SUMITOMO ELECTRIC TECHNICAL REVIEW
No. 100 · APRIL 2025
AUTOMOTIVE

Road-Crosser Detection Sensor with Radar & Camera for Expanding Detection Area and Enhancing Identification

Shunya ASANO*, Wataru FUJITA, Eiji MOCHIDA, and Kengo KISHIMOTO

In the U.S., the increase in pedestrian fatalities has led to a rising demand for pedestrian detection sensors in systems aimed at preventing traffic accidents. However, many existing sensors lack sufficient vertical detection range since they are typically installed at high positions in actual intersections, creating blind spots directly below them. In cases where the installation point is adjacent to a pedestrian crossing, there are instances where parts of the crosswalk or waiting area may go undetected. To address this issue, we have prototyped a fusion sensor combining radar with a camera. The fusion sensor eliminates blind spots below it by detecting pedestrians at the spots with the camera. Furthermore, using object recognition AI with camera images enables the sensor to identify the type of detection targets, a challenge for radar alone. This paper introduces the prototyped fusion sensor configuration and presents the results of performance verification experiments.

Keywords: traffic safety, pedestrian detection, radar, camera, fusion sensor

1. Introduction

The number of pedestrian fatalities in the United States increased by 77% between 2010 and 2022,[1] prompting efforts to enhance pedestrian safety through advanced driver assistance systems and optimized traffic signal timing.[2] To support these initiatives, road-crosser detection sensors utilizing radar,*1 image sensors*2 (cameras, infrared cameras, etc.), LiDAR,*3 or other suitable technologies are in growing demand as roadside sensors. Radar, in particular, has the advantage of being unaffected by changes in light conditions between day and night as well as weather variations, such as rain and fog. However, securing installation locations is a significant challenge to the widespread adoption of road-crosser detection sensors. Installing new poles entails substantial costs, and at many intersections, constructing new poles is not feasible due to various constraints. Therefore, utilizing existing poles is a prerequisite for deploying road-crosser detection sensors at intersections. On the other hand, many existing sensors lack sufficient vertical detection range, making it challenging to detect road-crossers both at a distance and in the immediate vicinity.[3] Therefore, as shown in Fig. 1, when a pole is located adjacent to a crosswalk, it is impossible to detect road-crossers directly beneath the sensor, both those within the crosswalk and at its entry point.

In addition, compared to cameras and LiDAR, radar is less capable of recognizing shapes, making it difficult to identify the type of road-crosser, such as pedestrians, cyclists, wheelchair users, or white-cane users. The risk of road-crosser accidents varies by type; for instance, it has been reported that the mortality rate per 100,000 person-years for wheelchair users is 36% higher than that for pedestrians.[4] Therefore, being able to identify the type of road-crosser could enable the identification of those at higher risk of accidents.

To address this, the authors developed a prototype fusion sensor composed of a radar and a downward-facing camera. In this configuration, the downward-facing camera detects areas directly beneath the sensor, eliminating blind spots that occurred with radar alone. Furthermore, by applying object recognition AI to the camera's images, it becomes possible to identify the type of road-crosser, which is difficult to achieve with radar alone. This paper presents the configuration of the prototype fusion sensor and the results of verification experiments.

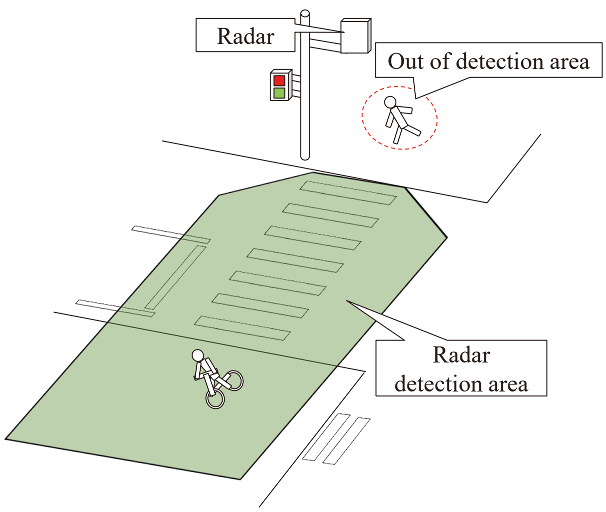


Fig. 1. Schematic diagram of radar installation

2. Fusion Sensor Specifications

Road-crosser detection sensors are typically installed at the corners of intersections, as shown in Fig. 2. The appearance of the prototype fusion sensor is also shown in the same figure. The specifications of the prototype fusion sensor are presented in Table 1.

