

Research Article

Analysis of Temporal Relationships between Eye Gaze and Peripheral Vehicle Behavior for Detecting Driver Distraction

Takatsugu Hirayama, Kenji Mase, and Kazuya Takeda

Graduate School of Information Science, Nagoya University, Nagoya 464-8601, Japan

Correspondence should be addressed to Takatsugu Hirayama; hirayama@is.nagoya-u.ac.jp

Received 24 September 2012; Revised 25 June 2013; Accepted 5 August 2013

Academic Editor: Klaus Bengler

Copyright © 2013 Takatsugu Hirayama 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.

A car driver's cognitive distraction is a main factor behind car accidents. One's state of mind is subconsciously exposed as a reaction reflecting it by external stimuli. A visual event that occurs in front of the driver when a peripheral vehicle overtakes the driver's vehicle is regarded as the external stimulus. We focus on temporal relationships between the driver's eye gaze and the peripheral vehicle behavior. The analysis result showed that the temporal relationships depend on the driver's state. In particular we confirmed that the timing of the gaze toward the stimulus under the distracted state induced by a music retrieval task using an automatic speech recognition system is later than that under a neutral state while only driving without the secondary cognitive task. This temporal feature can contribute to detecting the cognitive distraction automatically. A detector based on a Bayesian framework using this feature achieves better accuracy than one based on the percentage road center method.

1. Introduction

Driver distraction is a diversion of attention away from activities critical for safe driving toward a competing activity [1] and is a large risk factor that causes accidents [2]. Note that distraction differs from fatigue [3] which is defined as a state that disables one from continuing the activity [4]. Many researchers have developed driver distraction monitoring systems to maintain safety while driving by considering different types and levels of distraction [3]. The National Highway Traffic Safety Administration (NHTSA) classifies distractions into (1) cognitive distraction, (2) visual distraction, (3) auditory distraction, and (4) biomechanical distraction from the viewpoint of the driver's functionality [2]. Cognitive distraction can be considered as an internal state of the driver. It is difficult to sense this from outside. The other distractions are external factors that disturb the activity and can be observed more easily. We focus on cognitive distraction and seek novel findings to automatically detect it.

In the past few decades, a number of methods for detecting distraction have been proposed [3]. The methods fall into the following five categories based on the types of measures: (1) subjective report measures, (2) driver biological measures, (3) driving performance measures, (4) driver physical measures, and (5) hybrid measures. Among these measures, subjective report measures and driver biological measures are not suitable under real driving conditions. Driving performance measures as indicated by steering, braking behavior, and so forth are suitable for detecting visual distraction [5]. Even if a system can detect such overt behaviors that are more directly linked with risk, for maintaining safety, the timing may be too late to provide support to the driver after the detection.

Eye-gaze measure, which is one of the driver's physical measures, is a useful measurement of visual distraction especially for In-Vehicle Information System (IVIS) and Advanced Driver Assistance System (ADAS) assessment as specified in the existing standards ISO 15007-1 [6] and ISO/TS 15007-2 [7], and it has the potential for capturing symptoms of cognitive distraction [8]. An eye-gaze pattern could be used to discriminate driving while performing a secondary cognitive task from driving only [9]. Drivers under cognitive distraction had fewer saccades per unit time, which was consistent with less exploration of the driving environment [10]. Saccades may be a valuable index of mental workload [11]. Miyaji et al. reported that the standard deviations of eye movement and head movement could be suitable for detecting cognitive distraction that caused gaze concentration and

slow saccades when drivers looked at the roadway [12]. Kircher et al. indicated the percentage of time that the driver spent observing the road ahead, which is called the percentage road center (PRC) of gaze direction, was more than 92% under cognitive distraction in a field study [13]. Johansson et al. have reviewed the existing gaze-based techniques and metrics for analyzing visual and cognitive distractions [8].

