

Development of Hybrid Image Sensor for Pedestrian Detection

Hiroaki SAITO*, Kenichi HATANAKA and Toshikatsu HAYASAKI

To reduce traffic accidents and serious injuries at intersections, development of cooperative driving support systems and related sensors has been promoted. Along with this movement, the authors have developed a “hybrid image sensor” that consists of visible-ray cameras and far-infrared-ray (FIR) cameras to compensate each other. The images taken by these cameras are processed simultaneously, thereby covering various conditions including nighttime, shadows, and high temperature. The authors have also established an algorithm to detect pedestrians. This paper outlines the developed hybrid image sensor and pedestrian detection algorithm along with their evaluation results.

Keywords: pedestrian detection, far-infrared-ray camera

1. Introduction

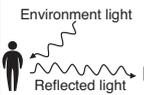
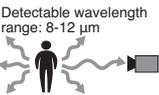
With the aim of reducing traffic accidents, vehicle-infrastructure cooperative driving safety support systems (Fig. 1) have been actively developed. These systems comprise a sensor installed on the roadside to detect nearby pedestrians, bicycles or other objects and warn approaching vehicle drivers. The sensor must accurately detect pedestrians or other objects in all environments 24 hours a day, 365 days a year.

The authors have worked for many years on developing a far-infrared-ray (FIR) camera^{*1} and a pedestrian detection algorithm^{*2 (1)}. The FIR camera uses a sintered ZnS lens made of our original far-infrared optical material to take images of pedestrian-emitted heat. With the aim of reducing traffic accidents at crossings, we have developed a hybrid image sensor combining FIR and visible-ray cameras, with a pedestrian detection algorithm that uses the image characteristics of both cameras. This paper outlines the newly developed hybrid image sensor, details the pedestrian detection algorithm, and the sensor’s pedestrian detection performance evaluation results.

2. Outline of Hybrid Image Sensor

Table 1 compares the working principle and performance of sensors used to detect pedestrians. Visible-ray cameras are widely used in the applications of detecting vehicles, for example traffic control systems; however, it is difficult for visible-ray camera to detect pedestrians at night when they are not illuminated (vehicles can be detected, with headlights on) and they can mistake the shadows of pedestrians for the pedestrians themselves. In contrast, since FIR cameras produce images by detecting the heat objects emit, this type of camera can be used even in the dark of night.

Table 1. Comparison of Pedestrian Detection Sensors

Type of camera	Visible-ray camera	Far-infrared-ray camera	Hybrid image sensor
Principle	Detects reflected environment light or headlight. 	Detects far-infrared-ray (heat) emitted from object. Detectable wavelength range: 8-12 μm 	Detects pedestrians via both thermal and visible-ray images.
Performance	Day	○	○
	Night	△-×	○
	Bad weather	△	○

FIR cameras can also take images even when water drops or snowflakes adhere to the lens; in various other adverse environments as well, they offer promise as pedestrian detection tools. However, one drawback of these cameras is that they cannot create an accurate image when there is insufficient thermal contrast between pedestrian and background (road surface), especially during the daytime in summer.

To compensate for these above shortcomings of visible-ray and FIR cameras, thereby enhancing pedestrian detection accuracy, we have developed a hybrid image sensor

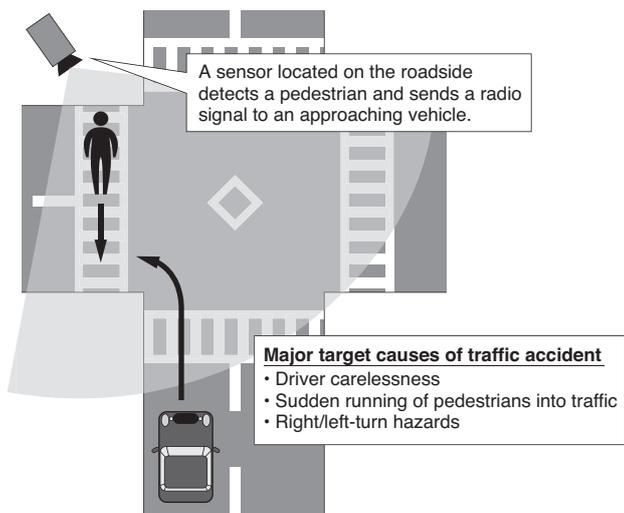


Fig. 1. Vehicle-Infrastructure Cooperative Driving Safety Support System

combining both types of cameras.

External view and internal configuration of the hybrid image sensor are shown in Fig. 2. The optical performance (such as refractive index and transmission factor) of the sintered ZnS (Photo 1) used as FIR camera lens is slightly inferior to that of Ge, a typical FIR transparent material. To overcome this disadvantage, our proven lens design technology was used to ensure necessary frequency resolution (MTF), distortion factor and other optical performance. Since sintered ZnS lenses reduce material and production costs, they reduce the market price of FIR cameras. The specifications of the hybrid image sensor are shown in Table 2.

