

## Research Article

# Identification of Cognitive Distraction Using Physiological Features for Adaptive Driving Safety Supporting System

Haruki Kawanaka,<sup>1</sup> Masahiro Miyaji,<sup>1</sup> Md. Shoaib Bhuiyan,<sup>2</sup> and Koji Oguri<sup>1</sup>

<sup>1</sup> Graduate School of Information Science and Technology, Aichi Prefectural University, 1522-3 Ibaragabasama, Aichi, Nagakute, Japan

<sup>2</sup> Department of Medical Information Science, Suzuka University of Medical Science, 1001-1 Kishioka, Mie, Suzuka 510-0293, Japan

Correspondence should be addressed to Haruki Kawanaka; [kawanaka@ist.aichi-pu.ac.jp](mailto:kawanaka@ist.aichi-pu.ac.jp)

Received 6 March 2013; Accepted 26 May 2013

Academic Editor: Martin Reisslein

Copyright © 2013 Haruki Kawanaka et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

It was identified that traffic accidents relate closely to the driver's mental and physical states immediately before the accident by our questionnaire survey. Distraction is one of the key human factors involved in traffic accidents. We reproduced driver's cognitive distraction on a driving simulator by means of imposing cognitive loads such as doing arithmetic and having conversation while driving. Visual features such as test subjects' gaze direction, pupil diameter, and head orientation, together with heart rate from ECG, were used in this study to detect the cognitive distraction. We improved detection accuracy obtained from earlier studies by using the AdaBoost. This paper also suggests a multiclass identification using Error-Correcting Output Coding, which can identify the degree of cognitive load. Finally, we verified the effectiveness of the multiclass identification by conducting a series of experiments. All these aimed at developing a constituent technology of a driver monitoring system that is expected to create adaptive driving safety supporting system to lower the number of traffic accidents.

## 1. Introduction

The number of traffic accidents in Japan are on a declining trend. Fatalities caused by traffic accident in Japan have also been gradually declining for the last 12 years, to reach 4,411 at the end of 2012 as shown in Figure 1. The major factor in this decline seems to come from use of devices as standard or optional equipment in vehicles for enhanced crashworthiness (e.g., airbag systems, seatbelts), intensive enforcement of the Road Traffic Law (e.g., forbidding drunk driving), and improving education of road traffic safety. Meanwhile, the number of nonfatal injuries in Japan caused from traffic accidents, though still in a decreasing trend, reached 824,539 at the end of 2012. Therefore establishing technologies which may prevent traffic accidents remains an important issue for the creation of a sustainable mobile society.

The driver support systems such as *Electronic Stability Control* (ESC), Lane Departure Warning, and Precrash Safety System have been installed into commercial vehicles to prevent from road traffic accidents. Several other safety

systems to lower traffic accidents have also been developed and installed in vehicles. One remarkable Precrash Safety System has a feature that can detect the direction of a driver's face or eyes. But, it is necessary that the current safety systems evolve more so that they can adaptively compensate for a driver's mental state in order to enhance the performance of those safety systems.

Accident investigation and analysis have been regarded as effective for helping to reduce traffic accidents. According to a report released by the Safety Engineering Committee of the Science Council of Japan, it is important to detect the signs of an imminent accident including the human factors in addition to investigation and analysis of traffic accidents [1].

According to the results of the traffic accident analysis carried out by the *National Police Agency in Japan* (NPAJ), traffic law violation counts for 77% of traffic accidents. These include desultory driving, not glancing at the road, no safety confirmation, and failure to predict distance. However, 90%

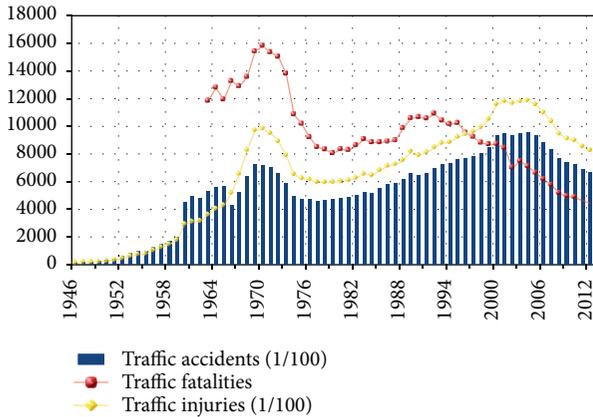


FIGURE 1: Number of road traffic accidents, fatalities, and injuries during 1946–2012 in Japan. Accidents and injuries were divided by 100 to put them all in the same scale.

or so of all traffic accidents are thought to be caused by human error [2].

An internet-based questionnaire survey has been carried out to collect information concerning traffic incidents while doing normal driving [3]. Please refer to the paper [3] for more information on this questionnaire. The original questionnaire survey was conducted in Japanese. Because the questionnaire consisted of eight sheets with A4 format and they are too many to be included in this paper, we translated a couple of questions to show as sample here. First, the respondents were shown a picture of traffic incidents including near-miss accidents and were asked whether they encountered any such experience within the past two to three years. Examples of the following two questions of the questionnaire items are shown in the following

(Q1) Please choose your driving conditions and actions immediately preceding your traffic incidents experience including near-miss accidents. Please select only one answer that you think is closest.

- I was falling asleep.
- I was deeply involved in thoughts and/or conversation.
- I was looking at (signs or landscape) out of the car.
- I was operating a device (car navigation, audio, and/or mobile phones).
- I did not check safety measures, or what I did was not sufficient.
- In an attempt to avoid another hazard, I made a wrong decision.
- I wrongly predicted the behavior of the other car.
- I made a mistake in driving sense (regarding speed, vehicle width, and/or distance).
- I made a wrong assumption in road geometry or road environment because of fog and icy roads.

- I was wrong in recognizing Traffic regulations for one-way, or lane change or regulations).
- I was wrong to recognize the traffic safety facilities (a signal, a guardrail, etc.).
- I made a wrong assumption about obstacles (parking vehicles etc.).
- I made a mistake in braking.
- I made a mistake in pressing either brake/accelerator pedal.
- I made a wrong operation of Steering Wheel.
- I correctly recognized the driving environment and did not make any mistake in judgment or in operations.
- I do not understand or I do not remember correctly.
- Other (write in more details).

(Q2) Please choose your psychosomatic state immediately preceding traffic incidents experience including near-miss accidents. Please select only one answer that you think is closest.

- I was in a rush.
- I was not at ease while driving.
- I was frustrated or angry for something.
- I had good things and bad things happening and therefore was in a bit of unstable condition because of these mixed feelings.
- I was not thinking clearly.
- I was sleepy or fatigued.
- My recent illness clearly left me in a state of decreased health.
- I was fully concentrating on driving and had no problem either mentally and physically.
- I do not understand or I do not remember correctly.
- Other (write in more details).

In the survey, we have asked test subjects about seven types of traffic incidents including near-miss accidents situations involving collisions. They are right turn, left turn, crossing path, person to vehicle, head-on, rear end, and lane change. These accidents were defined as potential accident risks in the *Advanced Safety Vehicle (ASV) Promotion Project* drafted by the Ministry of Land, Infrastructure, Transport and Tourism of Japan. In consideration of the effect of time on the respondents' memory of their own psychosomatic state, our survey only includes responses obtained from the respondents having traffic incidents including near-miss accidents traffic incidents experienced within the past two or three years.

The respondents' number was 2,000 (1,117 men and 883 women) and their average age was 41.1 years; their average driving experience was 19.9 years. We found that incidents of traffic incidents including near-miss accidents happened on an average of 2.39 times during the last three years from

the results of our questionnaire survey. The twenty-eight questionnaire items were set based on driver behavior immediately before the traffic incident, which is classified as one of the Road Traffic Law violations (e.g., no safety confirmation, inappropriate assumption, and desultory driving) defined by NPAJ. In the same manner, the ten items were set concerning psychosomatic states (e.g., haste, lowered concentration, and drowsiness).

Major answers for the driver behaviors were, in descending order, “No safety confirmation” (30.9%), “Inappropriate assumption” (23.2%), “Distracted driving” (12.5%), “Not looking ahead carefully” (3.7%), “Not looking movement carefully” (2.1%), and “Inappropriate operation” (1.2%). Different types of the violation of safety driving will be shown in the following list: the meaning of words such as special terms related to Road Traffic Law of Japan is explained.

#### (I) Delay in discovery

##### (1) Not looking forward

- (i) *Distracted driving*: the state of inattention (e.g., listening intently to the radio or driving carelessly) including drowsy driving.
- (ii) *Not keeping eyes on the road*: the state in which the eyes turned away from the traveling direction to carry out some other operation NOT directly related to the actual driving task (e.g., using mobile phone, looking for a guidance sign).

- (2) Failure to make safety check: the case when the driver overlooked confirming the safety checks (like looking front, back, right, and left directions), while slowing down to stop for some reason or while changing lanes.

#### (II) Judgment error

- (3) Failure to confirm traffic movement: not glancing to the traffic movements or disregarding and/or underestimating the danger by recognizing the other drivers.
- (4) Inappropriate assumption: making a mistake in driving sense such as vehicle width and speed or a mistake in correctly guessing the distance, speed, and behavior of the other party including a failure to predict hazard.
- (5) Traffic environment: false recognition and wrong judgment about the road conditions or road shape.

#### (III) Operational error

- (6) Improper steering and/or braking: too much delay or improper operation of the steering wheel to avoid the danger. Failure to avoid the danger by operating the brake.
- (7) Driving at unsafe speed: the case that it is not an appropriate safe speed for the situation even if it is within the speed limit.

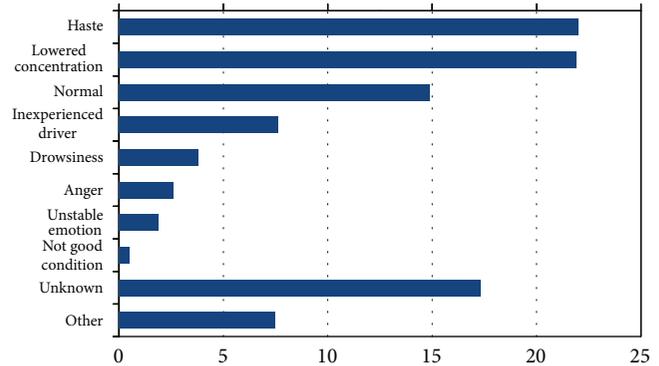


FIGURE 2: Answers to internet-based questionnaire survey of driver behavior.

