Guiding a Driver's Visual Attention Using Graphical and Auditory Animations

T. Poitschke, F. Laquai, and G. Rigoll

Technische Universität München
Institute for Human-Machine Communication
Theresienstrasse 90, 80333 Munich, Germany
{poitschke,laquai,rigoll}@tum.de

Abstract. This contribution presents our work towards a system that autonomously guides the user's visual attention on important information (e.g., traffic situation or in-car system status signal, etc.) in error prone situations while driving a car. Therefore we use a highly accurate head-mounted eye-tracking system to estimate the driver's current focus of visual attention. Based on this data, we present our strategies to guide the driver's attention to where he should focus his attention. These strategies use both graphical animations in form of a guiding point on the Graphical User Interface as well as auditory animation that are present via headphones using a Virtual Acoustics system. In the end of this contribution, we present the results from a usability study.

1 Introduction

This contribution presents our work on a system that autonomously guides the user's visual attention on important information (e.g., traffic situation or in-car system status signal, etc.) in error prone situations while driving a car.

According to investigations of U.S. NHTSA (National Highway Traffic Safety Administration) from the year 2000, the main reasons of 25% of all traffic accidents are driver inattention or distraction [10]. Also, Austrian and German accident statistics consider an attention deficit as main reason for at least 10% of all accidents with fatalities (e.g. [6]). This can mainly be justified by the high risk of distraction during the operation of today's infotainment systems compared to tasks like entering a telephone number or operating previous simple radio systems. For example, the entry of a navigation destination needs much more time. Furthermore, during those tasks the (visual) attention does not only decrease, but the view is even for a longer period of time not on the road but in the car. Thus, one goal during the development of new infotainment systems must be to focus the driver's attention – especially in a dangerous situation – quickly and safely to where they are most needed. Therefore, the task of recognizing a potential risk of accidents can be partially accomplished by the car itself, for example by steadily observing the distance to the vehicle in front or by recognizing lane borders.

D. Harris (Ed.): Engin. Psychol. and Cog. Ergonomics, HCII 2009, LNAI 5639, pp. 424–433, 2009.

2 Previous Work

To implement the presented concepts previous work from our institute and our research partners was integrated. Here, we used a high accurate eyetracking system to analyze the current focus of visual attention as well as a playback system for virtual acoustics using headphones.

Fig. 1. Head-mounted eyetracking system [13] and headphones for the virtual acoustics

2.1 Eyetracking

In order to efficiently guide a user's attention it is very important to know where the user is currently focusing his attention. The user's current focus of attention is highly correlated to its current line of sight and the corresponding fixated point. For the detection of the user's line of sight and gaze point, we used an eye tracking system (see fig. 1). To compute highly accurate and fast gaze data, we used the combination of the head-mounted gaze tracking system *EyeSeeCam* [13] and an external tracking [1]. Thereby, the gaze tracker computes the user's line of sight in its own coordinate system and is simultaneously tracked by a stereo camera system that also defines the world coordinate system. Afterwards, the gaze data is transformed into the world coordinate system using the position and orientation of the gaze tracker in 3D. From the current line of sight we can compute the user's current gaze point by simply subtending the gaze line with previously calibrated areas [1].

2.2 Virtual Acoustics

To present the warnings in a manner that is distinguishable from other ordinary warnings in the automobile, and to test whether a moving sound can have a stronger effect on the focus of attention compared to a fixed one, we used the virtual acoustics presented in [12] to replay the sounds. Thus, using headphones, it is possible to present a sound in any horizontal angle around the test person.

3 Concept and Implementation

To present the warning strategies in a most realistic environment, a rudimentary In-Vehicle Information System (IVIS) was implemented. The interaction system bases on a large central information display. Therefore, we used a 15 inches touchscreen that is mounted on a position comparable to current central displays in mass-production cars (see figure 3 and [8]). On this screen, the graphical user interface (GUI) of our infotainment system is visualized. For this contribution we did not implement the complete functionality but the user can enter navigation destinations or telephone numbers. The system also processes the data from the eyetracker and controls the virtual acoustics functionality.

3.1 In-Vehicle Information System

The GUI of the used prototypical IVIS is divided into a status bar at the bottom and an extensive control panel (see fig. 2). The status bar serves as return key for the menu navigation (e.g., returning to a superior menu level) and displays warnings. In the control panel the user can for example switch to the phone menu, a calendar function or select the radio. However, for this contribution only the menu items navigation and phone provided functionality, as these menus hold the largest potential for distraction in the car [11]. To ensure a better operability, the buttons are illustrated with icons that reflect their function, e.g., a telephone or a street.