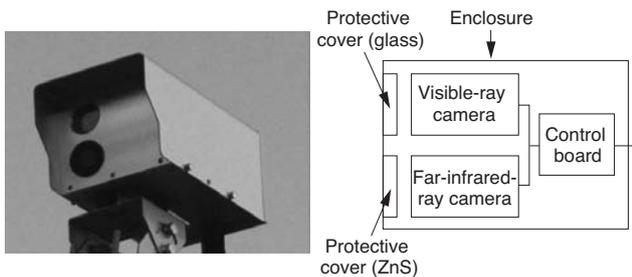


Fig. 2. External View and Internal Configuration of Hybrid Image Sensor



Photo 1. Sintered ZnS Lens

Table 2. Specifications of Hybrid Image Sensor

Item	Specification
View angle	H24[°] × V18[°]
Resolution	Far-infrared-ray camera: 320 × 240 [pixel] Visible-ray camera: 640 × 480 [pixel]
Gradation	8 bit (256 levels) gray scale
Dimension	H152 × W186 × D325 [mm]

3. Pedestrian Detection Algorithm

The basic process flowchart of the developed pedestrian detection algorithm is shown in Fig. 3. This algorithm not only processes the visible-ray and FIR images independ-

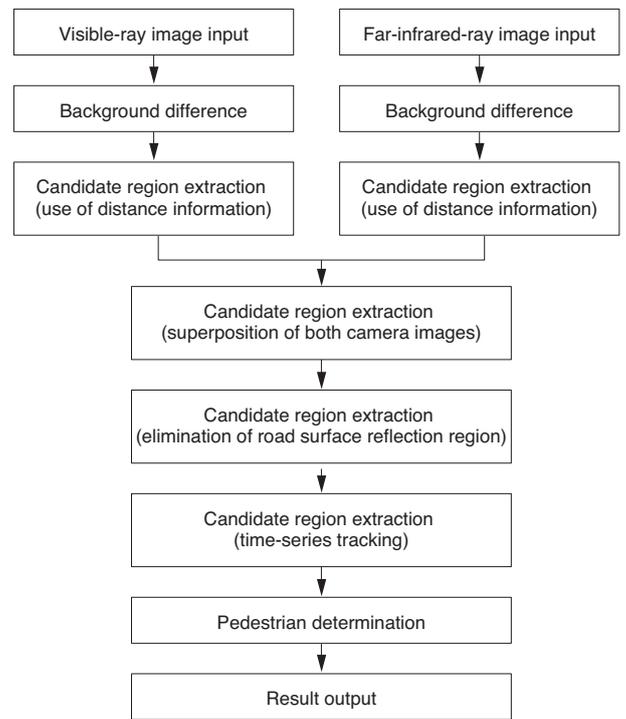


Fig. 3. Process Flowchart of Pedestrian Detection Algorithm

ently, it also unifies the image regions, taking into account the characteristic differences between both the images, to enhance pedestrian detection accuracy. The details of each process in this algorithm are described below.

3-1 Background difference

Background difference is used with input images to extract non-background regions. After the extraction of difference pixels through comparison between input images and background images, adjacent pixels are grouped together and the blanks in the grouped region are filled up in order to extract the region as a non-background region. Although time subtraction (interframe difference) is one possible method of extracting a non-background region, this technique was unsuitable for this study, since pedestrians move only a short distance per unit time as compared with vehicles, making it difficult to detect the movement of pedestrians walking far from the camera.

3-2 Candidate region extraction

The candidate pedestrian region is extracted by unifying the non-background regions that have been extracted using the background difference method and eliminating unnecessary regions. In addition to candidate region extraction based on the distance information, the following three data processing methods have been developed as means of improving candidate region extraction accuracy:

- Unifying non-background regions by superposing both camera images
- Eliminating unnecessary background region by detecting road surface reflection region
- Extracting candidate region by time-series tracking

The details of each method are described below.

