

Development of a Night Vision System (Vehicle-Pedestrian Collision Avoidance System)

Takayuki TSUJI*

Hiroshi HATTORI*

Masahito WATANABE*

Nobuharu NAGAOKA*

ABSTRACT

A night vision system has been developed to help reduce vehicle-pedestrian accidents occurring at night. High temperature objects presumed to be pedestrians are detected by processing the images from infrared stereo cameras mounted on the front of the vehicle, and the possibility of a collision is judged by calculating the position and relative moving vector of the pedestrian. In addition, voice guidance is provided and a highlighted infrared image of the pedestrian is displayed at the bottom of the front windshield using a head-up display (HUD) installed in the instrument panel.

As a result of confirming detection performance on a test course, it was found that the possibility of collisions with pedestrians on the road and crossing the road could be judged.

Further, as a result of confirming evasive action on the part of the driver with respect to the provided information, this system was found to be a useful means of reducing vehicle collisions with pedestrians. Further development is needed to bring this kind of system to the point of practical application.

1. Introduction

Approximately 10,000 traffic accident fatalities occur in Japan annually. Some 30% of these are a result of collisions with pedestrians, the vulnerable parties in traffic.⁽¹⁾

We are presently working on a practical application of a pedestrian injury-mitigation system in order to reduce the possible consequences resulting from vehicle-pedestrian accidents. This paper reports on our examination of systems for providing drivers with advance notice about the presence of pedestrians.

In order to examine the system functions and specifications, we first conducted a detailed investigation of accidents involving pedestrian fatalities. Fig. 1 shows the results of our investigation.

These results showed that:

- (1) As shown in graphs A), B), and C), the majority of accidents involving collisions with pedestrians on the road or crossing the road take place when vehicles are traveling straight on straight roads at night.
- (2) As shown in graph D), the majority of these accidents involve collisions with vehicles traveling at speeds of 40-80 km/h.

As a result, we developed a night vision system that would help detect pedestrians on the road or crossing the road that are difficult to see from vehicles driving at night

as shown in Fig. 2, and would notify the driver by providing voice guidance and displaying infrared camera images.

We also verified the usefulness of this system by means of tests to detect pedestrians crossing the road and obstacle avoidance tests.

2. System concept

2.1 System functions and configuration

The following target requirements were established for the purpose of setting system functions:

- (1) The information provided will be able to call the driver's attention to obstacles and will allow judgment of forward conditions.
- (2) The system will be capable of detecting objects ranging in size from a pedestrian's head to a full-body (20 x 20 cm to 2 x 2 m) while taking into consideration targets hidden by oncoming vehicles and other such factors on ordinary roads.
- (3) The system will provide a time margin (information provision time) of 3.5 seconds or more⁽²⁾ after information provision and before collision, and will provide the detection range needed for this.

We decided to satisfy requirement (1) by using voice guidance and a head-up display (HUD) of images of nighttime forward conditions from an infrared camera.

* Tochigi R&D Center

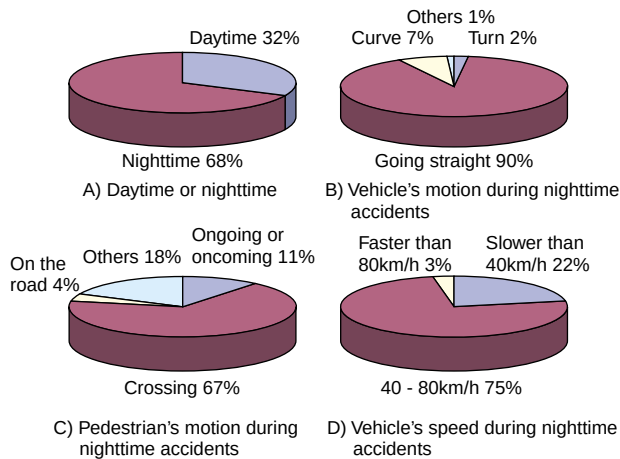


Fig. 1 Accident investigation of pedestrian fatalities

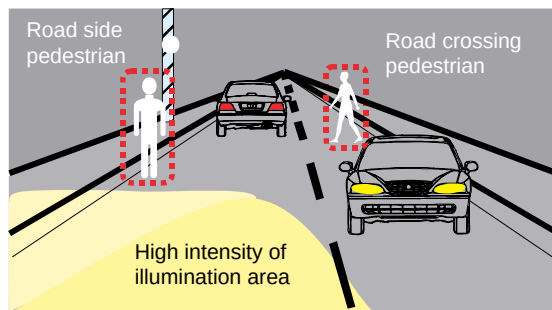


Fig. 2 Example of pedestrians who the driver should notice

We decided that detection capabilities satisfying requirements (2) and (3) could be obtained by using infrared stereo cameras, calculating the distance to high temperature objects that can be perceived as pedestrians, and thus judging the possibility of a collision. Fig. 3 shows the layout of the system on an actual vehicle, while Fig. 4 shows information provision in operation.

