Flotterod and Rohde [18]. Therefore, we will attempt to extend the model of Flotterod and Rohde for the traffic flow prediction of UTNs in the future and make a comparison with the results of this paper.

REFERENCES

[1] F.-Y. Wang, "Parallel control and management for intelligent transportation systems: Concepts, architectures, and applications," *IEEE Trans. Intell. Transp. Syst.*, vol. 11, no. 3, pp. 630–638, Sep. 2010.
[2] G. Xiong, S. Liu, X. Dong, F. Zhu, B. Hu, D. Fan, and Z. Zhang, "Parallel traffic management system helps 16th Asian games," *IEEE Intell. Syst.*, vol. 27, no. 3, pp. 74–78, May/Jun. 2012.
[3] I. Okutani and Y. J. Stephanedes, "Dynamic prediction of traffic volume through Kalman filtering theory," *Transp. Res. B*, vol. 18, no. 1, pp. 1–11, Feb. 1984.
[4] B. L. Smith, B. M. Williams, and R. K. Oswald, "Comparison of parametric and nonparametric models for traffic flow forecasting," *Transp. Res. C*, vol. 10, no. 4, pp. 303–321, Aug. 2002.
[5] M. M. Hamed, H. R. Al-Masaeid, and Z. M. Bani Said, "Short-term prediction of traffic volume in urban arterials," *J Transp. Eng.*, vol. 121, no. 3, pp. 249–254, May 1995.
[6] E. I. Vlahogianni, M. G. Karlaftis, and J. C. Golias, "Optimized and meta-optimized neural networks for short-term traffic flow prediction: A genetic approach," *Transp. Res. C*, vol. 13, no. 3, pp. 211–234, Jun. 2005.
[7] A. Stathopoulos, L. Dimitriou, and T. Tsekeris, "Fuzzy modeling approach for combined forecasting of urban traffic flow," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 23, no. 7, pp. 521–535, Oct. 2008.
[8] S. Sun, C. Zhang, and G. Yu, "A Bayesian network approach to traffic flow forecasting," *IEEE Trans. Intell. Transp. Syst.*, vol. 7, no. 1, pp. 124–132, Mar. 2006.
[9] E. I. Vlahogianni, M. G. Karlaftis, and J. C. Golias, "Spatio-temporal short-term urban traffic volume forecasting using genetically optimized modular networks," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 22, no. 5, pp. 317–325, Jul. 2007.
[10] W. Min and L. Wynter, "Real-time road traffic prediction with spatio-temporal correlations," *Transp. Res. C*, vol. 19, no. 4, pp. 606–616, Aug. 2011.
[11] S. Lin, Y. Xi, and Y. Yang, "Short-term traffic flow forecasting using macroscopic urban traffic network model," in *Proc. 11th Int. IEEE Conf. Intell. Transp. Syst.*, Beijing, China, 2008, pp. 134–138.
[12] S. Lin and Y. Xi, "An efficient model for urban traffic network control," in *Proc. 17th IFAC World Congr.*, Seoul, Korea, 2008, pp. 14 066–14 071.
[13] M. Van den Berg, A. Hegyi, B. De Schutter, and J. Hellendoorn, "Integrated traffic control for mixed urban and freeway networks: A model predictive control approach," *Europ. J. Transp. Infrastruct. Res.*, vol. 7, no. 3, pp. 223–250, Sep. 2007.
[14] *Traffic Software Integrated System Version 5.1 User's Guide*, Fed. Highway Admin., Washington, DC, USA, 2001.
[15] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *Int. J. Forecast.*, vol. 22, no. 4, pp. 679–688, Oct.–Dec. 2006.
[16] B. D. Greenshields, "A study of traffic capacity," in *Proc. Ann. Meeting Highway Res. Board*, 1935, vol. 14, pp. 448–477.
[17] M. J. Lighthill and G. B. Whitham, "On kinematic waves II: A theory of traffic flow on long crowded roads," *Proc. Roy. Soc. Lond. Ser. A*, vol. 229, no. 1178, pp. 317–345, May 1955.
[18] G. Flotterod and J. Rohde, "Operational macroscopic modeling of complex urban road intersections," *Transp. Res. B*, vol. 45, no. 6, pp. 903–922, Jul. 2011.

Studying the Effects of Driver Distraction and Traffic Density on the Probability of Crash and Near-Crash Events in Naturalistic Driving Environment

Renran Tian, Lingxi Li, *Member, IEEE*, Mingye Chen, Yaobin Chen, *Senior Member, IEEE*, and Gerald J. Witt

***Abstract*—Driver distraction detection and intervention are important for designing modern driver-assistance systems and for improving safety. The main research question of this paper is to investigate how the cumulative driver off-road glance duration can be controlled to reduce the probability of occurrences of crash and near-crash events. Based on the available data sets from the Virginia Tech Transportation Institute (VTTI) 100-car study, the conditional probability is calculated to study the chance of crash and near-crash events when the given cumulative off-road glance duration in 6 s has been reached. Different off-road eye-glance locations and traffic density levels are also evaluated. The results show that one linear relationship can be obtained between the cumulative off-road eye-glance duration in 6 s and the risk of occurrences of crash and near-crash events, which varies for different off-road eye-glance locations. In addition, the traffic density level is found to be one significant moderator to this linear relationship. Detailed comparisons are made for different traffic density levels, and one nonlinear equation is obtained to predict the probability of occurrences of crash and near-crash events by considering both cumulative off-road glance duration and traffic density levels.**

***Index Terms*—Cumulative driver off-road glance duration, driver distraction, human factors, naturalistic driving environment, traffic density.**

I. INTRODUCTION

DRIVER distraction is one of the most widely known prominent contributors to traffic accidents [1]. Based on the attention-competing model between road demand and concurrent secondary tasks [1]–[3], researchers tried to explain the underlying mechanism causing degraded driving performance and summarize empirical evidences of the effects of distraction on driving safety. According to one naturalistic driving study reported in [1], about 25% of police-reported cases [2] and 65%–80% of crash and near-crash cases have driver distraction and/or driver inattention involved as contributors. In [2]–[5], the relationship between external/internal driver distraction sources and driving safety/performance measures in different environments are investigated. It is clear that many kinds of distracted driving behaviors are associated with specified driving safety degradation.

Extensive research has been done to detect driver distraction and to design preventive systems to improve driving safety. According to the work in [5], the most commonly used inputs to detect driver distraction include driver biological measures [6], [7], driver physical measures [6]–[15], and driving performance measures [8]–[10], [15], [16]. It is noticeable that, for most of the comprehensive models, the detection of driver distraction is essentially based on the driver physical measures,

Manuscript received January 3, 2013; revised March 20, 2013; accepted April 20, 2013. Date of publication May 23, 2013; date of current version August 28, 2013. The Associate Editor for this paper was Dr. Q. Kong.

R. Tian, L. Li, M. Chen, and Y. Chen are with the Department of Electrical and Computer Engineering and also with the Transportation Active Safety Institute, Indiana University–Purdue University Indianapolis, Indianapolis, IN 46202 USA (e-mail: rtian@iupui.edu; ll7@iupui.edu; chen306@iupui.edu; ychen@iupui.edu).

