Analysis of Peripheral Vehicular Behavior in Driver’s Gaze Transition: Differences between Driver’s Neutral and Cognitive Distraction States

Takatsugu Hirayama\(^*\),1, Shota Sato\(^*\),2, Kenji Mase\(^*\),3, Chiyomi Miyajima\(^*\),4, and Kazuya Takeda\(^*\),5

Abstract—To support safe driving, numerous methods have been proposed for detecting distractions based on the measurements of a driver’s gaze. These methods empirically focused on certain driving contexts, and analyzed gaze behavior under particular peripheral vehicular conditions; therefore, numerous driving situations were not considered. To address this problem, we propose a data-driven approach that analyzes peripheral vehicular behaviors during gaze transitions of drivers, to compare their neutral driving state with a cognitive distraction state. The analysis results show that drivers, under the neutral conditions, turned their gaze to peripheral vehicles to be focused on; however, they did not do this consistently under the distracted conditions. In addition, we propose a simple classifier to discriminate between the distracted and the neutral states by analyzing peripheral vehicular behavior. The proposed classifier can manage various situations, and provide high discrimination accuracy, by focusing on gaze transitions from the front view toward other directions.

I. INTRODUCTION

Advanced Driver Assistance Systems (ADAS) and autonomous vehicles have attracted significant attention. However, researchers are concerned that drivers may develop an over-reliance on incomplete systems, which may lead to accidents. To prevent over-reliance, ADAS must effectively analyze the driver’s state as well as traffic situations. In this study, we focus on driver’s cognitive distraction, which is an observable state resulting from the over-reliance, and its relationships with driver’s eye-gaze and peripheral vehicular behaviors.

Driver distraction is a diversion of attention away from activities critical for safe driving toward a competing activity[1], and is a significant risk factor that can cause accidents[2]. Note that distraction differs from fatigue[3], which is defined as a state of exhaustion that disables a person from continuing an activity[4]. Numerous researchers have developed driver distraction monitoring systems that aim to promote driving safety by considering different types and levels of distraction[3]. The National Highway Traffic Safety Administration (NHTSA) classifies distractions into the following categories: (1) visual distraction; (2) auditory distraction; (3) biomechanical distraction; and (4) cognitive distraction from the viewpoint of the driver’s functionality[2]. Visual distraction, auditory distraction, and biomechanical distraction are caused by external factors that disturb the activity and can be readily observed. However, cognitive distraction is considered an intrinsic state; therefore, it is difficult to sense externally.

In the past few decades, numerous methods for detecting driver distraction have been proposed[3]. These methods can be divided into the following five categories, based on the factors being measured: (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’s biological measures are not suitable for actual driving conditions. Driving performance measures, as indicated by steering, braking, and other related driving behaviors, are suitable for detecting visual distraction[5]. Although a system may be able to detect overt behaviors that are more directly linked with risk, it may not be able to react in time to assist the driver after the detection.

In controlled settings, such as a driving simulator, the detection response task (DRT) is a promising method for measuring visual and cognitive distractions[6], [7], [8], [9]. It requires the subject to respond via a device such as a button to visual, tactile, or acoustic stimuli. The response time relates to the distractions. However, it is difficult to give actual drivers on the road this task without impacting driving safety. Eye-gaze measuring, which is included in the driver’s physical measures category, is a useful distraction measurement as specified in the existing standards ISO 15007-1[10] and ISO/TS 15007-2[11]. Johansson et al. have reviewed the existing gaze-based techniques and metrics for analyzing visual and cognitive distractions[12].

Most methods that detect cognitive distractions based on gaze measurement did not consider relationships with traffic situations. The others empirically focused on certain driving contexts, and analyzed gaze behavior under particular peripheral vehicular conditions; therefore, various driving situations were not considered. In contrast, we propose a data-driven approach that analyzes peripheral vehicular movements while the driver shifts the gaze from a certain direction toward another direction. As a result, this approach provides an exhaustive analysis of peripheral vehicular behavior during a driver’s gaze transition.

