Paper

Advanced Safety Vehicle (ASV) Technology Driver Support System Monitor using Three Onboard Camera

STEPHEN KARUNGARU*†   Non-member,   TOMOHIRO ISHINO†   Non-member
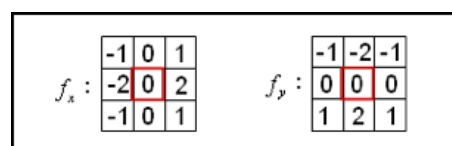KENJI TERADA†   Non-member

Abstract: This paper presents the development of safety driving support system using 3 onboard cameras. One camera monitors the driver to determine the driver's current gaze and the other is a front camera that detects pedestrians, running lane and vehicles in front. The rear camera detects pedestrians and approaching vehicles. The pedestrians and vehicles are detected using a specially trained HOG/Adaboost system. Lane detections uses edge detection and RANSAC. Information from the 3 cameras is then used to determine whether a situation is dangerous enough to warrant warning the driver. We have conducted experiments and the results confirm that this system has the potential to support automatic driving

Keywords: HOG feature value, human detection, vehicle detection, RANSAC, Adaboost

1. Introduction

In recent years, the number of occurrences of traffic accidents in Japan is decreasing due to progress and application of intelligent transportation systems [1]. However, more than 500,000 traffic accidents still occur each year, and the number of traffic accident deaths since 2005 has been steadily increasing to over 4,000 people every year. Moreover, the number of occurrences of traffic accidents by violation of laws and regulations by drivers (first parties) remains greater than that of mopeds. Of the causes, safety non-confirmation accounts for about 30.4%, inattentiveness operation about 16.7%, and vigorous and unconsciousness about 11.5%. These 3 violations account for over 58% of total accident causes. It can be confirmed that the causes of those accidents are the following law violations; unconfirmed safety, negligence, erroneous confirmation of safety, etc. Therefore, driver's carelessness is a main cause. Moreover, various kinds of road conditions occur due to road situations, for example, changing of lanes, vehicles in front and behind, intersections, sidewalks, etc. It is therefore also paramount to grasp the numerous situations during driving and to confirm the safety.

Several driver support systems have already been proposed under the Advanced Safety Vehicle (ASV) technology program. These include the LiDAR [2] and Mobileye [3]. The LIDAR system uses an infrared camera attached on the top of the vehicle. It can capture a surround image and calculate distance to objects. However, its drawbacks include difficulty in installation and high cost. The mobileye system uses one front dashboard camera to capture the dis-

Figure 1: Sobel operator

tance to objects in front of the vehicle. Although it is easy to install and inexpensive, its range of operation is narrow and therefore, not very effective for driver support.

In this paper, we propose a system similar to Mobileye but set up three cameras inside the vehicle to capture the front, the back, and the driver's face, and propose a more effective driving support system. The system uses HOG features and Adaboost for person and vehicle detection, detects driving lanes, driver features like nostril and pupils. This information is combined to design the driver support system.

The system first detects the lanes from the image acquired from the front and rear cameras. Subsequently, objects such as people and vehicles are detected using Adaboost and HOG feature quantities, and the lane of the host vehicle is also determined. We then detect the nostrils and pupils from the image acquired from the internal camera and use the degree of separation to judge the gaze direction of the driver. The information is used to support the driver.

The following sections will explain each part of the proposed method in details. Results of experiments and discussions are presented before concluding.
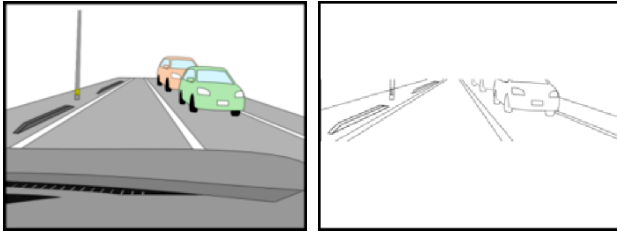
2. Lane Detection

2.1 Edge Extraction   After image capture, as a preprocessing step, edge detection is performed on the image. There are various operators for edge detection, among which the Sobel operator can best detect darker edges. We

* Corresponding author: karungaru@tokushima-u.ac.jp
† Tokushima University,
  2-1 Minami Josanjima, 770-8506, Tokushima, Japan.

