Simple gaze-contingent cues guide eye movements in a realistic driving simulator

Laura Pomarjanschi[ab] and Michael Dorr[c] and Peter J. Bex[c] and Erhardt Barth[a]

[a]Institute for Neuro-and Bioinformatics, University of Lübeck
[b]Graduate School for Computing in Medicine and Life Sciences, University of Lübeck
[c]Schepens Eye Research Institute, Harvard Medical School, Boston

ABSTRACT

Looking at the right place at the right time is a critical component of driving skill. Therefore, gaze guidance has the potential to become a valuable driving assistance system. In previous work, we have already shown that complex gaze-contingent stimuli can guide attention and reduce the number of accidents in a simple driving simulator. We here set out to investigate whether cues that are simple enough to be implemented in a real car can also capture gaze during a more realistic driving task in a high-fidelity driving simulator. We used a state-of-the-art, wide-field-of-view driving simulator with an integrated eye tracker. Gaze-contingent warnings were implemented using two arrays of light-emitting diodes horizontally fitted below and above the simulated windshield. Thirteen volunteering subjects drove along predetermined routes in a simulated environment populated with autonomous traffic. Warnings were triggered during the approach to half of the intersections, cueing either towards the right or to the left. The remaining intersections were not cued, and served as controls.
The analysis of the recorded gaze data revealed that the gaze-contingent cues did indeed have a gaze guiding effect, triggering a significant shift in gaze position towards the highlighted direction. This gaze shift was not accompanied by changes in driving behaviour, suggesting that the cues do not interfere with the driving task itself.

Keywords: Gaze guidance, driving assistance system, attention

1. INTRODUCTION

Humans are remarkably good at navigating complex visual environments. One of the key strategies to deal with the complexity of the natural visual world is selective attention, i.e. the ability to suppress irrelevant details of the visual input. This strategy is reflected in the anatomy of virtually all stages of the visual processing hierarchy; already in the retina, for example, cone density is highest in the centre with a steep falloff towards the periphery. As a consequence, humans move their eyes several times per second in order to serially sample the visual input. In later stages of the visual hierarchy, attention can change receptive field properties to enhance processing of an attended stimulus. While these selective mechanisms drastically reduce neuronal computational complexity and bandwidth requirements, e.g. in the optic nerve, they come at a perceptual cost: despite the impression that we constantly perceive everything in front of us at once, even large (and important) changes in a visual scene may go unnoticed if visual attention is deployed in the wrong location.[?] Clearly, biological evolution has ensured that human visual attention is good enough for most natural conditions, but today's visual environment often comprises unnaturally fast-changing objects (such as digital signage) that can move, at least relative to the observer, at unnaturally high speeds (for example, in traffic situations), and may overwhelm human observers simply by the number of objects that potentially require being attended. In safety-critical scenarios such as driving, wrong or suboptimal attentional deployment can have dire consequences; for example, common causes cited for traffic accidents are "failed to look" and "looked but failed to see" (LBFTS). LBFTS errors occur when a driver fails to detect another traffic participant despite having visually scanned the area where that participant was.[1,2] Such errors are particularly dangerous when approaching intersections: a 2009 study[?] estimates that

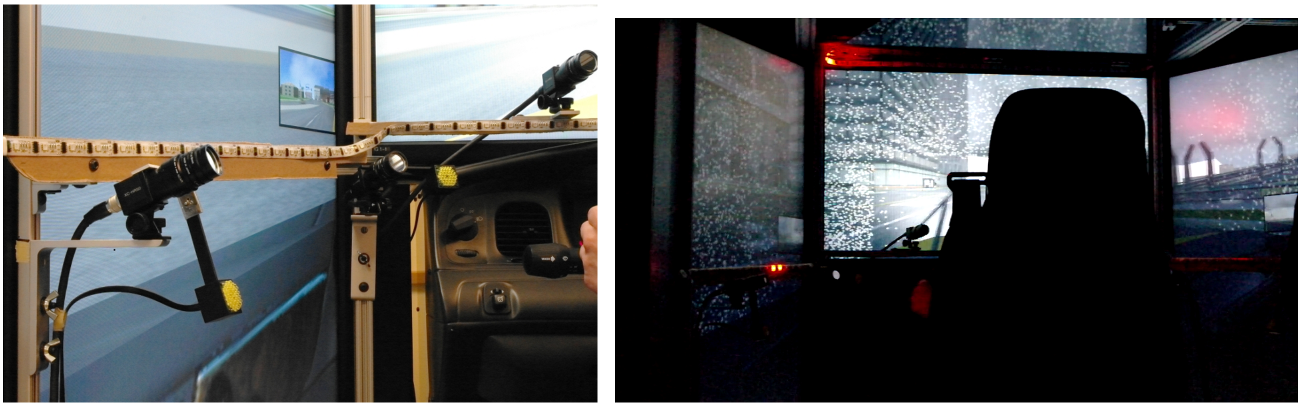
Further author information: (Send correspondence to Erhardt Barth)
Erhardt Barth: E-mail: barth@inb.uni-luebeck.de, Telephone: (+49) 451 500 5503