These approaches on driver's physics mainly measured only driver's gaze toward the road ahead or in-vehicle static objects without any regard for the peripheral traffic environment, which includes many scattering visual stimuli, or measured a rough correlation between spatial features of the gaze and the environment. They also need a long-term evaluation. To more flexibly support the driver, an improvement in the time resolution is required for the detection. We take account of short-term dynamics of cross media to detect cognitive distraction. In the field of human-computer interaction, some researchers have investigated a state of mind by analyzing the temporal relationships between eye movements and visual changes in the user interface [14-16]. The latent state is subconsciously exposed as a reaction reflecting it by external stimuli [17]. In controlled settings such as using a driving simulator, the detection response task (DRT) is an upcoming method of measuring visual and cognitive distractions [18-21], which asks the subject to respond via a device such as a button to visual, tactile, or acoustic stimuli. The response time relates to the distraction. However, it is difficult to give actual drivers on the road the task without disturbing the safety driving. In this work, our target is the temporal relationships between driver's gaze and peripheral vehicle behavior in a real driving situation. In particular it is the timing of the gaze toward the visual stimuli caused by the peripheral vehicle.

2. Analysis of Temporal Relationships between Driver Gaze and Peripheral Vehicle Behavior

2.1. Timing of Gaze Reaction to Overtaking Event. To analyze the temporal relationships, we focus on peripheral vehicle behaviors with a high level of visibility for the driver. When a peripheral vehicle (called the overtaking vehicle) overtakes the host vehicle driven by the driver, a visual change that occurs in front of the field of view can attract the driver's attention. We define this overtaking event as our target of analysis. The event has a base-point time $t_0(=0)$, a beginning time $t_b(=t_0 - T_d/2)$, and an ending time $t_e(=t_0 + T_d/2)$. t_0 is the time when the front position y_o of the overtaking vehicle in the direction of forward movement becomes equal to the front position y_h of the host vehicle. T_d is the duration of the overtaking event, which is a configuration parameter of the analysis and set in Section 4.2. Figure 1 shows the overtaking event.

We define saccade timing t_c and gaze timing t_g . The former is the time when the driver turns gaze toward the overtaking vehicle, whereas the latter is the time while the driver fixates the overtaking vehicle. The temporal relationships characterizing the gaze reaction to the overtaking vehicle are the time differences between the saccade timing t_c or the gaze



FIGURE 1: Overtaking event.

timing t_g and the base-point time t_0 of the event. Figure 2 shows the timing structure.

2.2. Hypothesis. Some researchers have investigated the correlations between gaze directions and traffic object positions, for example, road curvature, oncoming traffic, road signs, and pedestrians [22–24], and have shown that the correlations are high under the neutral state while driving only. Most of the researchers, however, dealt mainly with spatial correlation. On the temporal relationship between visual attention and external stimuli, Posner revealed that covert attention to them decreases the reaction time, and, conversely, distraction increases it [25]. We therefore propose the following hypothesis: the timing of when a driver gazes toward the overtaking event under a state of cognitive distraction is later than that under a neutral state.

3. Real-World Driving Database

We analyze a part of a database collected using the "NUDrive Vehicle" in Nagoya, Japan [26].

3.1. Data-Collection Vehicle. The "NUDrive Vehicle" was designed to synchronously record multimedia signals of driver performance (gas pedal, brake pedal, steering angle, velocity, acceleration, and position of the car), intervehicular distance, biological signals, videos, and audio signals. Various external sensors were mounted on a Toyota Hybrid Estima with a 2360 cc engine and automatic transmission and steering wheel on the right side. Figure 3 shows the data-collection vehicle. All the sensors used for recording were commercially available.

3.2. Participants. A total of 30 participants (10 males and 20 females) took part in the experiment. They were, on average, 39.0 years old (range of 29–52 years) and had held a driver's license for a mean period of 18.6 years (range of 8–32 years). They received 5000 Japanese yen as compensation for their participation.

3.3. Procedure. The participants first drove for a few minutes to get used to the vehicle and the sensors. Signals recorded



FIGURE 2: Timing of the driver's gaze toward the overtaking vehicle.

during the initial period were not used in this work. The experimenter monitored the experiment from the rear seat and indicated the route to the driver. During a particular period of driving, the participant performed a secondary hands-free task of retrieving and playing songs from a list of 635 titles from 248 artists using an automatic speech recognition system [27]. The secondary task artificially induced the cognitive distraction state. The experimenter instructed the participant to retrieve as many songs as possible; accordingly, within around 30 s of successfully retrieving each song, the participant had to retrieve another song. All experiments were performed on two-or three-lane highways. The experimental route was the same for all participants.