(a) Original image   (b) Edge detection

Figure 2: Edge detection result

(a) Edge detection   (b) Noise removal
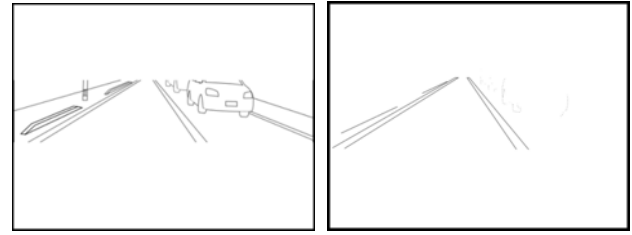
Figure 3: Noise removal result

choose this filter in this work. However, white and orange lines existing on the road surface are thin, and there are cases where it is difficult to capture the color change within the asphalt surface. In addition, from the brightened image, the orange line has a feature that makes it difficult to capture the color change as compared with the white line. For this reason, edge detection is performed separately using the Sobel operator from the R, G, and B channels of the acquired image, and those having the highest edge intensity among them are detected as edges.

Fortunately, the white lines used to determine the driving lane region exists only on the road surface. Therefore, by limiting the range of edge detection, we aim to eliminate noise that can be detected from artifacts such as buildings along side roads and pedestrian bridges that are likely to have linear edges. This step can greatly shorten the overall processing time. Consequently, the range to be detected is only the one to two-thirds area from the top of the captured region surroundings the vehicle. Figure 1 shows the Sobel operator, and Fig. 2 shows the edge detection result.

From the edge detection images, we remove noisy edges for white line detection. There are many edges other than the white lines, such as small edges existing on the road surface, white line edges of the adjacent lane, edges such as in front and oncoming cars, and curbstones on the detected edges. We aim to improve the accuracy of detection of white lines and the detection speed by eliminating these edges as noise. One useful feature of the small edges existing on the road surface is small edge strength. Therefore, when the strength of any edge is 80 or less, it is removed as noise. However, the white line edges of the adjacent lane, the edges of the front and the oncoming car, the curb stone, etc. will still remain. Fortunately, the edges other than the white line of the driving lane has a characteristic that the gradient direction has a value close to 270 degrees, so noise can also be reduced using the gradient direction. Only the edges that satisfy the condition shown in expressions (1) and (2) are used for white line detection with $|\nabla f|$ and θ representing the edge strength and gradient direction respectively. Figure 3 shows the edge detection image with noise removed.
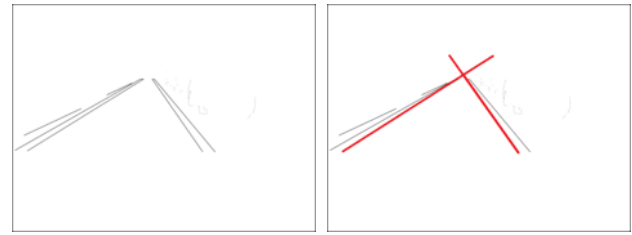
$$|\nabla f| \geq 80 \tag{1}$$

$$\begin{cases} 200 < \theta < 260 \\ 300 < \theta < 360 \end{cases} \tag{2}$$

(a) Edge detection   (b) Line detection result

Figure 4: Line detection result

2.2 White line detection   We aim to detect two straight lines from the detected edges. RANSAC [4] method is applied for straight line detection. RANSAC is a robust model generation method based on random sampling. Its concrete procedures are described below.

1. Randomly sample the data on the edge.

2. Select two points randomly from the sampled data.

3. Find the parameters of the straight line passing through the two selected points.

4. Vote on the number of data that exists within the fixed threshold of the obtained straight line.

5. Save the number of voted data and the two selected points.

6. Repeat from step 2.

This operation is performed 1000 times to detect two straight lines with positive and negative slopes and having largest data. Figure 4 shows an example of straight lines detection using RANSAC.

