https://doi.org/10.2352/ISSN.2470-1173.2018.17.AVM-163

Auto-simulator Preparation for Research into Assessing the Correlation Between Human Driving Behaviors and Fixation Patterns

Mingming Wang, Kate Walders, Martin E. Gordon, Jeff B. Pelz, Susan Farnand; Rochester Institute of Technology; Rochester, New York, USA

Abstract

Autonomous driving has the potential to positively impact the daily life of humans. Techniques such as imaging processing, computer vision, and remote sensing have been highly involved in creating reliable and secure robotic cars. Conversely, the interaction between human perception and autonomous driving has not been deeply explored. Therefore, the analysis of human perception during the cognitive driving task, while making critical driving decisions, may provide great benefits for the study of autonomous driving. To achieve such an analysis, eye movement data of human drivers was collected with a mobile eye-tracker while driving in a automotive simulator built around an actual physical car, that mimics a realistic driving experience. Initial experiments have been performed to investigate the potential correlation between the driving behaviors and fixation patterns of the human driver.

Introduction

Over the past few decades, autonomous driving has been experiencing rapid development. Advanced techniques such as imaging processing, computer vision, and remote sensing are widely implemented in propelling self-driving car. In 2015, Google made public its driverless project, which heavily depends on sensor detection. Although the sensor based approach seems to be efficient and robust toward achieving the ultimate goal of self-driving, it is not sufficient enough to convince the audience that autonomous driving is capable of making critical and pertinent driving decisions in complicated traffic situations. For example, the robotic car is more likely to be restricted to a full stop when the hard criterion of proceeding is not satisfied in complex driving scenarios, such as rush hour traffic. In order to ensure the safety of the driver and any passengers, it might disrupt the traffic flow and result in inconvenience. In contrast, looking back upon driving history, human drivers have been interacting with vehicles since they were first manufactured. Human drivers acquire extensive driving experience in appropriately dealing with complicated driving situations. Driving decisions are highly motivated by human perception. Therefore, the expectations from the public are that autonomous driving is able to perform reasonably and practically like human drivers do. Naturally, the interaction between human perception and autonomous driving is worth investigating further. The analysis of human perception during the cognitive driving task, while making critical driving decisions, can provide great benefits for the study of autonomous driving.

In 1999, Liu examined driver intentions based on the patterns of eye fixations [1]. Following the methodology of Liu, Underwood and Chapman selected inexperienced drivers as the experimental targets and compared them with experienced drivers [2] [3] [4]. The consequent conclusion verified that the incomplete visual attention (foveal and peripheral) of novice drivers leads to high-risk accident involvement, which addressed the importance of visual perceptions while driving. In later research, Crundall [5], and Lee [6]investigated eye movement activity within more cognitive driving tasks, which emphasized the necessity of visual search in understanding driving behaviors. Along with these on-road studies, some researchers utilized auto-simulators to inspect the connection between the visual pattern and driving performance. For example, Coeckelbergh, and Seya created virtual driving scenarios in learning the effect of visual field defects on steering and lane control [7] [8]. Konstantopoulos, and Underwood exploited similar manner to examine driver's visual spread search under multiple driving illuminations and weather [9].

Although the previous research achievements have indicated an intense connection between visual activities and driving performance, a few potential problems in the experimental design stand out as well. First, one major issue of on-road test scenes is that the environmental factors along the route are hardly maintained as the same. Especially in residential areas, dynamic features, such as pedestrians, pets, and parking vehicles, vary from time to time. Either element is able to affect the visual search in the driving route. The differences introduced by these variables will enormously decrease the ability to compare driving performance. Secondly, Tornros, Underwood, and Godley pointed out that the driving simulator is capable of recovering the actual driving condition and is being validated for assessing driving behavior [10] [11] [12]. However, one prerequisite is that simulation systems must be built up as being comparable with realistic driving situations. Thus, table-based simulation instruments, such as flat monitors and steering system, might be conducive to the sensation of video games in some prior studies, and can cause subjects to make unrealistic driving decisions. Third, the prior research was mostly concentrated on exploring the difference of visual activity by comparing different age groups or multiple cognitive tasks. Yet the practical and concrete goal of autonomous driving probably centers on accident prevention, which is supposed to be closer to the daily driving environment. Hence, the fundamental and routine traffic incidents in populated areas are worth emphasizing from the perspective of visual perception.

