensors (i.e., radar, Lidar, and camera) are widely adopted for object detection in front of vehicles [1–5]. Since there are limitations of single sensors, multi-sensor fusion technology can be used to compensate for the disadvantages of each single sensor [6,7].

In reference [8], by using background subtraction and a Haar wavelet translation, the foreground image was transformed into a second-order feature space. Then, based on the concept of a histogram of original gradients (HOG), horizontal and vertical high-frequency components were obtained. In a hierarchical SVM classifier architecture, the proposed system can classify pedestrians, automobiles, and two wheeled vehicles effectively. Yang et al. [9] used an optical flow method to calculate the motion

vectors of the objects. Subsequently, the focus of expansion (FOE) of each object was found by using voting. By using the concept of a hierarchical decision tree, false alarms for detection (e.g., shadows or ground marking lines, etc.) can be avoided. Finally, the collision time was calculated by using the motion vectors of the objects.

Millimeter-wave (MMW) radars detect objects by transmitting electromagnetic waves onto the objects and analyzing the reflected waves that are not affected by light and weather. These radars can measure the relative distances and speeds of objects. However, millimeter-wave radars are susceptible to noise and environmental interference. To address the issues related to the microwave radar noise, Park et al. [10] proposed applying a statistical model to the radar using hybrid particle filter to track the preceding vehicle.

The laser range finder is an electronic measuring instrument that uses a laser to accurately measure the distance to the target, which exhibits the advantages of high measurement accuracy and good stability. Nashashibi et al. [4] developed a method to detect, track, and classify multiple vehicles by means of a laser range finder mounted on a vehicle. The classification was based on different criteria: sensor specifications, geometric configuration, occlusion reasoning, and tracking information. The system was tested in highways and urban centers with three different laser range finders.

In contrast with range finder sensors, camera sensors are not only cost-effective but can also provide other useful information. Many novel vision-based object detection algorithms for the front of vehicles have been proposed in the past decade. Vehicle detection and vehicle distance estimation systems were proposed in reference [11]. By using the histogram of an oriented gradient (HOG) feature and support vector machine (SVM) classifier, the authors can segment the road area and identify the shadow area under the vehicle in which to detect the vehicle position. Guo et al. [12] used a two-stage detection algorithm for pedestrian detection. First, the candidate regions were decided from foreground image, then the edge features of object were identified in the second stage. The experiment result verified the accuracy of the proposed method.

Despite the advantages exhibited by all sensors, they have limitations that affect their object detection abilities. For instance, cameras are susceptible to light and environmental factors, and the radar stability is affected by the relative speed and surrounding environment. Hence, a sensor fusion mechanism is developed to compensate for the deficiencies of relying on a single sensor.

The series type fusion architecture based on laser and vision sensors was addressed in reference [13]. The proposed system can quickly find the region of interesting objects without a huge amount of computation time. The other advantage was that after the verification and comparison of each sensor, the overall false alarm rate was reduced. Wang et al. [14] proposed a system scheme for on-road obstacle detection by fusing an MMW radar and a monocular vision sensor. An experimental method to investigate the radar-vision point alignment was proposed. In addition, a region searching method for potential target detection was proposed to reduce image processing time. Wang et al. [15] proposed a tandem sensor fusion of series connection architecture that uses MMW radar to obtain the candidate position of the detected object. The position coordinates are converted into image coordinates that considered as regions of interest to reduce the number of window searches. Then the image is used to recognize and track the vehicle in candidate areas. A Kalman filter is used to compare the tracking trajectory of the radar and camera to improve the vehicle detection rate and reduce the false positive rate.

In the aforementioned references, using a single sensor to detect objects has significantly reduced the detection system cost; however, the system stability is still a challenge when considering special weather conditions. The main purpose of using series architecture in sensor fusion is to rapidly determine the candidate area via radar or Lidar and accelerate the image search process. Another advantage of using a second layer sensor is to reduce noise interference after verification and comparison. However, the entire tandem architecture system will fail when one of the sensors fails.

This paper extends our earlier vision based research work [16] and proposes a set of MMW radar and camera fusion strategies based on a parallel architecture that can compensate for the failure of a

single sensor and enhance the system detection rate using the complementary characteristics of the sensors. The radar subsystem provides noise filtering, tracking, and credibility analysis. The two-stage vision detection subsystem can rapidly identify the candidate area form image. The fusion strategy of parallel architecture systems depends on the confidence index of each sensor. Three kinds of scenario conditions (daytime, nighttime, and rainy-day) are implemented in an urban environment to verify the proposed system.

The contributions of this study include the following:

1. In order to solve the shortcomings of each single sensor, by using sensor fusion technology, we integrated the two sensor systems and improved the reliability of the systems.
2. For the fusion architecture of series type, any single sensor failure causes whole system failure. The proposed parallel architecture system depends on the confidence index of each sensor. The system can compensate for each other's sensors and avoid the limitations of series fusion architecture.
3. Three kinds of scenario conditions (daytime, nighttime, and rainy-day) were implemented in an urban environment to verify the proposed system's viability. The experiment results can provide the baseline of comparison for future research.

2. System Architecture

This study proposed a sensor fusion technology integrating MMW radar and camera for front object detection. The proposed system consists of three subsystems, including a radar-based detection system, vision-based recognition system, and sensor fusion system.

The image captured by the camera can easily be affected by lighting and weather conditions. Furthermore, the estimated distance of the front object derived from the camera image has a low precision. A sufficiently large velocity relative to the front object is necessary for the MMW radar to stably detect it. Accordingly, these two sensor subsystems were combined in a parallel connection to compensate for the limitations of each sensor and improve the robustness of the detection system. The overall architecture of the proposed detection and recognition system is shown in Figure 1.
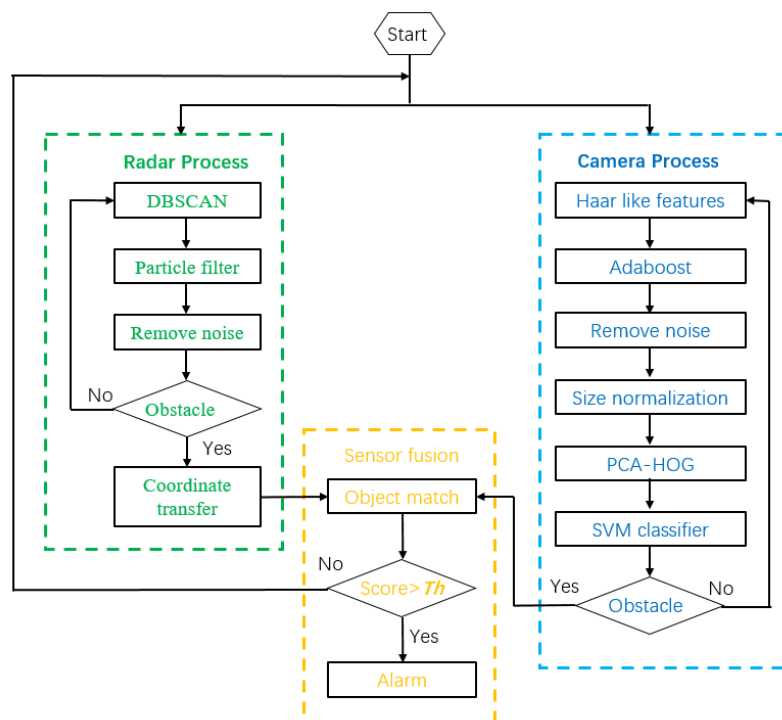

Figure 1. Overall architecture of an on-road obstacle detection system.