In the phone menu, the user can choose between the normal input of a telephone number and an address book entry. The number pad is arranged as known from a mobile phone. The navigation menu provides the user menu with four options to enter a navigation destination. However, for this contribution we only used the option "*New Destination*". The destination input is realized with an alphabetical keyboard. If the destination entering is finished, the guidance can be started. The user can switch between boxes at anytime to correct mistakes.

3.2 Graphical Animations

There arise several possibilities to guide a user's attention on a screen. For example, the user can be advised of messages that do not have to be directly taken into account, but as soon as possible, e.g., a new traffic message. According to the orientation reflex one can assume that a blinking light (=movement) in the peripheral field of view should attract sufficiently enough attention. Furthermore, a traffic jam message can be indicated by a blinking light in the status bar. Since the gaze direction is known through the eytracking system, a more detailed dialog may open if the user is looking at the message. Such messages can be easily acknowledged by tapping on the monitor.

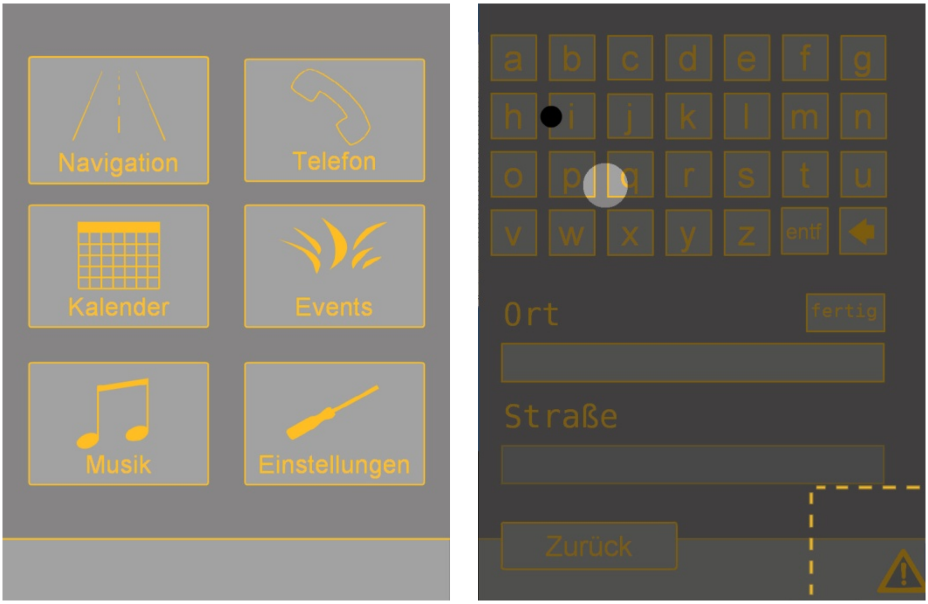

Fig. 2. Screenshots of the used interaction system. The left part shows the main menu of the implemented interaction system. The right part illustrates the graphical guidance of the attention using a guiding dot (light dot on the GUI). The small black dot represents the current gaze point, sand the orange line marks the target. Both are not visualized during system operation.

Fig. 3. System overview: mock-up and mounted displaying areas

However, there is also a second method to guide the user's attention if this concentrates on the road or the blinking on the monitor is simply missed. For such cases, we also implemented a guiding point. In addition to the blinking icon in the status bar, the entire screen is faded out, except of an area of about 30 pixels. Thus, this area attracts the user's attention (see [7]), and is designed to guide its attention to the actual target. Therefore, this area does not move autonomously around, but is positioned on the connecting line between the current gaze point and the actual target, where the user is supposed to look at. A distance of 150 pixels between the current gaze point the guiding point has proven useful. This phenomenon is comparable to a dust particle on the eye ball. The eye follows this dust particle involuntarily, and for example, the eye moves ever upwards. Since the movement is not automatic but in relation to the viewing direction, it is difficult to miss or ignore. The point disappears when the user's view arrived in the "*target area*", which is spread 300 pixels around the warning message. The area is chosen that large, because the target is in the border area of the working area of the used eyetracking system. An unintended triggering of the warning messages did not occur during the presented tests.