The main purpose of this paper is to inspect the visual perception of human drivers when approaching hazard conditions. To achieve such an analysis, the driving route has been designed to include multiple hazard situations in the pilot testing. We created an automotive simulator built around an actual physical car and a virtual scene with constant factors in order to optimize the mentioned issues of preceding methodologies. The pupil activity of human drivers from a mobile eye tracker has been collected when subjects are driving in the simulator. We believe such an experimental setup is able to mimic a realistic driving experience. This paper is divided into the following sections: the instrument composition section focuses on the introduction of instruments and their functionalities; the experimental design section depicts the proposed experiment procedures; the result and discussion section describes the experiment results and the current restriction; the conclusion and future work section summarizes the current work and the future plan.

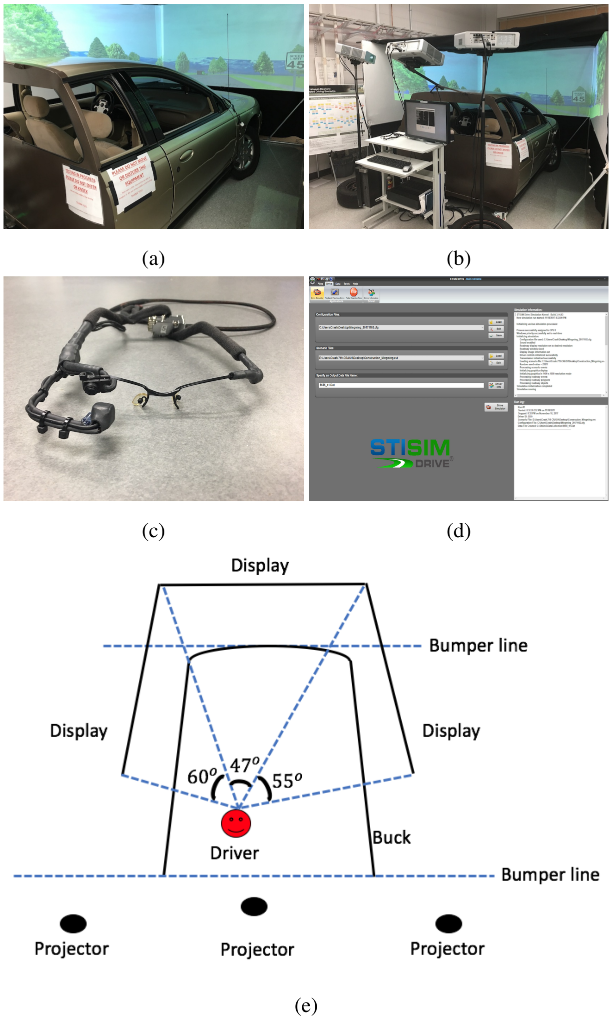

(a) (b) (c) (d) (e)

Figure 1: Instrument composition: (a) the auto-simulator, (b) the display system, (c) Positive Science eye-tracker, (d) STIMSIM simulation software, (e) the layout of the buck-display system.

Instrument Composition

The instruments for our project consist of the following major components shown in Figure 1(a)-(d): the auto-simulator, the display system, the simulation software, and the eye-tracker. For the simulator, a well-maintained 96' Saturn sedan was converted into the buck (frame of the simulator). The back passenger seats and the trunk of this vehicle were cut off to keep 2/3 of its original size. The mechanical control system including the steering wheel, accelerator and brake pedals, were replaced with an electronic system manufactured by Fanatec inc. The final size of the modified buck is 9.8 feet long by 5.2 feet wide. In the display system, three display-projector pairs were adopted to mimic the wide driving view, which includes two side-views and one central view. All three projectors are synchronized to project on the 7.6 feet long by 4.2 feet wide displays at the same resolution of 1920 by 1080. The mobile eye tracker package in the pilot testing, provided by Positive Science, uses two sensors to record the scene and pupil activities separately. The videos captured by these sensors have the same framerate, at 30 frame rate per second (fps), but a different resolution, eye at 240p while scene at 480p. Compatible with this eye-tracker, Yarbus software by Positive Science is utilized to calibrate and analyze the collected data.