According to the survey, the human error has occurred for at least 74% persons excluding traffic environments and others. The result means that it is important to detect the driver’s mental and physical states immediately prior to encountering a traffic accident for the purpose of lowering the number of traffic accidents. The survey also found that the psychosomatic states immediately before traffic incident were “Haste” (22.0%), “Lowered concentration” (21.9%), and “Normal” (14.9%) in descending order of frequency as shown in Figure 2. So, we have identified “Haste” and “Lowered concentration” (hereafter distraction) as key factors in terms of drivers physical and mental states which are likely to result in a traffic accident.

Klauser et al. have defined distraction as being caused by four subtasks: “Forward inattention”, “Looking or glancing at an external object”, “Operation of an on-board device unrelated to driving”, and “Drowsiness” [4]. Those factors accounted for a total of 77%. And an investigation of traffic accidents by NPAJ showed that “No safety confirmation”, “Not looking forward”, “Failure to predict hazard”, and “Distracted driving” account for 66% of violated safety driving duty. Their results reconfirm the results obtained from our survey in the same manner.

## 2. Previous Works

Based on the previous results, we focused on driver’s distraction which is suitable for analyzing the traffic incidents. And we should establish a driving support system which is optimized for the individual driver states.

Many studies have regarded changes in driver’s behavior based on physiological information and work performance in order to detect driver distraction [5, 6]. Physiological information consists of the gaze direction, head orientation, blinking, and so forth [7–14]. Work performance consists of steering, lane position, break, driving speed, following distance, and so forth [8, 9, 12, 13, 15–17].

2.1. AIDE. The *Adaptive Integrated Driver-Vehicle Interface* (AIDE) Project, a part of the EU 6th *Framework Programme* (FP6), is developing a system which can detect the driver

state and reduce low-priority information at times of driver distraction. In particular, as part of the EU AIDE Project, Kutila has detected the cognitive distraction which occurred when the driver's level of attention was lowered by factors such as conversation and distracting thoughts [10].

Distraction is classified primarily as either visual distraction or cognitive distraction in AIDE Project. Visual distraction indicates a state when the driver's attention is diverted from the direction of vehicle travel, such as when not looking forward. Cognitive distraction indicates a state of diminished vigilance with regard to the driving environment. The causes of cognitive distraction include conversation with other vehicle occupants and distracting thoughts. Kutila performed experiments to simulate cognitive distraction by providing tasks (cognitive loads) such as having conversation and doing arithmetic unrelated to the current task of driving to the drivers who were driving actual vehicles.

Based on physiological signals concerning eyes and head movements and the vehicle driving position, they used a *Support Vector Machine* (SVM) to detect cognitive distraction [18]. However, the accuracy was 62.4% for trucks and 86% for passenger cars, respectively and has not reached a level where practical application was possible.

**2.2. SAVE-IT.** A program that could help minimize the safety risk of distraction and improve the effectiveness of in-vehicle technologies was directed by *The National Highway Traffic Safety Administration* (NHTSA). It is *safety vehicle(s) using adaptive interface technology* (SAVE-IT), which is designed to mitigate distraction with effective countermeasures and enhance the effectiveness of safety warning systems. Volpe's Operator Performance and Safety Analysis Division manages the multiyear SAVE-IT program in support of NHTSA's Office of Vehicle Safety Research.

The SAVE-IT Project has developed a system which can detect the driver's distraction by means of physiological measures and work performance as features for pattern recognition on *Hidden Markov Model* (HMM) and SVM basis [19]. It has been said that the driver state could provide information to a system that decides the way to adapt the vehicle based on the demands of the task relative to the distraction state.

### 3. Organization of This Paper

Therefore, in this study, we focused on the detection of driver's cognitive distraction and proposed a new method of detecting cognitive distraction by AdaBoost [20]. Referring to the previous study [10], we reproduced driver's cognitive distraction by means of imposing mental loads to the subjects. As a clue of detection, gaze direction angles, head orientation angles, pupil diameters from an eye tracking camera system, and the intervals between *R-waves* (hereafter, RRI) from *electrocardiogram* (ECG) are collected and experimented on different combinations. And we have validated a performance comparison of this method with the SVM. Because of these, it was suggested that distraction detection performance was enhanced than ever before.

We described selections of features for recognition, test procedures using a driving simulator, a methodology for pattern recognition, and results of detection for driver's cognitive distraction in the following sections.

By the way, previous studies have identified just two states only (under ordinary or cognitive distraction state). They have not yet done multiclass identification of cognitive distraction. This paper suggested multiclass identification method based on pattern recognition for the degree of cognitive load. The *Error-Correcting Output Coding* (ECOC) [21] method was employed as the identification method, and two kinds of ECOC methods having different decoding rules were compared.

Additionally, we also describe possibilities of using this method in potential driving safety support systems in order to establish the constituent technology of monitoring driver's psychosomatic states. Multiclass identification could help to provide more adaptive driver assistance for the level of cognitive distraction.

## 4. Selected Features for Detection

**4.1. Visual Features.** Driving behavior occurs as a result of a process of driver's recognition, judgment, and operation. Then it has been said that approximately 90% of driver's perception used at the recognition stage comes from visual information[22]. And it has also been reported that the focus point of gaze and gaze direction may be an indicator for evaluating the driver's mental load and state of attention. For example, the optic nerve system may be affected by an acceleration of the sympathetic nerve when a driver is in a state of cognitive distraction by mental loads such as doing thinking and having conversation while at the steering wheel. Another thing is a pupil diameter that increases by acceleration of the sympathetic nerve when a driver has mental loads while driving. An earlier study by Victor et al. [23] has reported that arithmetic loads caused a fixation of frontal focal points at a narrow range and overall gaze direction to become concentrated in a particular zone. These suggest that detecting vision information may be effective in capturing changes in a driver's psychosomatic states and in particular in detecting cognitive distraction when the driver is engaged in conversation, thinking, and/or day dreaming.

We used the face LAB, which is a stereo camera system manufactured by seeing machines, to obtain the visual information as shown in Figure 3. It can measure physiological signals (gaze direction, eye ball location, face direction and location, blink, saccade, pupil diameter, *percent of the time eyelids are closed* (PERCLOS), etc.) from driver's face image by tracking some feature points of face and eye ball and several kinds of image processing at a sampling frequency of 60 Hz. Figure 4 shows the coordinate system for pitch angle and yaw angle of eye ball. The gaze direction and the head orientation are both output as a vertical rotation "pitch angle" component (up direction is positive) and a lateral rotation "yaw angle" component (rotation to the left direction is positive). The tracking rates of each of the eye movement and the head movement are also output as quality factors.

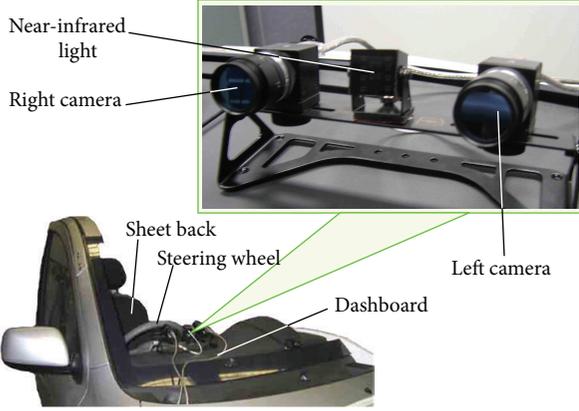


FIGURE 3: Stereo camera system to measure visual information.

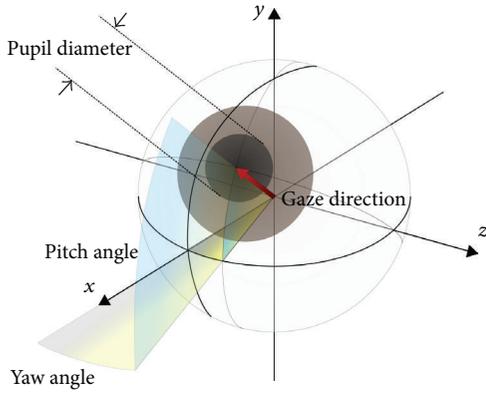


FIGURE 4: Coordinate system of eye movement.

According to Victor et al. [23], cognitive loads make an available gaze area narrow to the traveling direction. They focused on this point and detected cognitive distraction using the *standard deviation* (SD) of the combined gaze angle as a feature. In addition, because movement of the head acts in compensation for the movement of the eyes, it is believed that concentration of the gaze direction will result in changes in the head direction. Therefore, shown in Table 1, we selected the gaze angle, the head orientation angle, and tracking rate of them as well as pupil diameter for detection of cognitive distraction.

In this study as well as in Victor et al. study [23], we have used the standard deviation of the gaze and head orientation angle of test subjects in the same way. The standard deviation is calculated using (1) and (2) from immediately preceding the five seconds:

$$x(i) = \sqrt{x_{\text{pitch}}(i)^2 + x_{\text{yaw}}(i)^2}, \quad (1)$$

where  $x(i)$  is the combined gaze (or head orientation) angle,  $x_{\text{pitch}}(i)$  is the pitch angle, and  $x_{\text{yaw}}(i)$  is the yaw angle:

$$\sigma(i) = \sqrt{\frac{\sum_{j=i-4}^i (x(j) - \bar{x})^2}{4}}, \quad (2)$$

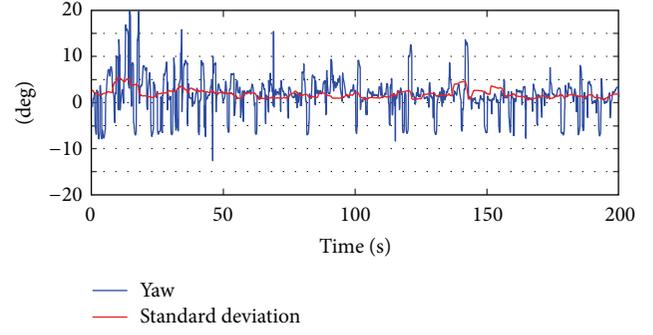


FIGURE 5: Example waveform of gaze yaw angle and the standard deviation.

TABLE 1: Amounts of eye and head movement.

	Left	Pitch
	Right	Yaw
Eye	Standard deviation of gaze angle	Pitch
	Right	Yaw
	Left	Gaze tracking quality
	Right	
	Left	Pupil diameter
	Right	
Head	Standard deviation of orientation angle	Pitch
		Yaw
		Position tracking quality

where  $\sigma(i)$  is the standard deviation of the gaze (or head orientation) angle. An example of gaze yaw angle is shown in Figure 5. Because the data for the amount of eye movement contains noise in the form of blink and saccade, the 13-point median filter was used to remove the noise.