Since there are cases where the two straight lines detected are not white lines, white line judgment is performed on those straight lines. When capturing the image using the front and the rear of the vehicle from the on-board camera, the white line is often straight. Therefore, the intersection of the straight lines on the white line exists near the vanishing point in the center of the image. In addition, when a correct white line cannot be detected, the angle formed by the straight line on the white line has a value close to 180° or 0°. Therefore, the white line is detected using the angle formed by the intersection point of the straight line on the white line.

Let the area sandwiched by white lines be the traffic lane area, the left side of the white line detected on the left side
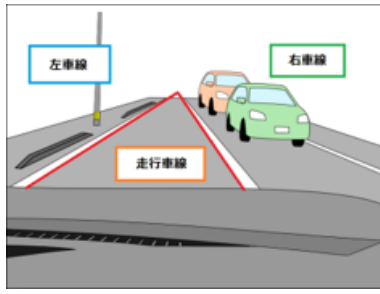
Figure 5: Lane area

of the image will be the left lane area, and the right side of the white line detected on the right side of the image will be the right lane area. Let the intersection point of the straight line on the white line be L_x, L_y and the angle formed be L_θ. If conditional expressions 3 and 4 are satisfied, the detected straight line will be a white line.

$$\begin{cases} 240 < L_x < 420 \\ 180 < L_y < 310 \end{cases} \tag{3}$$

$$10 < L_\theta < 110 \tag{4}$$

2.3 Dotted Line determination   In order to classify the lane that the host vehicle is on, the detected white and orange lines are classified as straight or dotted. Lane classification is performed by a combination of a straight white and dotted lines. If the left white line is straight and the right white line is dotted, the running lane will be the leftmost lane. Also, if the left white line is a dotted line and the left white line is a straight line, the running lane is the rightmost lane. If both are dotted or straight lines, it is the central lane.

For the determination of the dotted line and the straight line, the existence ratio E of the number of pixels of the edge present on the detected white line is used. Let the distance from the point of intersection of the straight lines to the edge of the vehicle be D and the number of pixels of the edge on the detected white line be X. The existence rate E of the white line is calculated by the formula 5. When the existence ratio E of the white line is 0.3 or more, the line is straight, otherwise it is a dotted (Fig. 6).

$$E = \frac{X}{D} \tag{5}$$

3. People and vehicle object detection

3.1 HOG Feature Value   In this work, images are captured using a USB camera mounted on the vehicle. Therefore, it is difficult to accurately detect vehicles using background difference, etc. Hence, the HOG feature quantity is used to capture the characteristics of the vehicle.

We introduce HOG feature quantities to detect vehicles from the captured images [5]. In order to obtain the HOG feature amount, we extract the HOG feature quantity using a cell size set as 8×8 pixels, the size of the block as 3×3 cells, and the histogram of the edge to be created having nine bins with 20 ˚ increments. Equations (6) and (7)
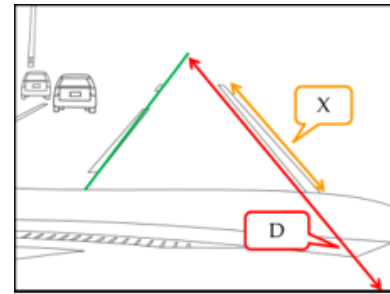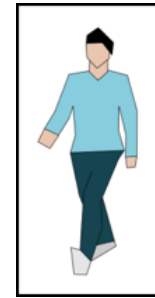

Figure 6: White line existence rate

Figure 7: Original image

are used for the gradient direction and strength calculation respectively. Equation 9 is used to normalize feature quantities for a certain n^{th} block.

$$m(x, y) = \sqrt{f_x(x, y)^2 + f_y(x, y)^2} \tag{6}$$

$$\theta(x, y) = \tan^{-1} \frac{f_x(x, y)}{f_y(x, Y)} \tag{7}$$

$$\begin{cases} f_x(x, y) = L(x + 1, y) - L(x - 1, y) \\ f_y(x, y) = L(x, y + 1) - L(x, y - 1) \end{cases} \tag{8}$$

$$\begin{cases} m(x, y) : Gradientstrength \\ \theta(x, y) : Gradientdirection \\ L(x, y) : Luminancevalue \end{cases}$$

$$v(n) = \frac{v(n)}{\sqrt{\left(\sum_{k=1}^{Q \cdot N} v(k)^2\right) + 1}} \tag{9}$$

$$\begin{cases} v(n) : Featurevalue \\ Q : Numberofcellsinblock \\ N : Gradientdirectionnumber \end{cases}$$

Figure 8 shows the result of calculating and converting the HOG feature amount for the image shown in Fig. 7.