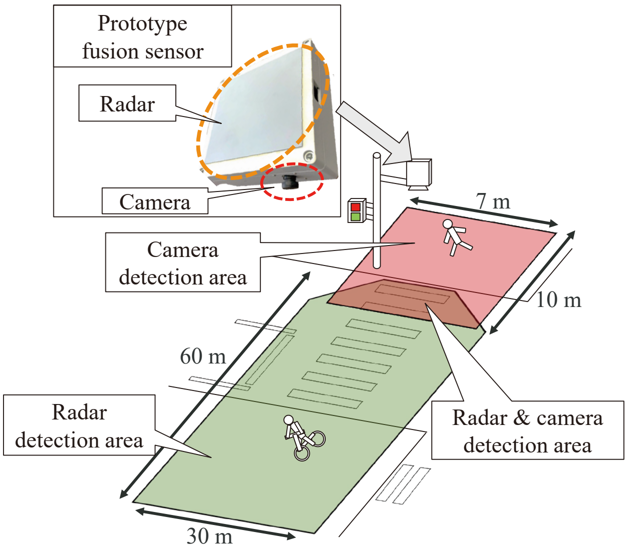


Fig. 2. Schematic diagram of fusion sensor installation

Table 1. Specifications of prototype fusion sensor

Modulation system	FMCW*4
Frequency range	60.0–61.0 GHz
Occupied bandwidth	250 MHz
Radar detection range	Front of the sensor: 4–60 m Side of the sensor: 30 m
Camera detection range	Front/rear of the sensor: 10 m Side of the sensor: 7 m

3. Road-Crosser Detection Method

This chapter introduces the sensor integration method necessary for detecting road-crossers, using the fusion sensor, as well as the road-crosser classification algorithm.

3-1 Sensor integration method

Figure 3 is a schematic diagram illustrating the detection system of the radar and camera in the fusion sensor presented in this paper. The position information of the detection target is obtained in the radar's plane coordinates (unit: meters) and the camera's image coordinates (unit: pixels) simultaneously. Since the coordinate systems of the position information differ, it is not possible to determine if the radar and camera have detected the same target, making detection integration unfeasible. Therefore, the authors used perspective projection transformation to convert the image coordinate system to the radar coordinate system and enable detection integration.

Perspective projection transformation is a model used to convert three-dimensional coordinates into image coordinates. It is generally used to simulate how objects existing in a three-dimensional coordinate system are projected onto an image.[5] Perspective projection transformation is represented by the following equation:

$$s\begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = \begin{bmatrix} f_{x_i} & 0 & c_{x_i} \\ 0 & f_{y_i} & c_{y_i} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} x_r \\ y_r \\ z_r \\ 1 \end{bmatrix} \cdots (1)$$

Where, f is the focal length of the camera, c is the center of the image coordinate system, r is the rotation vector, and t is the translation vector. Let the target's position in the image coordinate system (x_i, y_i) be the input and the position in the radar coordinate system (x_r, y_r) be the output. At this point, Eq. (1) consists of three equations with four unknowns (s, x_r, y_r, z_r), meaning the solution is not uniquely determined. Therefore, by assuming the target's height in the radar coordinate system z_r to be 0 m, Eq. (1) is reduced to three equations with three unknowns (s, x_r, y_r), making it possible to convert the position information from the camera coordinate system to the radar coordinate system. This allows for detection integration between the radar and the camera.

3-2 Pedestrian decision algorithm

This section explains the road-crosser decision algorithm. An algorithm for accurately detecting pedestrians using a radar has been presented in the literature.[6] This algorithm first removes stationary objects and extracts the position coordinates of objects with reflected wave intensities above a threshold as observation values. Next, it applies a continuously operating time-series filter, using these observation values as input. This time-series filter changes the status of the objects based on the status transition diagram shown in Fig. 4.