**4.2. Electrocardiogram.** Electrocardiogram (ECG) is a graphical record of changes in the heart's action potential over time. Typically an ECG has five deflections which are *P*-wave, *QRS* complex, and *T*-wave as shown Figure 6. The *P*-wave indicates activation of the atrium. The *QRS* complex indicates the onset of ventricular activation. And the *T*-wave indicates ventricular recovery from activation.

Because a heart activity is affected by the autonomous nervous system, it is closely related to mental activity. A heart rate is determined by the balance between the sympathetic and parasympathetic nervous systems. Ordinarily, the heart rate increases when the sympathetic nervous system is dominant, and the heart rate decreases when the parasympathetic nervous system is dominant. When the driver is exposed with mental stress at the wheel, the heart activity is affected by an acceleration of the sympathetic nerve caused by the autonomous nervous system. Kahneman et al. [24] reported that the heart rate increases in subjects at rest when they are provided with mental tasks.

When a driver is in a state of cognitive distraction, the effects of having conversation, doing thinking, or other factors besides driving are believed to appear in the heart rate.

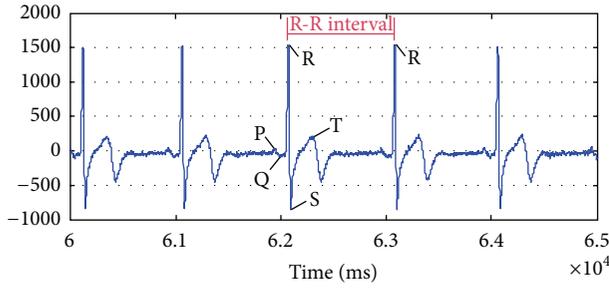


FIGURE 6: ECG waveform and QRS complex.

So, the RRI is decreased. Therefore, in this study, the heart rate RRI was used as one statistical feature for the detection of distraction.

We calculated the heart rate and the heart rate RRI by measuring ECG waveform shown in Figure 6. We used the standard limb lead (II) which is the voltage between the left leg electrode and the right arm electrode. And we used BioAmplifier (Polymate AP1000 produced by Digitex Laboratory Co., Ltd) to measure ECG. Its sampling frequency is 1000 Hz. The feature was acquired every five seconds, and sampling of data set was performed at 60 Hz. The noise was removed using band-pass filter (1–30 Hz).

Peak detection for calculation of RRI in the ECG waveform was performed as described later. The threshold value for  $R$ -wave peak voltage was set to  $R_{\min} = 400$  mV, and because the  $R$ -waves occur at intervals of 0.5–1.0 second, the minimum time before an  $R$ -wave appears was set to  $t_{\min} = 0.5$  second. Therefore, the conditions under which the ECG waveform  $f(t)$  is at the peak at time  $t$  are the following:

$$f(t) - f(t-1) > 0, \quad (3)$$

$$f(t) > R_{\min}. \quad (4)$$

When it is impossible to precisely determine  $t$  due to the effects of noise or other factors, a past value from  $f(t)$  is used and the conditions in (3) are changed so that the following formula is true for all  $i$ :

$$f(t-i) - f(t-(i-1)) > 0, \quad i = 0, \dots, n. \quad (5)$$

In this study,  $n = 3$ .

The time, which satisfies (4) and (5), is considered the  $R$ -wave peak candidate. If no new peak candidate is found during time  $t_{\min}$  following the initial peak candidate, the time, at which that peak candidate has occurred as the peak time  $P_n$ , is considered.  $RRI_n$  is then calculated as the difference between peak times using (6):

$$RRI_n = P_{n+1} - P_n. \quad (6)$$

In the same way as when identifying features in the gaze direction, RRI is calculated within a five-second time window, and the average value is found.

## 5. Driving Task and Cognitive Load Used in Experiments

**5.1. Driving Task.** We conducted an experiment using a mockup-type driving simulator as shown in Figure 7. The subjects were instructed to drive on a course that was projected on the screen in front of them. The selected driving course was a rural road without traffic signals, which allowed us to make a comparison with the results from the Intermediate Course of EU AIDE Project. This course has a road with two lanes in each direction with multiple buildings on the sides of the roads and continuous gentle curves. The maximum curve radius was approximately 700 meters and minimum curve radius approximately 80 meters. Such a driving course is reported to be suitable for detection of cognitive loads [10]. Sample images from a part of the driving course and the design of the driving course itself are shown in Figure 8.

**5.2. Cognitive Load.** Subjects had been instructed to pay attention to the speedometer and to control their speed at around 60 km/h. They were also instructed to drive as they ordinarily do while driving as part of their daily lives.

In order to simulate cognitive distractions (such as having conversation or doing thinking unrelated to the task of driving) for the subjects, we imposed two types of loads such as arithmetic tasks and conversation tasks, which followed the example of an earlier study by Kutila et al. [10]. The arithmetic tasks involved verbally subtracting prime number (for example 7) from 1,000 successively. The conversation tasks involved asking the subjects to describe a route which they regularly commuted (such as the road from school to home). At this time, they were asked to describe signals and landmark buildings in as much details as possible. Because it has been reported that cognitive loads produce a delay in reaction times [25], these methods were considered effective means of reproducing driver's cognitive distraction.

For driving on the simulator, the subjects first practiced to familiarize themselves with the driving, then drove without cognitive tasks, followed by driving with arithmetic tasks, and then drove with conversation tasks. The driving time was only five minutes for each segment, and there was a three-minute rest period between segments.

## 6. Verification of Features

**6.1. Gaze Direction and Head Orientation.** In order to examine the differences in features with and without the cognitive loads, we compared the standard deviations of the gaze angle and head rotation angle during ordinary driving and while driving with the cognitive loads.

The distribution of the frontal focal points during ordinary driving without cognitive load is shown in Figure 9 and with cognitive load imposed by arithmetic tasks is shown in Figure 10. During ordinary driving, the frontal focal points were scattered widely to the peripheral area. However, when the cognitive load was imposed, the frontal focal points were concentrated within a narrower range. Average value of the

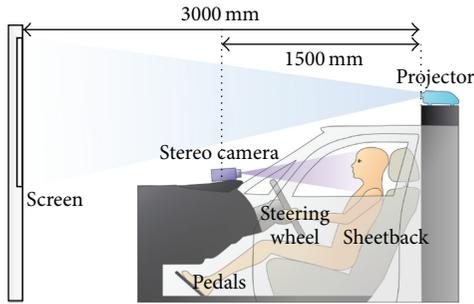


FIGURE 7: Driving simulator.

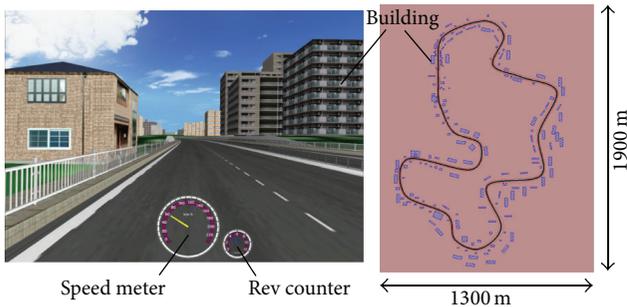


FIGURE 8: Sample images from a part of the driving course and the design of the driving course itself.

SDs for the gaze angles is shown in Figure 11. It demonstrates that the standard deviation of gaze angle decreased by 9% while driving with arithmetic load compared with ordinary driving (without any distraction). This matches the trends shown by earlier research [10]. It is believed that the increase under the cognitive load is what raised distraction among the test subjects.

The distribution of the head moving orientation during ordinary driving without cognitive load is shown in Figure 12, and with cognitive load imposed by arithmetic tasks is shown in Figure 13. Conversely an increase in standard deviation of the head rotation angle is accompanied by a decrease in the movement of the gaze angle when the cognitive loads were imposed. As shown in Figure 14, it demonstrates that an amount of increase in arithmetic was 54% compared with ordinary driving. It is believed that this is a compensatory action by which the driver attempts to obtain a field of wider view, which might be thought as *vestibuloocular reflex* (VOR) [26].

Based on these results, we have concluded that the standard deviations of gaze angle and head rotation angle are suitable to use as a feature for detection of cognitive distraction, and this is in agreement with research results from the EU AIDE Project.

**6.2. Pupil Diameter.** When cognitive loads such as doing arithmetic or having conversation were imposed to the test subjects, the pupil dilation occurred by acceleration of the sympathetic nerve as shown in Figure 15. And Figure 16 shows an example waveform of pupil diameter. Thereby the

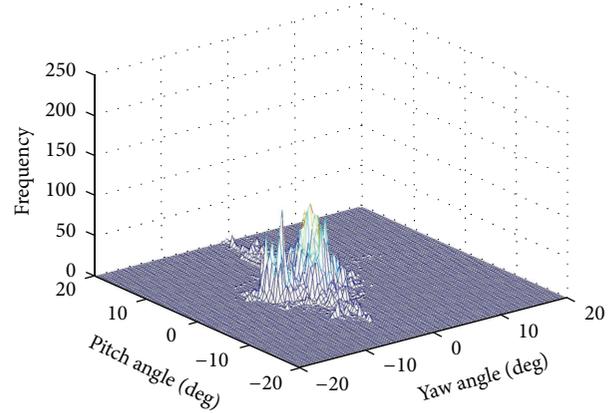


FIGURE 9: Distribution of the frontal focal points, scattered widely to the peripheral area, while doing ordinary driving.

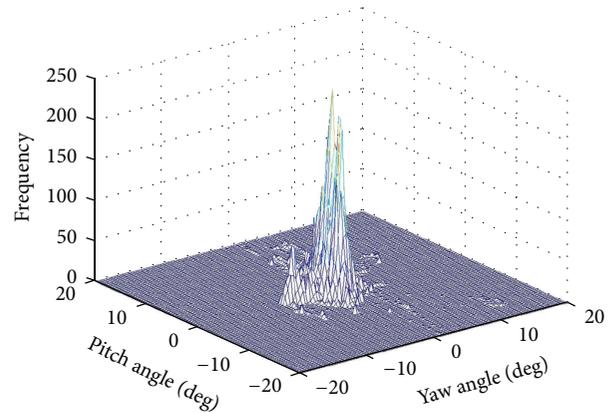


FIGURE 10: Distribution of the frontal focal points, concentrated within a narrower range, while driving with arithmetic load.

measured pupil diameter was increased. And the average value of the pupil diameter is caused by cognitive loads such as doing arithmetic increased as shown in Figure 17. It demonstrates that an amount of increase was 13.1% compared with ordinary driving. Tendency of the results agreed with the results from cognitive load experiments in another earlier study [24].