Figure 1: View of the state-of-the-art FAAC driving simulator. The system emulates the interior of a car, in which the windshield, covering a horizontal field of view of 225 degrees, is simulated using five LCD panels. The simulator is integrated with a six-camera Smart Eye PRO eye tracking system. The gaze-contingent warnings are implemented using LED arrays mounted horizontally; the LED arrays can be seen above the central monitor and on the horizontal bar slightly above steering wheel height.

43% of all road accidents that result in an injury in Europe occur at intersections, and that a large part of these are caused by the driver failing to properly scan the surrounding environment. In the more general case, even when no life is at risk, eye movement patterns are linked to task performance, for example in sports,[3] radiology,[4] and driving.[5]

Therefore, gaze guidance systems have been proposed that augment human vision by computer vision capabilities.[6] Such systems monitor the user's gaze and modify their display gaze-contingently in real time to direct the user's attention to critical locations; the gaze-contingent modifications may include enhancement of saliency (such as contrast increases, cues) at the intended location and suppression of saliency (e.g. by adding blur) to possible distractor locations. It is important to note that a crucial component in these systems is the gaze-contingent nature of the display, because the current attentional state of the user can be taken into account. Cueing that occurs regardless of the user's current gaze may be beneficial in simulator studies,[7] but may also be too intrusive to be accepted in practice.

Another obvious challenge is the choice of critical location to which the user is guided. This can be determined either by computer vision algorithms or by human experts; in the field of learning, studies have shown that the visualization of expert eye movements helped novices acquire perceptual skills faster in interpreting radiology images[8] and classifying fish locomotion patterns.[9] First gaze guidance systems have also been built for more interactive scenarios such as driving. Pomarjanschi et al.[10] used a driving simulator with an eye tracker to show that gaze-contingent cues can decrease the number of accidents with simulated pedestrians. In that study, pedestrians in high-risk situations were gaze-contingently highlighted to alert the driver. While encouraging, these first results cannot easily be transferred to real-world environments for two reasons. First, the PC-based driving simulator setup might have been too simple to represent realistic behaviour. For example, the lower cognitive load in a comparatively simple simulated environment might render gaze-contingent cues more effective; conversely, gaze-contingent cues might be ignored if the useful field of view were too narrow under higher cognitive load. Second, the gaze-contingent cues were embedded in the computer-generated simulator graphics and would hence be very difficult to implement in a real car. Therefore, the main goal of the current study was to investigate the effectiveness of gaze-contingent cues during more naturalistic driving conditions. LED arrays were mounted around the simulated windshield of a driving simulator with an integrated eye tracker. At the approach to certain intersections along predefined routes, these LED arrays were flashed as a function of the driver's gaze direction. In the following, we will report on the effect of such gaze-contingent stimulation on drivers' eye movements.



(a) (b)

Figure 2: Left: LED strip mounted on aluminium profiles across the bottom edge of the simulated windshield. Right: Stillshot of a gaze-contingent cue to the left during a data recording session. The subject was driving in an urban environment, in combined night and rain conditions. The cue started in the centre of the LED arrays and then rapidly moved leftward to capture the driver's attention.