II. ANALYSIS OF A DRIVER’S GAZE BEHAVIOR UNDER COGNITIVE DISTRACTION

Some researchers have demonstrated that eye-gaze measures have the potential to capture symptoms of cognitive distraction[12]. According to a study by Harbluk and Noy, drivers under cognitive distraction exhibited fewer saccades
per time unit, which was consistent with reduced exploration of the driving environment[13]. Saccades may be a valuable indicator of mental workload[14]. Miyaji et al. reported that the standard deviations of eye movements could be suitable for detecting cognitive distractions that caused gaze concentration while drivers viewed the roadway[15]. Kircher et al. indicated the percentage of time that drivers spent observing the road ahead, which is called the percentage road center (PRC) of gaze direction, was greater than 92% under cognitive distraction[16]. Angell et al. used an eye-gaze pattern to discriminate between driving while performing a secondary cognitive task and driving only[17]. These approaches primarily measured only driver’s gaze toward the road ahead; therefore, consideration was not given to the peripheral traffic environment, which typically includes many distracting visual stimuli.

As a study focusing on the relationship between gaze behavior and peripheral environment dynamics, Hirayama et al. analyzed the timing of a driver’s gaze toward the visual change that occurred when a peripheral vehicle passed the driver’s vehicle. This study confirmed that the timing under a cognitive distraction condition was later than that under a neutral driving condition[18]. They also proposed a Bayesian-based discriminator to distinguish between the distracted and the neutral states by using the temporal gaze distribution, which performed more accurately than a PRC-based one. However, the findings cannot be applied to every vehicle passing event encountered while driving. In addition, passing events occur somewhat infrequently. Such situation-dependent approaches, which identify gaze behavior characteristics in hypothetical traffic situations, are unsuitable for extracting useful findings required to create a versatile effective system.

Our proposed data-driven approach first detects intervals including a gaze transition from the driver’s gaze direction sequence. It then extracts characteristics of peripheral vehicular behavior during the intervals from various traffic environment dynamics. A comparison of the characteristics is made between various gaze transition patterns recorded during distracted conditions and neutral conditions, to identify their differences. We also propose a classifier to discriminate between the conditions, based on peripheral vehicular behavior co-occurring with each gaze transition pattern.

III. REAL-WORLD DRIVING DATABASE[18]

We analyzed a portion of a database collected using the "NUDrive Vehicle" in Nagoya, Japan[19].

A. Data-collection vehicle

The "NUDrive Vehicle" was designed to synchronously record multimedia driver performance signals (gas and brake pedal application, steering angles, velocity, acceleration, and vehicle position), inter vehicular distance, biological signals, videos, and audio signals. Various external sensors were mounted on a Toyota Hybrid Estima with an automatic transmission and a steering wheel on the right. All sensors used for recording were commercially available.

B. Participants

A total of 40 participants (20 males and 20 females) engaged in the experiment. They were, on average, 37.3 years old (with a range of 22 to 58 years old) and had held a driver’s license for a mean period of 17.3 years (with a range of 4 to 39 years). All participants signed an informed consent prior to their participation.

C. Procedure

The participants first drove for a few minutes to get acclimated to the vehicle and the sensors. Signals recorded during the initial period were not used in this study. The experimenter monitored the experiment from the rear seat and indicated the route to the driver. During a particular period of driving, the participants 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. The secondary task artificially induced the state of cognitive distraction. The experimenter instructed the participants to retrieve as many songs as possible; accordingly, within approximately 30 s of successfully retrieving each song, the participants had to retrieve another song. All experiments were performed on two or three-lane highways. The experimental route was the same for all participants. The duration of the neutral driving state (task condition $C_N$) and the cognitive distraction state ($C_M$) were 287 ± 43.2 s and 293 ± 21.1 s, respectively.

D. Measures

We analyzed the inter vehicular distance measured by the laser scanners, and the driver’s gaze direction extracted from the recorded video.

1) Inter vehicular distance recording: Two laser scanners,1 mounted on the front and back of the host vehicle, provided geometric information regarding the peripheral environment of the vehicle. The laser scanners covered 80-degree arcs at both the front and rear of the vehicle, providing an effective range of approximately 100 m to the front and 55 m to the rear. This configuration created 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. The dynamics of their positions and velocity relative to the host vehicle could be estimated, including instances when they were outside the laser range. The position was recorded on a horizontal plane, utilizing a coordinate system 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. The practical area to analyze was limited to a rectangular area with a length of 80 m, $-40 \leq y \leq 40$, and a width of 10 m, $-5 \leq x \leq 5$.