Based on the previous results, we have concluded that the average value of pupil diameter was suitable in use as a feature for detection of cognitive distraction.

**6.3. R-R Interval (RRI).** The average heart rate increased approximately by eight beats per minute when cognitive loads were imposed. The order of this result agrees with the one in previous studies [24]. Changes of heart rate RRI between ordinary driving state, doing arithmetic while driving state, and having conversation while driving state are shown in Figure 18. In addition, the heart rate RRI, which is caused by cognitive loads such as doing arithmetic and having conversation, was decreased as shown in Figure 19. It demonstrates a decreased rate of 10.6% compared with ordinary driving. This

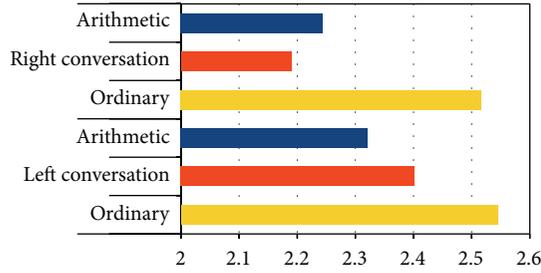


FIGURE 11: Average value of the standard deviation for various gaze angles.

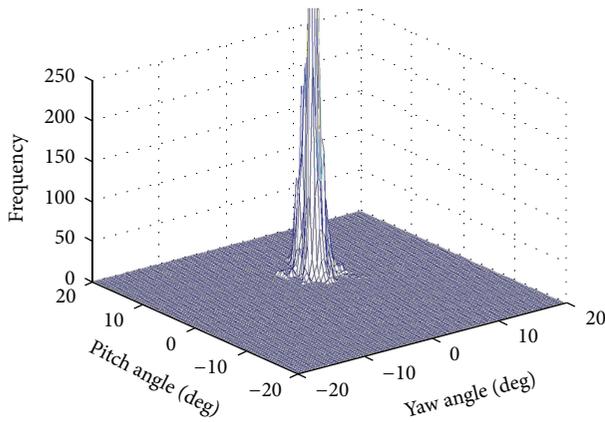


FIGURE 12: Distribution of the head moving orientation while doing ordinary driving.

is believed to be a result of the higher heart rate caused by the cognitive tasks, as described previously.

Based on the above results, we have concluded that the average value of the heart rate RRI was usable as a feature for detection of a state of cognitive distraction in drivers.

## 7. Machine Learning Algorithms to Detect Cognitive Distraction

**7.1. AdaBoost.** AdaBoost [20] is one of the widely used machine learning algorithms for pattern classification (called Boosting) that was proposed by Freund and Schapire. It has advantages such as high classification performance, rapid recognition process time, and extendability of recognition features. Indeed, potential advantages such as higher classification performance is the reason why we decided to apply the AdaBoost for detection of cognitive distraction.

Machine Learning of AdaBoost algorithm involves creating different classifiers while successively changing the weighting of each training data. A weighted majority decision is then made of these multiple classifiers in order to obtain the final classifier function. Each individual classifier is referred to as a “weak classifier,” while the combination of classifiers is a “strong classifier.” This process of successively learning simple and weak classifiers in order to boost classifier accuracy is known as “boosting.”

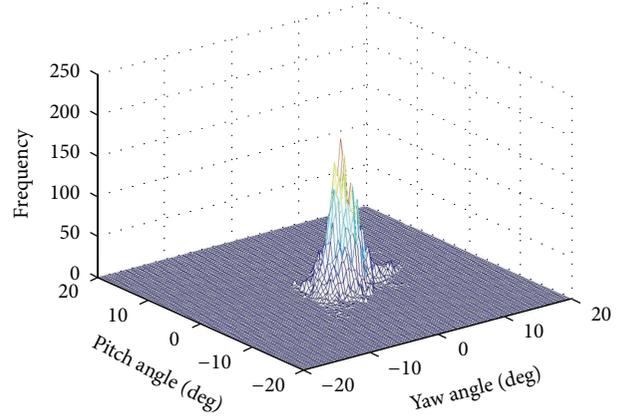


FIGURE 13: Distribution of the head moving orientation while driving with arithmetic load.

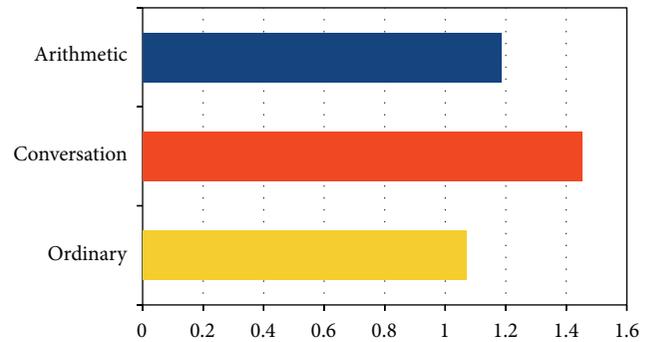


FIGURE 14: Average value of the standard deviations for various head rotation angles.

The combination of input  $x_i \in R^d$  for dimension  $d$  and the corresponding solution  $y_i \in \{\pm 1\}$  is used as the learning data. A quantity  $T$  of weak classifiers  $h_t(x)$ , ( $t = 1, \dots, T$ ) are weighted for reliability  $\alpha_t$  and joined together in order to create strong classifier  $H(x)$ .

The weighting at learning count  $t$  is  $D_t(i)$ . The initial weighting values are all the same:  $D_1(i) = 1/N$ . The weighting for data which could not be correctly identified is increased, so that learning is focused on the later weak classifiers.

In the case of a two-group classification problem, weak classifier  $h_t$  should be selected in order to reduce the error rate as much as possible based on distribution  $D_t$ :

$$\epsilon = P_{r_{D_t}} \{h_t(x_i) \neq y_i\} = \sum_{h_t(x_i) \neq y_i} D_t(i). \quad (7)$$

The specific AdaBoost algorithm for a two-group classification problem is the following. Assume that learning data  $(x_1, y_1), \dots, (x_n, y_n)$  is provided. However, assume that  $x_i \in X$ ,  $y_i \in Y = \{-1, +1\}$  (see Algorithm 1).

We used the GML Matlab Toolbox for the AdaBoost and adopted the stump which is usually used for a Boosting framework as a weak classifier. This stump is the simplest with only one classification node on each stump.

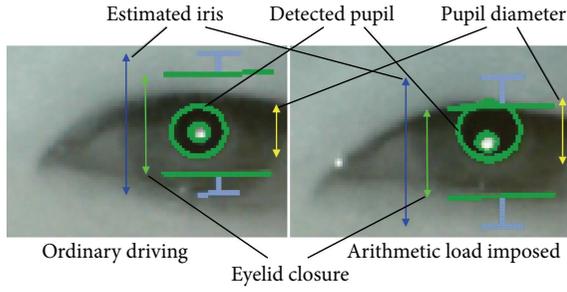


FIGURE 15: Detected pupil area and pupil diameter showing occurrence of pupil dilation.

**7.2. Multiclass Identification.** Many multiclass identification methods have already been proposed. These methods can be categorized into two main approaches. One approach uses a loss function which can handle more than three labels at the same time and minimizes it directly with some sort of method. Major examples of this approach are Neural Network and  $k$ -Nearest Neighbor ( $k$ -NN) algorithm. These approaches are theoretically easier to analyze. However, the computational cost is increased in the case for a large number of samples.

Another approach is based on combining the binary classifiers. The advantages of this approach is low computational cost and ease of implementation. Besides, it has a generalization capability as well as that of the previous approach [27, 28]. For example, error-correcting output codes, or ECOC, is one interesting form of ensemble learning that extends binary classifiers to multiclass identification. It divides multiclass classification problem into some binary classification problems by an encoding rule and decodes binary classification results to multiple classes by a decoding rule.

In this study, we decided to employ ECOC method to get high identification capability. ECOC can be extended easily by modifying the encoding or the decoding rule. Dietterich and Bakiri [29] have suggested the ECOC based on Hamming Decoding. This paper refers to it as the Hamming Decoding ECOC (HD-ECOC) [29] and Loss-based Decoding ECOC (LD-ECOC) [21], which is based on a loss function and is the method that extends the HD-ECOC.

Multiclass classification problem is divided into binary classification problems by using a code table. After learning of binary classifiers and identification for each binary classifier, the outputs are decoded according to the code table. As a result, it is possible to identify into multiclass.

**7.2.1. Hamming Decoding Error-Correcting Output Codes (HD-ECOC).** Let us assume multiclass classification of  $G$  classes ( $G > 2$ ). Firstly, training sample is defined as in (8):

$$(X, Y) \equiv \{x_i, y_i\}_{i=1}^n, \quad (8)$$

where  $x$  means input vector and  $y \in \{1, \dots, G\}$  means the number corresponding to its class label. The number of training sample is  $n$ , and  $i$  is its index.  $G$  dimensional vector  $s$  which has  $I(y = g)$  in  $g$ th element is defined as another

way of label  $y$ .  $S = (s_1, \dots, s_n)$  means set of  $s_i$ . Then  $I(R)$  is an indicator function with the following conditions for  $R$  as in (9):

$$I(R) = \begin{cases} 1 & \text{if } R \text{ is true,} \\ 0 & \text{otherwise.} \end{cases} \quad (9)$$

When  $p$  denotes the number of binary classification problems (binary classifiers), multiclass classification problem changes to binary classification problems with  $p \times G$  matrix  $W \in \{1, -1\}^{p \times G}$  which is called code table.  $W$  consists of 1 on the diagonal and  $-1$  on the rest. When  $G = 3$ ,  $W$  is defined as in (10):

$$W = \begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}. \quad (10)$$

If the components of  $W$  is 1, label 1 is allocated to learn and classify on binary classification problem. Using this code table  $W$ , class label is divided into  $p$  dimensional vector where  $z = Ws$  is called encoding label.

