Towards a Better Understanding of Gaze Behavior in the Automobile

Lewis L. Chuang
Max Planck Institute for
Biological Cybernetics
Spemannstrasse 38
Tübingen, 72076, Germany
lewis.chuang@tuebingen.mpg.de

Heinrich H. Bülthoff
Max Planck Institute for
Biological Cybernetics
Spemannstrasse 38
Tübingen, 72076, Germany
heinrich.buelthoff@tuebingen.mpg.de

Department of Cognitive and
Brain Engineering
Korea University
Seoul, Republic of Korea

AutomotiveUI'15, September 1–3, 2015, Nottingham, UK ACM 978-1-4503-3736-6.

Abstract
Gaze-tracking technology is used increasingly to determine how and which information is accessed and processed in a given interface environment, such as in-vehicle information systems in automobiles. Typically, fixations on regions of interest (e.g., windshield, GPS) are treated as an indication that the underlying information has been attended to and is, thus, vital to the task. Therefore, decisions such as optimal instrument placement are often made on the basis of the distribution of recorded fixations. In this paper, we briefly introduce gaze-tracking methods for in-vehicle monitoring, followed by a discussion on the relationship between gaze and user-attention. We posit that gaze-tracking data can yield stronger insights on the utility of novel regions-of-interests if they are considered in terms of their deviation from basic gaze patterns. In addition, we suggest how EEG recordings could complement gaze-tracking data and raise outstanding challenges in its implementation. It is contended that gaze-tracking is a powerful tool for understanding how visual information is processed in a given environment, provided it is understood in the context of a model that first specifies the task that has to be carried out.

Author Keywords
gaze-tracking; attention; information processing

ACM Classification Keywords
H.5.2 [User Interfaces]: Evaluation/methodology; H.1.2 [User/Machine Systems]: Human Factors

Figure 1: In this image, an gaze-tracker is constructed by combining two lightweight cameras: (1) a camera under the eye that tracks the pupil, (2) a camera that captures the user's viewing frustum. Photo: Courtesy of Valentin Schwind, University of Stuttgart.

Introduction
If *"the eye serves as a window to one's soul"*, it is no wonder that the advent of increasingly affordable and mobile eye-trackers has been welcomed with much enthusiasm, by the research and commercial community alike. In the context of driving, video-based gaze-trackers operate by tracking features of the camera-recorded eye (i.e., pupil) with a scene-camera recording. The scene-camera could either be head-mounted and represent the driver's changing view-frustum (e.g., SensoMotoric Instruments ETG), or have a rigid position in the vehicle (e.g., SmartEye, Seeing Machines). The former solution requires extensive manual labeling post-testing, while the latter assumes a highly accurate geometric model of the driver' environment on which it performs automatic labelling of the regions-of-interest (ROI).

Objectively speaking, eye-tracking technology allows one to estimate what someone is looking at, at which point in time, and for how long. The frequency and duration of fixations per ROI (i.e., instruments) are often assumed to respectively represent the ROI' importance to the task and the difficulty of interpreting them [9]. This provides usability researchers a means to evaluate the effectiveness (or distractibility) of in-vehicle information systems (IVIS) in terms of the frequency that a given device is looked at whilst driving. In other words, this approach treats gaze fixations as a proxy for driver's attention. Unfortunately, this may not be entirely accurate.

Gaze and Attention
Although unrestrained gaze has a range of $\pm 200^\circ$ [2][**?**], actual gaze-movements often demonstrates a much narrower range. In fact, 86% of all eye-movements has been shown to fall within a range of $\pm 15^\circ$ [1]. Somewhat disappointingly, the best prediction model for eye-movements is not one that is based on salient image properties, such as contrast or foreground objects [11]. Rather, the best model for gaze prediction has been shown to be one that assumes that the observer tends to dwell on the scene's center, presumably out of a natural tendency to conserve energy. In the context of driving, it could raise the question of whether a test driver might be looking at the center of the lane because of its relevance to the steering task or because it is merely the default gaze orientation. In fact, it is well established that visual objects can be attended to without the need for overt fixation [12]. We can selectively enhance the perceptibility of peripheral objects, even whilst keeping our eyes fixed to the center. Doing so, however, requires a large amount of effort. While it is generally agreed that fixated regions can be assumed to have benefited from a user's attention at some point in time [3], although it is less certain when this might have occurred. To complicate matters further, we can fail to notice the unexpected appearance of objects, even if they appear where we fixate [6]. This is termed *"inattentional blindness"*. Given that gaze-tracking cannot indicate the spatial and temporal allocation of covert attention, an over-reliance on overt fixation data itself, without a careful consideration of the traffic scenario or the relative placement of ROIs, could easily result in erroneous interpretations of the value of certain instruments.

Alternatives to Measuring ROI Dwells?
Instead of recording the number of fixations on a given ROI itself as an indicator of the ROI's value, it might be preferable to determine the extent to which novel ROIs, such as

new interface displays, attract gaze away from the default gaze position or scan pattern. Even though basic biases of eye-movements exist, such as the central bias, (novel) task demands can modify this default pattern of eye-movements, for better or for worse. Thus, we posit that a fundamental understanding of how eye-movements relate to the primary task itself (i.e., driving) is necessary prior to any meaningful interpretation of gaze movements across IVISs.