There are several reasons for using this eye-tracker in the pilot experiment. First, the configuration of the sensors is sufficient to afford the initial analysis of driving behaviors and visual patterns. Even though the resolution of each sensor is not HD level, the smaller size of collected video will effectively decrease the computational cost. Second, the previous research in applying this eye-tracker has equipped us with abundant experience and confidence for processing such tasks. As for the software part, the simulated scene was rendered by STIMSIM Drive software, which is able to deliver the customized virtual scenarios in different environmental conditions. Another eye-tracker from Pupil Labs is also adopted for the future research, although it was only employed to compare with the Positive Science eye-tracker in a separate test. This advanced device has two sensors for recording the binocular vision and one sensor for the scene. Compared with the Positive Science eye-tracker, the advantage of this eye-tracker is that the framerate increases up to 120 fps while the scene resolution reaches 1080p.

It is worth noting, that the geometry layout of the buck-display system was not optimized ideally in our pilot testing due to space restriction. The physical measurement in our experiment reflects that the average height of sitting subjects is 3.7 feet high above the ground and 8.6 feet long to the front display. Such layout lets the subject have a 47 degrees field of view for the central display while 60 and 55 degrees for the left and right display, respectively. STIMSIM recommends that them all be 45 degrees. The specific layout of buck-display system is shown in Figure 1(e).

Experimental Design

The experimental design was intended to answer two main questions. First of all, as we mentioned earlier, the essence of autonomous driving is to replace the human in performing daily driving tasks. Based on this notion, the fidelity of driver data is more convincing when the design of an experiment is closer to the routine driving task. Therefore, a residential area is desired, since many mature drivers have a rich residential driving experience in real life. In addition, the emphasis of this paper is to investigate how visual perception enables drivers to correctly deal with hazards. Fittingly, the residential area is a high risk place for traffic incidents.

Considering the above benefits and facts, a 2.7 mile route through a residential background was created. In order to increase the adapting level in the virtual scene, a short transition path of suburban area was added from the starting line before the driver enters into the residential area. In addition, the roadway is designed as two-lane two-way and subjects are not allowed to make turns. The recommended speed for subjects is 45 mile per hour (mph) in the suburban area and 25 mph in the residential area. Five common hazard incidents were created: Stop sign, pedestrian crossing, pet crossing, vehicle backing out from driveway to the street, and merging into the lane. These are illustrated in Figure 2. Before the driver starts the testing route, a short practice demonstration is provided to help the driver become familiar with the steering system. Drivers wear the Positive Science eye-tracker when running the scene in the auto-simulator. Perception data of human subjects have been collected for analysis.

Our data analysis in assessing the correlation between the driving behavior and fixation will heavily depend on the source data collection via the eye-tracker. When we associate the perception, like the fixation, with the driving response in a particular hazard incident, we are supposed to select a very narrow time range before the incident to investigate the visual pattern. Ideally, we expect that the time of the image taken by the pupil sensor is exactly the time when the scene sensor outputs the related scenario. In other words, it is desirable that the ongoing activity in the scene is strictly matching the pupil activity frame by frame for an accurate analysis. Unfortunately, although the scene and pupil sensors are synchronized in the post processing, the latency of the eye-tracker is difficult to avoid. The generation of such a delay happens when the eye-tracker processes the tasks from image capture, transfer, and rendering.

Figure 3 illustrates this latency process with an example of four consecutive frames. The target is drawn as a dot with an x. The cross represents the fixation calculated by the eye-tracker. In the left corner of each image, the solid black ellipse indicates the actual position of the pupil when the target is focused. Figure 3(a) suggests the pupil position of the eye, and the fixation point of eye-tracker on the left target in the first frame. Figure 3(b) reveals that the left target disappeared while the right target popped out in the second frame. However, both the pupil, and eye-tracker fixation point remain unchanged. In the third frame, Figure 3(c) shows that the fixation point of eye-tracker stays still, although the position of pupil implies the eye has moved to the right target. Figure 3(d) denotes the fixation point of the eye-tracker reached the point that the pupil just moved to in the last frame. The latency in this case is one frame, which is calculated by the frame number difference of two images represented by Figure 3(c)-(d).