2. METHODS

2.1 Driving simulator environment

The experiment took place in a high-fidelity, wide field-of-view DE-1500 driving simulator (FAAC Inc, Ann Arbor, MI). As no modifications could be made directly to the virtual environment, the cues were implemented using LED arrays fitted horizontally to the simulated windshield.

The simulator comprised an "open air" driving station, illustrated in Figure 1, whose windshield was simulated using a multidisplay system composed of five wide-screen LCD panels. The windshield covered a visual angle of approximately 225 by 38 degrees, and each LC display had a resolution of 1360×768 pixels. The controls of the simulator reproduced those of an automatic transmission vehicle. A real car – Ford Crown Victoria – was used to model the dashboard and the driving seat.

In addition to the visual component, the simulator also delivered auditory and haptic feedback. A set of speakers reproduced sounds such as engine and road noise, tire skidding, or braking noises. The force-feedback steering wheel offered variable resistance in function of the characteristics of the virtual road. Tactile feedback was also provided on curb strikes. The driving seat attempted to recreate the acceleration/deceleration feedback available in a real motor vehicle. Also, on collisions with objects or persons in the virtual environment, both auditory and haptic feedback was given.

The map of the virtual environment included a variety of terrains and settlement types such as urban, suburban, industrial, and rural areas, with a corresponding range of road types. In our experiment, other traffic participants were simulated using the simulator software's autonomous traffic mode, which populated the area around the subject driven car – the egocar – with other moving vehicles. The simulation parameters of traffic density and drivers' aggressiveness were set to high and moderate, respectively. The behaviour of these cars was non-deterministic, so that trials differed slightly in their traffic conditions.

The driving simulator was integrated with a six-camera SmartEye Pro remote eye tracking system, that could record eye movements over the entire windshield. The eye tracker delivered the gaze position on the multidisplay area at a sampling frequency of 60 Hz, and with a manufacturer-specified tracking accuracy of 0.5 degrees.

2.2 Gaze-contingent cues

Previous work with gaze-contingent stimuli in driving simulators directly accessed the simulated graphics;[10] for example, highlighting cues were easily realized by overlaying a texture onto pedestrians. The virtual environment used in the present study lacked an interface to the internal graphics, so we were challenged with the same problem that arises in a real car, i.e. how to combine the windshield's "contents" with additional gaze-contingent stimulation. Here, we used light-emitting diode (LED) arrays to implement horizontal directional cues to guide the attention of the driver towards a desired region of the environment. While these cues are much simpler and less location-specific than cues in previous work, such cues could also be implemented in a real car. A similar warning system, using rows of LEDs underneath the windshield to signal dangerous events to the driver, was implemented by Volkswagen AG[11] in a prototype car*.

A further limitation of the simulator engine that actually increased realism with respect to our experiment was the lack of scene and screen coordinates for other traffic participants. In principle, sophisticated computer vision algorithms would be required to track these other vehicles on the screen and to predict their behaviour, similar to what would be necessary in a real car. For simplicity, however, we here chose to cue at specific high-risk locations in the environment, namely intersections. As discussed earlier, many accidents at and near crossroads are caused by the driver's failure to properly scan all directions from which oncoming traffic might arrive, and thus the failure to identify potential threats. Therefore, gaze-contingent cues were presented either towards the left or the right of the driver during approach to a subset of intersections. In order to compare cues between trials, each intersection was assigned one fixed cue direction, and half of the trials were control trials where cues were not actually displayed. They were also suppressed when the driver looked in the desired direction already.

The LED arrays were fitted horizontally, one at the bottom of the windshield, covering the central 150 degrees of the visual field, and one at its top, covering only the 67 degrees corresponding to the centre screen. The arrays were mounted on a set of T-slotted aluminium profiles fitted transversally across the two vertically-set side displays (see Figure 2a). To create the LED arrays, we used three commercially available LED strips of one meter length each, daisy-chaining two of them to create the bottom array. Each strip comprised 32 individually programmable LEDs that could show up to 256 colours; for simplicity, we here used only an intense red colour (see Figure 2b). LEDs were toggled at a frequency of 5 Hz through a USB-to-serial converter board from an additional PC unit that collected simulation data from both the simulator setup and the eye tracker. These data streams were sent via UDP over Ethernet at a sampling rate of 60 Hz.