G. J. Witt is with the Delphi Electronics and Safety, Kokomo, IN 46092 USA (e-mail: Gerald.J.Witt@delphi.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TITS.2013.2261988

particularly eye/head movements or gaze analysis results [6]–[15]. Machine learning algorithms, such as support vector machines [8], [15] and AdaBoost [6], [7], are commonly used to understand the human behavior patterns. Some researchers combine other measures, such as driver input patterns and vehicle moving trajectories, along with the eye gaze patterns, to predict user distraction [8], [9], [15]. However, there are other references indicating that using only eye-glance information may be sufficient to detect large proportion of distraction or to explain distraction-related accidents [12], [13], [17], [18], [24], [25]. In [12], [13], and [17], algorithms were developed, including different measures of eye gaze or glance movements. In [17], [24], and [25], the researchers found that cumulative (total) eye-glance duration in the 5 s before the crash or near-crash events and 1 s during the event can serve as one sufficient measure of driver inattention. The main benefits of using driver gaze or eye-glance patterns for distraction detection include the following.

1) Eye-glance analysis can be completed in real time.
2) Eye-glance data are relatively easy to obtain in naturalistic driving environment using eye-tracking systems.
3) The methods used to collect eye-glance information are nonintrusive for the driving task itself.

Most of the studies in the current literature are completed in driving simulators and/or in designed experimental vehicles. As argued in [19], it is important to improve the known algorithms and models in a naturalistic driving environment. In [19], a series of methods to remove noise and to obtain the glance patterns in naturalistic driving environment are developed. Most studies about driver distraction or inattention in a naturalistic driving environment were based on the Virginia Tech Transportation Institute (VTTI) 100-car naturalistic driving data collection. In [17], different nonlinear algorithms are compared to detect driver distraction based on the eye-glance patterns and concluded that the eye-glance-based distraction is linearly and positively correlated with crash and near-crash probabilities. They also concluded that using the glance history does not help improve the detection of driver distraction [17]. In the other two reports from the National Highway Traffic Safety Administration (NHTSA) [24], [25], the driving risk based on the cumulative (total) off-road eye-glance duration was calculated and showed that 2 s (or longer) of cumulative (total) off-road eye-glance time during 6 s of driving time may significantly increase crash and near-crash probabilities.

After detecting the driver distraction, the next step is to apply effective intervention by providing warnings and to assist the avoidance of crashes [20]. Much effort has been made by researchers to investigate when to give the warnings [20], [22], how to provide warnings, and what types of warnings are effective [20], [21], as well as the responses to the warnings from distracted drivers [21], [23]. One important question after all these effects is the following: If the driver distraction can be limited to one specified level, how much improvement can be achieved in terms of reducing the chances of occurrences of crash and near-crash events? Research proved that using eye-glance patterns can efficiently detect driver distraction and inattention, particularly the cumulative driver off-road eye-glance duration in the 6 s of driving period. Thus, the *main purpose* of this study is to develop an effective model to calculate the probability of crash and near-crash events when certain cumulative off-road eye-glance duration in 6 s of driving time has been reached in naturalistic driving environment. To better understand the proposed relationship between the cumulative off-road eye-glance duration and the probability of crash and near-crash events, traffic density levels are also considered in the model.

The *contributions* of this paper are fourfold. First, this paper is different from other algorithms proposed in literature because we prove the *direct effect* of driver cumulative off-road eye-glance duration in 6 s on the probability of crash and near-crash events in naturalistic driving environment. Second, the results of this paper partly comply with the findings in one recent publication [17] that linear relationship can be achieved between the cumulative off-road eye-glance duration in 6 s and the risk of crash and near-crash events, but we also detect the effects of off-road eye-glance locations on the risk, which is not proven to improve distraction estimates in [17]. Third, the moderating effects of traffic density on risk estimation based on cumulative off-road eye-glance duration are proven, and we show that the model combining cumulative off-road eye-glance duration and traffic density can improve the power to predict the probability of crash and near-crash events. Finally, the proposed approach can facilitate the benefit analysis for the evaluation of active safety systems developed to intervene the driver off-road glance behaviors.

TABLE I
VARIABLES USED FROM THE VTTI 100-CAR NATURALISTIC DRIVING DATA

Event Dataset (Crash/Near-Crash Data)	Sequential eye glance locations in 30 seconds before the event
	Traffic density during the 30 seconds before the event
Baseline Dataset (Safe Driving Data)	Sequential eye glance locations in six seconds
	Traffic density during the six seconds safe driving period

II. SUMMARY OF VIRGINIA TECH TRANSPORTATION INSTITUTE 100-CAR STUDY EYE-GLANCE ANALYSIS RESULTS

This paper is based on the data collected from the VTTI 100-car naturalistic driving study and the subsequent eye-glance analysis results [24]–[26]. The VTTI 100-car naturalistic driving study is one large-scale data collection completed at VTTI. According to [26], a total of 100 cars were employed, and the driving data were continuously collected for about one year for each vehicle. After collecting all the data, the *event* data set and the *baseline* data set were constructed [26]. In the data sets released to the public, the event data set contains 30-s data for all the crash and near-crash events found in the data collection, and the baseline data set contains randomly selected 6 s of safe driving periods across all subjects along the one-year period to represent the general naturalistic driving patterns of all drivers. Eye-glance analysis has been completed for all cases in the event data set and around 5000 cases in the baseline data set to determine the driver eye-glance locations sequentially along the entire period [24], [25]. The variables used from the two data sets in this paper are listed in Table I.

In this paper, driver cumulative off-road eye-glance duration in a period of 6 s is used as the single variable to reflect driver distraction. This measure is mainly based on Chapter 6 of the NHTSA report related to the VTTI 100-car study [25] and is also supported by other studies [17] [24]. As reviewed earlier, a large amount of measures or inputs has been used to detect driver distraction, inattention, or other abnormal behaviors [5]; however, the studies [12], [13], [18], [27] also emphasized that driving gaze behavior analysis may be able to tell the driver distraction or abnormal driver physiological and psychological states sufficiently. Towards eye-glance analysis, different measures could be applied [25], such as 1) the total time that the eyes are off the forward roadway in 6 s, 2) the number of glances away from the forward roadway in 6 s, 3) the length of the longest glance away from the forward roadway in 6 s, and 4) the location of the longest glance away from the forward roadway [25, Tab. 6.1]. However, in three

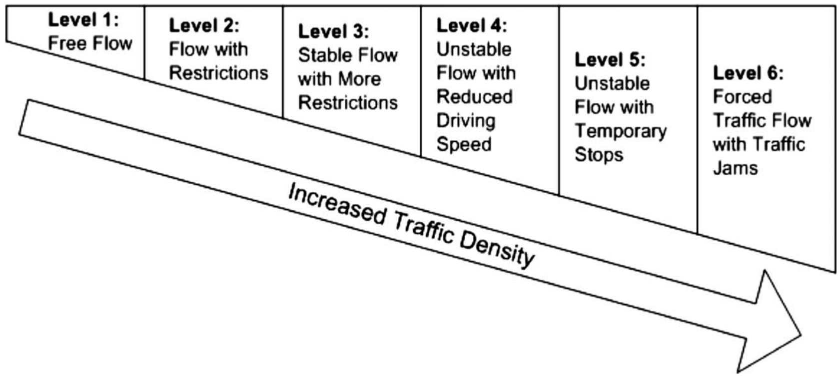

Fig. 1. Traffic density levels based on VTTI 100-car study data analysis report [26].

related studies [17], [25], [26], it was concluded or cited that the total time the eyes are off the forward roadway in 6 s is the most important measure affecting driving risk, and 2 s is proven to be the threshold based on their calculations that significantly increase driving risk.