Precisely because of the latency commonly existing in eye-trackers, we conducted a sub-experiment to inspect how the latency issue impacts our eye-tracker systems. In this experiment, a 27-inch display with a resolution of 1920 by 1080 was used to present a moving target that would only appear in either the left or the right side of the display. The observer was required to wear the eye-tracker to watch the target moving from left to right. A mirror was subtly placed between the display and the observer. Such an scheme allowed the mirror reflect the eye movement of the observer. With this arrangement, the scenario captured by the scene sensor displayed the actual eye movement in the mirror. The pupil and scene videos were synchronized in post-processing. As a comparison with the eye movement in the mirror, the movement of fixation point was also generated in the same scene video using the pupil video from the pupil sensor. We then determined the number of frames needed for the eye-tracker to catch up to the eye movement, as Figure 3 illustrates. As mentioned previously, the Positive Science eye-tracker was employed in our simulations. As a product manufactured eight years ago, the technical specification of this eye-tracker is relatively behind the current standard in the industry. For example, the resolution of the scene sensor is 480p with a framerate of 30 fps. The low resolution and framerate of the scene will lead to possible failures in identifying and comprehending potential important objects that help drivers to make critical driving behaviors. Comparing with Positive Science, the Pupil Labs eye-tracker is equipped with a resolution of 1080p and framerate of 60 fps, which resolves the potential issues mentioned above. In addition, the Pupil Labs eye-tracker has a taller field of view, which incorporates the entire scene regardless of the physical height of the driver. Therefore, for long term consideration, both Positive Science and Pupil Labs eye-trackers will be adopted in the auto simulation tests. Currently, it is worth examining the latency difference of these two eye-trackers. In the experiment, the driver completed ten trials for each eye-tracker in order to increase the fidelity of the observation data. The collected videos of the trials were processed to synchronize the scene and eye videos and then analyzed.

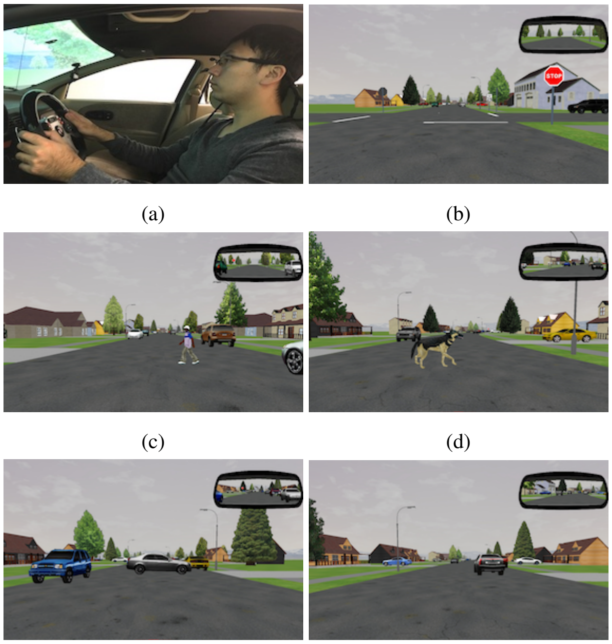

(a) (b) (c) (d) (e) (f)

Figure 2: The driving scene and simulated hazard incidents: (a) the physical driving setup, (b) the stop sign hazard, (c) pedestrian-crossing hazard, (d) pet-crossing hazard, (e) vehicle back-out hazard, (f) vehicle merging hazard. All the images are captured via STIMSIM software.

Results and Discussion

The auto simulation experiment was approved by the Human Subjects Research Office (HSRO) and the Institutional Review Board (IRB) at Rochester Institute of Technology. We invited seven human subjects to join our pilot testing. All the human subjects were fully informed of the potential risks and benefits.