3.2 Adaboost   The HOG feature quantity is learned using Adaboost [6]. Adaboost is a machine learning algorithm that can learn with less data. It calculates the sum of the product of the output of weak classifiers h_i allocated to the cell in the search window and the reliability α_i and discriminate between human and non-human in the input image. The expression of the weak classifier h_i is shown in
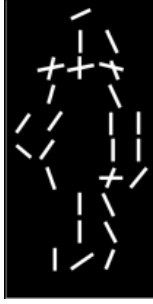
Figure 8: HOG

(a) Original image (b) Search area extraction result

Figure 9: Nostril Search Area

Figure 10: Nostril detection filter

Figure 11: Eye detection filter

expression 10. The expression of the reliability α_i is shown in the expression 11.

$$h_i(y, x) = \begin{cases} 1 : hog(y, x) \geq th \\ -1 : hog(y, x) < th \end{cases} \qquad (10)$$

th : Maximumseparationthreshold
$hog(y, x)$: TheHOGfeatureamount (y, x) oftheimagecell

$$\alpha_i = \log_e\left(\frac{1 - e_i}{e_i}\right) \cdot 0.5 \qquad (11)$$

e_i : Incorrectanswerrateofweakclassifiers

The weight of the learning sample is updated by the calculated reliability, class label, and response of the weak classifier. The formula for updating the weight is shown in Equation 12. The denominator of Eq. 12 is the sum of the weights of the samples. We also normalize the weights.

$$D_{t+1}(i) = \frac{D_t(i) \exp\left(-\alpha_t y_t h_t(i)\right)}{\sum_{i=1}^{N} D_t(i) \exp\left(-\alpha_t y_t h_t(i)\right)} \qquad (12)$$

4. Driving lane determination

We determine the traffic lanes of detected vehicles and location of people using information acquired by lane detection, human and vehicle detection. In the case where the midpoint of the lower side of the circumscribed quadrangle of the detected person or vehicle exists in the traveling lane area, it is assumed that the front object is the preceding vehicle, and if it is the rear image, the following vehicle.

5. Nostril and pupil detection

In this work, the position and size of the face greatly differs for each subject depending on the vehicle, the sitting height, the position of the chair, etc. Therefore, parameters of the face search region are set for each subject. The face search area is set so that its upper end is above the eyebrow, the lower end is near the jaw and the horizontal width is about 3 times the eye width in a front facing state.

For the purpose of speeding up the processing performed in this part, processing is performed on the luminance image already obtained. Next, in order to improve the accuracy of nostril and pupil detection and to reduce the processing speed, the detection region is limited. Features of nostrils
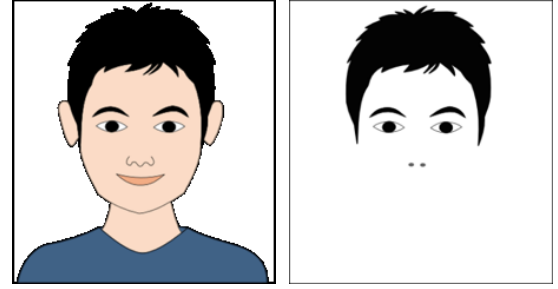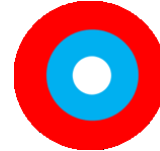
and pupils are characterized by low brightness values regardless of changes in sunshine. For this reason, we use only the area of the face search area whose luminance value is 70 or less. Figure 9 shows the nostril search area.