III. Methodology

A. Definitions of Random Events

In this paper, we are interested to identify the effects of cumulative off-road eye-glance duration in every 6 s and traffic density on the probability of occurrences of crash and near-crash events in a naturalistic driving environment. Three main random events are defined as follows.

- Event A: This is the crash or near-crash event in a naturalistic driving environment. $A = 1$ means the occurrence of crash and/or near-crash event A.
- Event EG: The cumulative driver off-road eye-glance duration in 6 s of naturalistic driving time is higher than or equal to a particular threshold. $\text{EG} \geq \alpha$ indicates that the event EG has occurred. The off-road glance is defined and based on the eye-glance analysis results of the VTTI 100-car study [24] [25]. There are 15 eye-glance locations defined in the analysis, and based on this, two levels of off-road eye glances are defined in this paper:
- Driver off-road eye-glance level I: Eye glances that are not in the "forward," "left forward," and "right forward" direction are defined as off-road eye glances. This is one absolute definition that directly discriminate driver eye behaviors based on whether it is on the road in front of the car or not.
- Driver off-road eye-glance level II: Eye glances that are not in the "forward," "left forward," and "right forward" directions and are also not in the "rear-view mirror," "left mirror," and "right mirror" directions are defined as off-road eye glances. The second definition of off-road eye glances is based on the consideration that the off-road glance duration in this paper is one measure used to predict driver distraction. Thus, although the visions focusing on the mirrors can represent the fact that the driver is not looking on the road in front of the car, there is a high probability that the driver is still focusing on driving and is just trying to check the situation around the vehicle. Thus, the eye glances to the mirrors may not be highly related to driver distractions. Comparisons are made between the two definitions of driver off-road eye glances in terms of the capability to affect the probability of crash and near-crash events.
- Event TD: This is the traffic density in the roadway where the car moving in naturalistic driving environment reaches one particular level. $\text{TD} = \beta$ indicates that event TD has occurred. The traffic density levels in this paper follow the definitions in [26], which were defined based on the NHTSA standards. Six levels of traffic density are shown in Fig. 1.

From levels 1 to 6, the traffic density in naturalistic driving environment keeps increasing continuously. (For detailed explanations of the six definitions, see the VTTI 100-car study report [26].) For traffic density levels 1, 2, and 3, although some restrictions have started to emerge in the stable flow requiring more driving maneuvers, the driving speed in these three levels are still maintainable. As for traffic density levels 2, 3, and 4, compared with level 1, these three density levels all have different amounts of restrictions that affect driving behavior; compared with levels 5 and 6, vehicles can still move constantly in the restrict flow, although the driving speeds may need to be reduced. For the highest two traffic density levels 5 and 6, long or short stoppages are required for the drivers to avoid crashes with other vehicles in the roadway. When the car is moving in this type of road condition, the relationship between cumulative off-road eye-glance duration and driver distraction, and between driver distraction and the risk of crash and near-crash, are greatly altered compared with normal driving conditions. This is mainly because the frequent stoppages and extremely low driving speed may cause the driver distraction to be not as risky as when the car is moving faster, which greatly increase the difficulty to predict eye-glance behaviors. Therefore, when we perform the regression to find the relationship between cumulative off-road eye-glance duration, traffic density levels, and the probability of the crash and near-crash events in this paper, we will only consider the traffic density levels from 1 to 4.

Based on the random events defined earlier, the corresponding probabilities include 1) $P(A = 1)$, which is the probability of crash and near-crash events in naturalistic driving environment; 2) $P(\text{EG} \geq \alpha)$, which is the probability that the cumulative off-road eye-glance duration in 6 s of naturalistic driving time is higher than or equal to threshold α; and 3) $P(\text{TD} = \beta)$, which is the probability that the traffic density of naturalistic driving environment has reached the level of β. In this paper, we assume that the probability of occurrences of crash and near-crash events in naturalistic driving environment $P(A = 1)$ is one constant value. Thus, our main objective is to calculate the conditional probabilities of the following.

1) $P(A = 1|\text{EG} \geq \alpha)$: The probability of crash or near-crash events, given the cumulative driver off-road eye-glance duration in 6 s of driving time in a naturalistic driving environment, is higher than or equal to α.
2) $P(A = 1|\text{EG} \geq \alpha, \text{TD} = \beta)$: The probability of crash or near-crash events, given the cumulative driver off-road eye-glance duration in 6 s of driving time in a naturalistic driving environment, is higher than or equal to α, and the traffic density level is equal to β.

B. Modeling the Effects of Cumulative Off-Road Eye Glance Duration on the Probability of Crash/Near-Crash Events

To calculate $P(A=1|\mathrm{EG}\geq\alpha)$, both the baseline data set and the event data set aforementioned are used. For cumulative off-road eye-glance duration from 0 to 6 s with the increment of 0.1 s, the total numbers of cases in the two data sets are counted as e_i and b_i with $(0\leq i\leq 60)$. Therefore, e_i refers to the total number of cases in the event data set that have $0.1^{*}i$ s of the total off-road eye-glance time in the 6 s before the event, and b_i is the total number of cases in the baseline data set that have $0.1^{*}i$ s of the total off-road eye-glance time in the 6 s of a normal driving period.

Note that, based on Bayes' theorem [28], we have

$$P(A=1|\mathrm{EG}\geq\alpha)=\frac{P(\mathrm{EG}\geq\alpha|A=1)\times P(A=1)}{P(\mathrm{EG}\geq\alpha)}. \tag{1}$$

We assume that the baseline data set is large enough (almost 5000 cases) to represent all the naturalistic driving cases; then. the probability of cumulative off-road eye-glance duration in every 6 s of driving in naturalistic environment to be higher than or equal to α can be calculated as follows using the b_i defined earlier:

$$P(\mathrm{EG}\geq\alpha)=\frac{\sum_{i=\alpha}^{60}b_i}{\sum_{i=0}^{60}b_i}. \tag{2}$$

Similarly, the conditional probability that, given a crash or near-crash event has occurred, the probability that the total off-road eye-glance duration in the 6 s prior to the event to be longer than or equal to α can be calculated as follows, using the e_i defined earlier:

$$P(\mathrm{EG}\geq\alpha|A=1)=\frac{\sum_{i=\alpha}^{60}e_i}{\sum_{i=0}^{60}e_i}. \tag{3}$$

Based on (1), the probability of the occurrences of crash and near-crash events given that the cumulative off-road eye-glance duration has reached one threshold α can be calculated as

$$P(A=1|\mathrm{EG}\geq\alpha)=\frac{\sum_{i=\alpha}^{60}e_i/\sum_{i=0}^{60}e_i}{\sum_{i=\alpha}^{60}b_i/\sum_{i=0}^{60}b_i}\times P(A=1). \tag{4}$$

Note that $P(A=1)$ is assumed to be one constant value; we can define the *single relative risk* RR_1 as

$$\mathrm{RR}_1(\alpha)=\frac{\sum_{i=\alpha}^{60}e_i/\sum_{i=0}^{60}e_i}{\sum_{i=\alpha}^{60}b_i/\sum_{i=0}^{60}b_i} \tag{5}$$

so that (4) can be simplified as

$$P(A=1|\mathrm{EG}\geq\alpha)=\mathrm{RR}_1(\alpha)\times P(A=1). \tag{6}$$

Since the value of $P(A=1)$ is assumed to be a constant $\forall_{\alpha_1}$ and $\alpha_2\in(0.1,0.2,0.3,\ldots,5.9,6.0)$, then, given two thresholds α_1 and α_2, the ratio of $\mathrm{RR}_1(\alpha_1)/RR(\alpha_2)$ may capture the different probability of crash and near-crash events between the corresponding driver eye-glance behaviors.