The Candidate Status is used to establish the error distribution of the time-series filter, and at this stage, it has not yet been determined whether the detected object is a road-crosser. The Confirmed Status represents the state where the detected object is confirmed as a road-crosser. A

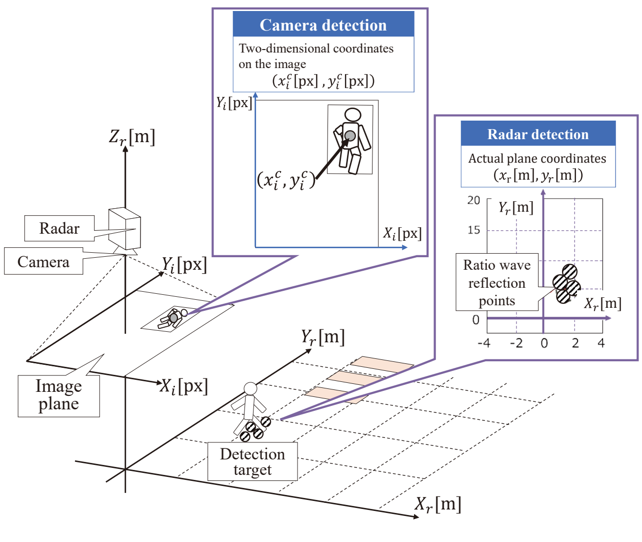


Fig. 3. Detection schematic of radar and camera

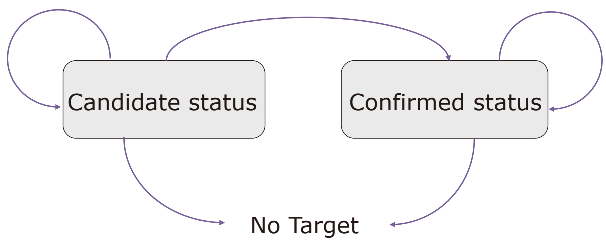


Fig. 4. Status transition diagram

transition from the Candidate Status to the Confirmed Status occurs when specific conditions unique to road-crossers are satisfied. If observation values are not continuously obtained for a certain period in both the Candidate Status and the Confirmed Status, the status transitions to No Target, ending the detection process of the time-series filter.

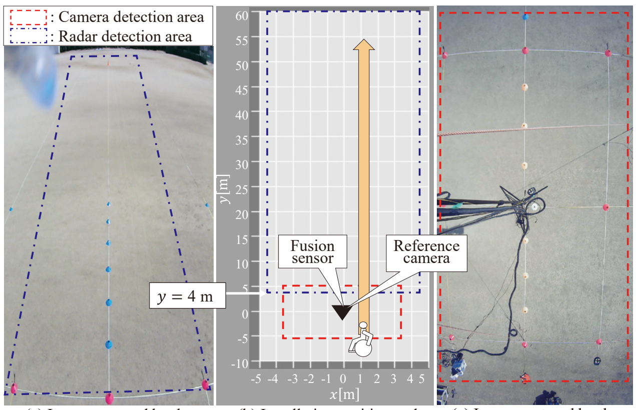
The fusion sensor presented in this paper includes detection by a camera, in addition to a radar. Since the position detection results from the camera can be converted into the radar coordinate system using perspective projection transformation, the time-series filter in the above algorithm can also be applied to the fusion sensor. However, in regions where the detection areas of the radar and camera overlap, integration of the detection information is required. Therefore, in the pedestrian detection process for the fusion sensor, an additional step was introduced to determine whether the same target is being detected. This step links the information detected by the radar with that detected by the camera when the positional difference between the radar and camera detection results is below a threshold. This enables the determination that the same target is being detected and allows continuous detection while retaining the classification information between the radar and camera.

4. Experimental Results

This chapter presents the results of the verification experiments.