(1) Unifying non-background regions by overlapping both camera images

A pedestrian region extracted from a visible-ray camera image is easily affected by the pedestrian shadow, road surface reflection and other factors; therefore the region often becomes larger than the actual size of the pedestrian. In addition, since a pedestrian shadow cast on the road is affected by the movement of the sun, the direction of pedestrian region expansion is indefinite. For FIR camera images, on the other hand, the background difference method cannot extract that portion of the pedestrian where thermal contrast with the background (road surface) is small. As a result, the size of the extracted non-background region is often smaller than the actual pedestrian region (Fig. 4). An exception is road surface reflection on a rainy day, when the extracted non-background region

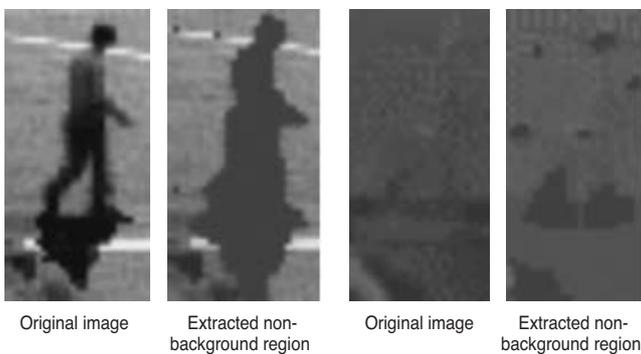


Fig. 4. Extraction of Non-Background Region
(Left: visible-ray camera images
Right: far-infrared-ray camera images)

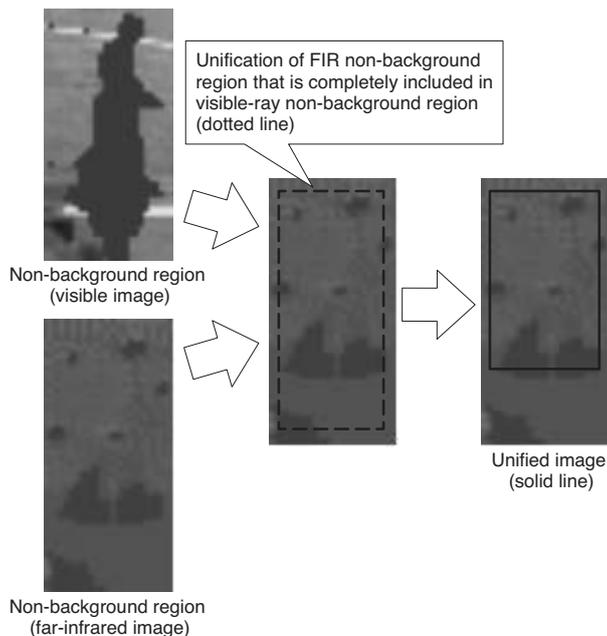


Fig. 5. Unification of Non-Background Regions by Superposing Camera Images

often becomes larger than the actual region of the pedestrian. Even in such an exceptional case, the region expands toward the pedestrian's feet, due to the characteristics of road surface reflection. The above discussions can be summarized as follows: If both the visible-ray and FIR cameras are fixed at nearly the same view angle and nearly the same shooting direction, it will be highly possible to extract a satisfactory candidate region for a pedestrian by grouping together the FIR camera's non-background regions within the visible-ray camera's single non-background region, and eliminating the lower part of the non-background region (see description in (2) below) as needed. A candidate region extracted according to this processing method is shown in Fig. 5.

(2) Eliminating unnecessary background region by detecting road surface reflection region

As described above, in such undesirable ambient conditions as rain, road surface reflection occurs in the lower part of the pedestrian images captured by visible-ray and FIR cameras. The presence of reflection, as well as its region and brightness, depend largely on rainwater film thickness distribution on the road surface. We took a lot of pedestrian images of FIR camera, and analyzed in detail with focus on the features common to these images. Analysis results showed that extracting the edge of a pedestrian's image, including reflected region, would effectively define the inflection point of edge intensity at the boundary between pedestrian and reflected region (= toe) (Fig. 6). This finding has enabled accurate elimination of a reflected region from a candidate region. To describe this process in more detail, the toe coordinates are assumed by the upper coordinates of the extracted region, and pedestrian height taken into account the distance from the camera. An edge

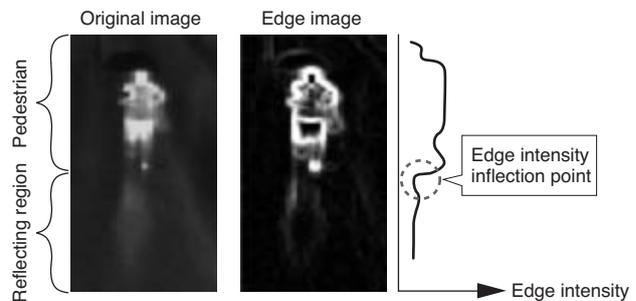


Fig. 6. Road Surface Reflection Region Detection Method
(Edge Intensity Inflection Point)

inflection point is then sought in the vertical neighborhood of the assumed toe coordinates, to eliminate the region below the coordinates of the reflection point.