C. Modeling the Effects of Cumulative Off-Road Eye Glance Duration and Traffic Density on the Probability of Crash/Near-Crash Events

To calculate probability $P(A=1|\mathrm{EG}\geq\alpha,\ \mathrm{TD}=\beta)$ for cumulative driver off-road eye-glance duration from 0 to 6 s with the increment of 0.1 s, and for each traffic density levels, the total numbers of cases in the baseline data set and the event data set are counted as $e_{i,j}$ and $b_{i,j}$ with $(0\leq i\leq 60$ and $j=1,2,3,4,5,6)$, where $e_{i,j}$ (respectively, $b_{i,j}$) is the number of cases with cumulative off-road eye-glance duration equals to $0.1*i$ s and with a traffic density level equals to j in the event (respectively, baseline) data set. Thus, two matrices can be defined as follows:

$$\begin{bmatrix} e'_{0,1} & \cdots & e'_{0,6}\\ \vdots & \ddots & \vdots\\ e'_{60,1} & \cdots & e'_{60,6}\end{bmatrix}\ \text{and}\ \begin{bmatrix} b'_{0,1} & \cdots & b'_{0,6}\\ \vdots & \ddots & \vdots\\ b'_{60,1} & \cdots & b'_{60,6}\end{bmatrix}.$$

Note that, based on the conditional probability equations [28], we have

$$\begin{aligned}&P(A=1,\mathrm{EG}\geq\alpha,\mathrm{TD}=\beta)\\&=P(A=1|\mathrm{EG}\geq\alpha,\mathrm{TD}=\beta)\\&\quad\times P(\mathrm{EG}\geq\alpha|\mathrm{TD}=\beta)\times P(\mathrm{TD}=\beta)\end{aligned} \tag{7}$$

$$\begin{aligned}&P(A=1,\mathrm{EG}\geq\alpha,\mathrm{TD}=\beta)\\&=P(\mathrm{EG}\geq\alpha|A=1,\mathrm{TD}=\beta)\\&\quad\times P(A=1|\mathrm{TD}=\beta)\times P(\mathrm{TD}=\beta).\end{aligned} \tag{8}$$

Clearly, (7) and (8) should be equal. By letting the right-hand side of (7) and (8) to be equal and by removing the same multiplier $P(\mathrm{TD}=\beta)$, we obtain

$$\begin{aligned}&P(A=1|\mathrm{EG}\geq\alpha,\mathrm{TD}=\beta)\\&=\frac{P(\mathrm{EG}\geq\alpha|A=1,\mathrm{TD}=\beta)\times P(A=1|\mathrm{TD}=\beta)}{P(\mathrm{EG}\geq\alpha|\mathrm{TD}=\beta)}\end{aligned} \tag{9}$$

Based on Bayes' theorem, we have

$$P(A=1|\mathrm{TD}=\beta)=\frac{P(\mathrm{TD}=\beta|A=1)\times P(A=1)}{P(\mathrm{TD}=\beta)}. \tag{10}$$

By substituting (10) into (9), we have

$$\begin{aligned}&P(A=1|\mathrm{EG}\geq\alpha,\mathrm{TD}=\beta)\\&=\frac{P(\mathrm{EG}\geq\alpha|A=1,\mathrm{TD}=\beta)\times P(\mathrm{TD}=\beta|A=1)}{P(\mathrm{EG}\geq\alpha|\mathrm{TD}=\beta)\times P(\mathrm{TD}=\beta)}\times P(A=1).\end{aligned} \tag{11}$$

All probabilities in (11) can be calculated using variables $e'_{i,j}$ and $b'_{i,j}$ as follows:

$$P(\mathrm{EG}\geq\alpha|A=1,\mathrm{TD}=\beta)=\frac{\sum_{i=\alpha}^{60}e'_{i,\beta}}{\sum_{i=0}^{60}e'_{i,\beta}} \tag{12}$$

$$P(\mathrm{TD}=\beta|A=1)=\frac{\sum_{i=0}^{60}e'_{i,\beta}}{\sum_{i=0}^{60}\sum_{j=1}^{6}e'_{i,j}} \tag{13}$$

$$P(\mathrm{EG}\geq\alpha|\mathrm{TD}=\beta)=\frac{\sum_{i=\alpha}^{60}b'_{i,\beta}}{\sum_{i=0}^{60}b'_{i,\beta}} \tag{14}$$

$$P(\mathrm{TD}=\beta)=\frac{\sum_{i=0}^{60}b'_{i,\beta}}{\sum_{i=0}^{60}\sum_{j=1}^{6}b'_{i,j}}. \tag{15}$$

By substituting the given expressions into (11), we obtain

$$\begin{aligned}&P(A=1|\mathrm{EG}\geq\alpha,\mathrm{TD}=\beta)\\&=\frac{\sum_{i=\alpha}^{60}e'_{i,\beta}\Big/\sum_{i=0}^{60}\sum_{j=1}^{6}e'_{i,j}}{\sum_{i=\alpha}^{60}b'_{i,\beta}\Big/\sum_{i=0}^{60}\sum_{j=1}^{6}b'_{i,j}}\times P(A=1).\end{aligned} \tag{16}$$

Based on the assumption that $P(A=1)$ is one constant value, we can define the *multiple relative risk* RR_2 as

$$\mathrm{RR}_2(\alpha,\beta) = \frac{\sum_{i=\alpha}^{60} e'_{i,\beta} \Big/ \sum_{i=0}^{60}\sum_{j=1}^{6} e'_{i,j}}{\sum_{i=\alpha}^{60} b'_{i,\beta} \Big/ \sum_{i=0}^{60}\sum_{j=1}^{6} b'_{i,j}}. \tag{17}$$

Therefore, the probability of crash and near-crash events given that the cumulative off-road eye-glance duration is longer than or equal to β and that traffic density level is equal to α can be calculated as

$$P(A=1|\mathrm{EG}\geq\alpha, \mathrm{TD}=\beta) = \mathrm{RR}_2(\alpha,\beta)\times P(A=1). \tag{18}$$

Considering that the value of $P(A=1)$ is assumed to be a constant in naturalistic driving environment for any α_1 and $\alpha_2 \in (0.1, 0.2, 0.3, \ldots, 6.0)$ and any β_1 and $\beta_2 \in (1, 2, 3, 4)$, the ratio of $\mathrm{RR}_1(\alpha_1, \beta_1)/\mathrm{RR}_1(\alpha_2, \beta_2)$ can capture as the difference of the probability of crash and near-crash events in terms of corresponding driver eye-glance behaviors and traffic density levels.