3.3 Auditory Animations

To ensure a fast and robust attention guiding, also an auditory signal can be reasonable, as this cannot be overlooked or occluded. Thus, such advises should be perceived more reliable. Therefore, the virtual acoustics system presented in [12] was used. This system allows it to replay a sound to a test person that is placed on a circle around him.

To design a sound that is best possible locatable, it is important that the sound has the widest possible spectrum. Since the sound should also enable an attention guiding from the lower mounted IVIS to the top (e.g., the road), the elevation has to be simulated additionally. Therefore, the sound's spectrum has to contain components at 8kHz, which are perceived as coming from above [5]. However, white noise should not to be used as a warning sound, because the human is not used to. Furthermore, a single sine at 8kHz is very unpleasant. Hence, more spectral components must be present. Thus, the basic warning sound is composed of three sine tones at a distance of one octave at 500Hz, 1kHz and 2kHz with a broadband noise of 0.1 to 8kHz. In addition, a narrow-band noise with a higher level at 8kHz is used. This sound mixture is good to locate and can direct the driver's attention to the top, without being unpleasant.

Similar to the guiding point, we implemented a sound that is moving with respect to the current gaze point. This is located between the current gaze direction and the target direction (e.g., the road). Thus, the tone is designed to direct a user's attention solely with its movement. In contrast to the animated sound, we also evaluated a fixed sound in a direction from the road, which might guide the attention faster because of the given immediacy.

3.4 Combined Animations

To achieve a maximum of benefits, the combination of visual and auditory attention guiding is possible. Thus, the guiding point described above was combined with the acoustic guidance. The entire concept does not only address the attention guiding on the screen, but particularly guiding the view from the screen to the road. Thus, the immediately visible point on the IVIS starts the movement of the user's view from the screen to the road. This movement will be perpetuated by the sound, even if the user's view has left the displaying area.

4 System Evaluation

The experiments were carried out in our driving simulation lab. The presented systems were integrated on an vehicle mock-up (see fig. 3), which is a rudimentary cockpit with a driver's seat, force feedback steering wheel and a touch screen to display the IVIS. In addition, the eye tracking system installed. Besides the eyetracker goggles the test persons (TP) wore the headphones for the virtual acoustics. To measure the driving performance of the TPs, the Lane Change Test (LCT, see [3]) is used as primary task measure, and the Peripheral Detection Task (PDT, see [9]) is used as secondary task measure.

4.1 Test Procedure

At the beginning of the test, the TP had to fill out a first questionnaire to collect demographic data. After each part of the test, another questionnaire composed of the SEA scale (self-report measure to evaluate the subjective mental workload, see [4]) and the System Usability Scale (SUS, see [2]) was handed out. At the end of the entire test run, there was a final interview. During the test, the reaction times and the missed points of the PDT as well as the lane keeping performance in the LCT were recorded for later analysis. Furthermore, the response times to alerts and warning, and the corresponding view position at the beginning of the system feedback were saved. First, the subjects had the possibility to acclimatize with the LCT and PDT. The next step was a test trip while a given navigation destination had to be entered into the navigation system without any guiding concept available. Afterwards, three runs with a simultaneous navigation input, and our warning strategies were arranged: display of graphical warnings on the IVIS and the moving and fixed the warning sound. The order varied from person to person to avoid learning effects from the results. The navigation inputs were chosen in a way that the TP was engaged as long as possible. At the beginning and the end of the experiment, a baseline ride without any secondary tasks were accomplished (only LCT and PDT) as a reference.

4.1.1 Graphical Warnings

The subjects were first informed that there will appear several warnings during the test rides, followed by a short demonstration of each graphical warning strategy (i.e., with and without guiding point). During the trip, the timing and manner of warnings were controlled from the test supervisor. Thus, it has been ensured that the messages with a guiding point occurred as often as those without. However, the timing was chosen in a randomly manner. If a warning message was not detected after 10 seconds, it was canceled and registered as disregarded.

4.1.2 Auditory Warnings

The auditory warning strategies (i.e., moving or fixed sound) were also presented before the test runs started. Thus, the volume could be adjusted to a pleasant value. Furthermore, the TPs were introduced, that there will appear several auditory warnings prior to the lane changes during the test runs, which indicate that the driver needs to focus its attention back on the road. These sounds were triggered by the test supervisor. Again, the warnings were uniformly distributed between the attention guiding with and without guiding point.