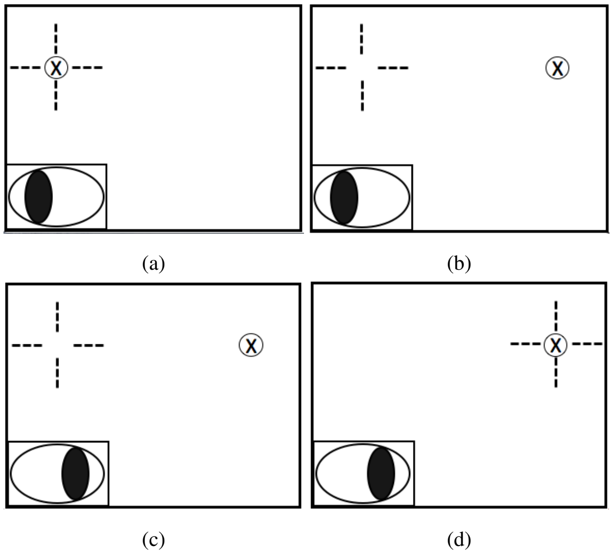


Figure 3: The latency diagram. Target is shown as a dot with an x. Eye-tracker fixation sample is drawn as a cross. In the left corner box of each subfigure, the black solid ellipse specifies the pupil position when eyes move to the target. (a) The pupil position of the eye, and the fixation point of eye-tracker on the left target. (b) The left target disappeared while the right target popped out in this frame. However, both the pupil, and eye-tracker fixation point remain unchanged. (c) The fixation point of eye-tracker stays still, although the position of pupil implies the eye has moved to the right target. (d) The fixation point of the eye-tracker reached the point that the pupil just moved to in the last frame.

The participants were between 24 and 32, and included 3 female and 4 male drivers. All of them acquired their driver licenses in United States with three years or more driving experience. All subjects specified that driving is a daily task. Before the test, they all passed a basic quiz of traffic rules that may be associated with the particular hazard incidents in the experiment. As mentioned earlier, a short practice demo was offered to the subjects. This demo included easy driving tasks, such as steering, acceleration, and deceleration, which let drivers become more familiar with controlling the hardware, and adapting to the virtual driving environment. Then, the test route was given to the drivers and all of them completed the whole route. Based on the recommended speed along the route, the calculated running time of this test was around 8 minutes. The statistical results indicate that the average driving time of subjects is around 8 minutes and 27 seconds, which was acceptably close to the estimated time.

Although the driving process proceeded as we planned, a non-negligible feedback issue from the subjects came to our attention. Five out of seven subjects mentioned that they experienced motion sickness during the driving. The symptoms consisted of dizziness, nausea, and inattention. One person among this group indicated that she experienced the same symptom after the test as well. Additionally, none of the subjects indicated that they regularly experience motion sickness when they drive or ride.

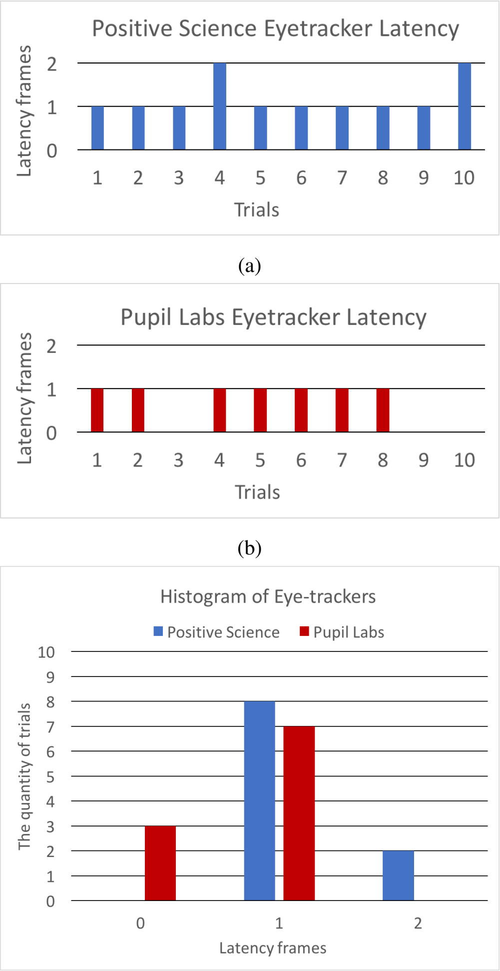
To fully examine reasons that caused the sickness, we collected more info from the subjects. The feedback indicated the following factors that induced the motion-sickness. First, and the most important reason, is the issue of non-ideal viewing angles. As we stated in the section of the instrument composition, the geometry layout of simulator-display system was not optimized due to the restriction of lab space. Even though the central filed of view (47 degrees) is close to the suggested viewing angle (45 degrees), the two side-viewing angles are 10 and 15 degrees, respectively, larger than the recommended angle (45 degrees). Such large viewing angles shorten the expected distance between the display and an observer. It then presents the observer ”zoomed-in” objects, which rapidly flow in the scene. Therefore, it easily causes sickness when subjects spent a long time driving. To verify this explanation, we invited a couple of the subjects who reported sickness previously to go through the test again. All the experimental factors were held constant except that only the central display was used. Feedback from this test shows that their sickness condition was effectively relieved.