A clustering algorithm and particle filter were applied to the MMW radar data to achieve noise removing and multi-object tracking. Then the object detected by the coordinate system of radar sensor was converted into an image coordinate. On the other hand, two-stage classifiers were implemented for the foreground segmentation and object recognition for the image data, respectively, then the object information could be obtained. Finally, a radial basis function neural network (RBFNN) was used to fuse the detected object information from the MMW radar and camera.

3. Radar-Based Object Detection

A 24 GHz short-range radar was adopted for front-end environment detection and a multi-object tracking method based on radar was proposed. This method can facilitate tracking multiple object simultaneously and removing noises, which were considered as non-real objects. The flow chart of the proposed radar-based detection subsystem is shown in Figure 2. First, the radar data were divided into different clusters using a clustering algorithm. The particle filter is then used for signal filtering and target tracking. Two kinds of probability scores will be evaluated in the particle filter process. The convergence of the particle swarm can reflect the quality of the tracking. For the stable tracking objects, the particles around the object have a higher weighting in the importance sampling step. Furthermore, these particles have a higher probability of survival in the resampling step. We define the range probability (P_r) as the survival probability of the particles within a radius of 1 m around the object to evaluate the quality of the tracking. On the other hand, the diversity of the particle swarm can cover of all the states of the object. We defined the available probability (P_a) as the survival probability of the particles after the resampling step. During the tracking process, in line with the value of P_a, the system adjusts the particle percentage of resampling to ensure the diversity of the particle swarm. In addition, the confidence index of the target object was derived from the range probability and probability of survival. This confidence index determines the credibility of the actual object. The relative velocity and distance between the vehicle and front object were provided by this subsystem.
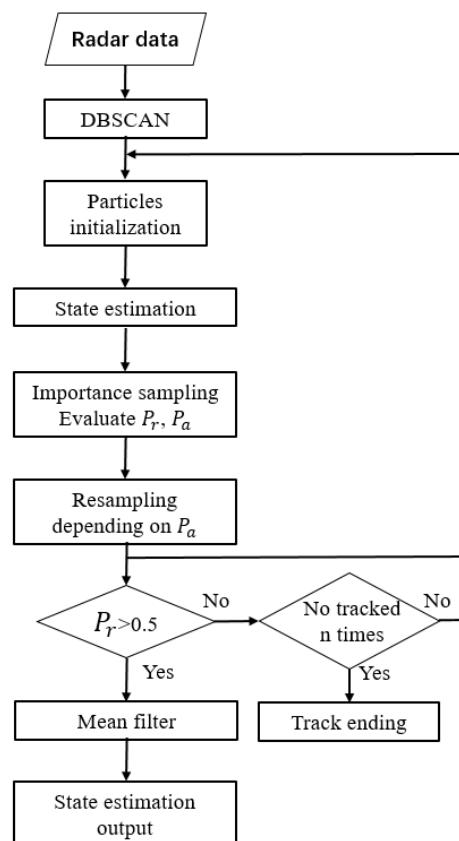
Figure 2. Flow chart of a radar-based object detection subsystem.

3.1. Radar Data Pre-Processing

The MMW radar signals are electromagnetic waves. Both reflection and refraction will occur when the electromagnetic waves occur on the medium. In addition to the reflected wave from the medium itself, some noise signals of non-real objects are also prone to appear. The relationship between relative distance and echo intensity information was statistically analyzed using a vast amount of data collected during experiments. The statistical results are shown in Figure 3. The statistical results of the signal distribution indicate that both real objects and noise show respective concentrations, and only a small part of the distribution of both overlaps. Accordingly, a noise filtering operation was performed. As shown in Figure 3a, after the signal on the left side of red curve was filtered, the subsequent target tracking and particle filter algorithm were performed. Density-based spatial clustering of applications with noise (DBSCAN) algorithm [17] was used to cluster the radar data, and the number of possible front objects was estimated.
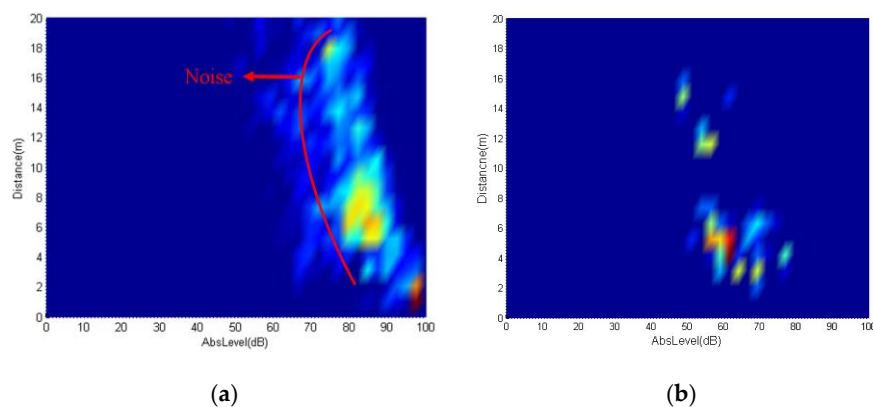

(**a**)　　　　　　　　　　　　　　(**b**)

Figure 3. Statistical results of the radar signal: (**a**) real object and (**b**) non-real object.

3.2. Particle Filter

A particle filter [18] is widely used in many fields, including object tracking, signal processing, and automatic control. In this study, particle filtering was used to filter the radar signal and track the objects in front of a vehicle. The particle filter algorithm uses a finite number of particles to represent the posterior probability of some stochastic process with partial observations. Each particle has the respective weight values that represent the probability of the particle being sampled from the probability density function. The procedure to implement a particle filter algorithm in this study was roughly divided into four steps as follows:

3.2.1. Particle Initialization

To cover all the potential object positions, n pieces of particles were randomly distributed within the radar detection area. Each particle represents a potential position of a real object, where the weight of the particle indicates the probability that the object is at this location.