D. Estimation of Probability of Crash And Near-Crash Events

The crash and near-crash probability is mainly affected by RR_1 and RR_2 values. For each given driver off-road glance duration α, RR_1 can be calculated, and if the traffic density β is also given, RR_2 can also be calculated. Thus, using the VTTI 100-car study data, the calculated RR_1 and RR_2 values $\forall\alpha \in (0.1, 0.2, 0.3, \ldots, 5.9, 6.0)$ and $\forall\beta \in (1, 2, 3, 4)$ can be modeled using exploratory regressions to find the relationship between the probability of crash and near-crash events, driver cumulative off-road eye-glance duration, and traffic density levels. Note that only traffic density levels from 1 to 4 will be studied. This is because of the different driver behavior patterns and road conditions in the levels of 5 and 6, as in Section II-A.

IV. Data Preparation

Four data sets from the VTTI 100-car study (which were released to the public) have been used for our analysis. All the data have been deidentified before being released to the public. The four data sets included are the following.

- Crash and near-crash event video reduced data. There were a total of 828 cases of crashes or near crashes. For each case, there was one variable of traffic density indicating the corresponding traffic density level.
- Crash and near-crash event video eye-glance analysis data. There were a total of 786 cases of crashes or near crashes. For 30 s of period prior to every event, eye-glance locations were sequentially recorded with the accuracy up to 0.1 s.
- Baseline video reduced data. There were a total of 19 616 safe driving cases in the data set, and traffic density level has been assigned to each case.
- Baseline video eye-glance analysis data. There were a total of 4950 6-s safe driving sessions being randomly selected from all drivers in the one-year period. Eye glance locations during each 6-s session have been sequentially recorded with the accuracy up to 0.1 s.

Fig. 2 shows the data processing steps toward the baseline and crash databases released from the VTTI 100-car study. In the four raw data sets, the event video glance data set contained 30-s cases. Thus, the first step was to cut the last 6 s for every case in the event video glance data set. Then, cumulative off-road driver eye-glance duration in a 6-s period has been calculated for each case in the baseline and shortened event video glance data sets, following the definitions of the aforementioned off-road eye-glance levels I and II. Because the total

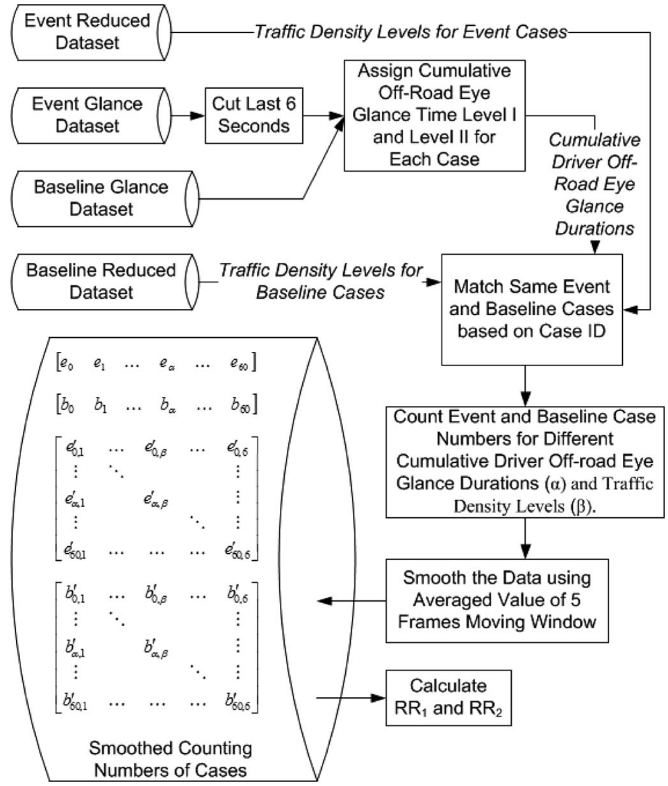

Fig. 2. Process of data analysis.

number of cases in the video glance data sets and reduced data sets for both event and baseline were not the same, the matching process was performed to combine the data sets together based on the case IDs. The combined data set contained both the driver cumulative off-road eye-glance duration and the traffic density level for every event and baseline cases.

The next step was to count the total number of cases in the combined data sets based on two conditions:

1) cumulative driver off-road eye-glance duration from 0.1 to 6 s;
2) the combinations of cumulative driver off-road eye-glance duration from 0.1 to 6 s and different traffic density levels.

The counted numbers were smoothed to generate the data to calculate RR_1 and RR_2 values. Smoothing the raw data could reduce the noise generated during naturalistic data collection [19]. In this paper, the counted numbers of cases for different cumulative off-road eye-glance duration (with or without considering different traffic densities) were supposed to change smoothly when the duration was increasing at a relatively small unit of 0.1 s. However, in the raw data, it is noticeable that the oscillation of the data dots exists along with a middle line, which may be caused by different errors or deviations during the data collection process and environment. The smoother used in this paper is one moving-window-based five-frame averaging method. As one example, Fig. 3 shows the smoothing results of the number of cases in the baseline data set with different driver cumulative off-road eye-glance duration across the range from 0.3 to 5.8 s. Because the smoothing method used the averaged value of five frames, the first two frames and the last two frames of the data were dropped automatically. As shown in Fig. 3, the data smoother could reduce the noise and generate a clearer trend of the data, following the middle line of the raw data.

V. Data Analysis Results

Toward the smoothed data, RR_1 and RR_2 values are calculated for all driver cumulative off-road eye-glance duration across the range of

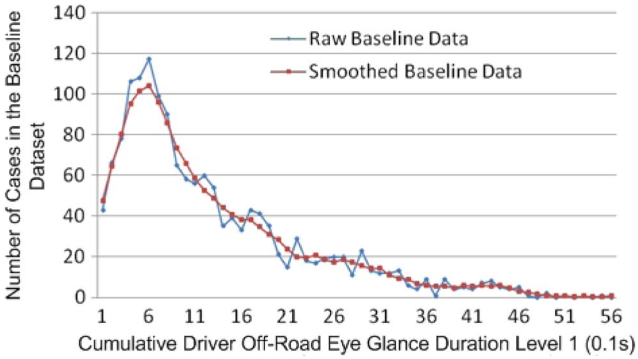

Fig. 3. Data smoothing results of the baseline data set for cumulative driver off-road eye-glance level 1.

TABLE II
LINEAR REGRESSION RESULTS FOR THE EFFECTS OF CUMULATIVE DRIVER OFF-ROAD EYE GLANCE DURATION ON RR_1 USING TWO LEVELS OF OFF-ROAD GLANCE DEFINITIONS

Off-Road Glance Level	*p*-value	R^2	Adjusted R^2
Level I	<0.0001	0.73	0.73
Level II	<0.0001	0.77	0.77

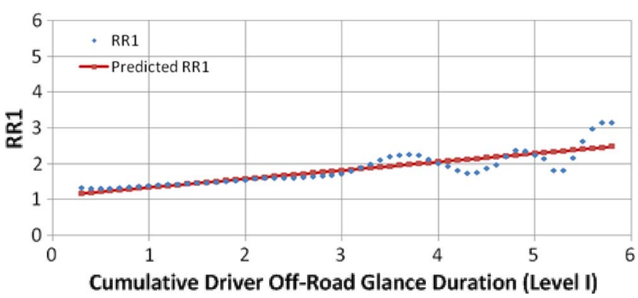

Fig. 4. Fit plot for the estimation of RR_1 values using cumulative driver off-road eye-glance duration calculated following the first definition of off-road glances.