Each packet from the driving simulator data stream received by the master script contained all the driving statistics, including the position in the virtual city, and the speed of the egocar. In parallel, the master had access to (and ran through) a list with all the cued locations. If the distance between the egocar and the current critical location was smaller than a predefined constant, the controller could take the decision to trigger a warning. In order to compensate for large speed variations between different drivers, the distance measure was combined with an additional criterion regarding the estimated time before reaching the critical location.

Once it was established that the egocar had reached a cued location, the first eye movement sample was checked, in order to determine which LEDs would be activated during the warning. On the top array, it was always the last two LEDs in the predefined cueing direction that were briefly flashed. However, on the bottom array, the cue was dynamic, and it followed a "chasing lights" pattern: each of the LEDs between the start position of the cue and the end of the array would light on and off in a rapid succession. The start LED was computed as the LED that corresponded to the horizontal component of the last read gaze sample, to which a predefined offset in the direction of the cue, equal to half the resolution of the centre screen, was added. If the subject's gaze was already in the cued direction, than no LED was lit. The total latency between the reception of a data packet from the driving simulator data stream and the triggering of the cue was on average 184 μs.

2.3 Experimental sessions

Thirteen subjects participated in the experiment. All had held a driver's license for more than 5 years and reported normal or corrected-to-normal vision. Experimental sessions began with a manufacturer-provided 5-point calibration of the eye tracking system. Then, subjects drove along a short training route, where no data was collected, to acclimate to the simulator environment. This route started with mostly straight and flat roads, and only gradually introduced bends and altitude differences. Presumably due to the slightly incorrect geometric perspective of the simulated display and the latency of the driving seat motion feedback system, several subjects

*For press coverage (in German), see
`http://www.spiegel.de/auto/aktuell/neue-warnsysteme-lichtorgel-statt-piepshow-a-562943.html`

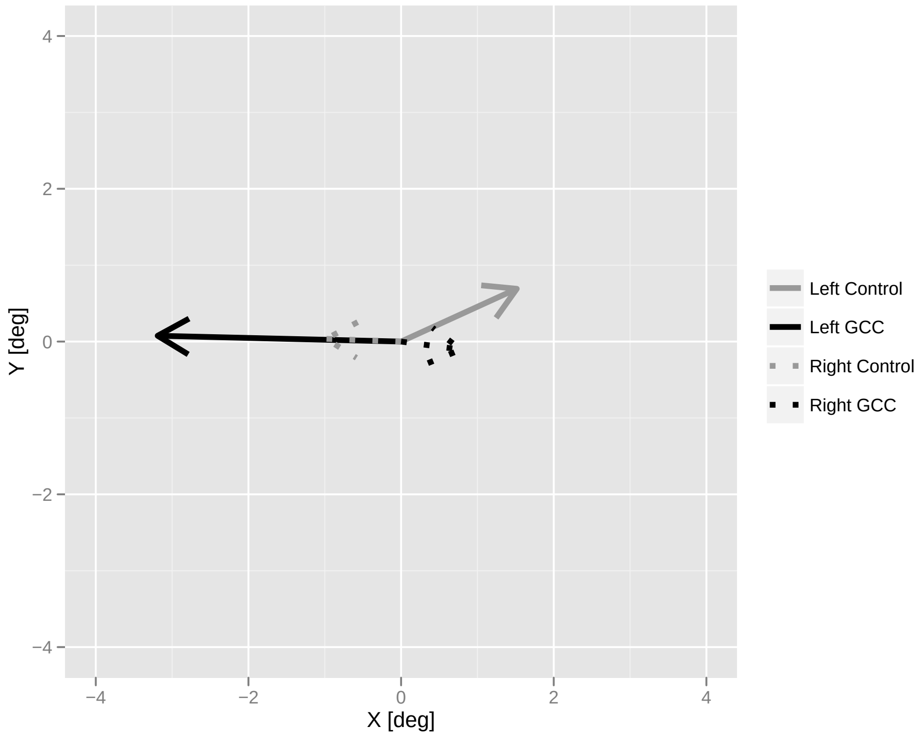


Figure 3: Mean vectors of the first saccade after cue triggering. For cued locations, the direction of the mean saccade vector roughly corresponds with the direction of the cue. The origin of the coordinate system corresponds to the centre of the "windshield", and for visualisation purposes, the saccade vectors have been translated to the origin.

reported symptoms of simulator sickness.[12] Despite the acclimatization route, one subject therefore had to terminate the experiment early without collecting any data.

