Towards an error-tolerant HMI approach to counteract registration errors in augmented reality

Vitalij Sadovitch, Michael Wittkämper, Fares Wallis, & Ina Petermann-Stock
Volkswagen AG, Wolfsburg
Germany

Abstract

An augmented reality head-up display (AR-HUD) enhances the driver's view by a superimposition of virtual content onto the real environment to support the driving task. A frequently addressed technical issue in augmented reality are deviations within the registration of virtual elements in 3D space. In the context of an AR-HUD these registration errors are leading to a reduction in driving performance and subjective usability assessment. In this contribution, an error-tolerant approach for Human-Machine-Interfaces (HMI) is described to counteract occurring registration errors. The approach was evaluated in a stereoscopic driving simulator with 62 participants performing a driving and two secondary tasks. The subjective perception of augmented reality accuracy was slightly higher, when real objects were virtually marked by extended or diffuse visualizations rather than conventional visualization types with sharp edges. Meanwhile there were no negative effects observed on driving performance and mental workload due to the error-tolerant HMI solutions.

Introduction

Augmented reality (AR) is considered as one of the most recent developments in human computer interaction technology (Billinghurst et al., 2015). It aims to combine reality and virtuality by aligning, respectively registering threedimensional virtual content with the real environment, allowing interaction in real time (Azuma, 1997). There are three distinguishable displaying techniques to realize visual AR. Optical see-through (OST) superimposes the real environment with virtual content directly in user's field of view by using semi-transparent mirrors as projection planes. In contrast, the real environment is shown in video see-through (VST) as a digital video feed on a separate display with an AR. These techniques can be realized as a stationary, a handheld or as a head-worn display (HWD), which is most common (Van Krevelen & Poelman, 2010).

Besides the recently growing usage of augmented reality in the field of consumer applications (e.g. smartphone games), Sanna and Manuri (2016) show some examples how AR is already used professionally in several industries (e.g. medicine, architecture, automotive). In the automotive industry engineers can be supported at

In D. de Waard, K. Brookhuis, D. Coelho, S. Fairclough, D. Manzey, A. Naumann, L. Onnasch, S. Röttger, A. Toffetti, and R. Wiczorek (Eds.) (2019). Proceedings of the Human Factors and Ergonomics Society Europe Chapter 2018 Annual Conference. ISSN 2333-4959 (online). Available from http://hfes-europe.org

the product development process, as well as customers at the driving task. Modern vehicles are often equipped with a head-up display to show virtual information to the driver, directly in line of sight onto the real environment. To this purpose, the semi-transparent windshield is used as an optical see-through technology to reflect a generated image into driver's eyes, who then perceives a virtual image behind the windshield. Originally, head-up displays were used for screen-fixed (unregistered) content like speed and navigation cues in the form of abstract icons. In contrast several publications show also the opportunity to use a HUD to present world-fixed (registered) augmented reality content in the field of navigation and other assistance systems (Gabbard et al., 2014). An exemplary comparison of a conventional HUD and an augmented reality head-up display (AR-HUD) is shown in figure 1.

Figure 1. Comparison of a conventional HUD and an AR-HUD showing speed information, a navigation cue and an obstacle warning.

The biggest benefit of head-up displays over traditional head-down displays is that there is no need to take the gaze away from the street to look at displayed information (Gish & Staplin, 1995). This leads to a reduction of reaction times in unexpected traffic situations (Horrey & Wickens, 2004; Kiefer, 1998). Furthermore, correctly superimposed augmented reality content is supposed to reduce mental effort and increase situation awareness, because the interpretation of information becomes easier due to the spatial context (Bengler et al., 2015).

However, an AR-HUD has much higher technical requirements than a conventional HUD. First, the virtual image needs to be placed in further distance to ensure that the driver is able to perceive real and virtual content simultaneously without the need of changing the focus constantly. While the virtual image distance of a conventional HUD is usually at 2 - 3 m, an augmented reality image plane should be placed in a distance of at least 7.5 m to the drivers' eyes (Jachens et al., 2017). Furthermore, a bigger field of view of the head-up display is usually needed for augmented reality applications. Both requirements lead to a big increase in the installation space of the display unit.