3.2.2. State Prediction

The state of the object changes over time. Discrete time was used to calculate the object state, and the state of the particle at next moment was predicted by the state and motion model at time $k-1$. Then the prior probability $P(x_k|x_{k-1})$ was obtained. The equation used to predict the object state is expressed as follows [19]:

$$X_k = FX_{k-1} + GW_k = \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{k-1} \\ y_{k-1} \\ \dot{x}_{k-1} \\ \dot{y}_{k-1} \end{bmatrix} + \begin{bmatrix} \frac{T^2}{2} & 0 \\ 0 & \frac{T^2}{2} \\ T & 0 \\ 0 & T \end{bmatrix} \begin{bmatrix} W_x \\ W_y \end{bmatrix} \tag{1}$$

with

$$F = \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \; G = \begin{bmatrix} \frac{T^2}{2} & 0 \\ 0 & \frac{T^2}{2} \\ T & 0 \\ 0 & T \end{bmatrix} \tag{2}$$

where T is the sampling time of the radar sensor, $X_k = \begin{bmatrix} x_k & y_k & \dot{x}_k & \dot{y}_k \end{bmatrix}^T$ denotes the state vector, and x_k and x_{k-1} denote the relative lateral distances between the target object and the sensor at the current time and the previous moment, respectively. y_k and y_{k-1} are the relative longitudinal distances between the target object and the sensor at the current time and the previous moment, respectively. $\dot{x}_k$ and $\dot{y}_k$ represent the lateral and longitudinal relative speeds of the target and the sensor, respectively. W_k is zero-mean Gaussian white noise.

3.2.3. Importance Sampling

This step is based on the concept of a Bayesian filter. The particles that are obtained during the state prediction stage and the information obtained from MMW radar are used to estimate the target position. The Bayesian theorem is used to update the prior probability then obtain the posterior probability. In this step, each particle is assigned a weight. Based on the assumption that the radar measurement area is $M \times N$ blocks, each block unit is 1 m^2. The measurement model of the radar sensor is expressed by Equation (3),

$$z_k^{(i,j)} = h_k^{(i,j)}(x_k) + v_k^{(i,j)} \tag{3}$$

where $v_k^{(i,j)}$ is the measured noise in (i, j) block and its Gaussian white noise with the means equal to 0 and the variance σ^2, while $h_k^{(i,j)}(x_k)$ is the signal strength of the object in the (i, j) block and its point spread function [20] is expressed as follows:

$$h_k^{(i,j)}(x_k) = \frac{\Delta_x \Delta_y I_k}{2\pi\Sigma^2} \cdot \exp\left(\frac{(i\Delta_x - x_k)^2 + \left(j\Delta_y - y_k\right)^2}{2\Sigma^2} \right) \tag{4}$$

where Δ_x and Δ_y are the block sizes, I_k is echo strength of the MMW radar, Σ is the blurring degree of the sensor, and the weight value of the particle can be obtained by the following equation:

$$w_k^{\tilde{i}} = \exp\left(\frac{h_k^{(i,j)}(x_k^{\tilde{i}})\left(h_k^{(i,j)}(x_k^{\tilde{i}}) - 2h_k^{(i,j)}\right)}{2\sigma^2} \right). \tag{5}$$

The weight of each particle in the space region is normalized. The normalization method is based on dividing the weight of each particle by the sum of all particle weights, as shown by Equation (6):

$$\hat{w}_k^{\tilde{i}} = \frac{w_k^{\tilde{i}}}{\sum\limits_{i=1}^{n} w_k^{\tilde{i}}}. \tag{6}$$

After the weight of each particle is obtained, the relative position of the object detected by the MMW radar can be estimated. The expected value of the target estimation is expressed as follows:

$$E\left(x_k \middle| y_k\right) = \sum_{i=1}^{n} w_k^{\tilde{}i} f\left(x_k^{\tilde{}i}\right).$$

(7)

3.2.4. Resampling

The method of estimating according to the weight of each particle is referred to as the sequential importance sampling (SIS) particle filter [18]. However, this method involves particle degradation, leading to insignificant weight values of most particles after several iterative operations. This triggers the system to perform unnecessary calculations on these particles. Thus, the real target position may not be covered by the remaining particles. The resampling method was used to address this issue. In each iteration process, the particles with smaller weight values were discarded and replaced by particles with larger weight values. After resampling, the weight values of all particles was set at $\frac{1}{n}$, then the next iteration was performed with new particles. The expected value of the target estimation is expressed as follows:

$$E\left(x_k \middle| y_k\right) = \sum_{i=1}^{n} \frac{1}{n} f\left(x_k^{\tilde{}i}\right).$$

(8)

3.3. Experimental Verification

A lot of object information was lost while the MMW radar information was processed by internal algorithms. Therefore, the original unprocessed data was obtained from the MMW radar in this study. The proposed particle filter algorithm was used to track the front object and address the issue of losing too much information.

To verify the feasibility of the algorithm proposed in this study, a laser range finder with high precision was used. The measurement error of the adopted lase finder was ±10 mm to record the center position of the frontal object. The experimental equipment installed to verify the radar tracking system is shown in Figure 4. Three verification conditions were set to avoid dark objects and lack of relative speeds, which can lead to losing laser range finder and radar information, as follows: metal and light-colored moving objects, a relative velocity of ±15 km/h or more, and objects moving from far away to nearby.
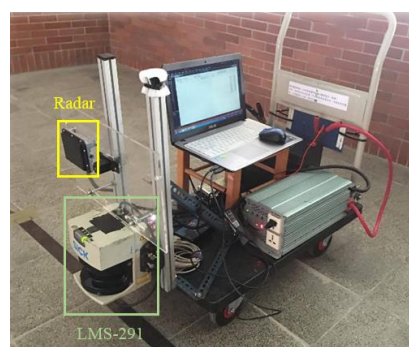

Figure 4. Equipment setup for radar tracking system verification.

The position of the object measured by the laser range finder is considered as the ground truth, which is illustrated by the blue line seen in Figure 5. The red line represents the tracking result obtained by the proposed particle filtering algorithm. The result of the internal algorithm of the radar sensor is illustrated by the green line. An offset between the detected and actual positions of the object may be observed owing to the characteristics of the radar sensor.
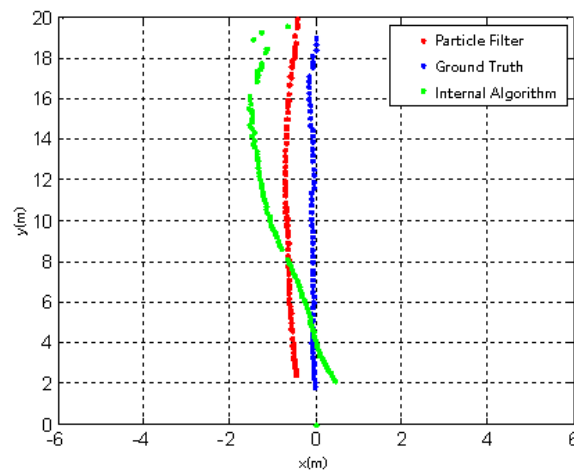
Figure 5. Estimated trajectories of moving objects using different methods.