The second factor was that the time gap between each hazard incident was not sufficiently long. The drivers indicated that they were likely to be sick when they have to accelerate and decelerate in a short distance. Since our simulator is not equipped with an inertial feedback system, acceleration and deceleration lack the essential physical response, such as body leaning backward and forward. The absence of these inertial features goes against the expectations of driving experience of people, which leads to uncomfortable physiological reactions. Some collected data associated with the hazard incidents among the subjects who reported motion sickness showed inconsistent fixation patterns. This indicates that motion sickness markedly weakened the fidelity and reliability of driver performance, compromising any data analysis of the correlation between driving behavior and perception patterns. We will provide a thorough evaluation with an enhanced experiment in the future.

In the latency experiment, observers wore both Positive Science and Pupil Labs eye-trackers to examine the latency difference. The subject completed ten trials for each eye-tracker to observe the target moving from left to right. Figure 4 presents the statistical analysis. For the Positive Science eye-tracker, eight out of ten trials resulted in 1 frame latency while two trials had 2 frames of latency each. On average, this eye-tracker has 1.2 frame latency. As the framerate is 30 fps, the latency of the eye-tracker is 40 milliseconds. Comparing with Positive Science eye-tracker, seven out of ten trials of the Pupil Labs eye-tracker share 1 frame latency and three trials hold no latency at all, for a 0.7 average frame latency. Thanks to its faster framerate 60 fps, the Pupil Labs eye-tracker has a much lower latency of 12 milliseconds. The experimental analysis proves that the eye-tracker of Pupil Labs has shorter delays in associating the visual perception with the ongoing activity in the scene.

Conclusion and Future Work

The primary goal of this study was to examine the visual perception of human drivers when they deal with hazard incidents and to provide a database of eye-tracking data of drivers negotiating potentially hazardous conditions. In order to complete this task, a creative automotive simulator, which was built around an actual physical car, was used. We have designed a virtual driving route to mimic a realistic driving experience including five hazard situations in the pilot testing. Pupil activity was recorded by a mobile eye-tracker while subjects drove the simulator. Additionally, a sub-experiment was conducted to understand the latency



(a)

(b)

(c)

Figure 4: Latency of eye-trackers. (a) The statistical results for the Positive Science eye-tracker trials. (b) The statistical results for the Pupil Labs eye-tracker trials. (c) Histogram of Latency comparing the two eye-trackers.

of the eye-trackers, which significantly affects the assessment in analyzing the correlation between the driving behavior and fixation patterns. In the auto-simulation experiment, we analyzed the potential reasons that lead to motion sickness. The experimental data indicated inconsistent fixation patterns due to driving performance impaired by motion sickness. Correlation analysis between fixation patterns and driving performance will be deferred to future work. Moreover, the sub-experiment reflected that the Pupil Labs eye-tracker was able to provide a better association between the visual perception and the ongoing activity in the scene, since the device presented less latency.

The future plan will focus on two major aspects. First, as we realized that improper viewing angles contributed to driver motion sickness, a more precise and strict layout of the display-buck system is required to be setup. Although we also expect that the inertial feedback system might decrease the simulator-sickness, the expected improvement does not justify the complicated engineering implementation and the required funding support. The dedicated arrangement of hazard incidents is expected to provide similar motion sickness relief as the inertial feedback system. The Pupil Labs eye-tracker has demonstrated multiple advantages in our tasks, however, it requires significantly more effort to process the collected data that is needed for exploring gaze patterns. Thus, the software compatible issue will be another issue to resolve.

References

[1] A. Liu, "Towards predicting driver intentions from patterns of eye fixations," *Vision in vehicles*, vol. 7, pp. 205–212, 1999.

[2] G. Underwood, D. Crundall, and P. Chapman, "Selective searching while driving: the role of experience in hazard detection and general surveillance," *Ergonomics*, vol. 45, no. 1, pp. 1–12, 2002.

[3] P. Chapman, G. Underwood, and K. Roberts, "Visual search patterns in trained and untrained novice drivers," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 5, no. 2, pp. 157–167, 2002.