Next, detection of the nostrils or pupils is performed using the degree of separation. In the interior of the vehicle, the brightness fluctuates due to the influence of sunshine and it is difficult to extract skin color. For this reason, sight line classification is carried out using a region such as a nostril or pupil where there is little change in color due to variation in brightness. When detecting the nostrils, detection is performed using the nostril detection filter in Fig. 10. When detecting the pupil, use the pupil detection filter in Fig. 11. The pupil detection area is from the midpoint of the nostril up to a fifth of the height of the face search area.

Calculation of the degree of separation is performed using the outer and the inner regions both in the nostrils and pupils. At first we find the degree of separation between the red and the blue regions. The obtained degree of separation is calculated as a normalized value and is close to the maximum value 1.0 when the two regions are completely separated. On the other hand, when it is not separated, it approaches the minimum value 0.0. Equation 13 shows the equation for obtaining the degree of separation μ.

$$\eta = \frac{\sigma_b^2}{\sigma_T^2} \qquad (13)$$

$$\sigma_b^2 = n_1(m_1 - m)^2 + n_2(m_2 - m)^2$$

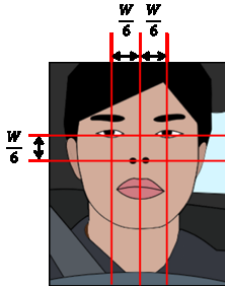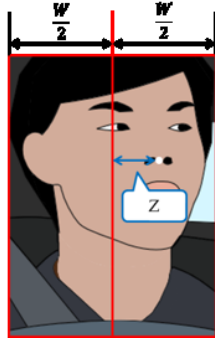$$\sigma_T^2 = \sum_{x \in C}(x - m)^2$$
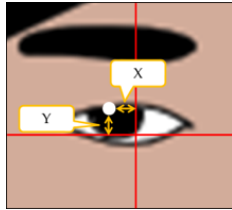
Figure 12: W

Figure 13: Z

Figure 14: X, Y

For the nostril detection, two maximum values of the degree of separation calculated from the nostril search region are used. However, there are cases where the detected nostrils are double-counted by selecting two maximum values. Therefore, when the maximum value of the degree of separation is detected, nostrils are not detected from the region less than 5 pixels in radius.

6. Gaze direction determination

In this work, we set the position and size of face in advance and use it to classify the line of sight. Let W be the width of the face, Y the distance in the vertical direction from the coordinates obtained by shifting from the nostril horizontally and vertically by one-sixth of W to the pupil, and X the distance in the lateral direction. Figures 12, 13, and 14 shows W, Z, X, and Y.

Eye direction widths θ_x, and θ_y are calculated using expression (14). Let Z be the distance in the horizontal direction from the center of the face area to the midpoint of the nostril. The orientation θ of the face in the horizontal direction is calculated by the formula (15).

$$\theta_y = \arctan \frac{2Y}{W}, \ \theta_x = \arctan \frac{2X}{W} \tag{14}$$

$$\theta = \arctan \frac{Z}{W} \tag{15}$$

The gaze direction is classified in eight directions. If the face angle is large to the left or right, it is judged that the subject is looking at the left or the right window, and if not, looking at the front. Further, when the pupil is facing downward, it is determined that the right or the left mirror is being viewed. When facing the front and the pupil is upward, it is judged that the subject is looking at the room mirror. If it is judged that the subject is looking at the front and if the pupil is facing the right or left, the subject is watching the right lane or sidewalk or the left lane.

We classify our eyes using conditional expressions (16), and (17).

$$\begin{cases} Roommirror : 75 < \theta < 110, \ 30 < \theta_x < 70, 85, \theta_y \\ Rightmirror : 90 < \theta, \ 101 < \theta_x, \ 102 < \theta_y \\ Leftmirror : \theta < 80, \ \theta_x < 70, \ 102 < \theta_y \\ Front : 80 < \theta < 130 \end{cases} \tag{16}$$

$$\begin{cases} LeftLane : 75 < \theta_x < 90, \ \theta_y > 94 \\ DrivingLane : 90 < \theta_x < 100, \ \theta_y > 94 \\ RightLane : 100 < \theta_x < 110, \ \theta_y > 94 \end{cases} \tag{17}$$

7. Dangerous object determination

We determine the danger posed by an object captured by the cameras using the distance between the detected object and the host vehicle.