The error and standard deviation of our proposed particle filter tracking algorithm and the internal algorithm of the radar sensor were compared to the ground truth to verify the tracking results. The error is defined as the absolute value of the estimated position from the algorithm and the ground truth. The average error is the sum of the errors divided by the number of times of detections. As shown in Table 1, the proposed algorithm had better performance considering the average error, the maximum error, and the standard deviation of error of the longitudinal or lateral direction. In addition, the number of times the proposed algorithm effectively detected objects was also greater than that obtained by the sensor internal algorithm.

Table 1. Error of each tracking methods (unit: centimeter).

Method	Average Error		Standard Deviation		Maximum Error	
	Lateral	Longitudinal	Lateral	Longitudinal	Lateral	Longitudinal
The proposed particle filter tracking algorithm	44.48	32.32	18.53	22.35	88.86	99.96
Internal algorithm of the radar sensor	52.66	33.66	42.08	26.47	169.2	116.8

4. Vision-Based Object Recognition

The two-stage vision-based object recognition system was similar to in our earlier work [16]. In the first stage, the Haar-like features algorithm was used to identify the candidate regions of object from foreground segmentation. The second stage is responsible for object recognition. Three kinds of objects (i.e., pedestrians, motorcycles, and cars) can be identified by SVM classifiers. The scheme of the two-stage vision-based object recognition process in shown in Figure 6. The object recognition results are shown in Figure 7.

The distance estimation of image object can be determined by using the polynomial model as expressed in Equation (9):

$$f(y_{im}) = g_0 y_{im}^5 + g_1 y_{im}^4 + g_2 y_{im}^3 + g_3 y_{im}^2 + g_4 y_{im} + g_5 \tag{9}$$

where $f(y_{im})$ is the estimation of distance, while y_{im} denotes the object coordinates v of the image.
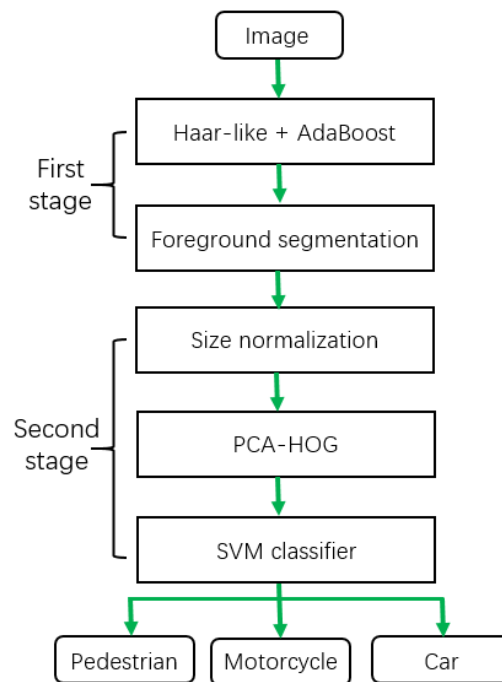
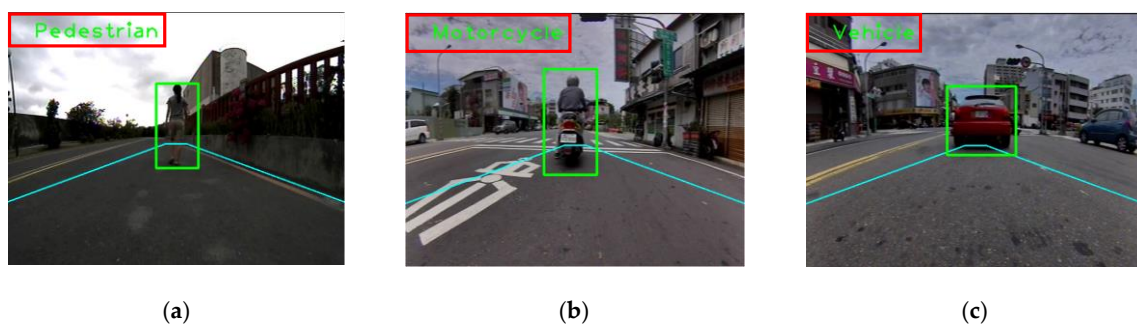**Figure 6.** Vision-based object recognition subsystem.

Figure 7. Classification results of vision-based recognition: (**a**) pedestrian, (**b**) motorcycle, and (**c**) car.

5. Sensors Fusion and Decision Mechanism

A single sensor system can operate independently; however, a parallel architecture was adopted in this study to fuse two different sensors. The main purpose of this is to improve the detection rate that can be achieved by a single sensor. The sensor fusion was divided into three parts. First, the two-dimensional coordinate information of the MMW radar was converted into the coordinate of the image. Afterwards, the information obtained by the two sensors was integrated into the same coordinate system. Next, the object information needed to be matched to determine whether the same object information had been obtained by both the MMW radar and camera, and to integrate the detection results of the two systems. Finally, the trusted sensor was determined based on the confidence index of the sensor.

5.1. Coordinate Transformation

The supervised learning algorithms was used to learn the relationship between the MMW radar coordinate and image coordinate system. Before the coordinate transformation, the radar coordinates (x, y) and image coordinate (u, v) needed to be recorded synchronously to be considered as training samples for offline learning. An MMW radar uses electromagnetic waves as a medium, and it exhibits better reflective property to metal objects. Hence, a triangular metal reflector was used as a target object to gather data obtained from the radar and the camera, as shown in Figure 8. A metal reflector

was randomly placed in a straight lane at a distance which ranged from 1 m to 12 m in front of the experimental vehicle, and a total of 280 training samples were established.
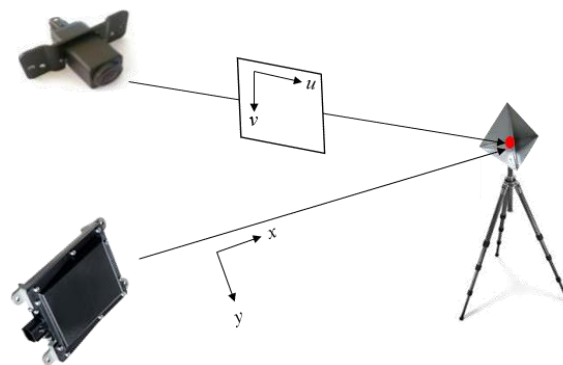

Figure 8. Coordinate transformation of data from the radar and camera.

The camera was installed at an angle parallel to the horizon. When the target object moved from far away to nearby, the position of its center point slightly changed near the center point of the image in the vertical direction. Thus, the variation in the image v-direction coordinate was not obvious. Therefore, the fusion system primarily enabled the neural network to learn the relationship between the MMW radar coordinate (x, y) and the image coordinate (u, v).

From the collected training samples, the longitudinal and lateral distances from the radar were considered as the input of the RBFNN, and the corresponding u coordinate of horizontal direction in the image was considered as an output. This network architecture allows for obtaining the coordinate conversion relationship between these two sensors. The network architecture is shown in Figure 9.
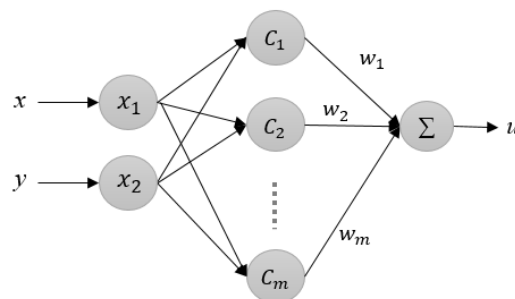

Figure 9. Diagram of radial basis function neural network (RBFNN) architecture.