4-1 Ground experiment

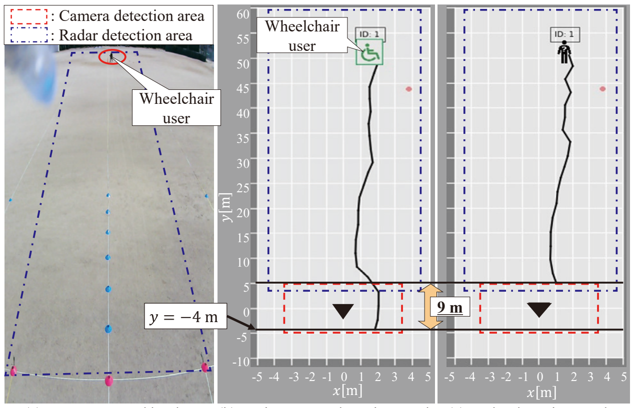
Performance verification experiments were conducted on the ground. An image of the installation environment is shown in Fig. 5 (b). The fusion sensor was installed at a height of 4 meters. Additionally, a reference camera was set up to observe the movement of the detection target. A comparison was then made between the detection results of the fusion sensor and those of the radar alone when a wheelchair user moved along the trajectory shown in Fig. 5 (b).



(a) Image captured by the reference camera (b) Installation position and movement of the detection target (c) Image captured by the downward-facing camera

Fig. 5. Installation environment of the ground experiment

Figure 6 shows the results of road-crosser detection on the ground. Figure 6 (a) is an image captured by the reference camera, showing the presence of the wheelchair user as the detection target. Figure 6 (b) and Figure 6 (c) respectively present the road-crosser detection results from the fusion sensor and the radar alone.



(a) Image captured by the reference camera (b) Fusion sensor detection results (c) Radar detection results

Fig. 6. Ground experiment results

In this case, the fusion sensor was able to begin detection at y = -4 m, whereas the radar alone began detection at y = +5 m. Therefore, the fusion sensor was able to detect the road-crosser at a point 9 meters closer than where the radar alone began detection, confirming that it resolves the blind spot directly beneath the sensor, which is a limitation of the radar alone. In addition, the detection results of the fusion sensor correctly displayed the wheelchair user's icon as the detection target, confirming that the system can continue detection even within the radar's detection range while retaining the classification information obtained through the camera.

4-2 Field experiment

Next, the fusion sensor was installed at an intersection, and performance verification experiments were conducted. An image of the installation environment is shown in Fig. 7. The fusion sensor was installed at a height of 4 meters. In addition, a reference camera and a distant downward-facing camera were installed. The distant downward-facing camera was positioned on the opposite sidewalk, across from the sidewalk where the fusion sensor was installed, enabling the identification of road-crossers entering the crosswalk from the opposite side.

The results of road-crosser detection at the actual intersection are shown in Fig. 8 (c) and Fig. 9 (c). Figure 8 (a) and Figure 9 (a) are images obtained from the reference camera. Figure 8 (b) and Figure 9 (b) show images from the downward-facing camera attached to the fusion sensor. Figure 8 (d) and Figure 9 (d) are images captured by the distant downward-facing camera.

In the scene of Fig. 8, three individuals—pedestrian, cyclist, and white-cane user—were present within the detection range of the fusion sensor. At this time, the icons in the detection results of Fig. 8 matched the actual road-crosser types, confirming that the road-crosser types were also able to be identified in the field. Furthermore, since the white-cane user was directly beneath the sensor, an area outside the radar's detection range, it was confirmed that the downward-facing camera was able to detect targets directly beneath the sensor, just as in the ground experiment.