Let  $Z = WS$  denote set of the encoding labels, and let  $z_{ij}$  denote the  $j$ th element of the encoding labels  $z_i = Ws_i$  of the  $i$ th class label  $y_i$ . When the code table is given, the set of input vectors  $X$  and the set of encoding labels  $Z_j \equiv (z_{1j}, \dots, z_{nj})$  which means the  $j$ th row of  $Z$  are used for learning. By this time, encoding labels are 1 or  $-1$ . Allwein et al [21]. have suggested an encoding method which adds a 0 to the code table. In the method, code table is redefined as  $W \in \{1, -1, 0\}^{p \times G}$ . The code table ( $G = 3$ ) is shown by (11):

$$W = \begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix}. \quad (11)$$

Learning was conducted for sample data if the encoding label  $z_{ij}$  is either 1 or  $-1$ . If the encoding label is 0, it is not used. We applied this code table  $W$ . Binary classifier which is created by learning is defined as hypothesis  $h$ , and  $H$  is the set of hypotheses. Figure 20 shows the encoding and the decoding with ECOC method when three-class classification is conducted.

Classification for evaluation data is conducted by the decoding based on the Hamming distance. Evaluation data and its class label are defined as in (12):

$$(\widehat{X}, \widehat{Y}) \equiv \{\widehat{x}_i, \widehat{y}_i\}_{i=1}^m. \quad (12)$$

The number of evaluation data is  $m$ . The Hamming distance  $d_i$  is defined as in (13):

$$d_i(W(r), H(\widehat{X})) = \sum_{j=1}^p \left( \frac{1 - \text{sign}(W(r, j) h_j(\widehat{x}_i))}{2} \right). \quad (13)$$

Step 1: Initialize with  $D_1(i) = 1/N$ .

Step 2: Perform the following for  $t = 1, \dots, T$ .

- (1) Learn the weak classifier based on distribution  $D_t$ . In other words, minimize the value for Formula (7), and find  $h_t: X \rightarrow Y$ .
- (2) Use the error rate and calculate reliability  $\alpha_t \in R$  as shown below.
 
$$\alpha_t = \frac{1}{2} \ln \{(1 - \epsilon_t)\epsilon_t^{-1}\}$$
- (3) Use the following formula to update distribution  $D_t$ .
 
$$D_{t+1}(i) = D_t(i) \exp \{-\alpha_t y_i h_t(x_i)\} Z_t^{-1}$$

$$Z_t = \sum_{i=1}^n D_t(i) \exp \{-\alpha_t y_i h_t(x_i)\}$$
 $Z_t$  is a normalization factor so that  $\sum_{i=1}^n D_{t+1}(i) = 1$ .

Step 3: For the final classifier, the strong classifier  $H(x)$  is found by the following equation after making a weighted majority decision using the reliability of all the weak classifiers.

$$H(x) = \text{sign} \left( \sum_{i=1}^T \{\alpha_i h_i(x)\} \right)$$

ALGORITHM 1: AdaBoost for a two-group classification problem.

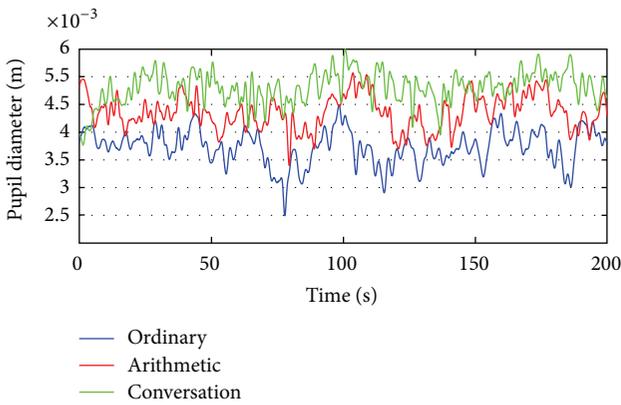


FIGURE 16: Example waveform of pupil diameter.

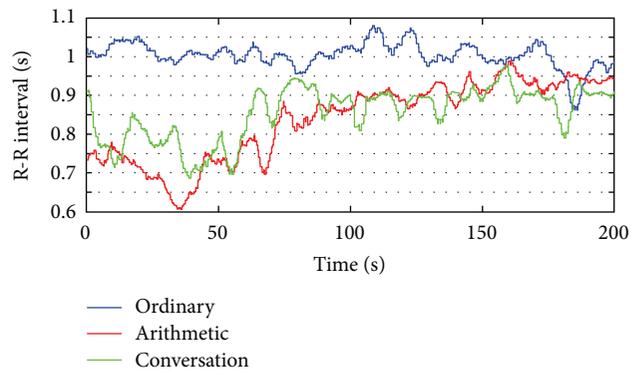


FIGURE 18: Change of heart rate R-R interval. It gets lowered when under cognitive distraction compared to ordinary driving.

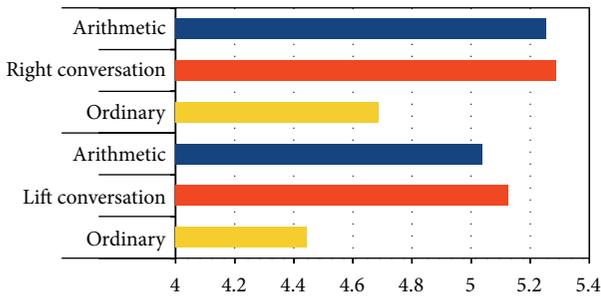


FIGURE 17: Average values of the pupil diameter. It increases while the driver is under cognitive distraction as compared to ordinary driving.

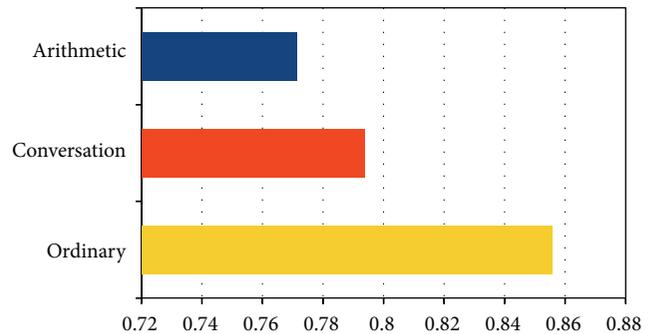


FIGURE 19: Average value of heart rate R-R interval. Doing arithmetic or having conversation lowered it.

$W(r)$  means  $r$ th column of  $W$ , and  $h_j(\hat{x}_i)$  means the hypothesis  $h$  which corresponds to  $z_{ij}$ . As (14) shows,  $r$ , which minimizes the Hamming distance, is the class label to obtain:

$$\hat{y}_i = \arg \min_r (W(r), H(\hat{X})) \tag{14}$$

7.2.2. Loss-Based Decoding Error-Correcting Output Codes (LD-ECOC). Output value from the binary classifier (like SVM or AdaBoost) is either 1 or -1 because of the binarization by signum function. Here, the original output value from classifier means Euclidean distance from the separating

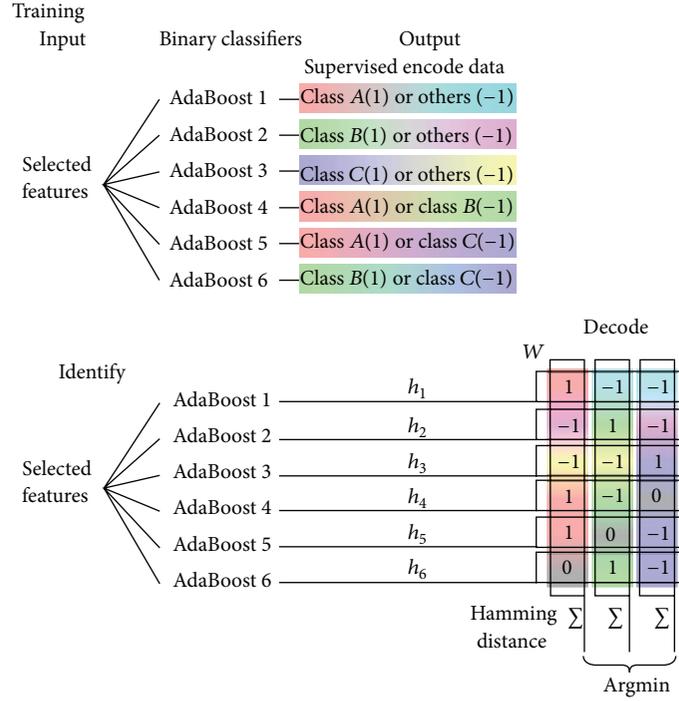


FIGURE 20: Encoding and decoding with error-correcting output codes or ECOC ( $G = 3$ ).

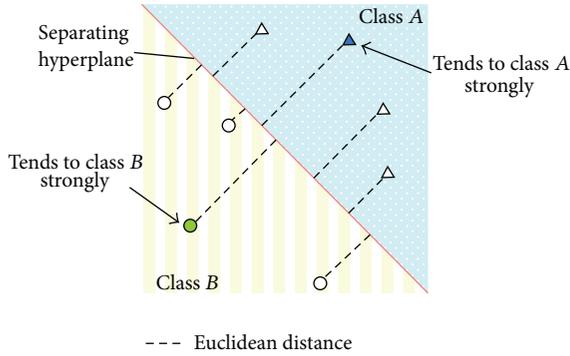


FIGURE 21: Example of class tendency of output from binary classifier. Bigger, positive value tends to class A; smaller, negative value tends to class B.

hyperplane which is created by learning. Thus, it can be interpreted as reliability in the class which is classified. For example, let the label of class A be 1 and the label of class B be -1. The bigger positive the output value is, the stronger it tends to class A. Similarly, the smaller negative the output value is, the stronger it tends to class B (see Figure 21).