To date, some models have been proposed that specify the basic eye-movements that are necessary for steering. For example, [8] emphasizes the importance of alternating fixations between the centre of the road and the tangent point of a road curve in order to determine the appropriate steering angle for keeping the automobile in the centre of the road. When implementing novel IVISs, it would be necessary to determine whether fixations on the novel ROIs are those that would otherwise be dedicated towards vehicular control. In this example, the presence of fixations to novel ROIs (e.g., GPS display, Twitter update feed) that occur intermittent to fixations of the near and far region ought to be interpreted and formalized as a novel task process that consumes cognitive resources in competition to the primary task. The example provided for steering has been integrated as a process within the ACT-R framework [7]. Thus, proponents of novel interface displays could benefit from formalizing the cognitive demands of their devices in like terms, namely process rules. Such an approach would result in more meaningful models for eye-movements, which will themselves be representative of cognitive process rules. In other words, they will more more clearly indicate how turn-taking is achieved between different tasks in the context of driving and serve as a more insightful tool.

Coupling Gaze-tracking and EEG

Although gaze-tracking is not a reliable indicator of the spatial and temporal allocation of covert attention, this might be overcome its combined use with electroencephalography ((EEG). EEG recording devices are lightweight, compared to other neuroimaging methods with comparable temporal resolution e.g., magnetoencephalography (MEG). Certain characteristics of the EEG signal in the time and frequency domain, such as event-related difference potentials and alpha desynchronization, have been associated with heightened attention. When coupled with gaze-tracking, such measures could help with evaluating the extent to which a fixated region-of-interest was attended to. Nonetheless, allowing eye-movements during EEG recording presents some challenges. First, the eye-movements can cause noticeable positive voltage gradient shifts across measured scalp activity. Several techniques can correct for this obvious artifact (see [4] for a review). However, other ocular-induced artifacts such as saccade-induced oscillatory activity might escape notice and have been confounded for cognitive activity (e.g., object memory [5]). In addition, EEG activity can be related to the planning of eye-movements. Such activity is cortically induced and it not always clear how to dissociate such activity from high-level cognitive function (i.e., visual attention) [10].

Conclusions and Outlook

Gaze-tracking is a powerful tool in discerning how a user seeks out and process visual information in a given task environment. Nonetheless, it ought to be employed with caution. The intuitive assumption that gaze is the equivalent of user attention can be easily falsified. In order to arrive at a meaningful interpretation of gaze activity, it is necessary to start with a fundamental model of the driving task and to define the observed movements in terms of the information that supports the successful execution of the task. The tendency to reduce observed gaze to areas-of-interest is an over-simplification that can yield little insight into the actual utility of novel IVISs.

Acknowledgements

The authors would like to thank the German Research Foundation (DFG) for financial support within SFB/Transregio 161.

REFERENCES

1. A. Terry Bahill, Deborah Adler, and Lawrence Stark. 1975. Most naturally occurring human saccades have magnitudes of 15 degrees or less. *Investigative Ophthalmology and Vision Science* 14, 6 (1975), 468–469.

2. Jasper Chen, Alan B Solinger, Jacques F Poncet, and Charles A Lantz. 1999. Meta-analysis of normative cervical motion. *Spine* 24, 15 (1999), 1571–1578.

3. Heiner Deubel. 2008. The time course of presaccadic attention shifts. *Psychological Research* 72, 6 (2008), 630–640.

4. Olaf Dimigen, Werner Sommer, Annette Hohlfeld, Arthur M Jacobs, and Reinhold Kliegl. 2011. Coregistration of eye movements and EEG in natural reading: analyses and review. *Journal of Experimental Psychology: General* 140, 4 (2011), 552–572.

5. Alon S Keren, Shlomit Yuval-Greenberg, and Leon Y Deouell. 2010. Saccadic spike potentials in gamma-band EEG: characterization, detection and suppression. *Neuroimage* 49, 3 (2010), 2248–2263.

6. Mika Koivisto, Jukka Hyönä, and Antti Revonsuo. 2004. The effects of eye movements, spatial attention, and stimulus features on inattentional blindness. *Vision Research* 44, 27 (2004), 3211–3221.

7. Dario D Salvucci. 2006. Modeling driver behavior in a cognitive architecture. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 48, 2 (2006), 362–380.

8. Dario D Salvucci and Rob Gray. 2004. A two-point visual control model of steering. *Perception* 33, 10 (2004), 1233–1248.

9. John W Senders. 1983. *Visual Scanning Processes*. Netherlands, University of Tilburg Press.

10. W Skrandies and K Laschke. 1997. Topography of visually evoked brain activity during eye movements: lambda waves, saccadic suppression, and discrimination performance. *International Journal of Psychophysiology* 27, 1 (1997), 15–27.

11. Benjamin T. Vincent, Roland Baddeley, Alessia Correani, Tom Troscianko, and Ute Leonards. 2009. Do we look at lights? Using mixture modelling to distinguish between low- and high-level factors in natural image viewing. *Visual Cognition* 17, 6 (2009), 856–879.

12. Helmholtz H. von. 1896/1989. Physiological Optics (1896 - 2nd German Edition, translated by M. Mackeben, from Nakayama and Mackeben, Vision Research, 29(11), 1631–1647, 1989). (1896/1989).