0.3 to 5.8 s or for the combinations with different traffic density levels. Exploratory regressions are performed to search for the underlying relationship.

A. Estimation of the RR_1

Linear regression has been completed to study the effects of driver cumulative off-road eye-glance duration on RR_1 values. The results are shown in Table II. In general, the linear regression analysis returns significant linear equations for both levels of off-road glances defined earlier. The p-values for both models are much less than 0.0001, which indicates the existence of a significant relationship. The R^2 and adjusted R^2 values for both models are in the range of 0. 7–0.8. When using the first (respectively, second) definition of off-road eye glances, the RR_1 can be calculated using the following:

$$\mathrm{RR}_1 = 1.102 + 0.237 * \alpha \tag{19}$$
$$\mathrm{RR}_1 = 1.134 + 0.584 * \alpha \tag{20}$$

where α stands for the driver cumulative off-road eye-glance duration $(0.3 \le \alpha \le 5.8)$.

The fit plots of the two models are shown in Figs. 4 and 5, respectively. It is clear that no matter which definition of off-road eye glances is used, one linear relationship between driver cumulative off-road eye-glance duration and the RR_1 value can be found. When we increase the driver cumulative off-road eye-glance duration, the risk of crash and near-crash events is always increasing. Therefore, when we can control the glance behavior in terms of driver off-road eye duration, we can use the ratio of calculated RR_1 values to estimate the change of driving risks.

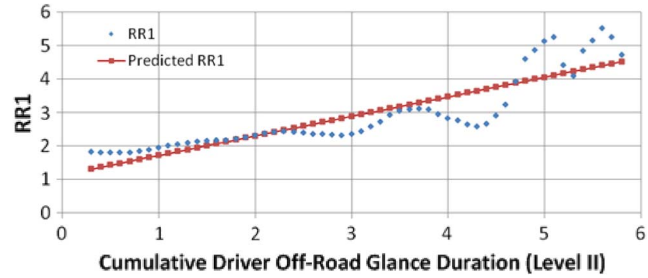

Fig. 5. Fit plot for the estimation of RR_1 values using cumulative driver off-road eye-glance duration calculated following the second definition of off-road glances.

TABLE III
COMPARISON OF THE FITTED SLOPES IN THE REGRESSION EQUATIONS TOWARD THE TWO DIFFERENT OFF-ROAD EYE GLANCE DEFINITIONS

Off-Road Glance Level	Slope	Slope Upper Bound (95%)	Slope Lower Bound (95%)
Level I	0.237	0.276	0.198
Level II	0.584	0.671	0.497

TABLE IV
LOGARITHMIC LINEAR REGRESSION RESULTS FOR THE EFFECTS OF CUMULATIVE DRIVER OFF-ROAD EYE GLANCE DURATION AND TRAFFIC DENSITY LEVELS ON RR_2 USING TWO LEVELS OF OFF-ROAD GLANCE DEFINITIONS

Off-Road Glance Level	*p*-value	R^2	Adjusted R^2
Level I	<0.0001	0.85	0.85
Level II	<0.0001	0.83	0.82

Two noticeable differences between Fig. 4 and Fig. 5 include the following.

1) The slope of the fit plot in Fig. 5 is larger than the slope of the fit plot in Fig. 4.
2) The RR_1 values calculated using the second definition of off-road eye glances are much higher than those values calculated using the first definition, particularly for longer duration.

The comparisons of the slopes shown in Table III indicate that the two slopes using different definitions of off-road eye glances are significantly different. Using more strict definition, level II will result in much faster increase of risk with increasing off-road glance duration. For shorter off-road eye glances, the difference is not big in terms of the probability of crash and near-crash events, no matter whether the off-road eye glance is to the mirrors or somewhere else. However, when the off-road glance duration becomes longer, the results show that the glances to the mirrors are much safer compared with the glances to other places.

B. Estimation of the RR_2

Logarithmic linear regression has been performed to study the effects of driver off-road glance duration and traffic density on RR_2 using two different definitions of off-road glances. The regression results are shown in Table IV.

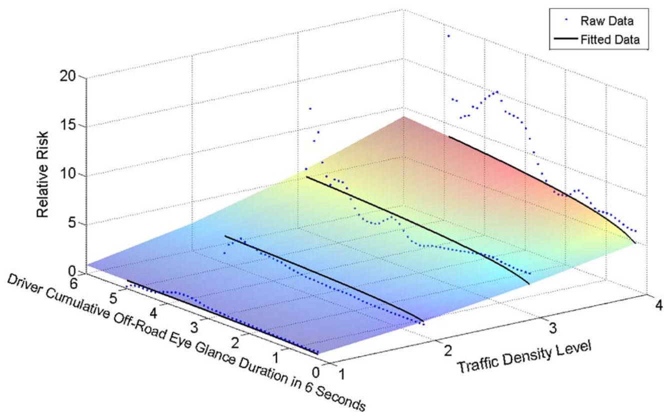

Fig. 6. Fit plot for the estimation of RR_2 using cumulative driver off-road eye-glance duration calculated following the first level of off-road glances and traffic density.

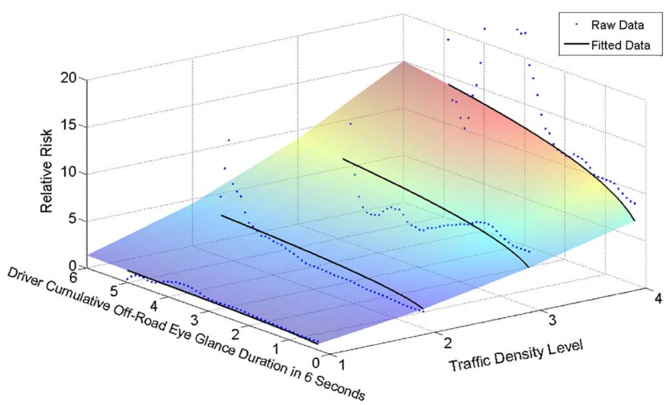

Fig. 7. Fit plot for the estimation of RR_2 using cumulative driver off-road eye-glance duration calculated following the second level of off-road glance and traffic density

The results show that both of the two models are significant ($p-$values < 0.0001) with good R^2 values. Compared with the results shown in Table II, it is clear that, by combining the traffic density with driver off-road glance duration in the model, we can increase the predictive power of the driving risk with increased R^2 values. The fitted regression equation to calculate the RR_2 using driver cumulative off-road eye-glance duration $\alpha(0.3 \le \alpha \le 5.8)$ and traffic density level $\beta(\alpha \le \beta \le 4)$ based on the first and second definitions of off-road glances are given, respectively, in the following:

$$\mathrm{RR}_2 = 10^{-0.448} * \alpha^{1.693} * \beta^{0.216} \tag{21}$$

$$\mathrm{RR}_2 = 10^{-0.353} * \alpha^{1.739} * \beta^{0.27}. \tag{22}$$

The fitted plots for the two models are shown in Fig. 6 and Fig. 7, respectively. The surfaces in the two plots represent the fitted equations earlier. The solid lines represent the fitted values for the collected data, and the dots are the raw data. Comparing these two plots, it is noticeable that, similar to the findings during the estimation of RR_1 values, the level II definition of off-road glances results in much higher RR_2 estimations for the same length of driver cumulative off-road eye-glance duration under the same traffic density levels. This difference is not very large when the traffic density level is low or the cumulative off-road eye-glance duration is short; however, the difference becomes much more noticeable when the traffic density level becomes higher and cumulative off-road eye-glance duration becomes longer. At the most extreme point of traffic density level 4 and cumulative off-road