5.2. Object Match

The MMW radar detection and image recognition systems operate independently, and the two systems obtain information about the detected objects, respectively. To fuse the information of the two systems, the object information must be matched first to determine whether the same object information has been detected by the two sensors. Coordinates shown in the same image may correspond to several different radar coordinate information, as illustrated by the green points shown in Figure 10. In addition, the distance estimated from the image coordinates may be inaccurate owing to the bumpy road surfaces that can cause the vehicle to shake; thus, it is difficult to match the object information and effectively determine whether the same object is detected.

Another RBFNN is used to match the object information and determine whether the same objects are detected by the two sensors. Six factors were entered as the network inputs, which affect the object match, including image coordinate u, object width, object height, object distance estimated from image, object distance measured by the radar, and the u coordinate converted from the radar to the image. Either "match" or "non-match" were obtained as the network output.
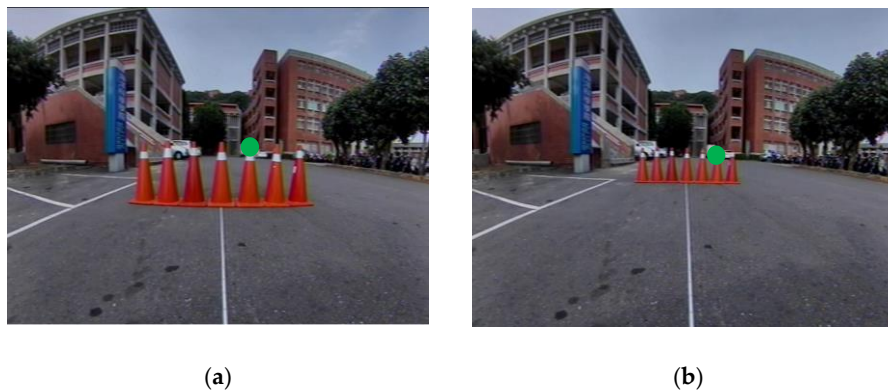
(**a**)　　　　　　　　　　　　　　　　　　　(**b**)

Figure 10. Diagram of same image coordinates correspond to different radar coordinates (**a**) Radar detection distance is 2.2 m and (**b**) Radar detection distance is 4 m.

5.3. Decision Strategy

If a single sensor in the sensor fusion of cascade architecture fails, then the entire system will inevitably fail. Meanwhile, the sensor fusion of parallel architecture determines which sensor should be trusted based on the decision mechanism. Although one of the sensors might not detect an object or gives a false alarm, if the other sensor correctly detects the object, then the confidence index of each sensor can be calculated via a scoring mechanism, and a credible subsystem can be determined based on the confidence index.

The confidence index of the radar subsystem was calculated as follows:

$$Score_R = P_r + P_a + A_{rn} \times \eta_r \tag{10}$$

where A_{rn} is the number of times the object tracked by particle filter. η_r is a constant.

The confidence index of the image subsystem was calculated as follows:

$$Score_I = S_d + A_{in} \times \eta_i + \lambda \tag{11}$$

where S_d denotes the distance from the input data point to the SVM hyperplane, A_{in} is the number of times the object tracked in image subsystem, and η_r and λ are constants.

The confidence index of the sensor fusion system was expressed as follows:

$$Score = Score_R + Score_I. \tag{12}$$

When the confidence index *Score* is greater than the set threshold *Th*, the reliability of the system is extremely high, and the output result obtained by the system represents the real situation. If the confidence index of each subsystem is greater than the threshold *Th*, then the subsystem with the highest score is responsible for the entire system decision making process.

6. Experiments

6.1. Experimental Platform and Scenarios

Three kinds of scenario conditions (daytime, nighttime, and rainy-day) were implemented to verify the proposed system. All the scenarios were carried out on urban roads. The MMW radar and camera were mounted on the front bumper of the experimental car, as shown in Figure 11.

Considering the effect of pavement puddles and shadow environment, the daytime scenarios included direct sunlight, pavement puddles, and shadow environments, as shown in Figure 12.

In the nighttime experiment, the scenarios included flashing brake lights of front vehicles, headlight reflections, and poor lighting environments, as shown in Figure 13.

Figure 11. Experimental car and sensors setup.

| (**a**) | (**b**) | (**c**) |

Figure 12. Daytime scenarios: (**a**) sunlight, (**b**) pavement puddle, and (**c**) shadow.

| (**a**) | (**b**) | (**c**) |

Figure 13. Nighttime scenarios: (**a**) brake light, (**b**) headlight reflection, and (**c**) poor lighting.

In order to reproduce the actual road conditions, we designed a rainy-day scenario too. As the sensors are mounted on the front bumper, the raindrops often adhered to the camera lens during the rainy day experiment, as shown in Figure 14.
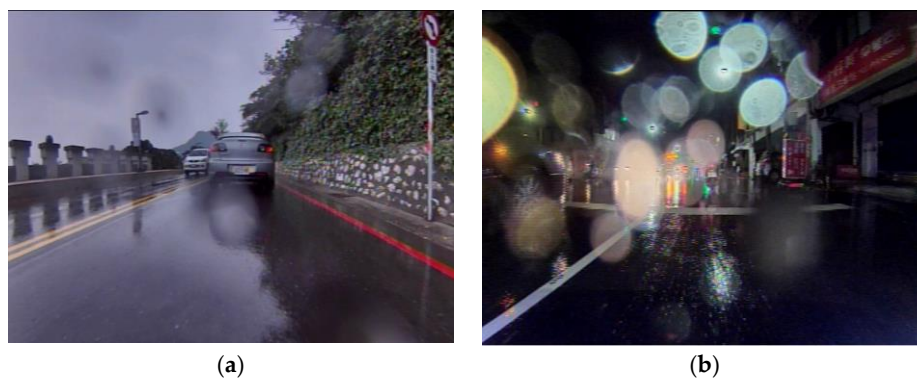

| (**a**) | (**b**) |

Figure 14. Rainy day scenarios: (**a**) daytime and (**b**) nighttime.