Besides higher requirements for the display technology itself, augmented reality requires a constant stream of detailed information about the environment and a precise tracking of the user's respectively the vehicle's movement for an accurate registration of virtual content (Van Krevelen & Poelman, 2010). Modern vehicles are equipped with a detailed road map and several sensors for the self-localization (e.g. GPS) and the detection of other road users and attributes (e.g. camera and radar). Nevertheless, the precision of those systems is limited, which may lead to deviations (registration errors) up to several meters in the superimposition of the environment with virtual objects (Bengler et al., 2015).

In general several types of registration errors can be differentiated. There are position errors which describe translational deviations in lateral, longitudinal or altitudinal direction. Additionally rotational deviations in yaw, pitch and roll are referred to as orientation errors (Holloway, 1997). Position errors result in a constant offset between a real and a virtual object in any of the three directions. In contrast when it comes to orientation errors, the degree of the offset varies with the distance to the virtual object. Furthermore, latency and signal noise might occur. Latency leads to a smooth swimming of the virtual object in up to all six degrees of freedom, whereas noise results in an unsmooth jitter (Sadovitch & Wittkämper, 2018).

Pfannmüller et al. (2014) investigated the effects of inaccurate positioned augmented reality navigation cues in a driving simulator study. The participants chose more often a wrong turn when the virtual arrow on the street had a position error of more than 3 m to the ideal position. Intuitively, the usability and other subjective assessment items were rated significantly better, when there was no or just a smaller registration error. In another empirical study participants watched movies of real driving scenarios with an AR-HUD, filmed out of the driver's perspective. Significantly negative effects on subjective system acceptance were reported with orientation errors of around 0.5° in yaw and 0.2° in pitch (Sadovitch & Wittkämper, 2018).

To improve registration precision and reduce latency several tracking and rendering approaches exist. To take advantage of the benefits of various tracking technologies the hybrid use of several tracking sensors is advised (Van Krevelen & Poelman, 2010). For example, the localization of a vehicle via GPS can be optimized by a sensor fusion with the data of a camera (Jachens et al., 2017). Another approach is to use the sensor data to predict the future position and orientation (Azuma, 1995). Besides an optimized tracking, rendering should be improved as well to reduce latency. Due to a long rendering pipeline the tracking data on which basis an image is rendered, is not the latest anymore at the time the image is displayed. Zheng (2015) describes several approaches to reduce rendering latency such as image displacement. Here, a larger image is rendered than displayed to be able to shift the image according to the latest tracking data just before displaying. Another approach suggests even to hide the image for a few frames, when it comes to very high frequented orientation changes of the vehicle (Tasaki et al., 2013).

In addition to technical tracking and rendering improvements, it is also possible to adapt the interface design of augmented reality visualizations to counteract registration errors. One possible approach is the use of contextual cues to decrease ambiguity of inaccurate registered information. In a laboratory study imprecise registered virtual objects were rather recognized as belonging to a specific real object, if the objects were similar in their form and color (Wegerich, 2012). Furthermore, some scientists assume that the design of virtual elements may lead to a lower or higher saliency of registration errors. In a collaborative AR application sharp edges were avoided in the design of virtual objects. Authors used blurred visualizations instead to make a possible overlap of virtual and real object less visible (Fuhrmann et al., 1999). In case of a virtual bounding box or a frame to mark a real object, overlapping due to registration errors may even be prevented by extending the size of the virtual element (MacIntyre & Coelho, 2000).

While the improvement of technical approaches is a main focus in augmented reality science, there is little scientific effort done yet to explore non-technical solutions for the compensation of registration errors. Unfortunately, the literature offers just a few proposals that are not even evaluated yet. In the following sections a simulator study will be described to investigate some of the introduced error-tolerant HMI solutions and some further ideas to counteract registration errors in augmented reality head-up displays.