2.2 Operating conditions for information provision

The target area for information provision by this system is the region that is difficult to see with headlights and so is subject to a visual enhancement system. This was defined as the region extending outward from the high intensity of illumination area of the headlights (which extends up to 30 m forward).

Based on result (2) from the investigation of pedestrian accidents, it was decided that the information provision system would operate at cruising speeds of 40 km/h and above. The objective was to satisfy the requirement of an information provision time of 3.5 seconds or more for driving speeds up to 80 km/h. This is the 30-80 m detection area shown in Fig. 5.

2.3 Detection and judgment conditions

The two detection areas and the judgment conditions shown in Fig. 5 were established in order to assist in the detection of pedestrians subject to possible collision.

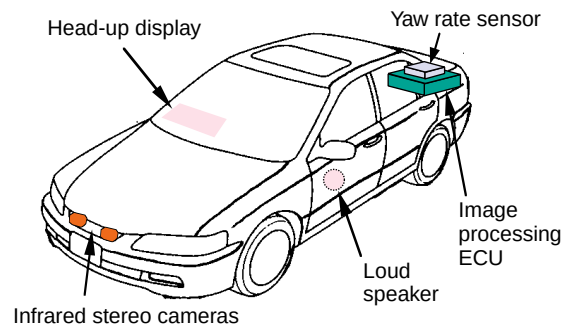


Fig. 3 System component locations

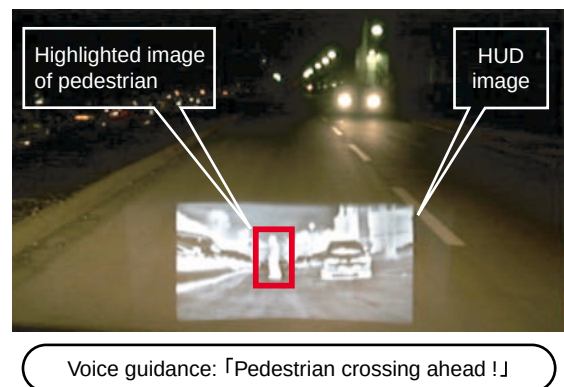


Fig. 4 Driver's view of the HUD image and voice guidance

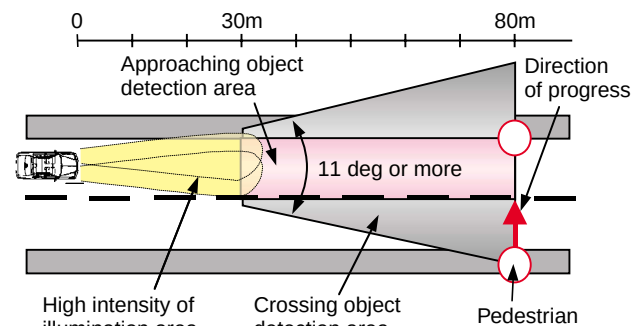


Fig. 5 Pedestrian detection area

2.3.1 Detection area and judgment of road-side pedestrians

The system detects pedestrians walking along the road at a rate of 5 km/h or less within an area the body width of the vehicle with an additional margin of 1 m to either side.

When the subject vehicle is turning, the system uses the yaw rate and vehicle speed to estimate the predicted course of the vehicle, and changes the shape of the detection area accordingly.

2.3.2 Detection area and judgment of crossing pedestrians

The condition for a collision between the subject vehicle and an object crossing the road is that, when seen from the subject vehicle, the object azimuth θ (in degrees) is constant.⁽³⁾ The azimuth θ is determined by the speed of

one's own vehicle V_j and the crossing speed V_h of the object according to the following equation:

$$\theta = \tan^{-1} \frac{V_h}{V_j} \quad (1)$$

Assuming that the pedestrian is crossing at the rate of 1 m/s and that the minimum operating speed for carrying out information provision is 40 km/h, then the detection area for crossing objects will be the total imaging range of the infrared cameras across an angle of 11 degrees or more.

Objects with a greater crossing speed will only appear for a short time as images in the infrared camera's angular field of view. Although they can be detected, the time for information provision will be reduced.

3. Detection processing algorithm

3.1 Definition of coordinate system

This system uses two coordinate systems, the 3-dimensional coordinate system based on the vehicle body as shown in Fig. 6, and the image coordinate system.

The origin O of the 3-dimensional coordinate system is the lens center of the camera on the right. Right-hand rotation of the axis is treated as positive. The image coordinate system is the plane in the 3-dimensional coordinate system where $Z=f$ (where f is the camera focal length), and the origin o is the point of intersection with the Z -axis.