7.1 Approach determination In this section, we will describe approach determination of nearby vehicles and people to establish the objects that the driver must pay attention to while driving.

We estimate the distance (in pixels) change between the host vehicle and the surrounding vehicles by measuring the distance between them. In the case of a vehicle on the right or left adjacent lane, the distance from the nearest point on the white line detected from the center of gravity of the detected vehicle to the lower end of the image along the white line is the distance to the vehicle. In the case of the preceding vehicle existing in the driving lane, the distance from the center of gravity of the vehicle to the lower end of the image is the distance to the vehicle.

The distance acquisition of the vehicle is performed in every frame, and when the acquired distance decreases from the previous frame, the approach count of the vehicle is decremented by one. Also, if not approaching, increment the approach count by +1. When the approaching count of the vehicle falls below −3, it is determined that the vehicle is approaching.

7.2 Warning judgment The monitoring and warning determination are made to judge whether or not the driver is performing safety confirmation using the information on the approaching determination result of the gaze direction.

First, if the driver is not gazing at the vehicle when approaching is determined for the vehicle detected on the left or the right lane, it is judged that the driver is not performing safety confirmation and the warning is sounded. In addition,

for vehicles or people detected on the driving lane, when the distance to the detected object is less than 240 pixels for a person, a warning is sounded because there is a possibility of a collision.

8. Results and Discussions

8.1 Data capture environment   A rear facing camera in the rear of the vehicle, a front facing camera in front of the driver and a camera oriented to face the driver on the dashboard are installed to capture the data required.

The data set used for HOG vehicle and human detection is from INRIA Person Dataset [7] and CBCL Pedestrian Database [8] respectively. The driver-facing camera is oriented such that the driver's face is at the center of the frame. Also, cameras capturing the front and rear of the vehicle are installed so to shoot above the position of approximately 235 pixels. The image size is 640×480. Three cameras with one laptop are used. Same numbers are assigned to images taken at the same time and saved. Processing is performed on the saved three sets of consecutive number images. Shooting time was done during the day.

8.2 Experiment result accuracy   This section shows the accuracy of the experiments. Table 1 shows lane determination results, Table 2 shows gaze judgment, and Table 3 shows dangerous object judgment.

Table 1: Lane determination

	Left lane	Driving	Right lane	all
Scene1	-	57.2%	89.4%	73.3%
Scene2	97.4%	-	-	97.4%
All	97.4%	57.2%	89.4%	91.3%

Table 2: Gaze determination

	All
Subject A	86.9%
Subject B	79.3%
All	82.3%

Table 3: Dangerous object determination

	Warning	Approach	All
Scene 1	70.0%	72.7%	71.3%
Scene 2	16.6%	63.1%	39.9%
Scene 3	72.7%	55.9%	64.3%
Scene 4	54.5%	69.2%	61.9%
Scene 5	-	54.9%	54.9%
All	53.6%	63.2%	58.4%

8.3 Lane determination   In scene 1, processing was performed on a scene in which the left white line is straight and the right white line is a dotted. In scene 1, it was judged that the driving lane was the left end for 380 frames out of
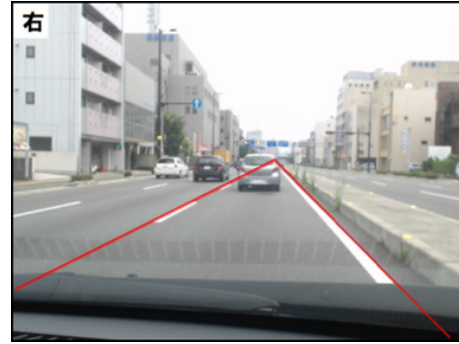
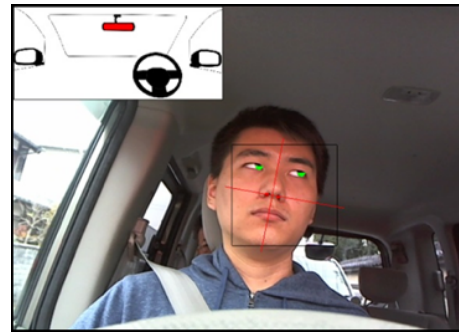Figure 15: Right lane determination image

Figure 16: Room mirror decision image

390 frames processed, and it was possible to obtain a highly accurate result, so it can be said that generally good results were obtained.