3.4. Measures. In this work, we analyzed the intervehicular distance measured by laser scanners and the driver's gaze direction extracted from the recorded video. The following are the details of the analyzed data and omit the account of other data.

3.4.1. Intervehicular Distance Scanning. Two laser scanners (front: RIEGL LMS-140i-80; rear: RIEGL LMS-Q120i), mounted on the front and back of the host vehicle shown in Figure 3, provided geometric information about the peripheral environment of the vehicle. The laser scanners covered 80-degree arcs at both front and back of the vehicle, to an effective range of about 100 m to the front and 55 m to the rear, but had blind areas at the left and right sides of the vehicle. The data were acquired at a sample frequency of 10 Hz. For tracking peripheral vehicles in the blind areas, we applied a Kalman filter to the data [28]. The dynamics of their position and velocity relative to the host vehicle could be estimated even if they were outside the laser range. The position was on a horizontal plane whose coordinate system was comprised of a moving directional axis y and its orthogonal axis x with origin (x_0, y_0) at the center of the frontal laser scanner. We did not take velocity into account in this work. The practical area to analyze is limited to a rectangular area with a length of $80 \text{ m}, -40 \le y \le 40$, and a width of $9.9 \text{ m}, -4.95 \le x \le 4.95$.

3.4.2. Video Recording. The driver's face was captured by a camera (Sony 1/2 inch CCD video camera DXC-200A) mounted on the dashboard. The data were acquired at a resolution of 692 pixels in width and 480 pixels in height at a sample frequency of 29.4 fps.



FIGURE 3: Data-collection vehicle.

3.4.3. Driver Gaze Tracking. The driver's gaze direction was manually labeled using ELAN, (http://www.lat-mpi.eu/tools/ elan) which is a tool for the creation of complex annotations on video and audio resources, by an annotator. We prepared five gaze labels according to ISO 15007-1 [6] and three additional labels to detect gaze toward overtaking vehicles and toward upper traffic signs that could be extracted from the low-resolution video as follows:

- (*g*₀) *right side* (gaze toward right mirror and right window by head turning);
- (*g*₁) *right front* (gaze rightward from *front*, including gaze toward overtaking vehicles in the right lane);
- (g_2) rear (gaze toward rear-view mirror);
- (*g*₃) *front* (gaze to road scene ahead, the reference direction);
- (*g*₄) *left side* (gaze toward left mirror and left window by head turning);
- (*g*₅) *left front* (gaze leftward from *front*, including gaze toward overtaking vehicles in the left lane);
- (*g*₆) *up* (gaze upward from front, including gaze toward upper traffic signs);
- (*g*₇) *down* (gaze downward from *front*, including gaze toward instrument panel).

The annotator detected the beginning of the saccade of each gaze behavior as the beginning of interval with the gaze label. We consider that the gaze direction can be labeled more stably and accurately by using commercially supplied eye-tracking systems (e.g., FaceLab). Figure 4 shows a sample set of face images that were given each label.

4. Timing Analysis of Driver Gaze under Cognitive Distraction toward Peripheral Vehicle Behavior

4.1. Extracting Overtaking Events. We extracted 274 overtaking events from the intervehicular distance data of the neutral task (task condition \mathcal{C}_N) and 81 events from that of the music-retrieval task (task condition \mathcal{C}_M).

International Journal of Vehicular Technology



FIGURE 4: Sample set of gaze labels and face images.