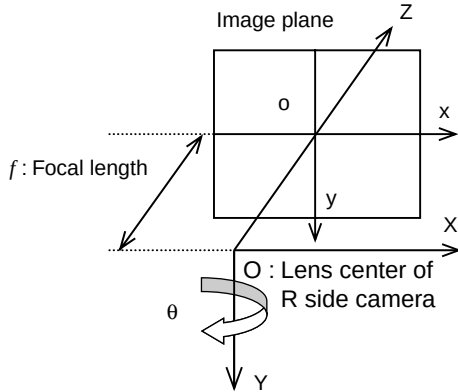


Fig. 6 Definition of coordinate system

3.2 Pedestrian isolation and tracking over time

Measurement of the difference between atmospheric temperature and pedestrian skin surface temperature in all seasons showed that a temperature difference of about 5°C occurs at atmosphere temperatures up to about 30°C. The measurement results are shown in Fig. 7.

Consequently, high temperature objects such as pedestrians appear in the infrared image with a higher brightness than background objects.

We confirmed the possibility of rendering pedestrians in images by driving on ordinary roads in every season, collecting image data, and performing two-level renditions of it. Fig. 8 shows an infrared image taken in the summer and the results of the two-level rendition.

The system on an actual vehicle generates a histogram of the brightness of the right-hand image from the infrared

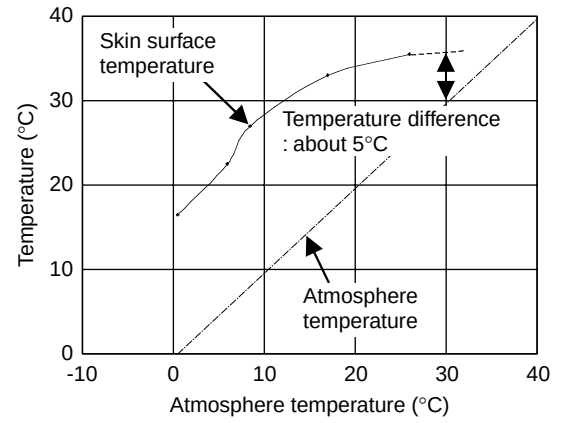


Fig. 7 Temperature differences in skin surface and atmosphere during all seasons



Fig. 8 Infrared image and result of two-level rendition under summer time conditions

stereo cameras and sets the intermediate value in the frequency distribution as the threshold value. The pedestrian in the right-hand image is elicited in this way.

The equality of the object in the right-hand image is judged at each point in time according to such conditions as the position of its center of gravity and characteristics of its geometry to track the object over time. Fig. 9 shows the tracking status of the object in the right-hand image over time.

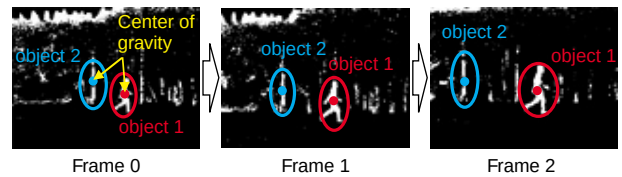


Fig. 9 Object tracking

3.3 Calculation of the object's 3-dimensional position coordinates

As shown in Fig. 10, the left and right images from the infrared stereo cameras display a difference of lateral position (disparity) of dn (in pixels) according to the distance to the object.⁽⁴⁾ The area around the object in the right-hand image is extracted and used to perform a correlation calculation with the left-hand image. In this way, the disparity dn is calculated. Next, equation (2) is used with the thus calculated disparity to derive the distance to the object Z_i (in meters).

The center of gravity position $q=(x_i, y_i)^T$ in the image

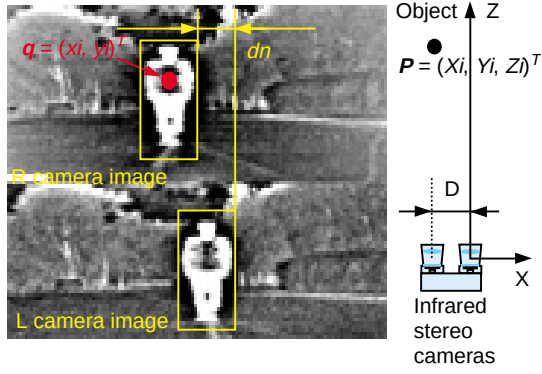


Fig. 10 Disparity of infrared stereo cameras

and the calculated value Z_i (in meters) are then used in equations (3) and (4) to calculate the 3-dimensional position of the object $P=(X_i, Y_i, Z_i)^T$.