The variation of features depends on the level of the driver's cognitive distraction. It is possible to capture the variation in more details by considering the class tendency of output value from each of the binary classifiers. Therefore, it is regarded that applying class tendency is effective to improve the identification capability of cognitive distraction. We suggest employing LD-ECOC which interprets hypothesis  $h$  as loss function. LD-ECOC is the method which extends the

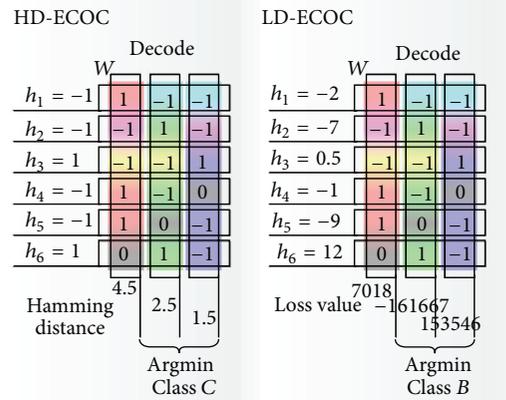


FIGURE 22: Error-correcting output codes or ECOC; left showing Hamming Decoding ECOC (HD-ECOC), and the right showing Loss-based Decoding ECOC (LD-ECOC).

decoding rule of the HD-ECOC. Loss value  $l_i$  is defined as in (15):

$$\begin{aligned}
 l_i(W(r), H(\hat{X})) &= \sum_{j=1}^p \begin{cases} \exp(|h_j(\hat{x}_i)|) & \text{if } W(r, j) h_j(\hat{x}_i) < 0 \\ -\exp(|h_j(\hat{x}_i)|) & \text{if } W(r, j) h_j(\hat{x}_i) > 0 \\ 0 & \text{if } W(r, j) h_j(\hat{x}_i) = 0 \end{cases} \quad (15) \\
 &= \sum_{j=1}^p -\text{sign}(W(r, j) h_j(\hat{x}_i)) \exp(|h_j(\hat{x}_i)|).
 \end{aligned}$$

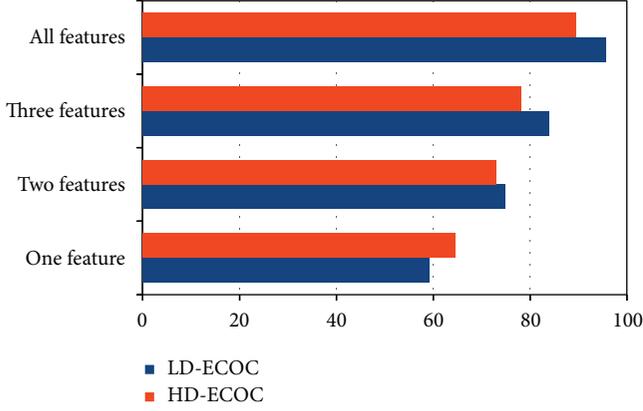


FIGURE 23: Average of the identification capability by the number of features.

Using exponential function, if the product of the code table  $W$  and the hypothesis  $h$  is positive,  $l_i$  becomes negative. Similarly, if the product of  $W$  and  $h$  is negative,  $l_i$  becomes positive. Additionally, if the product is 0,  $l_i$  becomes 0 and it is not used to classify. If the calculated loss value is small, it is possible to classify into proper class. Similarly in the case of the Hamming distance,  $r$ , which minimizes the loss value, is the class label to obtain as in (16):

$$\hat{y}_i = \arg \min_r l_i(W(r), H(\hat{X})). \quad (16)$$

Each example of the decoding process is shown in Figure 22. The left part shows the decoding of the HD-ECOC and the right part shows the decoding of the LD-ECOC method. For example, let output from binary classifiers be assumed as  $h_j = \{-2, -7, 0.5, -1, -9, 12 \mid j \in [1, 6]\}$ . When signum function is used, output is converted to  $\text{sign}(h_j) = \{-1, -1, 1, -1, -1, 1 \mid j \in [1, 6]\}$ . In the case of HD-ECOC, the Hamming distance 1.5 between 3rd column of  $W$  and output  $\text{sign}(h_j)$  is the minimum and it is classified as class  $C$ . However, in the case of LD-ECOC, loss value of 2nd column is the minimum and it is classified as class  $B$ .

**7.3. How to Evaluate the Overall Classification Performance.** In this study, the features which were used for detection of driver's cognitive distraction when cognitive loads were imposed on the drivers were standard deviation of the gaze angles, its tracking rates (quality index), standard deviation of head rotation angles, its tracking rates (quality index), the average value of pupil diameter, and the average value of heart rate RRI. These were used as input to SVM and AdaBoost, and learning and evaluation of the classification were performed with the output being whether a state of cognitive distraction was present or not (binary value: +1 or -1).

We used a twofold cross-validation method, which was generally used as a method for evaluating classification accuracy of unknown data. Ordinary driving (nongenerative) was defined as positive data (+1) and driving with a cognitive load was defined as negative data (-1). Each test subject data was divided into two sets, with test set  $X_1$  used to evaluate

TABLE 2: Comparison of capability with SVM and AdaBoost.

Selected features	Learning method	Cognitive load	Average accuracy	$F$ value
Gaze + head	SVM	Conversation	84.2	84.5
		Arithmetic	77.1	74.5
	AdaBoost	Conversation	86.1	87.4
		Arithmetic	81.6	83.1

performance when test set  $X_2$  was used for learning. In the same way, test set  $X_2$  was used for evaluation when test set  $X_1$  was used for learning.

In this study,  $O_p$  is defined as the positive output data,  $O_n$  is defined as the negative output data,  $T_p$  is defined as true data of the positive, and  $T_n$  is defined as true data of the negative. Then we used *Accuracy*, *Precision*, *Recall*, and the overall classification performance index  $F$  as the classification indexes following equation (17) to calculate those classification indexes:

$$\begin{aligned} \text{Accuracy} &= \frac{(O_p \cap T_p) \cup (O_n \cap T_n)}{T_p \cup T_n} \\ \text{Precision} &= \frac{O_p \cap T_p}{O_p} \\ \text{Recall} &= \frac{O_p \cap T_p}{T_p} \\ F &= \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}. \end{aligned} \quad (17)$$

In an experiment of later section, we evaluated the performance of identification using mainly *Accuracy* and  $F$  value.

## 8. Results and Discussions

**8.1. Enhancing Performance by Adding Extra Features.** Our test subjects included seven males and three females, all of whom gave their informed consent to participate in the series of experiments.

In accordance with the previous study, we used the SVM<sup>light</sup> as a software. The Gauss kernel was used as the kernel function. A parameter  $C$  indicates the relaxation of the restriction conditions in the soft margin SVM. And a parameter  $\sigma$  indicates the spread of the Gauss Kernel distribution. These parameters were determined by a grid search. In the experiment described later, because the quantity of positive data was twice the quantity of negative data, a cost factor  $cf = 0.5$  was set.

Table 2 shows the results from only two visual features, gaze and head direction. In order to verify the methodology of this study, SVM (conventional) and AdaBoost (proposed) are used to detect distraction. The capability is compared with *Accuracy* and  $F$  value.

The average *Accuracy* for detecting cognitive distraction using visual features by SVM was 77.1% when the driver was

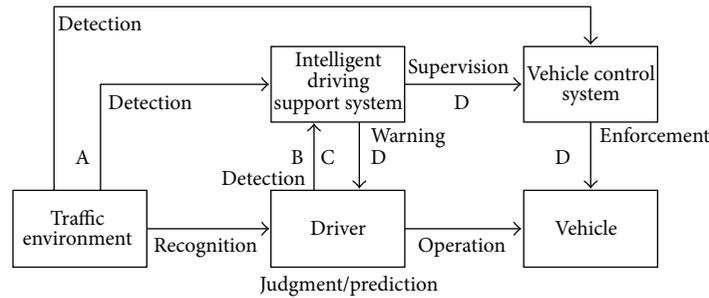


FIGURE 24: Intelligent driver support system.

TABLE 3: Enhancement of performance by adding another feature to visual information.

Selected features	Learning method	Cognitive load	Average accuracy	F value
Gaze + head + pupil	SVM	Conversation	89.4	89.9
		Arithmetic	85.6	86.7
	AdaBoost	Conversation	89.4	90.6
		Arithmetic	89.8	87.7
Gaze + head + RRI	SVM	Conversation	90.1	90.2
		Arithmetic	88.1	87.9
	AdaBoost	Conversation	88.3	90.0
		Arithmetic	89.3	90.3
Gaze + head + pupil + RRI	SVM	Conversation	<b>92.4</b>	92.2
		Arithmetic	91.8	92.0
	AdaBoost	Conversation	91.0	<b>92.3</b>
		Arithmetic	<b>92.0</b>	<b>93.1</b>

also doing arithmetic. So, it was demonstrating the same level of *accuracy* as the EU AIDE Project [10] which achieved an average *accuracy* of 74.2% by SVM. From the previous results, it can be said that this methodology and learning are appropriate in use for the detection of driver's cognitive distraction as well as a previous study. By AdaBoost, the average *accuracy* for detecting cognitive distraction using visual information with arithmetic was 81.6% and its *F* value was 83.1%. Judging from the *F* value in this study, AdaBoost has higher cognitive detection performance than that of SVM.

Next, Table 3 shows a comparison of the results when the average value of heart rate RRI and the pupil diameter was added to the only two learning features of visual information, gaze, and head direction, used in Table 2. Presence of additional features improved the detection performance such as the *F* value increased more than that of only the two pieces of visual information regardless of the kind of cognitive loads and learning methods. And the detection performance of AdaBoost was better when compared with visual information only as well as for SVM too.

Based on the previous results, adding both pupil diameter and the average value of heart rate RRI as recognition features can be said to be more effective to further enhance the detection performance under cognitive distraction. And the average value of heart rate RRI and the pupil diameter are said to be more effective features than visual information.

When a new feature is added, SVM also requires a support vector (SV) for use to classify in that dimension. In this study,

although the average heart rate RRI was added, the amount of change was smaller than the standard deviations (SD) of the gaze direction and head orientation, and one explanation is that learning of SV, which would properly maximize the margin, was not possible. On the other hand, with AdaBoost, we can explain that because the weak classifiers are learned successively, it is possible to perform learning with data that cannot be learned with SVM. This results in high detection performance and in even higher detection performance when the average heart rate RRI and pupil diameter were added.

In addition, there are differences in the learning methods used by SVM and AdaBoost. With a boosting algorithm such as AdaBoost, learning proceeds so that weight vectors are expressed with the smallest possible number of features. As a result, classification is performed with few features, and it is possible to analyze features that have a high contribution rate. On the other hand, SVM attempts to express the weight vectors using the smallest possible number of training samples, making it difficult to perform analysis of the features from the learning space. When a new feature such as pupil diameter or the average heart rate RRI is different from other features on time variation basis, SV may not be created easily in using SVM. AdaBoost may be easier in classification for features varying relatively in longer time span as shown in Figure 16 than other feature such as visual information as shown in Figure 18.

This means that the optimization of the classification may not be easy for the SVM. However, the AdaBoost, which is one of ensemble learning method, can be said to be easier to

TABLE 4: Comparing the detection capability with various combinations of individual features using AdaBoost.