2) Video recording: The driver’s face was captured by a camera2 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 29.41 fps.

1front:RIEGL LMS-140i-80; rear:RIEGL LMS-Q120i
2SONY 1/2 inch CCD video camera DXC-200A
3) **Driver gaze labeling:** The driver’s gaze direction was manually labeled by an annotator. We prepared eight gaze labels based on ISO 15007-1[10] that could be extracted from the low-resolution video as follows:

- **g₀:** right-side (gaze toward right mirror and right window by head turning);
- **g₁:** right-front (gaze slightly rightward from the front);
- **g₂:** rear (gaze toward rear-view mirror);
- **g₃:** front (gaze to road ahead, reference direction);
- **g₄:** left-side (gaze toward left mirror and left window by head turning);
- **g₅:** left-front (gaze slightly leftward from the front);
- **g₆:** up (gaze upward from the front, including gaze toward high-mounted traffic signs);
- **g₇:** down (gaze downward from the front, including gaze toward instrument panel).

The annotator detected the beginning of the saccade for each gaze behavior as the beginning of the labeled interval. Figure 1 displays a sample set of face images that were assigned for each label.

IV. ANALYSIS OF PERIPHERAL VEHICULAR BEHAVIOR IN A DRIVER’S GAZE TRANSITION

A general approach for analyzing gaze behavior first classifies the visual environment dynamics, then describes the characteristics of the behavior in each of the classified dynamics. As mentioned previously, the proposed approach first classifies the gaze behavior and then analyzes the environmental dynamics while the classified gaze behavior occurs. In this study, we detect intervals that include a gaze transition from a gaze direction sequence (IV-A), create a heat map of each interval, which represents spatial distribution of peripheral vehicles (IV-B), extract characteristics of the maps using principal component analysis (IV-C), and compare the map created under the distracted condition with the map created under the neutral condition (IV-D).

A. **Detection of gaze transition intervals**

We focus on primitive gaze transitions and detect the interval of the transition. We define a primitive gaze transition as an event where the driver maintains the gaze toward the road ahead (g₃) for more than τ second(s), then shifts the gaze from g₃ to another direction gᵢ, and then another direction gⱼ (i ≠ 3, j ≠ i). Therefore, the gaze transition interval is defined as a period of time from tᵢ − τ to tⱼ. tᵢ and tⱼ represent the times when the driver shifts the gaze to gᵢ and gⱼ. Figure 2 displays the gaze transition interval. In this study, τ was set to two seconds. We did not consider gaze behaviors occurring in some traffic scenes along curves and including lane changes, because they would have a particular relationship with visual environment dynamics. TABLE I lists the frequency of the detected gaze transitions for different directions, under the neutral and the cognitive distraction conditions. The duration of the detected interval was 2.58 ± 0.391 s.

B. **Creation of peripheral vehicular map**

The peripheral vehicular map is a heat map that depicts traffic situations where peripheral vehicles were present around the host vehicle during a gaze transition interval. The values in the map are computed according to how frequently
the peripheral vehicles were present, and the velocity differences between the host vehicle and the peripheral vehicles. In this study, the analysis area was sampled at one meter intervals. Therefore, the map is represented as a 10 × 80 dimensional vector.

We first create the map based on mixtures of Gaussians at each time $t$ within a gaze transition interval as follows:

$$ p(x_t) = \sum_{c=1}^{C} |v_{c,t}| \exp \left\{ -\frac{1}{2} (x_t - x_{c,t})^T \Sigma^{-1} (x_t - x_{c,t}) \right\}, \quad (1) $$

where $x = (x, y) \in \mathbb{Z}^2$ represents the sample points on the map, $C$ represents the number of peripheral vehicles that were present in the analysis area during the interval, $x_c$ and $v_c$ represent the position and the relative velocity of peripheral vehicle $c$, respectively. We assume that the velocity difference regardless of its sign influences driver’s visual attention. This is the reason why we calculate the absolute value of the relative velocity. An identity matrix is applied to the variance-covariance matrix $\Sigma$. We then accumulate the created maps within the gaze transition interval, and normalize the range of values in the accumulated map. We refer to the normalized map as the peripheral vehicular map $p(x)$. Figure 3(a) displays an example of the map. To analyze the differences between the segments, we segment them into 21 subregions as displayed in Fig.3(b).