We selected four courses inside the simulated world, each set in diverse environments: two were situated mostly in urban and suburban areas, while the other two focused mostly on highway sections. The route length ranged between 3 and 8 km. In addition to varying the complexity of the environment, we also varied the weather conditions. Impaired visibility increases attentional load, and therefore only one course was set in daylight conditions on a clear day. One course simulated a day with moderate to severe fog, and the remaining two courses were both set in combined night and rain conditions.

In each session, only approximately half of the intersections were associated with a visible cue. The other half served as a baseline to control for intersection-specific directional biases. For these cases, gaze-contingent cues were triggered and recorded internally, but not displayed via the LED arrays. In a balanced design across subjects, routes were then mirrored so that every intersection served as the location of gaze-contingent cues and control cues the same number of times. Overall, each complete trial consisting of the four pre-programmed routes contained approximately 40 control and 40 GCC locations.

3. RESULTS

Between all subjects, we collected more than 5 hours of driving data. The corresponding eye movement data comprised approximately 45000 saccades. Of particular interest were the 705 driving episodes around the critical intersections where gaze-contingent or control cues were triggered.

We first analysed the effect of gaze-contingent cues on driving parameters such as speed, acceleration, brake usage, or steering wheel inclination. No statistically significant differences were found between cued and non-cued

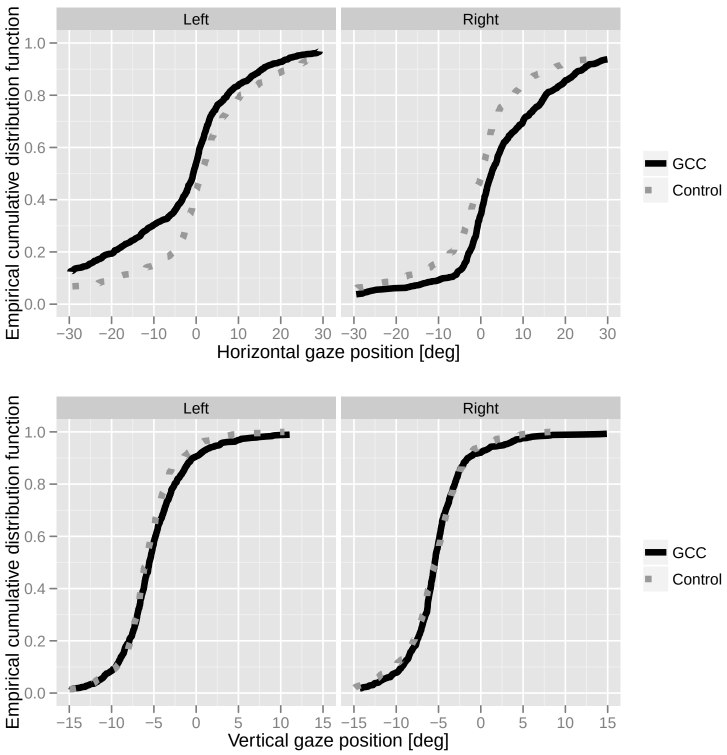


Figure 4: Empirical cumulative distribution functions for horizontal (top) and vertical (bottom) components of saccade landing points in the two seconds after a gaze-contingent cue. Relative to controls, the saccades are drawn towards the horizontal direction of the cue. In the vertical domain, saccades following 'left' cues showed a small bias towards the upper part of the screen (note the different x-axis scales on top/bottom).

locations regarding these parameters.

Naturally, we were most interested in whether gaze-contingent cues succeeded in guiding the gaze of the drivers in the desired direction. First, we therefore extracted the first saccade immediately following each of the gaze-contingent cue or control events. The mean saccade vectors are plotted in Figure 3, which shows that the average post-cue saccades go in the opposite (cued) direction from the control saccades. In order to assess the statistical significance of this effect, we used the two-dimensional variant of the Kolmogorov-Smirnov test.[13] For a data-driven estimate of the distribution of the K-S test statistic D, we resampled 5000 distributions of the control saccades and computed the pair-wise D values. For the left-ward cues, the distribution of saccade vectors was highly significantly different from the control cues ($p << 0.001$); for the right-ward cues, the effect almost reached significance ($p = 0.0524$).