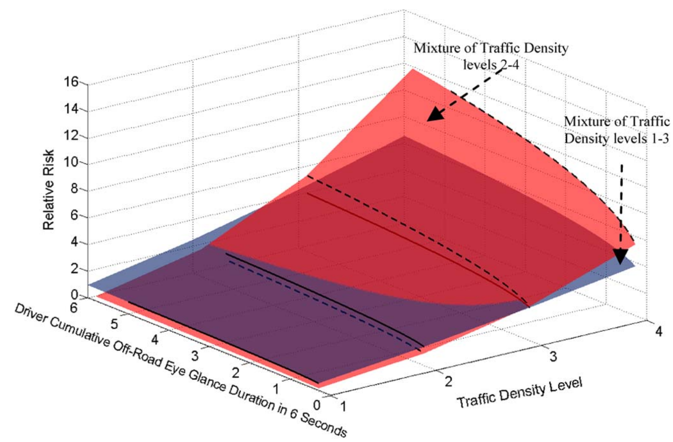

Fig. 8. Comparison of the fit plots about the two examples of mixtures of different traffic density levels.

eye-glance duration close to 6 s, the difference between estimated RR_2 values increased to 50%.

Although the two levels of off-road glance definitions are associated with different RR_2 estimations, similar trends that can be found in the two surfaces include the following.

- With the increase in driver cumulative off-road eye-glance duration and traffic density levels, the estimated RR_2 value (risk) is increasing in both dimensions.
- The effects of driver cumulative off-road eye-glance duration on RR_2 is limited for traffic density level 1, but the effect becomes much stronger as the traffic density level increases, particularly when the level II definition of off-road glances is used.

VI. Discussion

One important finding of this paper is that traffic density β has significant moderating effects on the relationship between driver cumulative off-road eye-glance duration α and the probability of crash and near-crash events RR_2. The moderating effect has been proven by using the following regression model when the coefficient α_4 for the interaction multiplier $a \times \beta$ is significantly different from 0 with a p-value of < 0.001 for both levels of off-road glance definitions:

$$\mathrm{RR}_2 = \alpha_1 + \alpha_2 \times \alpha + \alpha_3 \times \beta + \alpha_4 \times \alpha \times \beta. \tag{23}$$

In naturalistic driving environment, the traffic density can be highly dynamic. In many cases, it may not be that easy to assign one traffic density level, particularly along a certain period of time. Thus, instead of using one traffic density level, the driving environment may actually be a combination of different traffic density levels. Based on the traffic density definitions earlier, we use the "mixture of traffic density levels 1–3" and the "mixture of traffic density levels 2–4" as two examples to simulate the naturalistic driving environment. The "mixture of traffic density levels 1–3" mainly considers that vehicles can still move in the desired speed, and the "mixture of traffic density levels 2–4" mainly considers that the movement of vehicles is no longer totally free of other obstacle objects. These two driving environments share two common traffic density levels.

Fig. 8 shows the fit plots of the two different driving conditions about the effects of the driver cumulative off-road eye-glance duration and the traffic density on RR_2 values. It seems that, at lower traffic density levels, the differences are not as much as those from higher traffic density levels between the two driving conditions. The slopes of the fitted logarithmic linear regression equations show that, in both traffic environments, cumulative off-road eye-glance duration

TABLE V
COMPARISON OF THE SLOPES IN THE FITTED LOGARITHMIC LINEAR REGRESSION EQUATIONS BETWEEN THE "MIXTURE OF TRAFFIC DENSITY LEVELS 1–3" AND THE "MIXTURE OF TRAFFIC DENSITY LEVELS 2–4"

	Slopes	**95% Upper Level**	**95% Lower Level**
Off-Road Eye Glance Duration (α)	TD 1-3: 0.19* TD 2-4: 0.269	TD 1-3: 0.268 TD 2-4: 0.325	TD 1-3: 0.113 TD 2-4: 0.213
Traffic Density Level (β)	TD 1-3: 1.399* TD 2-4: 2.564*	TD 1-3: 1.523 TD 2-4: 2.709	TD 1-3: 1.276 TD 2-4: 2.419

* p-value < 0.001
TD 1-3: Mixture of Traffic Density 1-3
TD 2-4: Mixture of Traffic Density 2-4

plays equally significant effects on driving risks because, although the corresponding slopes are both significantly different from 0, they are not significantly different from each other. However, the effects of traffic density levels in these two driving conditions are not only significantly different from 0 but significantly different from each other as well, as shown in Table V. This result suggests that 1) driver off-road glance duration is one significant factor to affect driving safety, although this effect may not always be significantly moderated by the driving environment, particularly when multiple traffic density levels are shared between the conditions; and 2) in the more complicated driving environment, increased traffic density levels may increase the driving risks even faster.

VII. CONCLUSION

In this paper, the effects of driver cumulative off-road eye-glance durations in 6 s and traffic density levels on driving risks in terms of the probability of crash and near-crash events have been modeled and analyzed using naturalistic driving data. Two levels of off-road glance definitions are proposed and discussed. The developed models allow the estimation of changes of driving risks for different eye-glance behaviors and for the combinations of different traffic density levels. The driver cumulative off-road eye-glance duration is shown to have significantly linear effect on the risk of crash and near-crash events, and traffic density is one significant moderator to this relationship. Combining the driver glance behavior and traffic density together can increase the predication power of the probability of crash and near-crash events. The significant nonlinear relationship between them has also been proposed in this paper. Off-road eye-glance locations are also proven to have significant effect on the relationships. In particular, for longer duration, off-road eye glances to the vehicle mirrors are proven to be significantly safer compared with the glances to other off-road locations. In general, increased driver off-road glance duration and traffic density will increase the probability of crash and near-crash events, and the effects are stronger for higher traffic density levels and longer off-road glance duration.

One limitation of this paper is the sample size. Although the 100-car study completed at VTTI is one of the largest naturalistic driving data collection completed in the world, the usable event data sets are still relatively small, particularly when more variables are included. This may reduce the power of the data to represent the entire picture in the naturalistic driving environment. With more data collected in the future, more influential factors and moderators may be added into the predication model. Another limitation of this paper is that, although two levels of off-road glance definitions are used, only the cumulative duration is used to measure the driver eye behavior. Driver distraction may be better predicted using more complicated algorithms. In addition, the comparisons between findings in naturalistic driving with laboratory-based experimental results may further consolidate the models.

REFERENCES

[1] J. D. Lee, K. L. Young, and M. A. Regan, "Defining driver distraction," in *Driver Distraction: Theory, Effects, and Mitigation*, K. Young, J. D. Lee, and M. A. Regan, Eds. Boca Raton, FL, USA: CRC, 2009, pp. 31–40.