Methods

Experimental Design

The experiment was conducted as a 4 (event type) x 5 (visualization type) within subject design in a static driving simulator with one primary task and two secondary tasks. As the main task subjects had to monitor an Adaptive Cruise Control (ACC) assistance system while driving a vehicle on the highway. The system adjusted the vehicle velocity automatically to maintain a safe distance from vehicles ahead. For that a preferred target speed and distance had to be set. Although, the driver usually sets these settings, in this case the target distance was predefined and the speed was determined by traffic signs automatically. The subjects had tosteer and monitor the assistance system, which is categorized as level 1 automation (SAE J3016, 2016).

Most vehicles were detected accurately by the ACC system and thus the vehicle regulated the velocity and distance. However, sometimes the system was erroneous and vehicles were not detected, which lead to the system not reducing velocity, despite a slower vehicle driving ahead. Most of the later situations did not necessitate an intervention from the driver. But every trial there was one critical event where the vehicle ahead suddenly braked. To avoid a collision, subjects had to recognize the system failure and brake manually. Summarizing there were four types of events in every trial:

- Event type no. 1: No ACC relevance
- Event type no. 2: Correct ACC behaviour
- Event type no. 3: ACC failure in safe distance
- Event type no. 4: ACC failure in critical distance

Braking reaction time in critical ACC failure events was analysed and considered as an indication for driving performance. To avoid collision scenarios in the stereoscopic simulation setting, the upfront vehicle accelerated automatically if the distance got shorter than 10 m. Bad driving performance was interpreted as an evidence for driver distraction due to the type of augmented reality visualizations, which are described in the following paragraph.

Participants were supported in the monitoring task by an augmented reality visualization in a simulated head-up display, indicating which vehicle was detected by the ACC system. Vehicles ahead with a relevance for the ACC adjustments of velocity were marked with an orange frame directly in driver's field of view. To simulate registration errors the augmented reality marking was always swimming around the vehicle with orientation errors varying dynamically in between of 0 - 0.5° in pitch and 0 - 1° in yaw. There were five different visualization types used for the augmented reality frame (see figure 2). The *basic frame* is based on a simulator study, where Pfannmüller (2017) observed negative effects of registration errors in an AR-HUD on driving performance and subjective system assessment. The basic frame consisted of sharp borders and was not optimized towards an error-tolerant HMI. The other four visualization types were assumed to be more error-tolerant by using an *extended*, a *diffuse*, a *fragmented* and a *reduced frame* to mark the detected vehicle ahead. The subjective evaluation of registration precision was measured by several items of a questionnaire.

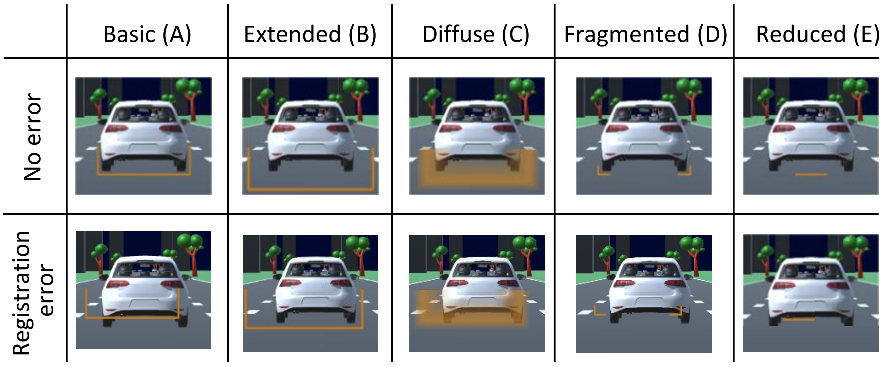

Figure 2. Types of augmented reality visualizations with and without a registration error.