D is the distance (in meters) between cameras, while f and p are the focal length (in meters) and pixel pitch (in m/pixel) of the cameras.

$$Z_i = f \times \frac{D}{dn \times p} \quad (2)$$

$$X_i = \frac{x_i \times Z_i \times p}{f} \quad (3)$$

$$Y_i = \frac{y_i \times Z_i \times p}{f} \quad (4)$$

3.4 Calculation of the object's relative moving vector

The calculated position of the object along the X-axis can be changed significantly by slight changes in the turning angle of the vehicle while it is moving. Therefore, it is difficult for the system when operating from a traveling vehicle to identify the difference in the 3-dimensional position of a crossing pedestrian along the X-axis.

Consequently, as shown in Fig. 11, the system uses the vehicle speed and yaw rate sensors to calculate the distance traveled (T) and turning angle (q) of the subject vehicle. Equations (5) and (6) are then used to compensate for the difference.

However, $P'[t]=(Xa[t], Ya[t], Za[t])^T$ and $P[t]=(Xc[t], Yc[t], Zc[t])^T$ are the 3-dimensional positions A and C, respectively, of the object as calculated at time t .

$$\begin{bmatrix} Xc[t] \\ Yc[t] \\ Zc[t] \end{bmatrix} = \begin{bmatrix} \cos \theta & 0 & -\sin \theta \\ 0 & 1 & 0 \\ \sin \theta & 0 & \cos \theta \end{bmatrix} \begin{bmatrix} Xa[t] \\ Ya[t] \\ Za[t] \end{bmatrix} + \begin{bmatrix} Tx \\ 0 \\ 0 \end{bmatrix} \quad (5)$$

$$Tx = \frac{360 \times T}{2 \times \pi \times \theta} \times (1 - \cos \theta) \quad (6)$$

Fig. 12 shows the group of object position $P[t]=(Xc[t], Yc[t], Zc[t])^T$ during the period from time $t=0$ (the past) to $t=n$ (the present) with compensation for the turning angle of the subject vehicle using equations (5) and (6).

The straight-line directional vector $L=(Lx, Ly, Lz)^T$ is obtained by performing principal component analysis⁽⁵⁾ on the center of gravity position coordinate $Q=(\bar{X}, \bar{Y}, \bar{Z})^T$ and the matrix of cross-correlation coefficient for the group of the object position ($P[t]=, t=0, 1, \dots, n$) shown in equation

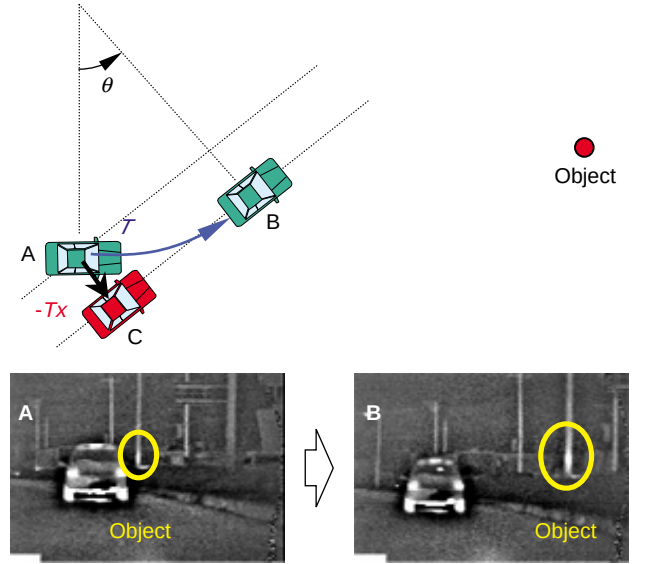


Fig. 11 Compensation for the object's 3D position

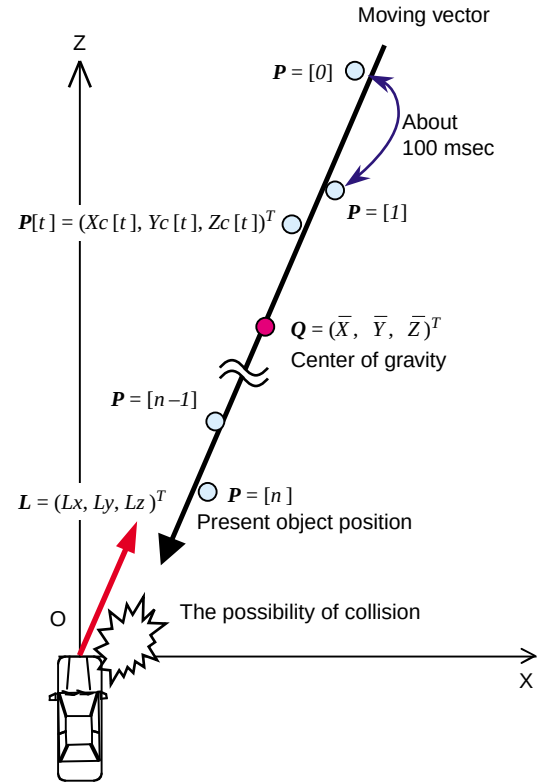


Fig. 12 Predicted 3D moving vector

(7). In this way, the relative moving vector for the object is determined as the straight line shown in equation (8).