**C. Extraction of peripheral vehicle behavior characteristics**

We extract peripheral vehicular behavior characteristics from the maps for each gaze transition pattern. In this study, we apply singular value decomposition[20] to the maps and extract eigenvectors with higher contribution rates. Each of the eigenvectors is regarded as a basis vector for peripheral vehicular behavior. Figure 4 displays the average peripheral vehicular map and the eigenvectors with the first, second, and third highest contribution rates. We refer to them as the first, second, and third eigenvectors. The eigenvector elements have normalized values that range from $-1$ to $+1$. We regard the absolute value of the vector element as an influence on the driver’s gaze transition. In this paper, we analyze the eigenvectors in the gaze transition from the front to the right-side, right-front, left-side, or left-front.

**D. Comparison of peripheral vehicular behavior**

We compared the eigenvectors of the maps under the distracted condition with those under the neutral condition.

1) **Gaze transitions from front to right-side ($g_3 \rightarrow g_1$)**: In Fig.4(a), we can confirm that the first eigenvector for the neutral state $C_n$ has a larger influence in $a_{1,3}$, $a_{2,3}$, and $a_{3,3}$ (see A in the figure) and the second and the third eigenvector have a large influence in $a_{5,3}$ (see B). This indicates that peripheral vehicles were frequently present in the right lane ahead, and at the rear of the host vehicle, when the drivers shifted their gaze toward the right mirror under the neutral condition. Conversely, the eigenvectors for the cognitive distraction state $C_m$ have a larger influence in $a_{2,3}$ and $a_{3,3}$ (see C) but do not have influence in $a_{5,3}$ (see D). The first eigenvector also has a slightly larger influence in $a_{3,1}$ (see E). We surmise that the drivers did not shift their gaze toward the right-side in synchronization with peripheral vehicles under the distracted condition.

2) **Gaze transitions from front to right-front ($g_3 \rightarrow g_1$)**: Figure 4(b) illustrates that the eigenvectors for both states have a larger influence in $a_{3,3}$, $a_{4,3}$, and $a_{5,3}$ (see F). The contribution rate of the first eigenvector for the neutral state (24.9%) is higher than that for the distracted state (10.5%). The peripheral vehicles generally were present at the right of the host vehicle, when the drivers shifted their gaze slightly toward the right. This is especially true in the neutral state. In addition, the first eigenvector for the distracted state has a larger influence in $a_{1,3}$ and $a_{2,3}$ (see G). We regard this characteristics as a gaze timing delay in response to a passing vehicle, which supports the existing findings[18].

3) **Gaze transitions from front to left-side ($g_4 \rightarrow g_1$)**: The primary objective of a gaze to the left-side is to check whether any vehicles are present in the left lane to the rear of the host vehicle, using the left mirror. However, as shown in Fig.4(c), the first eigenvector for both states has a larger influence in $a_{4,3}$, $a_{5,3}$, and $a_{6,3}$ (see H). We surmise that the vehicles in the left lane behind the host vehicle did not have a large influence on the driver’s vision, because most of them were moving slowly and backed away from the host vehicle. As mentioned previously, we did not focus on certain scenes where the host vehicle performed a lane change.

4) **Gaze transitions from front to left-front ($g_3 \rightarrow g_5$)**: Figure 4(d) illustrates that the first eigenvector for the neutral state has a large influence in $a_{2,1}$ and $a_{3,2}$ (see I) but not for the distracted state. The drivers would shift their gaze toward the left-front in synchronization with peripheral vehicles under the neutral condition. In addition, the eigenvectors for both states have a larger influence in $a_{2,3}$ and $a_{3,3}$ (see J). We plan to identify the cause in future work.
We confirmed that the drivers shifted their gaze toward the right-side, right-front, and left-front in synchronization with peripheral vehicles under the neutral condition, but not under the distracted condition. We also identified some differences in the eigenvectors of peripheral vehicular maps between up, down, and rear gazes, under both conditions. However, the drivers would not look at the peripheral vehicles in the gaze transition intervals. They might shift their gaze toward traffic signs, the instrument panel, or other visual objects.