[2] K. Young and M. Regan, "Driver distraction: A review of the literature," in *Distracted Driving*, I. J. Faulks, M. Regan, M. Stevenson, J. Brown, A. Porter, and J. D. Irwin, Eds. Sydney, NSW, Australia: Australasian College Road Safety, 2007, pp. 379–405.

[3] T. Horberry, J. Anderson, M. A. Regan, T. J. Triggs, and J. Brown, "Driver distraction: The effects of concurrent in-vehicle tasks, road environment complexity and age on driving performance," *Accident Anal. Prev.*, vol. 38, no. 1, pp. 185–191, Jan. 2006.

[4] B. Megan, K. L. Young, and M. A. Regan, "Sources of distraction inside the vehicle and their effects on driving performance," in *Driver Distraction: Theory, Effects, and Mitigation*, K. Young, J. D. Lee, and M. A. Regan, Eds. Boca Raton, FL, USA: CRC, 2009, pp. 191–213.

[5] Y. Dong, Z. Hu, K. Uchimura, and N. Murayama, "Driver inattention monitoring system for intelligent vehicles: A review," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 2, pp. 596–614, Jun. 2011.

[6] M. Miyaji, H. Kawanaka, and K. Oguri, "Driver's cognitive distraction detection using physiological features by the adaboost," in *Proc. 12th Int. IEEE Conf. Intell. Transp. Syst.*, St. Louis, MO, USA, Oct. 2009, pp. 1–6.

[7] M. Miyaji, M. Danno, H. Wawanaka, and K. Oguri, "Driver's cognitive distraction detection using adaboost on pattern recognition basis," in *Proc. IEEE Int. Conf. Veh. Electron. Safety*, Columbus, OH, USA, Sep. 2008, pp. 51–56.

[8] Y. Liang, M. L. Reyes, and J. D. Lee, "Real-time detection of driver cognitive distraction using support vector machines," *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 2, pp. 340–350, Jun. 2007.

[9] L. Yekhshatyan and J. D. Lee, "Changes in the correlation between eye and steering movements indicate driver distraction," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 1, pp. 136–145, Mar. 2013.

[10] T. Jirayama, K. Mase, and K. Takeda, "Detection of driver distraction based on temporal relationship between eye-gaze and peripheral vehicle behavior," in *Proc. 15th Int. IEEE Conf. Intell. Transp. Syst.*, Anchorage, AK, USA, Sep. 2012, pp. 870–875.

[11] M. Kutila, M. Jokela, G. Markkula, and M. R. Rue, "Driver distraction detection with a camera vision system," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2007, pp. VI-201–VI-204.

[12] M. Sodhi and B. Reimer, "Glance analysis of driver eye movements to evaluate distraction," *Behav. Res. Methods, Instrum. Comput.*, vol. 34, no. 4, pp. 529–538, Nov. 2002.

[13] A. Azman, Q. Meng, and E. Edirisinghe, "Non-intrusive physiological measurement for driver cognitive distraction detection: Eye and mouth movements," in *Proc. 3rd Int. Conf. Adv. Comput. Theory Eng.*, 2010, pp. V3-595–V3-599.

[14] M. Su, C. Hsiung, and D. Huang, "A simple approach to implementing a system for monitoring driver inattention," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, Taipei, Taiwan, Oct. 2006, pp. 429–433.

[15] M. Wollmer, C. Blaschke, T. Schindl, B. Schuller, B. Farber, S. Mayer, and B. Trefflich, "Online driver distraction detection using long short-term memory," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 2, pp. 574–582, Jun. 2011.

[16] F. Tango, C. Calefato, L. Minin, and L. Canovi, "Moving attention from the road: A new methodology for the driver distraction evaluation using machine learning approaches," in *Proc. 2nd Conf. Human Syst. Interact.*, Catania, Italy, May 2009, pp. 596–599.

[17] Y. Liang, J. D. Lee, and L. Yekhshatyan, "How dangerous is looking away from the road? Algorithms predict crash risk from glance patterns in naturalistic driving," *Hum. Factors*, vol. 54, no. 6, pp. 1104–1116, Dec. 2012.

[18] N. Edenborough, R. Hammoud, A. Harbach, A. Ingold, B. Kisacanin, P. Malaway, T. Newman, G. Scharenbroch, S. Skiver, M. Smith, A. Wilhelm, G. Witt, E. Yoder, and H. Chang, "Driver state monitor from Delphi," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2005, vol. 2, pp. 1206–1207.

[19] C. Ahlstrom, T. Victor, C. Wege, and E. Steinmetz, "Processing of eye/head-tracking data in large-scale naturalistic driving data sets," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 2, pp. 553–564, Jun. 2012.

[20] M. Smith, G. Witt, D. L. Bakowski, D. Leblanc, and J. D. Lee, "Adapting collision warnings to real-time estimates of driver distraction," in *Driver Distraction: Theory, Effects, and Mitigation*, K. Young, J. D. Lee, and M. A. Regan, Eds. Boca Raton, FL, USA: CRC, 2009, pp. 501–518.

[21] J. D. Lee, J. D. Hoffman, and E. Hayes, "Collision warning design to mitigate driver distraction," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, Vienna, Austria, Apr. 2004, pp. 65–72.

[22] J. D. Lee, D. V. McGehee, T. L. Brown, and M. Reyes, "Collision warning timing, driver distraction, and driver response to imminent rear-end collisions in a high-fidelity driving simulator," *Hum. Factors*, vol. 44, no. 2, pp. 314–334, 2002.

[23] G. Abe, M. Itoh, and T. Yamamura, "Visual distraction when driving and its influence on driver response to a forward collision warning system," in *Proc. SICE Annu. Conf.*, Tokyo, Japan, Sep. 2011, pp. 179–183.

[24] S. G. Klauer, J. Feng Guo, and T. A. Sudweeks, "An analysis of driver inattention using a case-crossover approach on 100-car data: Final report," Nat. Highway Traffic Safety Admin., Washington, DC, USA, Tech. Rep. DOT HS 811 334, May 2010.

[25] S. G. Klauer, T. A. Dingus, V. L. Neale, J. D. 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," Nat. Highway Traffic Safety Admin., Washington, DC, USA, Tech. Rep. DOT HS 810 594, Apr. 2006.

[26] T. A. Dingus, S. G. Klauer, V. L. Neale, A. Petersen, S. E. Lee, J. Sudweeks, M. A. Perez, J. Hankey, D. Ramsey, S. Gupta, C. Bucher, Z. R. Doerzaph, J. Jermeland, and R. R. Knipling, "The 100-car naturalistic driving study, Phase II—Results of the 100-car field experiment," Nat. Highway Traffic Safety Admin., Washington, DC, USA, Tech. Rep. DOT HS 810 593, Apr. 2006.

[27] L. Li, D. Wen, N. Zheng, and L. Shen, "Cognitive cars: A new frontier for ADAS research," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 1, pp. 395–407, Mar. 2012.

[28] R. Yates and D. J. Goodman, *Probability and Random Processes: A Friendly Introduction for Electrical and Computer Engineers*, 2nd ed. Hoboken, NJ, USA: Wiley, 2004.