The relative velocity (V_s) is calculated from the changes occurring over time in the object position.

$$\begin{bmatrix} \bar{X} \\ \bar{Y} \\ \bar{Z} \end{bmatrix} = \frac{1}{n+1} \begin{bmatrix} \sum_{i=0}^n Xc[i] \\ \sum_{i=0}^n Yc[i] \\ \sum_{i=0}^n Zc[i] \end{bmatrix} \quad (7)$$

$$\frac{X - \bar{X}}{L_x} = \frac{Y - \bar{Y}}{L_y} = \frac{Z - \bar{Z}}{L_z} \quad (8)$$

3.5 Processing to judge the possibility of a collision

The two detection judgments described in section 2.3 regarding detection judgment conditions are carried out with respect to an object that is up to 2 m in height above the road plane and that satisfies the conditions in equation (9) for the calculated relative velocity (V_s) and the driving velocity V calculated from the speed sensor.

$$|V - V_s| \leq \frac{V}{2} \quad (9)$$

3.6 Processing flow

In order to realize real time processing, the system separates processes (1) and (2). These are carried out on independent operation cycles, with process (1) occurring at an NTSC frame rate of 33 msec, and process (2) occurring every 100 msec.

Process (1): Using only the right-hand image from the infrared stereo cameras, this process performs two-level rendition, tracking, and image display of the object.

Process (2): This process calculates the 3-dimensional position of the object from the stereo images, estimates the relative vector, and judges the possibility of a collision.

The pedestrian detection processing flow is shown in Fig. 13.

4. Hardware

4.1 Infrared stereo cameras

The infrared stereo camera system is made up of two 320 x 240-pixel uncooled infrared cameras that have a horizontal field of view of 12 degrees and a vertical field of view of 9 degrees. The distance between the cameras is 360 mm.

The characteristics of the infrared stereo cameras must meet the following four conditions in order to measure the distance to the object correctly:

- (1) The two cameras must have synchronous detection.
- (2) The two cameras must have equivalent image brightness characteristics.
- (3) The two cameras must be in alignment.
- (4) The optical systems of the cameras must not have any distortion.

Synchronous detection was realized by fabrication of prototype synchronous detecting circuits. The cameras in the present prototype system demonstrated sufficiently similar image brightness characteristics.

Corrections for the alignment of the cameras and any distortion in their optical systems were made in the software.

Fig. 14 shows the layout of the prototype infrared stereo cameras when installed on an actual vehicle.⁽⁶⁾

4.2 Yaw rate sensor

The system uses an optical fiber gyroscope with a dynamic range of ± 50 deg/s and a minimum detection resolution of about 0.03 deg/s to calculate the turning angle of the vehicle in order to estimate the relative moving vector and change the shape of the detected area.