In addition to the driving performance and the subjective perception of registration precision the mental workload was measured through a secondary *tactile Detection Response Task (tDRT)*. For this purpose, a vibration motor was fixated at the wrist of participant's dominant hand. Every three to five seconds the motor started to vibrate. The task was to press a switch, which was attached at the thumb, against the steering wheel, as fast as possible. If there was no response for 2.5 seconds the vibration stopped automatically and the event was logged as a miss. This task is suited for sensitive measurements of mental workload in primary tasks with high mental load (Bruyas & Dumont, 2013). Since the primary task in this study was very monotonous it was necessary to implement another secondary task to increase the overall task load. Therefore the participants performed an auditive N-back task (N = 1) additionally (Kirchner, 1958). The following letters were played in three seconds intervals*: c, g, h, k, p, q, t* and *w*. If the same letter was played two times in sequence, subjects had to signalize it verbally by saying the german word "*Jetzt*" (English translation: "*now*"). A verbal response was choosed instead of a tactile one, because participants had already to react to the vibrations of the Detection Response Task with a tactile response.

Material

To realize a realistic visual experience of the augmented reality head-up display, the experiment was conducted in a stereoscopic driving simulator. Participants had to wear a pair of shutter glasses to perceive a 3D image of the projected Virtual Test Drive simulation. The head-up image, which was rendered with the game engine Unity 3D, was shown as an overlay with a second stereoscopic projector. However, the simulation was limited to a single front view without side views. The authors consider this circumstance as acceptable since the driving task was concentrated on the traffic in front of the participants. The front view had a field of view of 60° and the projection surface was approximately 3.80 m away of driver's eyes.

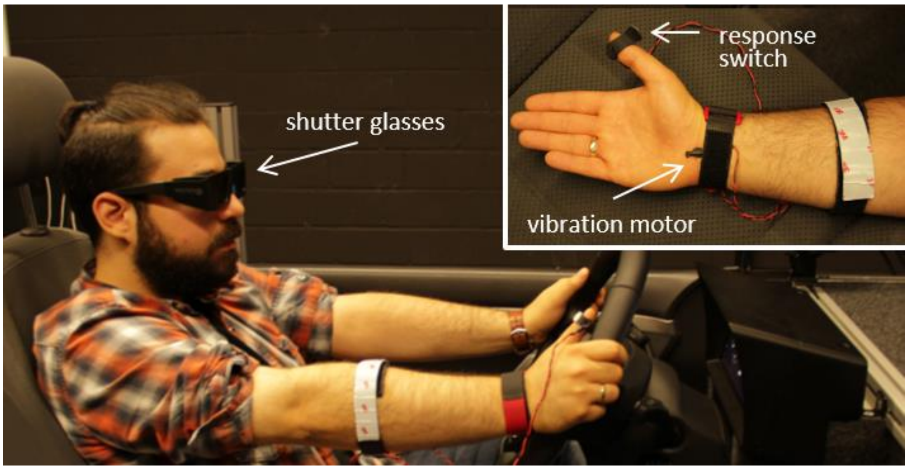

Figure 3. Experimental setup with a participant wearing shutter glasses and tDRT apparatus.

Procedure

In the beginning, the participants were informed about the experiment procedure and the handling of personal data, which was saved separately. All data was anonymized and only used for the purpose of the present study. It was ensured that no conclusions about the person are possible from the results. After agreeing voluntarily, participants were asked to fill in a sociodemographic questionnaire and perform the *Butterfly Stereo Acuity Test* to gather information about stereo vision capabilities for post-hoc analyses.

Afterwards the participants were introduced to the ACC assistant system and the augmented reality visualizations. The main task as well as the two secondary tasks were trained separately and together. The actual experiment contained five simulated driving scenes with different HMI conditions. The order of the conditions was balanced according to the latin square (Bortz & Schuster, 2010). After finishing the experiment the participants were debriefed and received a simple gift as an expense allowance. The experiment was designed to take 90 minutes.