4.2 Results

The following section summarizes the results of the questionnaires, the PDT and the LCT. The experiments were conducted with 15 test persons (TP, 14 male and one female). The average age was 28.6 years. Seven TPs stated that they have already operated a navigation system with touchscreen, five used only systems without a touchscreen, and three had no experience with navigation systems. The level of interest in technology within the sample was high, and almost all expressed their interest in technical innovations. All subjects reported to use their PC daily. Therefore, it is unlikely that differences in the test results are caused by different abilities in the operation of the IVIS.

During the baseline runs, all subjects responded to all points of PDT with an average response time of 766ms. The mean lane deviation was 0.51m.

During driving and competitive navigation inputs without any warnings, only an average of 81% of the PDT points was recognized with an average response time of 1043ms. Thus, the reaction was 36% slower (average 277ms) compared to driving without load. The lane deviation was 1.03m (twice as high as during the reference runs). These results are all statistically significant.

During a test without driving task, all TPs could differentiate between a fixed sound and a sound that was rotating around them. However, during driving only four of 15 TPs stated that they recognized a difference between the fixed and the moving sound. Thus, a separated evaluation is addressed only very briefly. This difference to the test runs is probably caused by the high load during the driving experiments (i.e., LCT and PDT combined with navigation input). One TP even stated that he did not perceive the moving of the sound, because he was not able to pay attention. The TPs who perceived a difference, rated the task with the fixed sound as more strenuous (105 to 90 on the SEA-scale). Further, these TPs had a larger lane deviation then during the runs with the moving sound (1.03 m to 0.95 m). These differences are not significant.

In contrast to runs without warning, we achieved a significant improvement of the lane keeping performance if the warning strategies were enabled: The lane deviation (with warning, fixed sound) averaged only 0.83 m, which is only slightly more than 1.6 times as much as for the reference run. During runs without warning, the TPs had a lane deviation which was more than twice as much as during the reference runs. This represents an improvement of about 40%. The data of trips with a moving sound were slightly worse. However, this difference was not significant. The remaining additional lane deviation was caused by the poor lane keeping while typing the navigation destination when the user's gaze rested on the IVIS. However, using the auditory warning strategies we achieved a reduction of late lane changes and missed road signs.

Furthermore, the warning strategies did not have a significant effect on the PDT response times (1064ms to 1043ms). Here, an effect was not expected, since the timing for the warnings was tuned to the road signs, but not to the PDT. Interestingly, the detection rate decreased by 10% (from 80.92% to 70.06%). This is probably related to the fact, that the TPs relied on the auditory warning and therefore did not check the PDT points as often as without warnings. This is also related to the test setup, as the top priority task for the TPs was the lane changing. Furthermore, during normal driving warnings will appear with a much lower frequency than in the presented experiment. Therefore, habituation to the sound is not as strong as in the given experiment. However, this effect is still not negligible. Summing up, all test persons stated that the auditory warning attracts their attentiveness very fast and consequently can guide it very quickly and safely back on the road.

For the attention guiding using graphical animations on the screen (i.e., the guiding point and flashing warnings), the response time for the PDT is something higher compared to runs, in which only an input into the navigation system was to make (1109ms to 1043ms). This difference is significant but small. Analyzing the PDT detection rates, this effect gets more obvious. The detection rate is only at 66% compared to 81%. This is probably due to the fact, that the driver's view is often focused on the IVIS. The lane deviation is with 1.18m also slightly higher. However, this difference is not significant. Therefore, it can be assumed that the warning provides only a minimal additional burden. A similar result provides also the evaluation of subjectively experienced effort. With 97 points during the runs with guidance point and 98 points for the rides without a point, this is in the same range as for the runs with auditory warnings.

Since the primary task (LCT) is always processed first, the evaluation of the response times was not very enlightening. It was apparent that in the average 37% of the flashing warnings had to be aborted, since it was not detected after 10 seconds. This might be partly linked to the placement of the warnings. During our experimental procedure, this proved to be adverse, since the subjects occluded the warnings several times with their own arms, even though they were instructed that this problem might occur. However, they were often too busy with the primary task that they could not always respond to the flashing warning. Nevertheless, the attention guiding using the guiding point always led to the desired goal.