[4] G. Underwood, P. Chapman, N. Brocklehurst, J. Underwood, and D. Crundall, "Visual attention while driving: sequences of eye fixations made by experienced and novice drivers," *Ergonomics*, vol. 46, no. 6, pp. 629–646, 2003.

[5] D. Crundall, P. Chapman, G. Underwood, E. van Loon, and G. Chapman, "Developing simulator-based visual search and hazard perception training," in *Behavioural research in road safety 2006: Sixteenth seminar*, 2006, pp. 213–224.

[6] S. Lee, S. Klauer, E. Olsen, B. Simons-Morton, T. Dingus, D. Ramsey, and M. Ouimet, "Detection of road hazards by novice teen and experienced adult drivers," *Transportation Research Record: Journal of the Transportation Research Board*, no. 2078, pp. 26–32, 2008.

[7] T. R. Coeckelbergh, W. H. Brouwer, F. W. Cornelissen, P. Van Wolffelaar, and A. C. Kooijman, "The effect of visual field defects on driving performance: a driving simulator study," *Archives of ophthalmology*, vol. 120, no. 11, pp. 1509–1516, 2002.

[8] Y. Seya, H. Nakayasu, and P. Patterson, "Visual search of trained and untrained drivers in a driving simulator," *Japanese Psychological Research*, vol. 50, no. 4, pp. 242–252, 2008.

[9] P. Konstantopoulos, P. Chapman, and D. Crundall, "Driver's visual attention as a function of driving experience and visibility. using a driving simulator to explore drivers eye movements in day, night and rain driving," *Accident Analysis & Prevention*, vol. 42, no. 3, pp. 827–834, 2010.

[10] J. Törnros, "Driving behaviour in a real and a simulated road tunnela validation study," *Accident Analysis & Prevention*, vol. 30, no. 4, pp. 497–503, 1998.

[11] S. T. Godley, T. J. Triggs, and B. N. Fildes, "Driving simulator validation for speed research," *Accident analysis & prevention*, vol. 34, no. 5, pp. 589–600, 2002.

[12] G. Underwood, D. Crundall, and P. Chapman, "Driving simulator validation with hazard perception," *Transportation research part F: traffic psychology and behaviour*, vol. 14, no. 6, pp. 435–446, 2011.

Acknowledgement

We would like to thank Dr. David W. Messinger for his support of this research. We would also like to thank Han Li, Hao

Xie, and Xiangzhen Kong for their assistances in experiments.

Author Biography

Mingming Wang received his MS in Electrical Engineering from Rochester Institute of Technology. He is currently a PhD candidate in the program of Imaging Science from Rochester Institute of Technology. His interests include image processing, computer vision, human perception, color processing, and autonomous driving.

Kate Walders received her MS in Applied and Experimental Engineering Psychology from Rochester Institute of Technology in 2012. She is a Staff Research Associate in the Multidisciplinary Vision Research Laboratory in the College of Imaging Science, and an Adjunct Psychology Instructor in the College of Liberal Arts at Rochester Institute of Technology. She also works and volunteers as a Research Consultant both within, and outside of, RIT.

Prof. Martin E. Gordon, PE, DFE is a Registered Professional Engineer (PE) and Diplomate Board Certified Forensic Engineer (DFE). His specialty is Accident Reconstruction and Highway Safety. He is Professor and Program Director of Undergraduate Studies at the Rochester Institute of Technology. Gordon is the 2018 president of the National Academy of Forensic Engineers. He has consulted or provided expert testimony in close to 150 traffic accident reconstruction and product liability cases some of them very high-profile.

Jeff B. Pelz is the Frederick Wiedman Professor in the Chester F. Carlson Center for Imaging Science at the Rochester Institute of Technology. He holds a B.F.A. in Photography and an M.S. in Imaging and Photographic Science from RIT and a Ph.D. in Brain and Cognitive Science from the University of Rochester. His research focuses on the development and application of wearable eye-tracking devices in the study of complex tasks.

Susan Farnand received her BS in engineering from Cornell University and her MS in imaging science PhD in color science from the Rochester Institute of Technology. She is currently an Assistant Professor in the Program of Color Science at the Rochester Institute of Technology. She is the Publication Vice President of the Society for Imaging Science and Technology.

Free access to this paper is brought to you with the generous support of ON Semiconductor.

All research funding for this paper is referenced in the text; unless noted therein, no research funding was provided by ON.