4.2. Clustering of Peripheral Environment Data. We do not analyze traffic simulation data, we but focus on real traffic data. There are some peripheral vehicles in addition to the overtaking vehicle that act as triggers of random events in a real traffic scene. The peripheral vehicles are the visual stimuli. From the standpoint of the visual target search task, the overtaking vehicle and the other peripheral vehicles are regarded as target and visual distractors, respectively. The performance of the target search task depends on the traits of the visual distractors [29]. In this work, a trait of the peripheral environment needs to be defined. We classify the traffic scenes spatiotemporally according to time while the peripheral vehicles exist in the subperipheral area within the interval of the overtaking event, T_d . The peripheral area is divided into six subareas corresponding to the six gaze labels g_i (j = 0, ..., 5). Figure 5 shows the six subperipheral areas. Each of the six areas is assigned a Boolean variable a_i . A vector, **e**, comprises the six variables: $\mathbf{e} = (a_5, a_4, a_3, a_2, a_1, a_0)$, which represents the trait of the peripheral environment. In this work, if the cumulative duration while the peripheral vehicles exist in subperipheral area a_i goes over 50% of the targeted time interval, $a_i = 1$; otherwise, $a_i = 0$. We focus on the dynamics of the driver's gaze on the base point of the overtaking event, t_0 , as a reference point of analysis. Therefore, the interval of the overtaking event, T_d , is divided into the first half \mathbf{e}_1 from t_b to t_0 and the second half \mathbf{e}_2 from t_0 to t_e . Practically, we set T_d to 10 s ($t_0 = 0, t_b = -5, t_e = 5(s)$) as a sufficient time duration to analyze in consideration of the analysis area with the length of 80 m and the maximum velocity difference (=40.9 km/h) between the host vehicle and the overtaking vehicles in the area.

From all data including the overtaking events: 274 events for \mathcal{C}_N and 81 events for \mathcal{C}_M , we could not extract uniformly distributed vectors \mathbf{e}_1 and \mathbf{e}_2 . In addition, we eliminated some traffic scenes along curves and including lane changes because they would induce a specific gaze behavior. Adequate samples were extracted for two types of state transitions of peripheral environment as follows: environmental state transition \mathcal{C}_A from $\mathbf{e}_1 = (0, 0, 0, 0, 0, 1)$ to $\mathbf{e}_2 = (0, *, 0, *, 1, *)$, 75 events for task \mathcal{C}_N and 23 events for task \mathcal{C}_M , and

FIGURE 5: Definition of six subperipheral areas.

environmental state transition \mathcal{C}_B from $\mathbf{e}_1 = (0, 0, 0, 0, 1, 1)$ to $\mathbf{e}_2 = (0, *, 0, *, 1, *)$, 43 events for task \mathcal{C}_N and 13 events for task \mathcal{C}_M . 43.0% for \mathcal{C}_N and 44.4% for \mathcal{C}_M of all overtaking events were fallen into either \mathcal{C}_A or \mathcal{C}_B . "*" denotes any binary value. The former represents when the overtaking vehicle runs in the right lane and the other peripheral vehicles do not exist for more than 2.5 s (=50% of $T_d/2$) before it overtakes the host vehicle, whereas the latter represents when the overtaking vehicle exists in the *right-front* area for more than 2.5 s before it overtakes the host vehicle. We analyze the relationships between the extracted events for two environmental conditions \mathcal{C}_A and \mathcal{C}_B and the gaze data below.

4.3. Testing of Hypothesis Based on Temporal Gaze Distribution. The relative frequency distributions of the gaze directions (temporal gaze distributions) that were measured in the peripheral environment \mathscr{C}_A varied in terms of time, as shown in Figure 6. Since the relative frequency of the *right-front* gazes increased quickly after the base point of the overtaking event, t_0 , regardless of the task, the participants frequently gazed toward the overtaking vehicle in the right lane when the number of visual distractors was small. We verified



FIGURE 6: Time variations of relative frequency distributions of gaze directions that were measured in the peripheral environment \mathscr{C}_A .

the hypothesis (see Section 2.2) of the temporal difference between the relative frequency distributions for the musicretrieval task \mathscr{C}_M and the neutral task \mathscr{C}_N by performing a statistical test. Let us limit the interval for testing to the time from $t_s = t_0 (= 0)$ to $t_f = 2(s)$ following the work of Merat et al. [19]. The average saccade timings \bar{t}_c for the two conditions \mathscr{C}_N and \mathscr{C}_M were 0.82 s and 1.07 s (SD = 0.52 s and 0.42 s), respectively. The saccade timing for the \mathscr{C}_N condition was shorter than that for the \mathscr{C}_M condition, but we did not obtain adequate samples for performing any statistical test to the saccade timings. The average gaze timings \bar{t}_g for the two conditions \mathscr{C}_N and \mathscr{C}_M were 1.09 s and 1.46 s (SD = 0.51 s and 0.43 s), respectively. The two relative frequency distributions within the limited interval had normality and homoscedasticity. The *t*-test revealed that the gaze timing for the \mathcal{C}_N condition was significantly shorter than that for the \mathcal{C}_M condition, t(183) = -3.76, P < .001.