Selected features	Cognitive load	Average accuracy	F value
Gaze	Conversation	67.6	68.8
	Arithmetic	66.7	67.7
	Both	72.8	80.5
Head	Conversation	84.5	86.2
	Arithmetic	78.3	81.2
	Both	84.0	89.0
Pupil	Conversation	85.0	84.2
	Arithmetic	86.2	86.5
	Both	87.0	90.0
RRI	Conversation	72.3	71.4
	Arithmetic	77.3	75.3
	Both	78.3	84.3
Gaze + head	Conversation	86.1	87.4
	Arithmetic	81.6	83.1
	Both	85.8	89.9
Gaze + pupil	Conversation	86.0	86.3
	Arithmetic	84.5	84.6
	Both	86.8	90.2
Gaze + RRI	Conversation	76.1	77.6
	Arithmetic	79.1	78.5
	Both	79.0	84.3
Head + pupil	Conversation	89.7	90.8
	Arithmetic	87.1	88.7
	Both	89.7	92.6
Head + RRI	Conversation	87.8	89.6
	Arithmetic	87.6	88.4
	Both	89.6	92.9
Pupil + RRI	Conversation	89.7	90.0
	Arithmetic	88.1	87.6
	Both	90.2	92.8
Gaze + head + pupil	Conversation	89.4	90.6
	Arithmetic	89.8	87.7
	Both	90.2	93.0
Gaze + head + RRI	Conversation	88.3	90.0
	Arithmetic	89.3	90.3
	Both	90.3	93.3
Gaze + pupil + RRI	Conversation	89.9	90.1
	Arithmetic	89.0	89.4
	Both	91.1	93.5
Head + pupil + RRI	Conversation	90.6	92.0
	Arithmetic	89.8	91.2
	Both	92.5	94.7
Gaze + head + pupil + RRI	Conversation	<b>91.0</b>	<b>92.3</b>
	Arithmetic	<b>92.0</b>	<b>93.1</b>
	Both	<b>92.9</b>	<b>95.1</b>

classify the features. In the future, when many features were selected from among physiological signals, AdaBoost will be superior because of its high ability to efficiently select features and the ease of evaluating the features' contribution rates.

*8.2. Contribution to Average Accuracy from Individual Features.* We compared the influence of individual features (which were standard deviation (SD) of gaze angle, SD of head rotation angle, pupil diameter, and heart rate RRI) by means of using the AdaBoost to detect a state of driver

TABLE 5: Identification capability for each feature set.

Set	Selected features	Cognitive load	Hamming decoding	Loss-based decoding
A	Gaze	No load	70.6	42.5
		Conversation	45.6	45.0
		Arithmetic	56.4	43.0
		Average	57.5	43.5
B	Head	No load	74.8	59.7
		Conversation	52.8	55.8
		Arithmetic	69.3	54.8
		Average	65.6	56.8
C	Pupil	No load	88.3	84.1
		Conversation	56.0	71.7
		Arithmetic	76.1	76.8
		Average	73.5	77.6
D	RRI	No load	74.7	65.7
		Conversation	42.5	48.6
		Arithmetic	69.0	64.0
		Average	62.1	59.4
E	Gaze + head	No load	81.8	70.4
		Conversation	58.2	64.5
		Arithmetic	63.2	56.7
		Average	67.7	63.9
F	Gaze + pupil	No load	89.0	83.5
		Conversation	59.6	74.1
		Arithmetic	75.0	76.1
		Average	74.5	77.9
G	Gaze + RRI	No load	79.9	70.2
		Conversation	53.0	61.7
		Arithmetic	69.1	67.0
		Average	67.3	66.3
H	Head + pupil	No load	91.8	87.1
		Conversation	66.1	81.6
		Arithmetic	75.3	78.6
		Average	77.7	82.4
I	Head + RRI	No load	83.8	76.2
		Conversation	59.6	70.1
		Arithmetic	75.2	75.5
		Average	72.8	73.9
J	Pupil + RRI	No load	91.2	88.3
		Conversation	63.1	83.6
		Arithmetic	80.3	83.9
		Average	78.4	85.3
K	Gaze + head + pupil	No load	92.5	87.9
		Conversation	66.9	84.0
		Arithmetic	74.5	80.2
		Average	78.0	84.0
L	Gaze + head + RRI	No load	86.3	80.2
		Conversation	60.4	74.8
		Arithmetic	75.0	75.5
		Average	73.7	76.9
M	Gaze + pupil + RRI	No load	91.5	88.7
		Conversation	66.5	86.3
		Arithmetic	80.5	83.7
		Average	79.5	86.2
N	Head + pupil + RRI	No load	92.5	89.9
		Conversation	71.1	88.1
		Arithmetic	81.2	87.4
		Average	81.6	88.5

TABLE 5: Continued.

Set	Selected features	Cognitive load	Hamming decoding	Loss-based decoding
O	Gaze + head + pupil + RRI	No load	<b>95.5</b>	<b>95.0</b>
		Conversation	<b>82.5</b>	<b>96.8</b>
		Arithmetic	<b>89.6</b>	<b>95.5</b>
		Average	<b>89.5</b>	<b>95.8</b>

distraction as the driver was subjected to cognitive loads. The results of average accuracy when the driver was subjected to either conversation or arithmetic load and also under both the cognitive loads are shown in Table 4.

When just one feature was used to detect distraction, in descending order of individual contribution, the average accuracy was pupil diameter, standard deviation of head rotation angle, heart rate RRI, and standard deviation (SD) of gaze angle. And it was in the same order for  $F$  value too. In cases when two features were combined, the average accuracy was SD of head rotation angle plus heart rate RRI, pupil diameter plus heart rate RRI, SD of head rotation angle plus pupil diameter, SD of gaze angle plus pupil diameter, visual information, and SD of gaze angle plus heart rate RRI, which was the same order for  $F$  value. The top four common results in order for average accuracy were combination of all the four features, SD of head rotation angle plus pupil diameter plus heart rate RRI, pupil diameter plus heart rate RRI, and SD of head rotation angle plus heart rate RRI, which were also demonstrating the same order in  $F$  value. From these results, it can be said that some combination of four recognition features may enhance the capability to detect a state of driver's cognitive distraction according to a specification of a driver safety supporting system.

**8.3. Multiclass Identification for Degree of Cognitive Load.** There were ten human subjects (ages from 21 to 23 years, the average being 21.9) in the experiment. All subjects gave their informed consent before the experiment. To identify cognitive distraction, both HD-ECOC and LD-ECOC methods were used and AdaBoost was employed as the binary classifier.

Identification capability of each feature set is shown as in Table 5. When LD-ECOC was used for set which consists of all features, identification capability was 95.8%.

Figure 23 shows the average of the identification capability by the number of features. Comparing with sets  $A$  to  $D$  of Table 5 that apply just a single feature each and set  $E$  or later sets that apply multiple features, identification capability was improved with increasing the number of features in both HD-ECOC and LD-ECOC methods. Therefore, one conclusion can be derived that the multiclass identification results in higher accuracy by applying multiple features.

When just a single feature was applied to identify, the identification capability of the set  $C$  (pupil diameter) was the highest. Thus it is quite likely that the contribution of the pupil diameter for identification is the highest. Comparing between the sets  $F$ ,  $H$ , and  $J$ , each of whom applied the pupil diameter, and set  $J$  which also applied the RRI has

the highest identification capability. For this reason, it is likely that the second highest contribution is from the RRI. Additionally, comparing the sets  $M$  and  $N$ , we can see that the capability of the set  $N$  is higher. Thus, it is likely that the head rotation angle has higher contribution than the gaze angle. Consequently, the order of the contribution for identification is pupil diameter, RRI, head rotation angle, and gaze angle. So our experience is that the pupil and heartbeat information are more effective than the gaze and the head information that have been referred to be as dominant by Kuttila et al. earlier study [10].

Comparing HD-ECOC and LD-ECOC methods, identification capability of LD-ECOC is generally higher than HD-ECOC as described in [21]. However, identification capability of LD-ECOC tended to decrease when less features were applied. Especially, there was the obvious difference between HD-ECOC and LD-ECOC in the identification of normal driving of the set  $A$ . Output value from classifier of the LD-ECOC means class tendency. Therefore, if the classifier which strongly declares wrong class tendency exists, the possibility of wrong identification could increase.

In the case of set  $A$  and set  $B$  that have low contribution mentioned previously, it is likely that some classifiers declare wrong class tendency strongly. Meanwhile, there is a possibility that the degree of wrong identification is reduced even if wrong classifier exists in the HD-ECOC. Because the decoding of the HD-ECOC equally handles the wrong tendency and others with signum function. However, when several features including pupil diameter which contributes to identify properly were applied, the identification capability of LD-ECOC is getting higher than HD-ECOC. Therefore, LD-ECOC method is effective for multiclass identification.

Incorrect identification tends to occur, because separating hyperplane of binary classifier tends to be unstable in the case of less features. In consequence, it is possible that the capability of HD-ECOC is better than LD-ECOC in the case of only a single feature. When features were added, the capability of LD-ECOC tended to get higher than HD-ECOC. It is likely that the separating hyperplane became stable by increasing the number of features, and reliability of loss value which reflects the tendency of the class was improved.

In addition, comparing identification of a single feature and multifeatures, there was high capability in the identification of multifeatures. Therefore, applying multifeatures can be effective to identify the level of cognitive distraction.

## 9. Suggestion of Potential Driver Support System

We identified the relationship among driver's behavior, psychosomatic states, and expected intelligent safety systems in Table 6. The relationship was derived from our previous survey [3].

For example, when a driver's psychosomatic state monitoring system could detect a driver's distraction, the system could provide appropriate information to the driver, give warnings or intervene in the driver's operation. And it is helpful to minimize the risk of incidents in combination with

TABLE 6: Expected intelligent safety systems.

Driver's behavior	Psychosomatic states	Expected intelligent safety systems
Inappropriate assumption	Normal	Providing information from the roadside infrastructure (1) ITS service (AHS) (2) Driving safety support system (3) Traffic info collected by prove cars
	Haste and distraction	Monitoring of surroundings (1) Precrash safety system
Failure to make safety check	Haste and distraction	(2) Night view system
Distracted driving	Distraction	(3) Rear-end monitoring system
Not looking forward	Distraction	(4) Side-view monitoring system (blind spot monitoring)

information surrounding the vehicle provided by the surrounding monitoring system. If the driver's psychosomatic state is normal but driver makes inappropriate assumptions while driving, traffic safety information from road infrastructures such as represented by intelligent transportation systems services (e.g., VICS in Japan) and the driving safety support system (e.g., DSSS in Japan) could be employed. The realization of intelligent driver support systems activated by detecting each driver's psychosomatic states is expected to be a way to help minimize road traffic safety risks.