In scene 2, processing was carried out for scenes where the white line on the left side is dotted, the white line on the right side runs on a straight lane, and the white lines on both sides are straight. In scene 2, among the 238 frames processed, the traveling lane was judged to be the center for 71 frames out of the 124 frames for the center lane when the left and right white lines were straight. In addition, the driving lane was judged to be the right end for 102 frames out of the 114. Generally good results were obtained. The cause of erroneously determining of the central lane as the rightmost lane is that the white line on the left was hidden and the white line existence ratio decreased as it approached the preceding vehicle at the time of signal change wait. In order to improve the accuracy, it is necessary to determine whether or not the white line is occluded by a vehicle, and to add exception processing accordingly.

The right lane determination result is shown in Fig. 15.

8.4 Fixation judgment   In gaze determination experiments, experiments were conducted on two subjects, A and B.

For subject A, we handled the scene where the driver gazes to the front, right front, right, driving mirror and right side mirror. For subject A, the driver's gaze direction was judged to be frontal for 930 frames out of the 1225 frames (for which the driver was facing the front for 1002 frames). In addition, the driver's gaze direction was judged to be right-front for 64 out of the 74 frames. Also, the driver's gaze direction of the driver was judged to be the right side

mirror for 26 out of 36 frames. The gaze direction of the driver was rightward for 97 out of 102 frames. Moreover, it can be said that generally good results were obtained because the driver's gaze direction was determined to be a room mirror at 9 out of 10 frames.

For subject B, we handled the scene where the driver gazed at the front, the driving mirror, the right side mirror, and the left side mirror. The driver's gaze direction was determined to be frontal for 180 frames of 194 frames. In addition, the driver's gaze direction was judged to be the driving mirror for 10 of the 16 frames. In addition, the driver's gaze direction was judged to be the right side mirror for 10 of the 10 frames. Also, the driver's gaze direction was judged to be the left side mirror for 7 of the 10. In addition, the driver's gaze direction was judged to be rightward for 5 of 7 frames. Good results were generally obtained for the front and right mirrors. There was a case that the gaze direction was erroneously determined as the front when the driver was looking at the room mirror. It is thought that erroneous judgment occurred because the feature looking at the front mirror is similar to the room mirror, the angle calculation method of the current face in the horizontal direction and the angle calculation of the eye. Therefore, it is necessary to examine features that are more likely to capture changes in eye direction.

Fig. 16 shows the room mirror determination result.

8.5 Dangerous object determination   In scene 1, processing was performed on the frames that include a pedestrian. It was possible for the driver to approach the pedestrian for 8 of the 11 frames. In addition, the driver was warned about the pedestrian in 7 of 10 frames. Errors occurred when detecting the person. It can be considered that a detection method corresponding to a change in size is necessary.

In scene 2, processing was carried out for a scene with an overtaking car and a passing bicycle. For the bicycle, the driver gazed at it for 9 of the 13 frames. Also, we were able to warn of the bicycle 6 of the 11.

Also, we made a decision on the oncoming vehicle. The driver gazed at the oncoming vehicle for 7 of the 23 frames. Also, it was not possible to make any judgment for 14 of 17 frames after passing the oncoming vehicle. It is necessary to improve the method so that smaller and large vehicles can be detected.

In scene 3, the vehicle being overtaken by the motorcycle at the intersection was processed. An accuracy of 12 of 19 frames was achieved. Also, it was possible to determine the approach of the motorcycle for 6 of 36 frames when the driver focused on the motorcycle. We could not detect the motorcycle because the approach determination could not be made after being overtaken. It seems that the detection failed because the side of the motorcycle was reflected. In order to cope with it, it is necessary to create a learning image corresponding to the orientation of a person or a motorcycle.