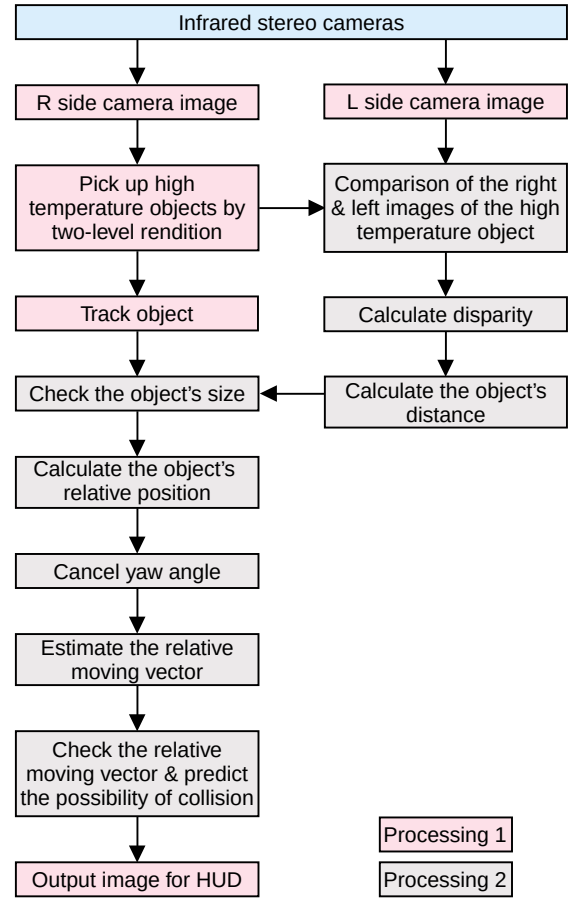


Fig. 13 Processing flow of pedestrian detection

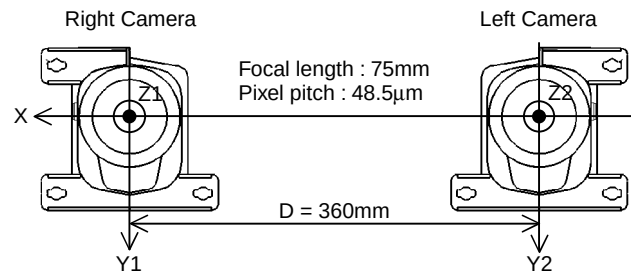
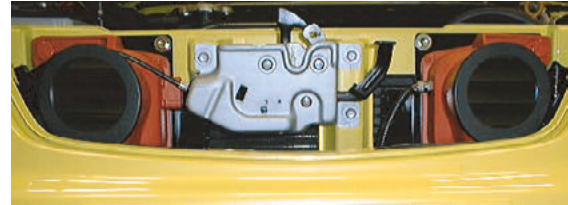


Fig. 14 Infrared stereo cameras Installed in the car

4.3 Image processing ECU

The left and right images from the infrared stereo cameras are output as NTSC signals. These are merged in a 640 x 480-pixel image split horizontally into two (left and right camera images each in a 640 x 240-pixel format) from which 8-bit gray-scale images are sampled at a video rate (30 frames/second). (See the infrared stereo cameras screen image in Fig. 10.)

Speed, yaw rate, and other such information about the vehicle are read in a 33 msec cycle, along with the images. The configuration of the image processing ECU is shown in Fig. 15.

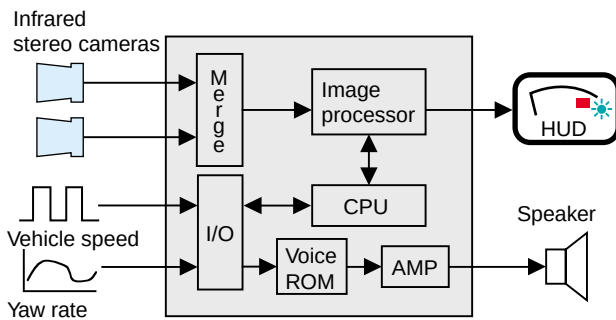


Fig. 15 Image processing ECU

4.4 Head-up display (HUD)

The system displays an 11- x 4-degree (horizontal x vertical) field of view image in order to be capable of displaying the detection range for pedestrians crossing the road and a full-body view of a pedestrian 30 m ahead.

The HUD visible to the driver has a 10- x 3.5-degree (horizontal x vertical) display. This size was selected because it is sufficient to allow a natural-feeling perception of a pedestrian's direction and distance when a pedestrian ahead of the vehicle is displayed as described above.

The display is located in the lower front of the driver, so that the angle when looking down is 5 degrees. The focus point of the image is approximately 2 m forward, near the

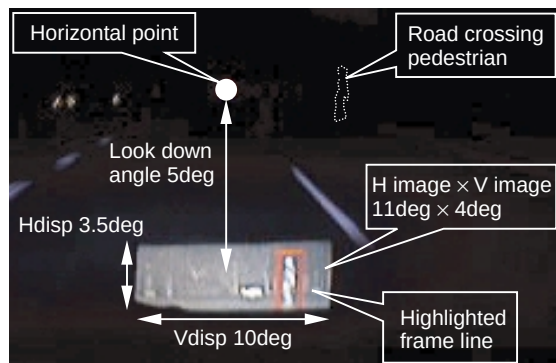


Fig. 16 Projected images of HUD



Fig. 17 Appearance of the HUD

front edge of the hood.

The brightness of the visual image is 0-50 cd, and the color of the detected pedestrian's highlighted image is an easily-recognizable red. (See Fig. 16.) An exterior view of the prototype HUD is shown in Fig. 17.⁽⁶⁾

5. Test results

5.1 Object detection results

Two verification tests were carried out to determine whether the prototype system satisfied the detection performance parameters as they had been set.

5.1.1 Relative velocity calculation test

Fig. 18 shows the results of relative velocity calculations conducted when driving at 60 km/h toward a stationary object. (The horizontal axis shows the distance to the object, and the vertical axis shows the relative velocity error.)

Moving at a speed of 60 km/h, the forward distance to the object at 3.5 seconds before collision is approximately 60 m. Objects to be detected should have a relative velocity within 60 ± 30 km/h, as determined by equation (9).