6.2. Radar-Based Detection Subsystem

The radar detection subsystem uses MMW radar to perceive the environment ahead. The proposed multi-object tracking algorithm with a particle filter can effectively track the objects in front and remove

non-object noise. The radar subsystem experiments tested three different categories of objects under different conditions. The detection results are shown as green circles in Figures 15 and 16. The tests primarily involved a single target in a lane. If there were multiple targets, the alert was reported for closest target to the experimental vehicle. Other targets continued to be tracked.
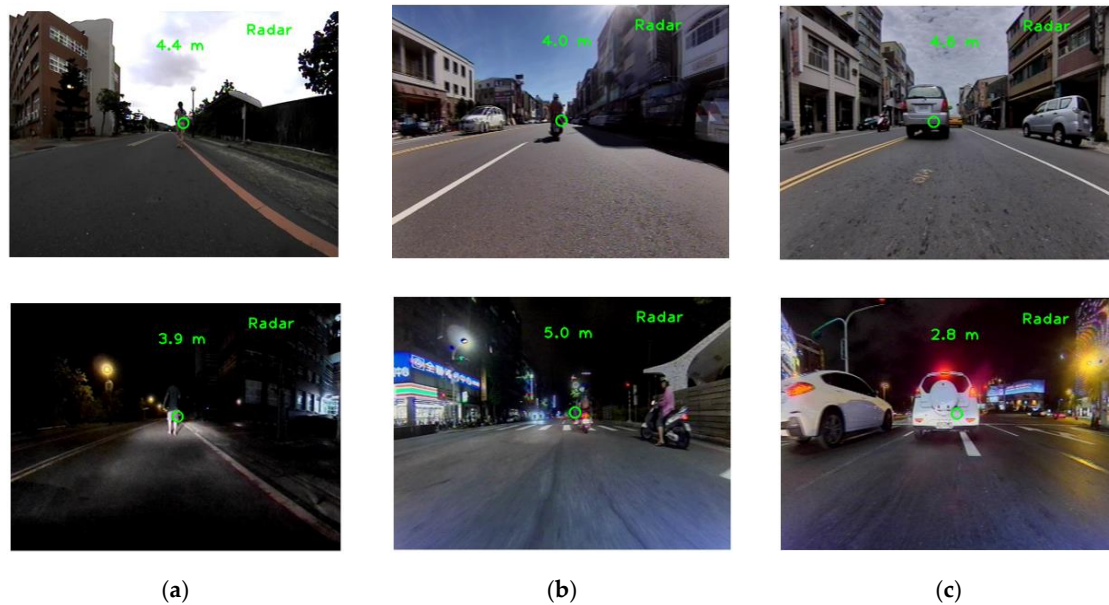

| (a) | (b) | (c) |

Figure 15. Detection results of radar subsystem, upper row: daytime, lower row: nighttime (**a**) a pedestrian, (**b**) motorcycle, and (**c**) car.
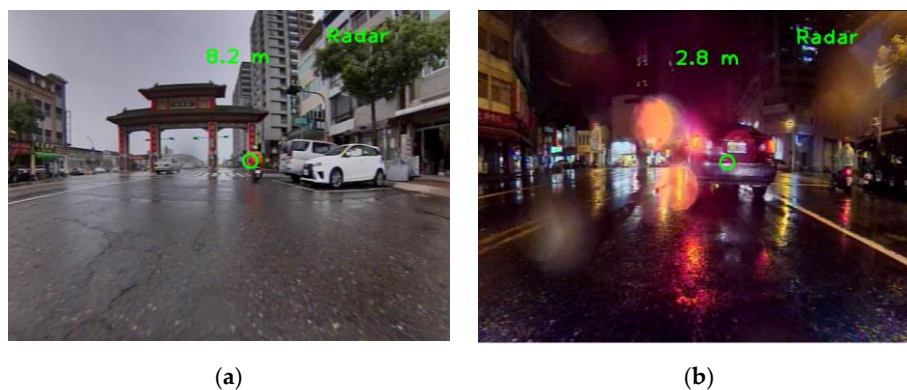

| (a) | (b) |

Figure 16. Detection results of radar subsystem for a rainy day, (**a**) motorcycle and (**b**) car.

A detection rate exceeding 60% was maintained by the radar detection system during daytime, nighttime, and rainy days. The experimental tests performed under different weather conditions verified that the radar detection system is not affected by weather conditions. The experimental results are listed in Table 2.

Table 2. Detection results of radar systems under different weather conditions.

Condition	Total Frame	Correct Detection	Misinformation	Misjudgment	False Alarm Rate	Detection Rate
Daytime	17,346	11,254	27	6065	0.2%	64.9%
Nighttime	7022	4338	0	2684	0%	61.8%
Rain day	11,193	8135	0	3058	0%	72.6%
Total	35,561	23,727	27	11,807	0.01%	67.0%

6.3. Vision Recognition

The advantages of two-stage vision-based object recognition system are as follows: By using Haar-like features, the first-stage classifier can detect efficiently candidate areas. Unfortunately, the Haar-like algorithm suffers from higher false positive rates (see the purple rectangles in Figure 17). Therefore, the second-stage PCA-HOG algorithm classifier was utilized to compensate for the higher false positive rates of the first-stage result.

Figure 17. Error detection of the Haar-like algorithm.

The detection results of the vision-based object recognition subsystem are shown as yellow rectangles in Figure 18. The results of the rainy-day experiment are shown as green rectangles in Figure 19.
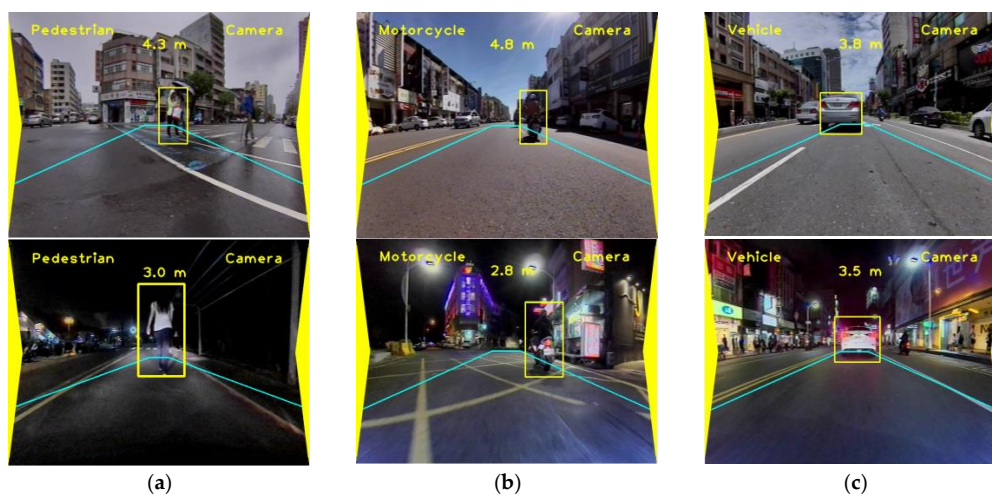

Figure 18. Detection results of a vision-based subsystem, upper row: daytime, lower row: nighttime (**a**) a pedestrian, (**b**) motorcycle, and (**c**) car.

All the experiments performed under different weather conditions involved three classifications of objects: pedestrians, motorcycles, and cars. The detection results of vision-based systems are listed in Table 3.

Table 3. Detection results of vision-based systems under the different weather conditions.