Furthermore, we analysed potential effects beyond the first saccade only and therefore looked at saccade behaviour in the 2 s interval following each gaze-contingent cue. In total, this yielded more than 3500 saccades to analyse. First, for each saccade landing point, we computed the distance between its horizontal coordinate and the centre of the multidisplay system. The top panels in Figure 4 clearly show a shift in gaze position in the direction of the warning for cued locations; these differences are highly significant (Kolmogorov-Smirnov test, $p << 0.001$). The bottom panels show the same analysis for the vertical component; whereas there is no significant effect on vertical landing point for 'right' gaze-contingent cues, 'left' cues induce a small upward bias ($p = 0.002$). This means that sometimes the subjects looked at the gaze-contingent cues directly. It is also noticeable that in all conditions, gaze is vertically centred at approximately 5 degrees below the centre of the windshield.

For visualisation purposes, we further divided the 4 s intervals surrounding each gaze-contingent cue into 200 ms bins and plotted the mean horizontal position for all saccade landing points contained in each bin. The

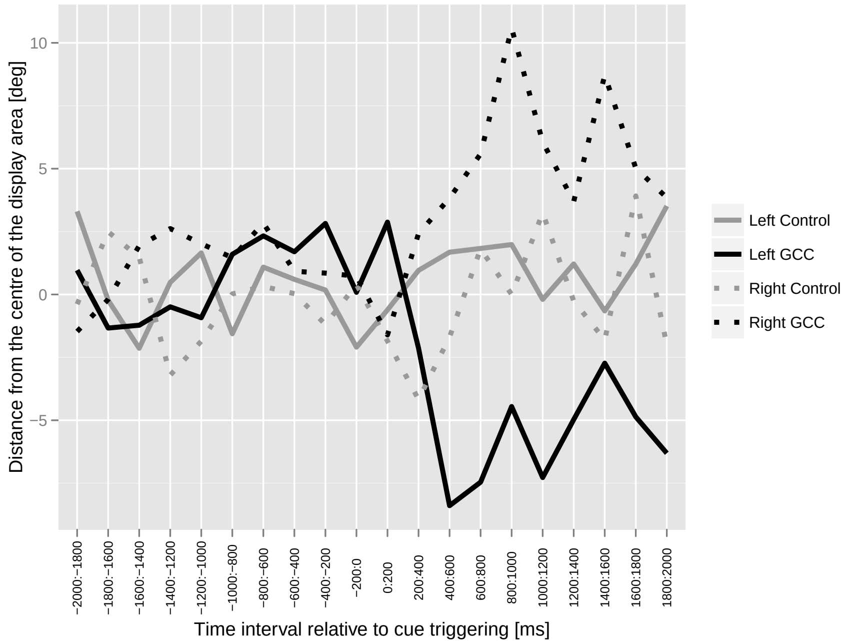


Figure 5: Mean horizontal saccade landing point distribution two seconds before, and two seconds after event triggering. For both left- and right-ward cues, there is a noticeable gaze shift towards the cued direction.

plot, shown in Figure 5 suggests a clear guiding effect beginning 400 ms after the cue.

4. DISCUSSION

Many everyday activities critically rely on attending to the right location at the right time. Because of its inherently visuo-motor nature, driving is particularly dependent on correct deployment of attention; hundreds of thousands of people are injured every year in traffic accidents that were ultimately caused by inattention. While self-driving cars might eventually take the driver out of the loop, it seems likely that human-operated vehicles will remain the norm for at least several decades, with more than 1 billion such cars on the world's roads today; certainly, one major issue for the first generation of self-driving cars will be how to rapidly and reliably bring the driver back into the loop when necessary.