The relative velocity calculation for a stationary object at a distance of 60 m has an accuracy within ± 25 km/h, as shown in Fig. 18. Therefore, a pedestrian walking at a velocity of no more than ± 5 km/h will have an estimated relative velocity calculation accuracy within ± 30 km/h. This will satisfy the detection condition of a relative velocity within 60 ± 30 km/h.

Similar tests were also conducted at cruising speeds of 40-80 km/h, and these confirmed that the system could detect a pedestrian 3.5 seconds before collision.

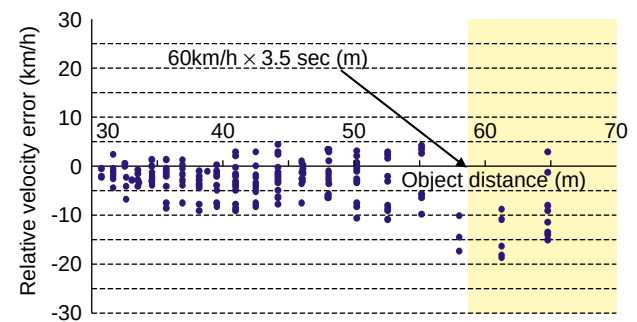


Fig. 18 Relative velocity error of the object distance

5.1.2 Relative moving vector calculation test

Fig. 19 presents the results from calculation of the relative moving vectors of a pedestrian crossing the road at approximately 1 m/s and a stationary object (a streetlight) approximately 60 m ahead of a vehicle traveling at 50 km/h.

From calculation of the relative moving vectors, the crossing velocity of the pedestrian was 1.02 m/s, and the velocity of the streetlight was 0.08 m/s. This shows that the lateral moving vectors are being calculated with an accuracy within ± 0.1 m/s.

The calculation of the relative moving vectors at a point 3.5 seconds before collision and approximately 50 m away yielded the result of 40 km/h. This satisfies the conditions of equation (9), and shows that the system is capable of detection and judgment of pedestrians crossing the road.

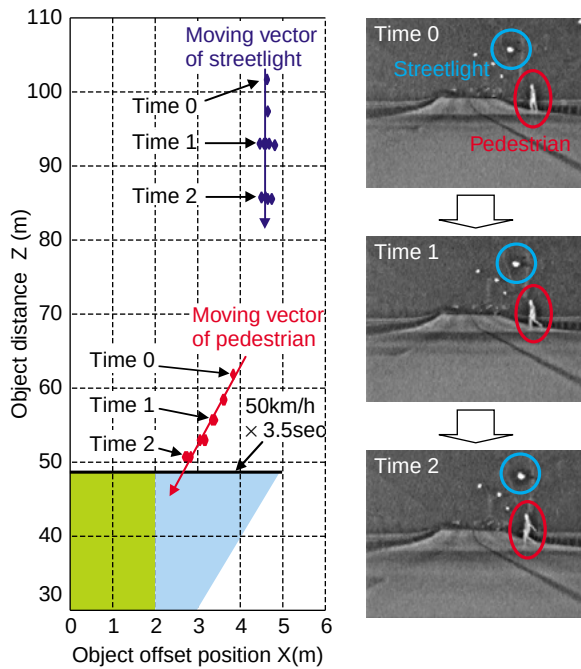


Fig. 19 Detected moving vector and infrared images

5.2 Evaluation of the information provision method

Nighttime driving conditions that make obstacles difficult to see were reproduced on the test course, and multiple techniques for providing information about obstacles were implemented. Comparisons were then done of the timing of the start of the collision avoidance action for each technique.

(1) Comparison of the information provision methods

Voice guidance only and voice plus highlighted infrared image

(2) Comparison of the image display positions

LCD on the instrument panel and the HUD (See Fig. 20.)

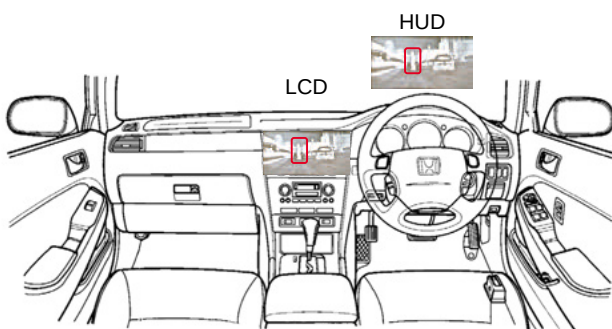


Fig. 20 HUD and LCD display position

5.2.1 Test methods

Fog lamps were used as the only illumination to limit the nighttime range of vision to 20-30 m, and the test subjects were instructed to cruise at a constant speed of 40 km/h. Information was then provided on the obstacles. Obstacles were placed randomly at two locations, one directly in front of the vehicle and one on the centerline, in order to avoid a simple reaction when the information was provided.