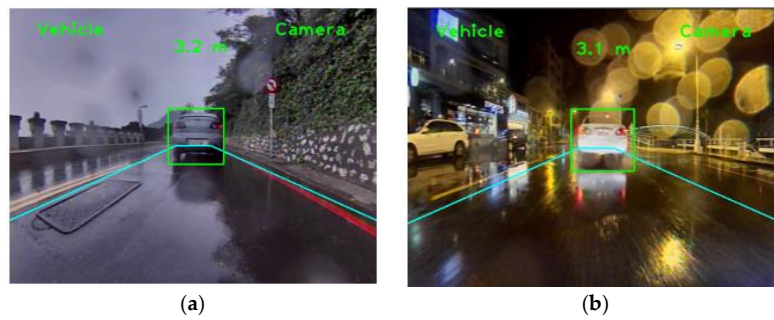| Condition	Total Frame	Correct Detection	Misinformation	Misjudgment	False Alarm Rate	Detection Rate
Daytime	17,392	14,909	46	2437	0.3%	85.7%
Nighttime	7043	4915	21	2107	0.3%	69.8%
Rain day	11,193	4335	141	6717	1.3%	38.7%
Total	35,628	24,159	208	11,261	0.5%	67.8%

Figure 19. Detection results by vision-based subsystem for a rainy day, (**a**) daytime and (**b**) nighttime.

Due to the high sensitivity to light sources, the performance of camera sensor depends on the condition of light sources. For example, suffering in an insufficient light source, the vision-based systems cannot extract completely the features of objects at night. On the other hand, in rainy weather experiments, the raindrops adhering to the camera lens block the object in front of the vehicle. Thus, the system cannot effectively identify the information of the target, leading to the failure of the image subsystem. Therefore, the worst detection rates are achieved at night and on rainy days.

6.4. Sensor Fusion System

This system integrates MMW radar and camera information and improves the scene when one of the detection systems fails by using the sensor fusion of parallel architecture. The system presents complementary characters. For example, as shown in Figure 20, the radar did not detect the front vehicle when the relative speed of the radar and object was relatively small; thus, the camera was used to compensate for the radar failure. On the other hand, when the raindrops adhering to the camera lens blocked the scene, leading to image detection failure, the radar compensated for this situation, as shown in Figure 21.
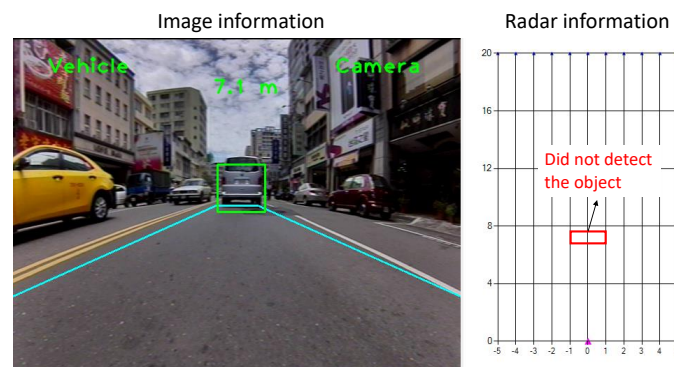

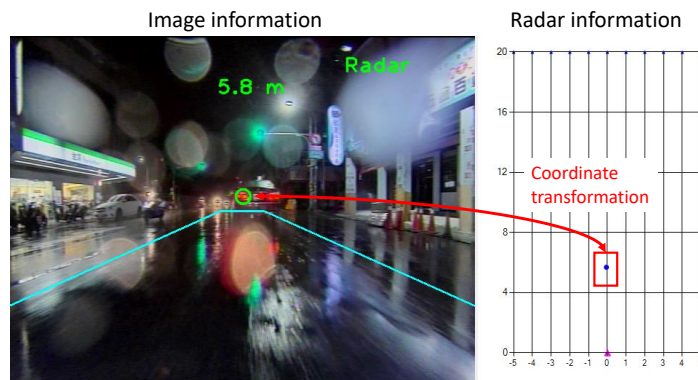**Figure 20.** Sensor fusion compensate radar failure.

Figure 21. Sensor fusion compensate for image failure.

In addition to compensating for single sensors failures, the system integrates the sensors' information when both the radar and camera detect objects simultaneously. The system relies on the coordinate transformation and object matching decision mechanism to determine whether the same objects are detected by the two sensors, as shown in Figure 22.
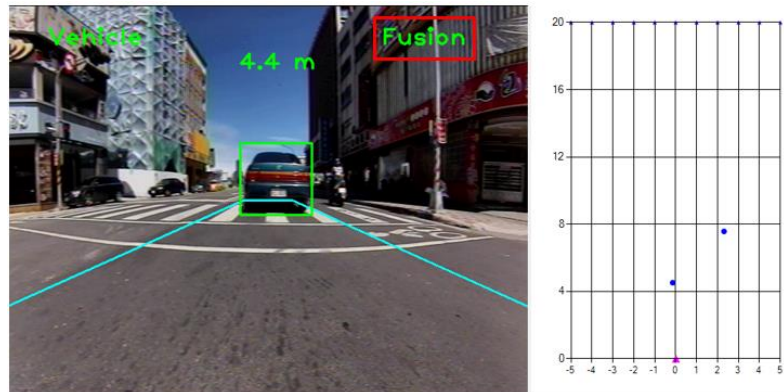

Figure 22. Detection result of sensor fusion system.

The parallel sensor fusion architecture proposed in this study exhibits the advantages of compensating for the disadvantages of relying on a single sensor. It improves the scene in case of subsystem failure and significantly increases the system detection rate and stability, as listed in Table 4. Regardless of the weather conditions, better detection rates were achieved by the sensor fusion system than those obtained when relying on a single subsystem.

Table 4. Detection results of each system under different weather conditions.

Condition	Sensor	Total Frame	Correct Detection	Misinformation	Misjudgment	False Alarm Rate	Detection Rate
Daytime	radar	17,392	11,254	27	6065	0.2%	64.7%
	image	17,392	14,909	46	2437	0.3%	85.7%
	fusion	17,392	16,414	46	978	0.3%	94.3%
Nighttime	radar	7043	4338	0	2684	0%	61.6%
	image	7043	4915	21	2107	0.3%	69.8%
	fusion	7043	6450	21	593	0.3%	91.6%
Rain day	radar	11,193	8135	0	3058	0%	72.6%
	image	11,193	4335	141	6717	1.3%	38.7%
	fusion	11,193	9413	141	9985	1.3%	84.1%

Table 5 lists the detection results of each system for the three object categories under different weather conditions. The sensor fusion system can achieve a detection rate of more than 90%.

Table 5. Detection results of each system.

Sensor	Total Frame	Correct Detection	Misinformation	Misjudgment	False Alarm Rate	Detection Rate
Radar subsystem	35,628	23,727	27	11,807	0.01%	66.6%
Image subsystem	35,628	24,159	208	11,261	0.6%	67.8%
Sensor fusion system	35,628	32,277	208	3143	0.6%	90.5%

We also compared our results with existing related works. The comparison results are listed in Table 6.