Because gaze and attention are typically tightly coupled, gaze guidance systems have been suggested to aid users in deploying their attention more efficiently.[6] In such systems, an eye-tracking device continuously monitors the user's gaze; if necessary, display contents are modified in real time to attract the user's attention to the desired location. The potential complexity and therefore efficacy of such modifications obviously depends on the display type. For computer-based applications such as the watching of training videos, where the screen contents can be modified almost arbitrarily, gaze guidance has been shown to be helpful.[9] Similarly, gaze-contingent cues decreased accident rates in a PC-based driving simulator;[10] however, the realism of that driving simulator was very limited.

In this paper, we have contributed a proof-of-concept towards more practically feasible gaze-guidance systems in motor vehicles, by implementing such a system in a driving simulator. Gaze-contingent cues were implemented by LED arrays mounted in the bezel of the (simulated) windshield and triggered during the approach to intersections. These gaze-contingent cues guided attention horizontally towards the cued hemifield without drawing attention vertically to the LED arrays. Driving parameters were not adversely affected. Because of technical limitations of the simulation environment, however, we could not programmatically induce traffic scenarios with a high risk of accidents; therefore, it remains for future work to investigate whether the efficacy of gaze guidance will translate into better driving behaviour.

ACKNOWLEDGMENTS

This work was supported by the Graduate School for Computing in Medicine and Life Sciences funded by Germany's Excellence Initiative [DFG GSC 235/1], and by NIH grants EY018664 and EY019281.

REFERENCES

[1] M.-B. Herslund and N. O. Jørgensen, "Looked-but-failed-to-see-errors in traffic," *Accident Analysis & Prevention* **35**(6), pp. 885–891, 2003.

[2] A. Koustanaï, E. Boloix, P. V. Elslande, and C. Bastien, "Statistical analysis of looked-but-failed-to-see accidents: Highlighting the involvement of two distinct mechanisms," *Accident Analysis & Prevention* **40**(2), pp. 461–469, 2008.

[3] M. Land and P. McLeod, "From eye movements to actions: how batsmen hit the ball," *Nature Neuroscience* **3**(12), pp. 1340–1345, 2000.

[4] H. L. Kundel and P. S. L. Follette, "Visual search patterns and experience with radiological images," *Radiology* **103**(3), pp. 523–528, 1972.

[5] 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* **46**, pp. 629–646, 2003.

[6] E. Barth, M. Dorr, M. Böhme, K. R. Gegenfurtner, and T. Martinetz, "Guiding the mind's eye: improving communication and vision by external control of the scanpath," in *Human Vision and Electronic Imaging XI*, B. E. Rogowitz, T. N. Pappas, and S. J. Daly, eds., *Proc. SPIE* **6057**, pp. 116–123, 2006. Invited contribution for a special session on Eye Movements, Visual Search, and Attention: a Tribute to Larry Stark.

[7] D.-Y. D. Wang, D. F. Pick, R. W. Proctor, and Y. Ye, "Effect of a side collision-avoidance signal on simulated driving with a navigation system," in *Proceedings of the 4th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, 2007.

[8] D. Litchfield, L. J. Ball, T. Donovan, D. J. Manning, and T. Crawford, "Viewing another person's eye movements improves identification of pulmonary nodules in chest X-ray inspection," *Journal of Experimental Psychology: Applied* **16**(3), pp. 251–262, 2010.

[9] H. Jarodzka, T. van Gog, M. Dorr, K. Scheiter, and P. Gerjets, "Learning to see: Guiding students attention via a models eye movements fosters learning," *Learning & Instruction* **25**(June), pp. 62–70, 2013.

[10] L. Pomarjanschi, M. Dorr, and E. Barth, "Gaze guidance reduces the number of collisions with pedestrians in a driving simulator," *ACM Transactions on Interactive Intelligent Systems* **1**(2), pp. 8:1–8:14, 2012.

[11] M. Kiss, "Driver assistance in the peripheral field of view." Science Beyond Fiction: European Future Technologies Conference, Prague. Session on Modelling and guiding attention in an increasingly complex world, 2009.

[12] J. Casali, "Vehicular simulation-induced sickness. Volume 1. An overview.," tech. rep., DTIC Document, 1985.

[13] G. Fasano and A. Franceschini, "A multidimensional version of the Kolmogorov-Smirnov test," *Monthly Notices of the Royal Astronomical Society* **225**, pp. 155–170, 1987.