In contrast, Figure 7 shows that the participants rarely looked at the overtaking vehicle in the peripheral environment \mathscr{C}_B while performing the music-retrieval task \mathscr{C}_M , whereas they followed the same behavior as \mathscr{C}_A for the neutral task $\mathcal{C}_N.$ In the same manner as $\mathcal{C}_A,$ the average gaze timings \overline{t}_c for the two conditions \mathscr{C}_N and \mathscr{C}_M were 0.88 s and 1.25 s (SD = 0.67 s and 0.60 s), respectively. The saccade timing for the \mathscr{C}_N condition was shorter than that for the \mathscr{C}_M condition, but we did not obtain adequate samples for performing any statistical test to the saccade timings. The average gaze timings \bar{t}_q for the two conditions \mathscr{C}_N and \mathscr{C}_M were 0.93 s and 1.34 s (SD = 0.51 s and 0.45 s), respectively. The *t*-test revealed that the gaze timing for the \mathscr{C}_N condition was significantly shorter than that for the \mathscr{C}_M condition, t(83) =-3.1, P < .005. These results support our hypothesis but are limited to evaluation for two conditions, \mathscr{C}_A and \mathscr{C}_B , of the peripheral environment.

5. Detection of Driver Distraction

ŀ

5.1. Discrimination between Distraction and Neutral State in a Bayesian Framework. To identify the class label (*distraction*, i.e., \mathcal{C}_M , or *neutral*, i.e., \mathcal{C}_N) of the driver state using the temporal gaze distribution, we use a naive Bayesian framework as follows:

$$P(C_t \mid G_t, \mathbf{E}_t) = \frac{P(C_t, \mathbf{E}_t) P(G_t \mid C_t, \mathbf{E}_t)}{P(G_t, \mathbf{E}_t)}$$

$$= \frac{P(C_t \mid \mathbf{E}_t) P(\mathbf{E}_t) P(G_t \mid C_t, \mathbf{E}_t)}{P(G_t, \mathbf{E}_t)},$$
(1)

where *t* represents the time from when the host vehicle was overtaken by the other vehicle, C_t represents whether a gaze belongs to the distraction class at time *t*, that is, the binary class label, G_t represents whether the direction of the gaze is *right-front*, that is, gaze label, and \mathbf{E}_t represents the condition of the peripheral environment. One of the important characteristics of the Bayesian framework is the capability to infer the state of an unobserved variable, given the state of the observed variables. In our case, we want to infer the driver's internal state, that is, cognitive distraction, given the gaze data and the peripheral environment.

To make a decision as to the class of discrimination is assigned to the gaze data, the equation can be iterated over a time interval related to an overtaking event. We can then accumulate the computed posteriors and choose the class of driver state with greater score based on Maximum a Posteriori (MAP) as follows:

$$Q\left(C \mid G, \mathbf{E}\right) = \sum_{t=t_s}^{t_f} P\left(C_t \mid G_t, \mathbf{E}_t\right), \qquad (2)$$

$$S = \underset{C}{\operatorname{arg\,max}} Q\left(C \mid G, \mathbf{E}\right). \tag{3}$$





FIGURE 7: Time variations of relative frequency distributions of gaze directions that were measured in the peripheral environment \mathcal{C}_B .

In this work, the priors $P(C_t | \mathbf{E}_t)$ and $P(\mathbf{E}_t)$ are assumed as uniform, that is, noninformative prior distributions, and C_t and \mathbf{E}_t are regarded as binary variables not depending on time $t: P(G_t | C, \mathbf{E}) \simeq P(G_t | C_t, \mathbf{E}_t)$, $\mathbf{E} = (\mathbf{e}_1, \mathbf{e}_2)$. Therefore, we can apply the temporal gaze distribution extracted in Section 4.3 to $P(G_t | C, \mathbf{E})$.