Information provision was implemented at a distance of 40 m (3.6 seconds before collision). The test conditions are shown in Fig. 21.

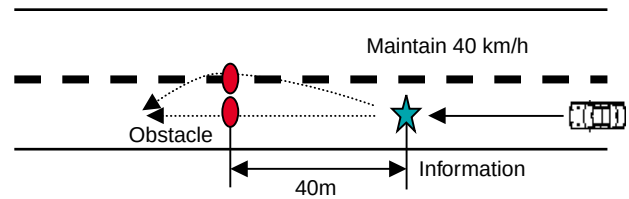


Fig. 21 Interface test condition

5.2.2 Test results

Collision avoidance actions varied with the test subjects. Therefore, the faster of steering avoidance and braking avoidance was selected as the timing for the start of collision avoidance action (response time). Fig. 22 shows the comparison results for the different information provision methods, and Fig. 23 shows the comparison results for the different display positions.

The results from the two tests show that collision avoidance could be initiated about 1 second in advance with voice plus HUD. With voice only and with LCD display, however, the timing for the start of collision avoidance was 1 second or more in many cases. This indicates that voice plus HUD providing a highlighted infrared image is more effective with respect to collision avoidance by the driver.

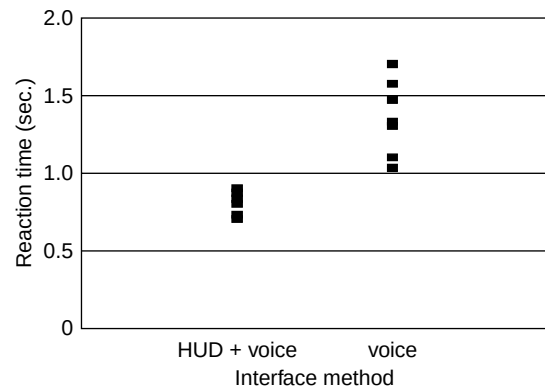


Fig. 22 Comparison of voice and IR image with voice

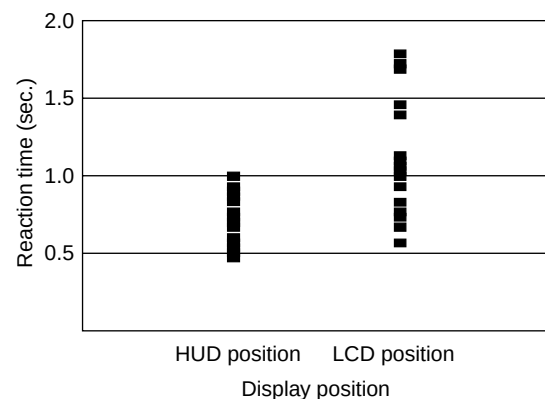


Fig. 23 Comparison of HUD and LCD positions

6. Conclusion

Based on an analysis of nighttime vehicle-pedestrian accidents, we created a system that uses infrared stereo cameras, an image processor, voice guidance and HUD to notify drivers who are driving at night of the presence of pedestrians on the road and crossing the road. This project yielded the following two results:

- (1) We created an algorithm to judge the possibility of collision with high temperature obstacles that can be perceived as pedestrians by means of infrared stereo cameras and image processing. We determined that this enables judgment of the possibility of a collision 3.5 seconds in advance at driving speeds of 40-80 km/h.
- (2) We created an interface that assists in the detection of pedestrians and initiates the possibility of collision avoidance within 1 second from information provision by using voice guidance and a HUD to display a highlighted infrared image that allows comparison of the virtual image to the real roadway scene observed by the operator.

Further development is needed to bring this kind of system to the point of practical application.

References

- (1) "The annual traffic accident statistical database, 1998 Edition," Institute for Traffic Accident Research and Data Analysis (ITARDA), (1999)
- (2) H. Uno and K. Hiramatsu, "Relationship between time margin and drivers' steering avoidance in an emergency situations," Journal of Ergonomics Vol. 35, No. 4, p. 219-227 (1999)
- (3) N. Uchida, K. Fujita, S. Katayama, "Intersection Collisions at High-Visibility Intersections," Journal of the Society of Automotive Engineers of Japan Vol. 30, 1, p. 133-138 (1999)
- (4) M. Okutomi and etc., "Computer Vision Technical Review and View," p. 123-137 (1998)
- (5) "Handbook of Image Analysis," University of Tokyo Press, p. 40-43 (1991)
- (6) N. Asanuma et al. "Introduction of HONDA ASV-2 (Passenger car)," Honda R&D Technical Review, Vol. 12, No. 2, p. 1-8 (2000)

■ authors ■



Takayuki TSUJI



Hiroshi HATTORI



Masahito WATANABE



Nobuharu NAGAOKA