V. DISCRIMINATION BETWEEN DRIVER STATES USING A PERIPHERAL VEHICULAR MAP

A. K-NN classifier

We propose a k-nearest neighbor (k-NN) classifier to discriminate between the distracted and the neutral conditions, using the peripheral vehicular map. We apply the classifier to a 10 × 80 dimensional discrimination space for each gaze transition pattern, which includes peripheral vehicular maps under both conditions. The classifier finds k nearest neighbors using the normalized Euclidean distance between the maps. The distance between map i and j, which normalizes the range of each dimension, is calculated as

\[ D(p_i(x), p_j(x)) = \sqrt{\sum_{x,y}^{10} \sum_{80} (p_i(x, y) - p_j(x, y))^2 / \sigma_{x,y}^2}, \]

where \( p(x) \) represents the peripheral vehicular map and \( \sigma_{x,y}^2 \) represents the variance of \( p(x, y) \) on all maps in the space.

B. Discrimination results

We applied leave-one-out cross validation for each data element of the gaze transition patterns listed in TABLE I to obtain the discrimination accuracy. TABLE II lists the accuracies in the case of \( k = 5 \) and the discrimination cycles. We employed the Bayesian-based discriminator, based on the timing analysis of a gaze toward a passing vehicle in the right lane[18] as a comparison method. We also employed a baseline method based on PRC[16], which discriminated between the distracted and the neutral states by thresholding the PRC. The duration for calculating the PRC was set to 60.0 s. The threshold was set to 86.0% by searching for an equal rate between the discrimination of the states.

We can confirm that the discrimination accuracy for every gaze transition pattern under the neutral condition was higher than the accuracy under the distracted condition. This result indicates that the drivers, under the neutral condition, shifted their gaze from the front toward other direction when peripheral vehicles were present in a spatial pattern; their gaze did not shift under the distracted condition. In this section, we focus on the gaze transition from the front to the right-side, right-front, left-side, or left-front, similar to Section IV. In TABLE II, \( g_3 \rightarrow g_0/g_1/g_4/g_5 \) displays the average accuracy and the discrimination cycle using the gaze transitions. The accuracy of the proposed method was higher than the accuracy of the existing methods for the neutral state, but lower for the distracted state. It is notable that the proposed method realized the discrimination cycle eight times as frequently as the existing methods. By using all gaze transitions (\( g_3 \rightarrow g_4 \)), the proposed method increased the tendencies.

VI. CONCLUSIONS

Hypothesis-testing approaches focus inordinately on specific situations. Our proposed data-driven approach can analyze a greater diversity of situations; it first detects intervals including a primitive gaze transition, then extracts characteristics of visual environmental contexts during the intervals. We compared the extracted characteristics of peripheral vehicular behavior during various gaze transition patterns of drivers under cognitive distraction conditions with those under neutral conditions. In addition, we discriminated between the distracted and the neutral states using a simple instance-based classifier that regards the extracted peripheral vehicular behaviors as instances. The results indicated that the drivers would adequately shift their gaze toward particular peripheral vehicles under the neutral condition; the drivers did not adequately shift their gaze under the distracted condition.

In future work, we are planning to analyze gaze transitions using a slightly longer interval. The analysis will include multiple gaze transitions, while accounting for individual differences in driver attention levels. We will also employ an eigenspace approach to discriminate between the states, as well as analyze the differences. More importantly, we must automate gaze tracking in various driving situations.

ACKNOWLEDGMENT

This work is partially supported by JST CREST Kashin-free project and JSPS KAKENHI Grant Number 26730119.

REFERENCES


<table>
<thead>
<tr>
<th>Neutral state (C)</th>
<th>Distraction state (C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean map</td>
<td>Eigenvector</td>
</tr>
<tr>
<td>1st</td>
<td>2nd</td>
</tr>
<tr>
<td>16.5%</td>
<td>11.2%</td>
</tr>
</tbody>
</table>

(a) Gaze transition from the front to the right-side.

(b) Gaze transition from the front to the right-front.

(c) Gaze transition from the front to the left-side.

(d) Gaze transition from the front to the left-front.

Fig. 4. Characteristics of peripheral vehicular behavior during gaze transitions.