5.2. Results. We matched the class label computed from equation (3) with the true label to discriminate between distraction state \mathcal{C}_M and neutral state \mathcal{C}_N . The experimental data were the same as the data for analysis of gaze timing (see Section 4): 75 events for \mathcal{C}_N and 23 for \mathcal{C}_M in \mathcal{E}_A , 43 events for \mathcal{C}_N , and 13 for \mathcal{C}_M in \mathcal{E}_B . We applied leave-one-out cross validation to obtain the discrimination accuracy.

Table 1 shows the accuracies of the two-class discrimination. As this is for testing our hypothesis (see Section 4.3), we limited the time interval for discrimination to from $t_s = t_0 (=$ 0) to $t_f = 2(s)$. Here, we employed a baseline method based on the percentage road center (PRC) [13], that is, proportion of total gaze duration toward road scene ahead to total time of sample events, which detected \mathscr{C}_M and \mathscr{C}_N by thresholding PRC at π_r ; that is, if PRC was larger than π_r , the method determined the data as distraction state \mathscr{C}_M ; otherwise, neutral state \mathscr{C}_N was assigned. The thresholds π_r for \mathscr{C}_A and \mathscr{C}_B were set to 79.0%, by searching for an equal rate between detection of \mathscr{C}_M and \mathscr{C}_N . We can confirm that the proposed method performed more accurately than the baseline one.

6. Discussion

We obtained better test results supporting our hypothesis and confirmed that the proposed discriminator performed more accurately than the baseline one. This approach takes advantage of the shorter time needed to detect cognitive distraction because it focuses on only the important scene for the detection but needs to trigger the gaze reaction to the overtaking event. The Bayesian rule-based method excels in application. It can be naturally integrated into the stateof-the-art method based on Bayesian networks using hybrid measures [30].

Figures 6 and 7 also suggest another difference between two tasks \mathscr{C}_M and \mathscr{C}_N . Note the frequency of gazing to a downward direction, that is, *down*. We consider that the participants frequently looked at the speedometer or navigation system while performing the neutral task. This behavior agrees with the prior findings of PRC [13]. In the temporal section without the overtaking event, PRC needs to be addressed to detect cognitive distraction.

Here, let us compare the average timing of the gaze to *right front* in condition \mathscr{C}_A with that in condition \mathscr{C}_B (see their values in Section 4.3). We can confirm that the latter was slightly shorter than the former. In condition \mathscr{C}_B , there was a vehicle that preceded the overtaking vehicle. The participants might still focus their attention on the preceding vehicle or *right-front* area and then effectively react to the next overtaking event. The event did not cause inhibition of return [31], which retards their reaction.

These results were verified under only two limited environmental conditions, that is, \mathscr{C}_A and \mathscr{C}_B , of peripheral vehicles because we could not analyze enough experimental data. We need to increase the number of clusters of peripheral environment data for wide-ranging analysis and to model the dynamics of peripheral vehicles based on, for example, time to collision (TTC), velocity, acceleration, and interaction among vehicles for deeply analyzing the temporal relationships and increasing the discrimination accuracy. We also have to analyze the other secondary tasks and differences among individuals.

7. Conclusions

The dynamics of the external environment can elicit reactions reflecting the human internal state, that is, make the latent

TABLE 1: Two-class discriminant accuracies.

	Peripheral environment \mathscr{C}_A		Peripheral environment \mathscr{C}_B	
	Distraction (\mathscr{C}_M)	Neutral (\mathscr{C}_N)	Distraction (\mathscr{C}_M)	Neutral (\mathscr{C}_N)
Proposed method	66.7%	64.7%	75.0%	61.5%
Baseline method (PRC)	50.0%	52.9%	50.0%	53.8%

state explicit. We showed that the temporal factor, that is, timing, of a reaction is important for understanding the state by focusing on cognitive distraction in a car-driving situation. The concrete contribution of this paper is twofold. First, we obtained test results supporting our hypothesis that the timing of when a driver gazes toward the overtaking event under cognitive distraction is later than that under the neutral state. Second, we confirmed that a Bayesian-based detection of distraction using the temporal gaze distribution performed more accurately than the PRC-based one. The findings of this work should be generalized through additional analysis in future work. We have built a large database of 500 drivers [26]. The generalized findings will suggest a risk of voice interactive navigation (hands-free navigation) using automatic speech recognition and a novel testing scenario without driver's extra workload for driver information system in real driving situation to the Alliance of Automobile Manufacturers (AAM) guidelines [32]. To put our approach into practical use, we will refer to peripheral vehicle-tracking systems based on computer vision techniques [24, 33] and replace the laser scanner with a camera-based system.