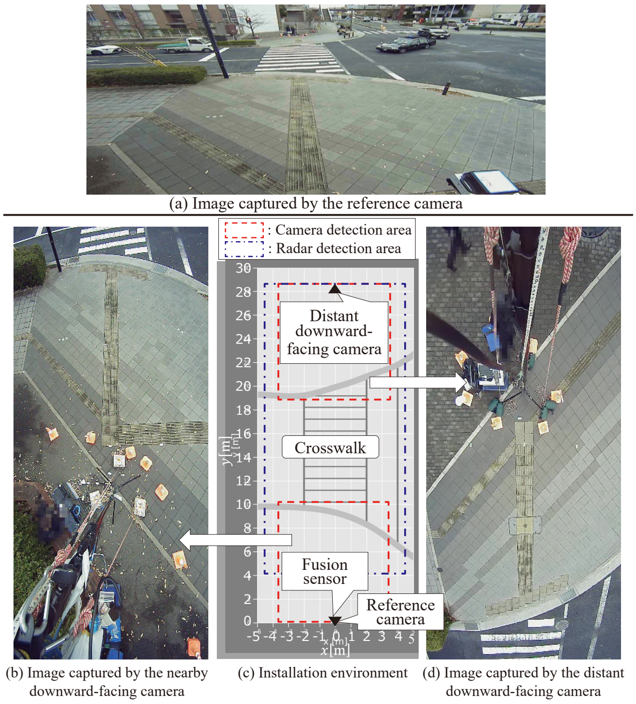


Fig. 7. Installation environment of the field experiment

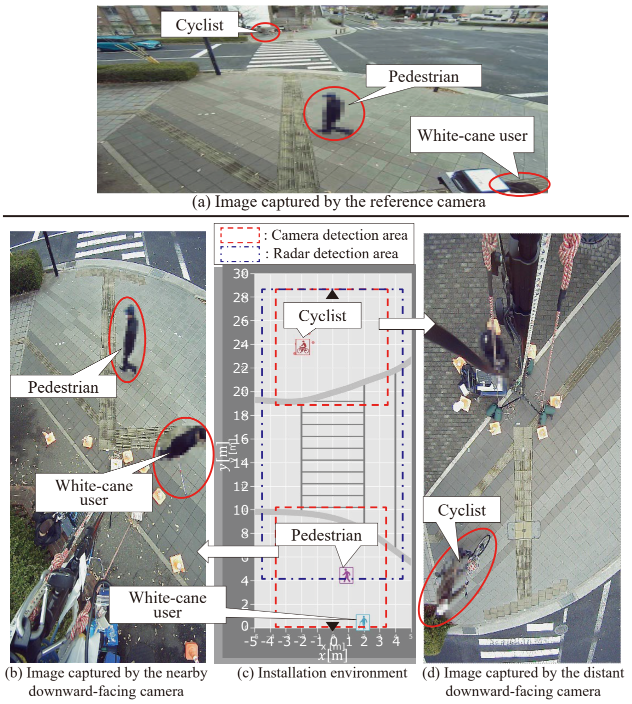


Fig. 8. Field experiment results [Scene 1]

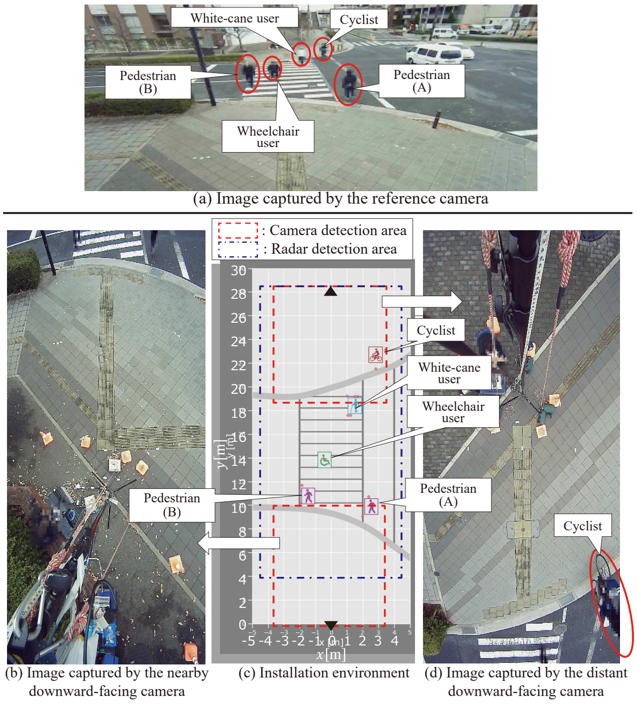


Fig. 9. Field experiment results [Scene 2]

In the scene of Fig. 9, five individuals—pedestrian (A), pedestrian (B), a wheelchair user, a white-cane user, and a cyclist—were present within the detection range of the fusion sensor. Pedestrian (A), pedestrian (B), the wheelchair user, and the white-cane user were outside the detection range of the camera and were detected solely by the radar. Even in this case, icons indicating the types of the detected targets were displayed, demonstrating that the system can continue detecting road-crossers even within the radar's detection range while retaining the classification information obtained through camera detection.