Participants

Twelve females and 50 males participated at the experiment. All of the 62 participants were employees of the Volkswagen AG, working in different departments and participating voluntarily during their free time.

The participants had an average age of 36.05 years (SD = 9.65). Although nobody achieved an inadequate result in the stereo vision test, the stereo vision of six people was slightly impaired (> 100 seconds of arc). Since there were no conspicuity noticed in the post hoc tests, these participants were not excluded from data analyses.

Results

Subjective error tolerance

The subjective error tolerance of the examined HMI solutions was addressed by three items in the questionnaire: *perceived offset*, *perceived swimming* and *visual impairment of registration errors* (see figure 4). A repeated measures ANOVA was conducted for each item. While there was no significant difference concerning the perceived offset (F (4, 244) = 1.95, p = .103), a main effect of the visualization type was observed regarding the perceived swimming (F (4, 244) = 3.93, p=.004). A post hoc test with Bonferroni correction showed that the effect was mainly based on a significant difference (p = .015) between the best rated HMI (diffuse frame) and the worst rated one (reduced frame). This did not lead to any significant differences in the visual impairment of registration errors (F (4, 244) = 1.79, p=.131).

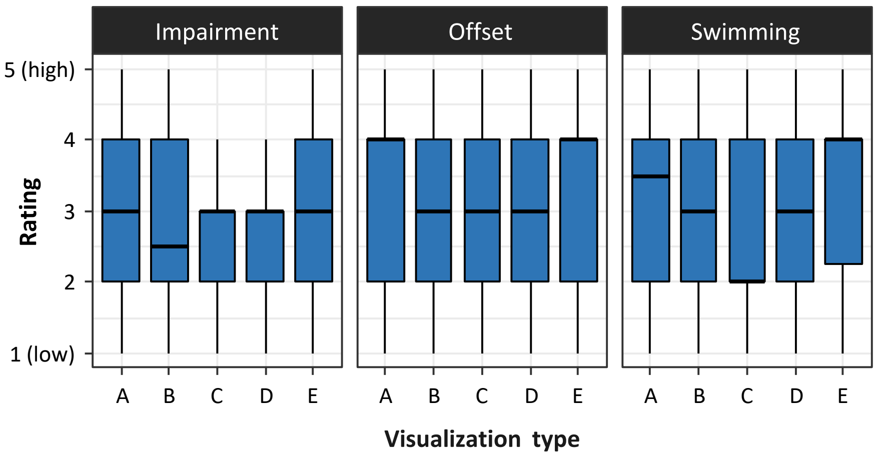

Figure 4. Subjective rating of visualization types regarding error tolerance.

After the experiment, subjects were asked to sort the visualization types in a ranking regarding the error tolerance. A significant difference was observed (χ^2 = 12.66, p = .013). Over 50% of the participants ranked either the extended or the diffuse HMI on the highest rank and the basic HMI on one of the two lowest ranks with the lowest error tolerance. In comparison with the basic visualization type the extended (p = .004) and the diffuse (p = .042) depiction were rated as significantly more error tolerant. The fragmented frame was rated just slightly, but not significantly better (p = .418) and the reduced HMI slightly worse (p = .879) than the basic visualization type.

Further subjective ratings

Additionally participants were asked to rate the visualization types regarding three positive attributes (*usefulness*, *design*, *detectability*) and three negative attributes (*distraction*, *masking*, *workload*). Mean results are shown in figure 5. A repeated measures ANOVA showed significant effects of the visualization type for every item except of distraction ($F(3.37, 205.42) = 0.767$, $p=.587$).

The extended HMI was rated as most useful and detectable. Furthermore, it demanded the least perceived mental workload. However, participants preferred the design of the fragmented frame, which masked less of the environment than the other visualization types. The highest masking was produced by the diffuse HMI, which was also harder to detect than the sharp edged designs. Overall, the reduced frame was rated worst. In average it was perceived as less useful and harder to detect than the other visualization types. Moreover, the reduced HMI caused the highest perceived mental workload and participants did not like its design.