(3) Extraction of candidate region by time-series tracking

Only a pedestrian image at a given instant is insufficient for accurate extraction of a candidate region, in the case of the image containing two or more pedestrians overlapping one another. To ensure high-accuracy extraction of a candidate region even in such an undesirable condi-

tion, a candidate region search technique that uses past pedestrian detection results is introduced, in addition to the above-described candidate region extraction technique. Specifically, the newly introduced technique defines a search range after taking into account the moving speed of pedestrians (including bicycles) while using the initially captured image within the pedestrian detection frame as a reference image, and performs a similarity calculation. A search image having the maximum correlation value exceeding a predetermined level is then selected as the candidate region.

3-3 Pedestrian determination

Whether the extracted candidate region is the detection target is judged according to the following procedures. After the candidate region has been roughly screened using the pedestrian's size condition, the support vector machine (SVM)^{*3} algorithm is used for the candidate regions in the visible-ray and FIR images to determine whether or not the regions represent a pedestrian. The region is finally determined to be a pedestrian only when either of the regions represents a pedestrian. The histogram of oriented gradients (HOG)^{(2)*4} is used as a feature value for SVM determination. Since an extracted candidate region may be slightly dislocated from the actual position of the pedestrian, new candidate regions are added to the extracted candidate region to prevent pedestrian detection error. Furthermore, the size of the extracted candidate region must be normalized before SVM determination. The aspect ratio of the candidate region of the standing pedestrian differs from that of the sideways bicycle. To normalize the size of the pedestrian's candidate region, we separated the pedestrian region from the sideways bicycle region with attention paid to the geometrical features.

4. Pedestrian Detection Performance Evaluation

To evaluate the effectiveness of the newly developed hybrid image sensor and pedestrian detection algorithm, pedestrian detection testing was carried out for images captured around Japan throughout the year. The detection rates used as performance evaluation criteria are defined in **Table 3**, while the evaluation results are given in **Table 4**. For every ambient condition, the false-negative rate and false-positive rate of the sensor were lower than 10%, within acceptable ranges. Most instances of false-negatives were attributable to (a) pedestrians in a group or (b) a pedestrian overlapped with another object (especially when overlapping continued immediately after their appearance in

Table 3. Definition of Detection Performance Criteria

Item	Definition
False-negative rate [%]	Number of detected pedestrians / actual number of pedestrians × 100
False-positive rate [%]	Number of detection errors / total number of detected pedestrians × 100

* Target: pedestrians, bicycles, wheelchairs

Table 4. Detection Performance Evaluation Result

	Sunny/ day	Fine/ twilight	Clear/ night	Rain/ day	Rain/ twilight	Rain/ night
False-negative rate [%]	4.8%	3.2%	2.9%	6.2%	6.3%	6.9%
False-positive rate [%]	8.0%	3.5%	7.1%	5.6%	8.7%	3.9%

the image). Considering that pedestrians enter a crossing from relatively specific positions and move in relatively limited directions, the additional use of this specific information is expected to improve the false-negative rate. False-positive rates were mostly attributed to capturing the image of part of a motorcycle or vehicle. It may be reduced by effectively using the engine temperature information acquired by the FIR camera.

5. Conclusion

A hybrid image sensor and its pedestrian detection algorithm have been developed to overcome the disadvantages of conventional pedestrian detection sensors. Pedestrian detection performance evaluation testing of the sensor has demonstrated acceptable false-negative and false-positive rates as low as less than 10%. Our future challenges are to further reduce the non-detection and detection error rates, and improve the accuracy of pedestrian direction and speed measurements.

Technical Term

- *1 Far-infrared-ray (FIR) camera: A camera that detects the heat (electromagnetic waves with wavelength 8-12 μm) emitted from an object. This camera offers the advantage of measuring the temperature of an object or taking its picture even in the dark.
- *2 Pedestrian detection algorithm: A technique for detecting pedestrian presence in a detection area by processing input sensor signals via software. Technique performance is evaluated using two indicators: (a) false-negative rate, representing percentage of non-detected pedestrians, and (b) false-positive rate, representing rate at which objects are mistaken for pedestrians.
- *3 SVM (support vector machine): A pattern recognition technique. This technique uses past experimental rules.
- *4 HOG (histograms of oriented gradients): An image feature value used for object detection. After an input image gradient has been determined, the gradient is divided for each local region to calculate a histogram.

References

- (1) H. Saito, T. Hagihara, K. Hatanaka and T.Sawai, "Development of Pedestrian Detection System Using Far-Infrared Ray Camera," SEI Technical Review, No.72, pp.112-117 (2011)
- (2) N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," IEEE Conf. on CVPR, pp.886-893 (2005)

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

#### H. SAITO\*

- Automotive Technology R&D Laboratories  
Engaged in the research and development  
of charging systems for electric vehicles



#### K. HATANAKA

- Group Manager, Automotive Technology R&D Laboratories

#### T. HAYASAKI

- General Manager, Automotive Technology R&D Laboratories