Table 6. Comparison with existing related works.

Sensor Type	Object	Fusion Type	Environment	Time Cost	Hardware
Camera [9]	X	X	Daytime	50 ms	Intel i7 3.4 GHz
Camera [12]	Pedestrian	X	Daytime	66–100 ms	Core 2 2.66 GHz
Camera & Lidar [13]	Pedestrian	Series	Daytime	66 ms	Dual-core PC
Camera & Radar [15]	Car	Series	Daytime	16 ms	Intel i7 3.0 GHz
Camera & Radar (the proposed approach)	Car Motor Pedestrian	parallel	Daytime Nighttime Rainy-day	60 ms	Intel i7 2.6 GHz

7. Conclusions

Two types of sensors, an MMW radar and a camera were integrated in this study to develop a frontal object detection system based on sensor fusion using parallel architecture. A particle filter algorithm was employed by the radar detection subsystem to remove noise from non-objects while tracking objects at the same time, and converting the target information into the image coordinates using RBFNN. On the other hand, the image object could be identified as one of three main categories (pedestrians, motorcycles, and cars) by the two-stage vision-based recognition subsystem. The information obtained by the two subsystems was integrated. The sensor with higher credibility was selected as the system output result. Three kinds of experiments (daytime, nighttime, and rainy-days) were performed to verify the proposed system. The experiment results show the detection rates and the false alarm rates of proposed system were approximately 90.5% and 0.6%, respectively. These detection rates are better than those obtained by single sensor systems.

Author Contributions: Conceptualization, Y.-H.L. and J.-W.P.; Data curation, K.-Q.Z.; Formal analysis, J.-W.P. and T.-K.Y.; Investigation, K.-Q.Z.; Methodology, Y.-H.L.; Software, K.-Q.Z.; Supervision, J.-W.P. and T.-K.Y.; Validation, Y.-W.H.; Writing—original draft, Y.-W.H.; Writing—review & editing, Y.-H.L. All authors have read and agreed to the published version of the manuscript.

References

1. Mukhtar, A.; Xia, L.; Tang, T.B. Vehicle detection techniques for collision avoidance systems: A Review. *IEEE Trans. Intell. Transp. Syst.* **2015**, *16*, 2318–2338. [CrossRef]
2. Blanc, C.; Aufrere, R.; Malaterre, L.; Gallice, J.; Alizon, J. Obstacle detection and tracking by millimeter wave radar. In Proceedings of the FAC/EURON Symposium on Intelligent Autonomous Vehicles, Lisboa, Portugal, 5–7 July 2004; pp. 322–327.
3. Cho, H.J.; Tseng, M.T. A support vector machine approach to CMOS-based radar signal processing for vehicle classification and speed estimation. *Math. Comput. Model.* **2012**, *58*, 438–448. [CrossRef]
4. Nashashibi, F.; Bargeton, A. Laser-based vehicles tracking and classification using occlusion reasoning and confidence estimation. In Proceedings of the Intelligent Vehicles Symposium, Eindhoven, The Netherlands, 4–6 June 2008; pp. 847–852.
5. Natale, D.J.; Tutwiler, R.L.; Baran, M.S.; Durkin, J.R. Using full motion 3D Flash LIDAR video for target detection, segmentation, and tracking. In Proceedings of the IEEE Southwest Symposium on Image Analysis & Interpretation (SSIAI), Austin, TX, USA, 23–25 May 2010; pp. 21–24.
6. Kmiotek, P.; Ruichek, Y. Multisensor fusion based tracking of coalescing objects in urban environment for an autonomous vehicle navigation. In Proceedings of the IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems, Seoul, Korea, 20–22 August 2008; pp. 52–57.

7. Premebida, C.; Monteiro, G.; Nunes, U.; Peixoto, P. A Lidar and vision-based approach for pedestrian and vehicle detection and tracking. In Proceedings of the Intelligent Transportation Systems Conference, Seattle, WA, USA, 30 September–3 October 2007; pp. 1044–1049.

8. Liang, C.W.; Juang, C.F. Moving object classification using local shape and HOG features in wavelet-transformed space with hierarchical SVM classifiers. *Appl. Soft Comput.* **2015**, *28*, 483–497. [CrossRef]

9. Yang, M.T.; Zheng, J.Y. On-road collision warning based on multiple foe segmentation using a dashboard Camera. *IEEE Trans. Veh. Technol.* **2015**, *64*, 4947–4984. [CrossRef]

10. Park, S.; Hwang, J.P.; Kim, E.; Kang, H.J. Vehicle tracking using a microwave radar for situation awareness. *Control Eng. Pract.* **2010**, *18*, 383–395. [CrossRef]

11. Huang, D.Y.; Chen, C.H.; Chen, T.Y.; Hu, W.C.; Feng, K.W. Vehicle detection and inter-vehicle distance estimation using single-lens video camera on urban/suburb roads. *J. Vis. Commun. Image Represent.* **2017**, *46*, 250–259. [CrossRef]

12. Guo, L.; Ge, P.S.; Zhang, M.H.; Li, L.H.; Zhao, Y.B. Pedestrian detection for intelligent transportation systems combining AdaBoost algorithm and support vector machine. *Expert Syst. Appl.* **2012**, *39*, 4274–4286. [CrossRef]

13. Oliveira, L.; Nunes, U.; Peixoto, P.; Silva, M.; Moita, F. Semantic fusion of laser and vision in pedestrian detection. *Pattern Recognit.* **2010**, *43*, 3648–3659. [CrossRef]

14. Wang, T.; Zheng, N.; Xin, J.; Ma, Z. Integrating millimeter wave radar with a monocular vision sensor for on-road obstacle detection applications. *Sensors* **2011**, *11*, 8992–9008. [CrossRef] [PubMed]

15. 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]

16. Hsu, Y.W.; Zhong, K.Q.; Perng, J.W.; Yin, T.K.; Chen, C.Y. Developing an on-road obstacle detection system using monovision. In Proceedings of the International Conference on Image and Vision Computing New Zealand (IVCNZ), Auckland, New Zealand, 19–21 November 2018.

17. Ester, M.; Kriegel, H.; Sander, J.; Xu, X. A density-based algorithm for discovering clusters in large spatial databases with noise. In Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, Portland, OR, USA, 2–4 August 1996; pp. 153–158.

18. Orhan, E. *Particle Filtering*; Center for Neural Science of New York University Sciences: New York, NY, USA, 2012.

19. Yu, M.; Oh, H.; Chan, W.H. An improved multiple model particle filtering approach for manoeuvring target tracking using airborne GMTI with geographic information. *Aerosp. Sci. Technol.* **2016**, *52*, 62–69. [CrossRef]

20. Musso, C.; Champagnat, F. Improvement of the Laplace-based particle filter for track-before-detect. In Proceedings of the International Conference on Information Fusion (FUSION), Heidelberg, Germany, 5–8 July 2016; pp. 153–158.