It is clear that future intelligent drive support systems should have functions such as recognizing potential risks during ordinary driving, processing in realtime, making correct driving decisions, and intervening into vehicle control system adaptively. Figure 24 shows the functional concept for such an integrated intelligent driver support system.

The system works as follows.

- (a) Detect and estimate risk factors in the environment.
- (b) Detect and estimate a state of the driver with regard to driver's behavior and psychosomatic states.
- (c) Estimate the reliability of the driver's decision concerning risk (presence of human error).
- (d) Evaluate the driver capacity for receiving information and warnings. If a driver's capacity is insufficient or the danger exceeds the human ability to react, the intelligent drive support system intervenes, either via the vehicle control system or directly, to operate the vehicle safety systems.

## 10. Conclusion

It was identified that one of the driver's psychosomatic states being involved in a traffic incident immediately before a traffic accident was distraction by means of using internet-based survey. Therefore we focused on the driver's state, especially distraction, and tried to detect the state. The states of cognitive distraction in the subjects were reproduced by subjecting them with arithmetic task and conversation task. Then, using previous test procedure, we established the methodology in higher performance for detection of driver's cognitive distraction based on both SVM and AdaBoost.

As a result, we were able to enhance the performance for detection of driver's cognitive distraction by adding the average value of pupil diameter from camera images and the average value of heart rate RRI from the ECG waveforms as the main pattern recognition features as opposed to standard deviations of gaze angle and head rotation angle as features which were adopted in Kutilla et al. earlier study. It should be noted that the conversation task given at the experiment of this study was to explain the effect on use of visuospatial working memory. But, the contents of the conversation (whether visuospatial or phonological) may affect the experimental results. It is a task we have left for a future work.

We suggested use of multiclass identification method of driver's cognitive distraction. We employed error-correcting output codes or ECOC method which has the advantage of low computational cost and ease of implementation. Two methods (HD-ECOC and LD-ECOC) were employed to identify cognitive distraction. The maximum identification capability was found to be 95.8% when the LD-ECOC method was used and when all the features (standard deviation of the gaze angles, its tracking rates (quality index), standard deviation of head rotation angles, its tracking rates (quality index), the average value of pupil diameter, and the average value of heart rate RRI) were applied. While, earlier studies concentrated only with the identification of cognitive distraction or not, multiclass identification of cognitive distraction becomes possible with this method. When we create a driver's psychosomatic state monitoring system for potential intelligent drive supporting system, it may be more effective to apply all physiological features of this study in order to obtain more accurate and more redundant detection performance for cognitive distraction. We also intend to seek possibilities for use of this method in potential intelligent driving support systems which may prevent a traffic accident effectively.

Future issues include improving a methodology in order to further enhance the detection accuracy for driver's cognitive distraction and identifying significant features from among new physiological signals such as the blinking rate, so that they can be used in an intelligent driver monitoring system. To identify even more details of cognitive distraction, it is necessary to increase the number of classes. But then,

how to deal with the computational cost should be considered because the required number of binary classifiers increases exponentially. Such a system is one important constituent technology of a potential driver supporting system at the area of the prevention safety aimed at lowering the number of the traffic accidents.

## Conflict of Interests

The authors of this paper have no conflict of interests to disclose.

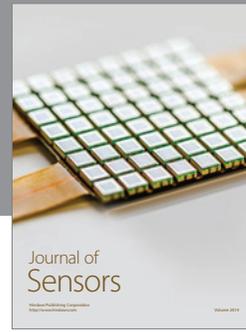
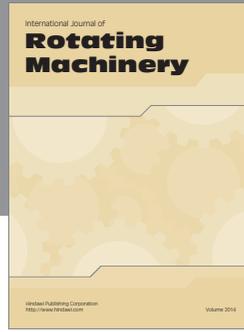
## Acknowledgments

The present work received generous support from the TOYOTA Info Technology Center Co., LTD. The authors are grateful to Kishimoto Yoshifumi and Atsushi Nagase, for their important contributions to the experiments. They also appreciate the computing services offered by the high-performance computing cluster at the Institute of Information Science and Technology at Aichi Prefectural University.

## References

- [1] "A suggestion by the science council of japan about accident investigation method," The expert committee for safety engineering in Science Council of Japan, 2000.
- [2] J. R. Treat, N. S. Tumbas, S. T. McDonald et al., "Tri-level study of the causes of traffic accidents: final report. Executive summary," Tech. Rep. DOT-HS-034-3-535-79-TAC(S), Transportation Research Institute, 1979.
- [3] M. Miyaji, M. Danno, and K. Oguri, "Analysis of driver behavior based on experiences of road traffic incidents investigated by means of questionnaires for the reduction of traffic accidents," *International Journal of ITS Research*, vol. 6, no. 1, pp. 47–56, 2008.
- [4] S. G. Klauer, T. A. Dingus, V. L. Neale, J. Sudweeks, and D. J. Ramsey, "The impact of driver inattention on near-crash/crash risk: an analysis using the 100-car naturalistic driving study data," Tech. Rep. DOT HS 810 594, National Highway Traffic Safety Administration, 2006.
- [5] K. Young, J. D. Lee, and M. A. Regan, *Driver Distraction: Theory, Effects, and Mitigation*, CRC Press, 2008.
- [6] C. Ahlström and K. Kircher, "Review of real-time visual driver distraction detection algorithms," in *Proceedings of the 7th International Conference on Methods and Techniques in Behavioral Research (MB '10)*, August 2010.
- [7] M. Sodhi, B. Reimer, and I. Llamazares, "Glance analysis of driver eye movements to evaluate distraction," *Behavior Research Methods, Instruments, and Computers*, vol. 34, no. 4, pp. 529–538, 2002.
- [8] Y. Zhang, Y. OwecHko, and J. Zhang, "Driver cognitive workload estimation: a data-driven perspective," in *Proceedings of the 7th International IEEE Conference on Intelligent Transportation Systems (ITSC '04)*, pp. 642–647, October 2004.
- [9] J. Engström, E. Johansson, and J. Östlund, "Effects of visual and cognitive load in real and simulated motorway driving," *Transportation Research Part F*, vol. 8, no. 2, pp. 97–120, 2005.
- [10] M. Kuttila, M. Jokela, G. Markkula, and M. R. Rué, "Driver distraction detection with a camera vision system," in *Proceedings of the 14th IEEE International Conference on Image Processing (ICIP '07)*, vol. 6, pp. VI201–VI204, September 2007.
- [11] J. L. Harbluk, Y. I. Noy, P. L. Trbovich, and M. Eizenman, "An on-road assessment of cognitive distraction: impacts on drivers' visual behavior and braking performance," *Accident Analysis and Prevention*, vol. 39, no. 2, pp. 372–379, 2007.
- [12] Y. Liang, M. L. Reyes, and J. D. Lee, "Real-time detection of driver cognitive distraction using support vector machines," *IEEE Transactions on Intelligent Transportation Systems*, vol. 8, no. 2, pp. 340–350, 2007.
- [13] Y. Liang, J. D. Lee, and M. L. Reyes, "Nonintrusive detection of driver cognitive distraction in real time using Bayesian networks," *Transportation Research Record*, vol. 2018, no. 2007, pp. 1–8, 2008.
- [14] A. Azman, Q. Meng, E. A. Edirisinghe, and H. Azman, "Non-intrusive physiological measurement for driver cognitive distraction detection: eye and mouth movements," *International Journal of Advanced Computer Science*, vol. 1, no. 3, pp. 92–99, 2011.
- [15] D. L. Strayer and F. A. Drews, "Profiles in driver distraction: effects of cell phone conversations on younger and older drivers," *Human Factors*, vol. 46, no. 4, pp. 640–649, 2004.
- [16] M. Blanco, W. J. Biever, J. P. Gallagher, and T. A. Dingus, "The impact of secondary task cognitive processing demand on driving performance," *Accident Analysis and Prevention*, vol. 38, no. 5, pp. 895–906, 2006.
- [17] A. Benedetto, A. Calvi, and F. D'Amico, "Effects of mobile telephone tasks on driving performance: a driving simulator study," *Advances in Transportation Studies*, no. 26, pp. 29–44, 2012.
- [18] V. N. Vapnik, *The Nature of Statistical Learning Theory*, vol. 12 of *Statistics for Engineering and Information Science*, Springer, 1999.
- [19] J. Lee, M. Reyes, and D. McGehee, "Safety vehicles using adaptive interface technology (task 5) final report: Phase 1," Tech. Rep., Research and Innovative Technology Administration Volpe National Transportation System Center, 2004.
- [20] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," in *Proceedings of the 2nd European Conference on Computational Learning Theory (EuroCOLT '95)*, pp. 23–37, 1995.
- [21] E. L. Allwein, R. E. Schapire, and Y. Singer, "Reducing multiclass to binary: a unifying approach for margin classifiers," *Journal of Machine Learning Research*, vol. 1, no. 2, pp. 113–141, 2001.
- [22] E. Hartman, "Driver vision requirement," Tech. Rep. 700392, Society of Automotive Engineers, 1970.
- [23] T. W. Victor, J. L. Harbluk, and J. A. Engström, "Sensitivity of eye-movement measures to in-vehicle task difficulty," *Transportation Research Part F*, vol. 8, no. 2, pp. 167–190, 2005.
- [24] D. Kahneman, B. Tursky, D. Shapiro, and A. Crider, "Pupillary, heart rate, and skin resistance changes during a mental task," *Journal of Experimental Psychology*, vol. 79, no. 1, pp. 164–167, 1969.
- [25] S. Sternberg, "High-speed scanning in human memory," *Science*, vol. 153, no. 3736, pp. 652–654, 1966.
- [26] J. D. Crawford and T. Vilis, "Axes of eye rotation and Listing's law during rotations of the head," *Journal of Neurophysiology*, vol. 65, no. 3, pp. 407–423, 1991.
- [27] C.-W. Hsu and C.-J. Lin, "A comparison of methods for multiclass support vector machines," *IEEE Transactions on Neural Networks*, vol. 13, no. 2, pp. 415–425, 2002.

- [28] R. Rifkin and A. Klautau, "In defense of one-vs-all classification," *Journal of Machine Learning Research*, vol. 5, pp. 101–141, 2004.
- [29] T. G. Dietterich and G. Bakiri, "Solving multiclass learning problems via error-correcting output codes," *Journal of Artificial Intelligence Research*, vol. 2, no. 1, pp. 263–286, 1994.



**Hindawi**

Submit your manuscripts at  
<http://www.hindawi.com>