The attention guidance using the guiding point also performed better regarding the results from the System Usability Scale. Here, the guiding point achieved 75 from 100 possible points, compared to 64 points for the flashing warnings. This result is statistically significant.

Even though an evaluation of the response times was not reliable due to non deterministic latencies that occurred randomly during the experiments, more than half of the subjects stated, that they could react faster, if they could simply follow the guiding point with their eyes. Anyhow, half of TPs felt restricted by this guidance concept. During normal driving, a warning should not occur as frequently as in the experiment and not always during the interaction with the IVIS. Therefore, one can assume that the sense of restriction will decrease.

Nearly all TPs stated that the guiding point guided their focus of attention to the monitor in every case. Further, 12 TPs (80%) rated the guidance point as safe and 10 (67%) as pleasant in the final questionnaire.

5 Conclusion and Outlook

The evaluation showed that while the guiding system was activated, the test persons missed less traffic signs and therefore accomplished the primary task with a small lane deviation. Also, the overall lane deviation was smaller if the system exhorted the drivers to reallocate the visual attention on the driving task. In our study, the auditory animations were more effective compared to the graphical animations.

Currently we are working on the expansion of the presented system. Therefore we are especially testing the benefits of the presented concepts in a multi-display setup (i.e., freely programmable instrument cluster, large area Head-Up Display and the presented Central Information Display). Therefore we are also working on the integration of our own remote eye- and gazetracking system, which should ensure a higher level of acceptance by a non-intrusive measurement of the user's focus of visual attention. Furthermore, we are working on a hardware setup, that allows a sound presentation comparable to the virtual acoustics system, but without the currently needed headphones. This should also ensure a higher acceptance of the system.

Acknowledgements

This work is supported in part within the DFG excellence initiative research cluster "*Cognition for Technical Systems - CoTeSys*", see also www.cotesys.org. Also, we would like to thank the student workers from TU München who supported the work presented in this contribution. Especially we would like to thank Miss Monika Forstner.

References

1. Bardins, S., Poitschke, T., Kohlbecher, S.: Gaze-based Interaction in various Environments. In: Proceedings of 1st ACM International Workshop on Vision Networks for Behaviour Analysis, VNBA 2008, Vancouver, Canada (October 31, 2008)
2. Brooke, J.: SUS - A quick and dirty usability scale. Technical report, Redhatch Consulting Ltd. (1996)
3. DaimlerChrysler, A.G.: Research and Technology. Lane Change Test 1.2 User Guide (2004)
4. Eilers, K., Nachreiner, F., Hänecke, K.: Entwicklung und Überprüfung einer Skala zur Erfassung subjektiv erlebter Anstrengung. Zeitschrift für Arbeitswissenschaft 40, 215–224 (1986)
5. Fastl, H., Zwicker, E.: Psychoacoustics, Facts and Models. Springer, Heidelberg (1990)
6. Kuratorium für Verkehrssicherheit. Verkehrsunfallstatistik (2007), `http://www.kfv.at`

7. Heinecke, A.: Mensch-Computer-Interaktion. Hanser Fachbuch Verlag, Leipzig (2004)
8. Laquai, F., Ablassmeier, M., Poitschke, T., Rigoll, G.: Using 3D Touch Interaction for a Multimodal Zoomable User Interface. In: Proceedings of the International Conference on Human-Computer Interactional HCI International 2009, San Diego, USA (2009)
9. Matthews, G., Desmond, P.: Stress and Driving Performance: Implications for Design and Training. In: Hancock, P., Desmond, P. (eds.) Stress, Workload and Fatigue, ch. 1.8. Lawrence Erlbaum Associates, Mahwah (2001)
10. National Highway Traffic Safety Administration – NHTSA. NHTSA Driver Distraction Research: Past, Present, and Future (July 2000), `http://www-nrd.nhtsa.dot.gov/departments/nrd-13/driver-distraction/PDF/233.PDF`
11. Stutts, J., Reinfurt, D., Staplin, L.: The Role of Driver Disraction in Traffic Crashes. In: AAA Foundation for Traffic Safety (2001)
12. Völk, F., Kerber, S., Fastl, H., Reifinger, S.: Design und Realisierung von virtueller Akustik für ein Augmented-Reality-Labor. In: Tagungsband Fortschritte der Akustik – DAGA 2007, Stuttgart, March 19-22, pp. 559–560. DEGA (2007)
13. EyeSeeCam Homepage, `http://www.eyeseecam.com`