Acknowledgments

This work is partially supported by JST CREST (Core Research for Evolutional Science and Technology of Japan Science and Technology Agency) Kashin-free project and a Grant in Aid for Scientific Research from MEXT (Ministry of Education, Culture, Sports, Science, and Technology) of Japan under Contract no. 23700168. The authors thank Professor Chiyomi Miyajima and Mr. Masataka Mori for collecting experimental data.

References

- M. A. Regan, J. D. Lee, and K. L. Young, "Defining driver distraction," in *Driver Distraction: Theory, Effects, and Mitigation*, chapter 4, pp. 42–54, CRC, 2008.
- [2] T. A. Ranney, W. R. Garrott, and M. J. Goodman, NHTSA Driver Distraction Research: Past, Present, and Future, National Highway Traffic Safety Administration, 2001.
- [3] Y. Dong, Z. Hu, K. Uchimura, and N. Murayama, "Driver inattention monitoring system for intelligent vehicles: a review," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 2, pp. 596–614, 2011.
- [4] H. Croo, M. Bandmann, G. Mackay, K. Rumar, and P. Vollenhoven, *The Role of Driver Fatigue in Commercial Road Transport Crashes*, 2001.
- [5] Y. Liang and J. D. Lee, "Combining cognitive and visual distraction: less than the sum of its parts," *Accident Analysis and Prevention*, vol. 42, no. 3, pp. 881–890, 2010.

- [6] ISO, 15007-1:2002, "Road vehicles—measurement of driver visual behavior with respect to transport information and control systems—part 1: definitions and parameters," 2002.
- [7] ISO/TS, 15007-1:2001, "Road vehicles—measurement of driver visual behavior with respect to transport information and control systems—part 2: equipment and procedures," 2001.
- [8] E. Johansson, J. C. Engström, C. Cherri et al., *Review of existing techniques and metrics for IVIS and ADAS assessment*, Adaptive Integrated Driver-Vehicle Interface, 2004.
- [9] L. Angell, J. Auflick, A. Austria et al., Driver Workload Metrics Project—Task 2 Final Report, U.S. Department of Transportation, National Highway Traffic Safety Administration, 2006.
- [10] J. L. Harbluk and Y. I. Noy, *The Impact of Cognitive Distraction* on Driver Visual Behavior and Vehicle Control, Ergonomics Division, Road Safety Directorate and Motor Vehicle Regulation Directorate, Ontario, Canada, 2002.
- [11] J. G. May, R. S. Kennedy, M. C. Williams, W. P. Dunlap, and J. R. Brannan, "Eye movement indices of mental workload," *Acta Psychologica*, vol. 75, no. 1, pp. 75–89, 1990.
- [12] M. Miyaji, H. Kawanaka, and K. Oguri, "Driver's cognitive distraction detection using physiological features by the AdaBoost," in *Proceedings of the 12th International IEEE Conference on Intelligent Transportation Systems (ITSC '09)*, pp. 90–95, October 2009.
- [13] K. Kircher, C. Ahlstrom, and A. Kircher, "Comparison of two eye-gaze based real-time driver distraction detection algorithms in a small-scale field operational test," in *Proceedings of the 5th International Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, pp. 16–23, 2009.
- [14] R. Ishii and Y. I. Nakano, "Estimating user's conversational engagement based on gaze behaviors," in *Intelligent Virtual Agents, Lecture Notes in Computer Science*, vol. 5208, pp. 200– 207, 2008.
- [15] T. Hirayama, J. Dodane, H. Kawashima, and T. Matsuyama, "Estimates of user interest using timing structures between proactive content-display updates and eye movements," *IEICE Transactions on Information and Systems*, vol. E93-D, no. 6, pp. 1470–1478, 2010.
- [16] R. Yonetani, H. Kawashima, T. Hirayama, and T. Matsuyama, "Mental focus analysis using the spatio-temporal correlation between visual saliency and eye movements," *Journal of Information Processing*, vol. 20, no. 1, pp. 267–276, 2012.
- [17] T. Hirayama, Y. Sumi, T. Kawahara, and T. Matsuyama, "Infoconcierge: proactive multi-modal interaction based on mind probing," in *The Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA/ASC '11)*, 2011.
- [18] N. Merat and A. H. Jamson, "Multisensory signal detection: a tool for assessing driver workload during IVIS management," in *Proceedings of the 4th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, 2007.
- [19] N. Merat, E. Johansson, J. A. Engström, E. Chin, F. Nathan, and T. W. Victor, *Specification of a secondary task to be used*