5. Performance Evaluation

Next, the evaluation results of detection accuracy in this experimental environment are discussed. The detection performance of the fusion sensor at the intersection described in Section 4-2 was evaluated. The evaluation indices used were the miss detection time rate, false detection time rate, and misidentification time rate. Each evaluation index is represented by the following formulas:

$$\text{Miss detection time rate} = \frac{\text{Miss detection time}}{\text{Total presence time of pedestrians}} \quad \cdots\cdots (2)$$

$$\text{False detection time rate} = \frac{\text{False detection time}}{\text{Sensor's detection time} + \text{False detection time}} \quad \cdots\cdots (3)$$

$$\text{Misidentification time rate} = \frac{\text{Misidentification time}}{\text{Sensor's detection time}} \quad \cdots\cdots (4)$$

Where the miss detection time rate represents the percentage of time the sensor failed to detect road-crossers within the detection range, the false detection time rate represents the percentage of time the sensor detected road-crossers who were actually not present, and the misidentification time rate represents the percentage of time the sensor detected road-crossers but incorrectly identified their type. From the results shown in Table 2, it was confirmed that the prototype fusion sensor achieved high detection performance and identification rates for road-crossers in the actual environment.

Table 2. Evaluation of accuracy in the field experiment

Miss detection time rate	3.39%
False detection time rate	3.36%
Misidentification time rate	4.48%

6. Conclusion

A prototype fusion sensor combining a radar and a camera was developed to detect road-crossers directly beneath the sensor and to identify their types. This fusion sensor significantly enhances the potential for installation on existing poles, such as those for traffic signals. Furthermore, its ability to identify the types of road-crossers allows for the recognition and detection of traffic-vulnerable individuals in crosswalks, suggesting its potential application in signal control systems that account for high-risk road-crossers.

The authors expect that the utilization of this new technology will accelerate the adoption of road-crosser detection sensors installed at intersections and contribute to enhancing road-crosser safety.

Technical Terms

*1 Radar: A sensor that emits radio waves and measures the distance, direction, and speed of objects based on the information obtained from the reflected waves.

*2 Image sensor: A sensor that detects the presence of objects and identifies their types based on images captured by a digital camera.

*3 LiDAR: The abbreviation for “Light Detection and Ranging,” referring to a sensor that emits light and measures the distance and shape of objects based on the information obtained from the reflected light.

*4 FMCW: The abbreviation for "Frequency Modulated Continuous Wave," referring to a modulation method in which the signal frequency continuously varies over time.

References

(1) Governors Highway Safety Association (2024), Pedestrian Traffic Fatalities by State: 2023 Preliminary Data https://www.ghsa.org/sites/default/files/2024-06/2023%20Pedestrian%20Traffic%20Fatalities%20by%20State.pdf

(2) United States Department of Transportation Federal Highway Administration (2006), Traffic Detector Handbook: Third Edition—Volume I https://www.fhwa.dot.gov/publications/research/operations/its/06108/06108.pdf

(3) Yusuke Yahata, et al., “Millimeter-Wave Radar with Wide Viewing Angle for Improved Installation Flexibility,” SUMITOMO ELECTRIC TECHNICAL REVIEW No.97 (October 2023)

(4) Kraemer JD, Benton CS. BMJ Open (2015). Disparities in road crash mortality among pedestrians using wheelchairs in the USA: results of a capture–recapture analysis, BMJ Open 2015;5:e0083 96, Doi : 10.1136/bmjopen-2015-008396

(5) Z. Zhang, A flexible new technique for camera calibration, in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 11, pp. 1330-1334, Doi: 10.1109/34.888718 (November 2000)

(6) Atsushi Higashi, et al., “A Pedestrian Detection Method based on 24-GHz Band Radar for Driving Safety Support Systems,” ITS world congress 2019, Singapore (October 2019)

Contributors The lead author is indicated by an asterisk (*).

S. ASANO*
• Information Network R&D Center

W. FUJITA
• Information Network R&D Center

E. MOCHIDA
• Ph.D.
Assistant Group Manager, Information Network R&D Center

K. KISHIMOTO
• Group Manager, Information Network R&D Center