In scene 4, processing was carried out for a scene that included an oncoming vehicle on a two-lane road. An ac-
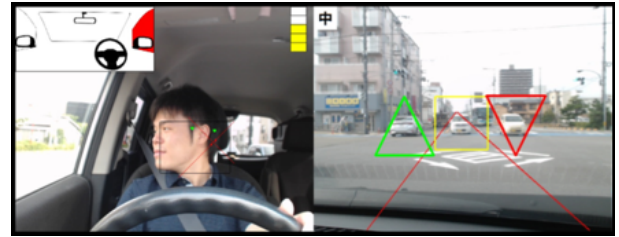

Figure 17: Warning judgment

curacy of 28 of 51 was achieved. In many cases, approach determination cannot be made when detecting an oncoming vehicle at a short distance. Although the vehicle is moving towards the lower end on the image, the area where the vehicle is detected does not follow the movement of the vehicle and the upper part of the vehicle is detect. It is necessary to improve detection of vehicle movement.

In scene 5, processing was performed for a scene when entering an intersection in front of an opposite car and turning right. An accuracy of 248 of 443 frames was achieved. In addition, when the driver was looking at the right, it was possible to trigger a warning of the oncoming vehicle approaching for 16 of the 22 frames. Also, after the oncoming car passed, it was possible to make a decision for 127 of 194 frames away from the vehicle. The warning system obtained generally good results.

A warning determination result is shown in Fig. 17.

9. Conclusion

In this paper, we proposed driver support system using 3 on board cameras. We capture the driver, the vehicle in front and behind using the 3 cameras.

After edge detection, the RANSAC method is used to extract straight lines. We detect the presence of a person of vehicle using HOG feature quantity and Adaboost. We also monitored the driver by detecting nostrils and pupils and judging gaze direction.

Experimental results confirmed that we can detect people and vehicles, estimate gaze direction, and detect hazardous objects using USB cameras during the day.

In future, we need to deal with dark scenes such as nighttime, tunnels, etc. In addition, when the driver wears glasses or sunglasses, it is necessary to improve the detection method because of reflection of light by the lens. Moreover, the detection of the pupil cannot be performed when the lens frame and pupil overlap. Finally, in order to put the system into practical use, it is necessary to improve processing speed using GPU or multithread processing.

References

[1] National Police Agency Traffic Bureau: "The Situation of the traffic Accidents Occurrence in 2015" `http://www.e-stat.go.jp/SG1/estat/List.do?lid=000001150496` (2016/12/19, Japanese)

[2] Velodyne LiDAR HDL-64E obstacle detection sensor, `http://velodynelidar.com/hdl-64e.html`   (Feb,

2017).

[3] Mobileye C2-270 distance detection sensor, j21 corporation, `http://www.imobile.bz/product/mobileyeC2270` (Feb, 2017, Japanese).

[4] Martin A. Fischler & Robert C. Bolles, "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography". *Comm. of the ACM*. Vol.24, No.6, pp.381-395, 1981.

[5] N. Dalal and B.Triggs, "Histogram of Oriented Gradients for Human detection", *Proc. of CVPR 2005*, Vol.1 pp.886-893, 2005

[6] Y. Freund and R. E. Schapire, "A decision-theoretic gener-alization of on-line learning and an application to boosting", *Journal of Computer and System Sciences*, Vol.55, No.119, 1997. DOI:10.1006/jcss.1997.1504

[7] LEAR group. "INRIA Person Dataset". `http://pascal.inrialpes.fr/data/human/` (2016/12/19)

[8] MIT "CBCL Pedestrian Database", `http://cbcl.mit.edu/cbcl/software-datasets/CarData.html` (2016/12/19)

Stephen Karungaru (Non-member) received a PhD in Information System Design from the Department of information science and Intelligent Systems, University of Tokushima in 2004. He is currently an associate professor in the same department. His research interests include pattern recognition, neural networks, evolutionary computation, image processing and robotics.

Ishino Tomohiro (Non-member) received a master's degree Information System Design from the Department of information science and Intelligent Systems, University of Tokushima in 2017.

Kenji Terada (Non-member) received the doctor degree from Keio University in 1995. In 2009, he became a Professor in the Department of Information Science and Intelligent Systems, University of Tokushima department. His research interests are in computer vision and image processing. He is a member of the IEEE, IEEJ, and IEICE.