in safety assessment of IVIS, Adaptive Integrated Driver-Vehicle Interface, 2007.

- [20] G. L. Rupp, Performance Metrics for Assessing Driver Distraction: The Quest for Improved Road Safety, SAE International, 2010.
- [21] L. Hsieh, R. Young, and S. Seaman, "Development of the enhanced peripheral detection task: a surrogate test for driver distraction," *SAE International Journal of Passenger Cars*, vol. 5, no. 1, pp. 317–325, 2012.
- [22] M. F. Land and D. N. Lee, "Where we look when we steer," *Nature*, vol. 369, no. 6483, pp. 742–744, 1994.
- [23] N. Apostoloff and A. Zelinsky, "Vision in and out of vehicles: integrated driver and road scene monitoring," *International Journal of Robotics Research*, vol. 23, no. 4-5, pp. 513–538, 2004.
- [24] L. Fletcher and A. Zelinsky, "Driver inattention detection based on eye gaze-road event correlation," *International Journal of Robotics Research*, vol. 28, no. 6, pp. 774–801, 2009.
- [25] M. I. Posner, "Orienting of attention," The Quarterly journal of experimental psychology, vol. 32, no. 1, pp. 3–25, 1980.
- [26] K. Takeda, J. H. L. Hansen, P. Boyraz, L. Malta, C. Miyajima, and H. Abut, "International large-scale vehicle corpora for research on driver behavior on the road," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 3, pp. 1–15, 2011.
- [27] S. Hara, C. Miyajima, K. Itou, and K. Takeda, "An online customizable music retrieval system with a spoken dialogue interface," *Journal of Acoustical Society of America*, vol. 120, no. 5, pp. 3378–3379, 2006.
- [28] Y. Li, C. Miyajima, N. Kitaoka, and K. Takeda, "Driving scene retrieval using integrated vehicle motion feature matching," in *Proceedings of the 5th Biennial Workshop on DSP for In-Vehicle Systems*, pp. 1–8, 2011.
- [29] J. M. Wolfe and T. S. Horowitz, "What attributes guide the deployment of visual attention and how do they do it?" *Nature Reviews Neuroscience*, vol. 5, no. 6, pp. 495–501, 2004.
- [30] L. Malta, C. Miyajima, N. Kitaoka, and K. Takeda, "Analysis of real-world driver's frustration," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 1, pp. 109–118, 2011.
- [31] M. I. Posner and Y. Cohen, "Components of visual orienting," in *Attention and Performance*, X. H. Bouma and D. Bowhuis, Eds., pp. 531–556, Lawrence Erlbaum Associates, Hillsdale, NJ, USA, 1984.
- [32] Alliance of Automobile Manufactures, Statement of Principles, Criteria and Verification Procedures on Driver Interactions with Advanced In-Vehicle Information and Communication Systems, 2006.
- [33] A. Doshi and M. Trivedi, "Investigating the relationships between gaze patterns, dynamic vehicle surround analysis, and driver intentions," in *Proceedings of the IEEE Intelligent Vehicles Symposium*, pp. 887–892, June 2009.



The Scientific World Journal

International Journal of Distributed Sensor Networks

VLSI Design



Journal of Electrical and Computer Engineering

Rotating Machinery

Hindawi

Active and Passive







International Journal of Chemical Engineering