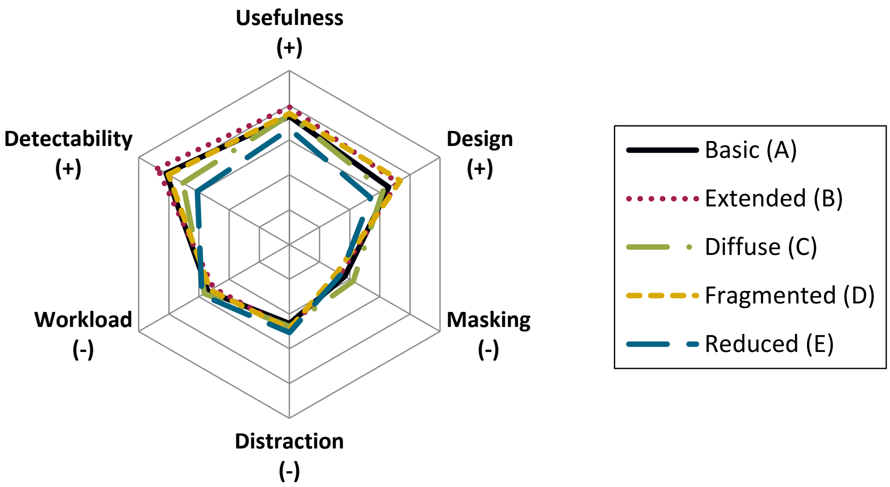

Figure 5. Further subjective ratings of visualization types.

Driving performance

The driving performance was measured by analyzing the break reaction time in critical ACC failure events. Reaction times faster than 300 milliseconds were excluded from further analysis, because they suggest a rather coincidental break activity. In hindsight to the automatic acceleration of the upfront vehicle to avoid collision scenarios, reaction times bigger than two seconds were defined as missed reactions. Since only four cases of missed reactions occurred in more than 300 trials, no further analyses were conducted.

The smallest mean reaction time of 1,001.41 ms (SD = 230.66) was observed for the reference frame. Reactions were just slightly slower in the reduced HMI (M = 1,006.41 ms, SD = 238.37) and the diffuse HMI condition (M = 1,022.39 ms, SD = 255.98). Even the highest mean reaction times of 1,050.30 ms (SD = 234.39) and 1,062.78 ms (SD = 224.70) with the visualization types fragmented and extended frame were just slightly worse than the fastest reactions. Based on this data a repeated measures ANOVA showed no significant main effect for the visualization type ($F(4, 180) = 0.83$, $p = .508$) regarding the driving performance.

Secondary task performance

The performance in the n back task was not recorded, because the task was only implemented to increase overall task load. For an adequate interpretation of the performance data in the tDRT, stimuli were assigned post hoc to the different types of events. Participants received most of the vibration stimuli in unspecified events and only few stimuli in critical failure situations (see table 1). Since the augmented reality marking of detected vehicles is rather designed to increase situation awareness in general than to warn in critical situations, the tDRT performance in the event type no. 4 was not further analyzed.

Table 1. Average amounts of vibration stimuli in each event type per visualization type.

	Event type 1	*Event type 2*	*Event type 3*	*Event type 4*
Basic HMI	35.51	8.08	3.43	1.10
Extended HMI	35.87	8.30	3.48	1.08
Diffuse HMI	36.08	8.72	3.41	1.07
Fragmented HMI	35.87	8.31	3.30	1.11
Reduced HMI	35.28	8.30	3.23	1.13

To analyze secondary task performance a two factors (event type x visualization type) repeated measure ANOVA was calculated. The results confirmed the sensitivity of the tDRT to measure differences in the mental workload within the present setting. A significant main effect of the event type was observed for the hit rate ($p < .001$) as well as the amount of false alarms ($p = .008$) and the reaction times ($p < .001$). The average reaction times for the separated types of events are shown in figure 6.

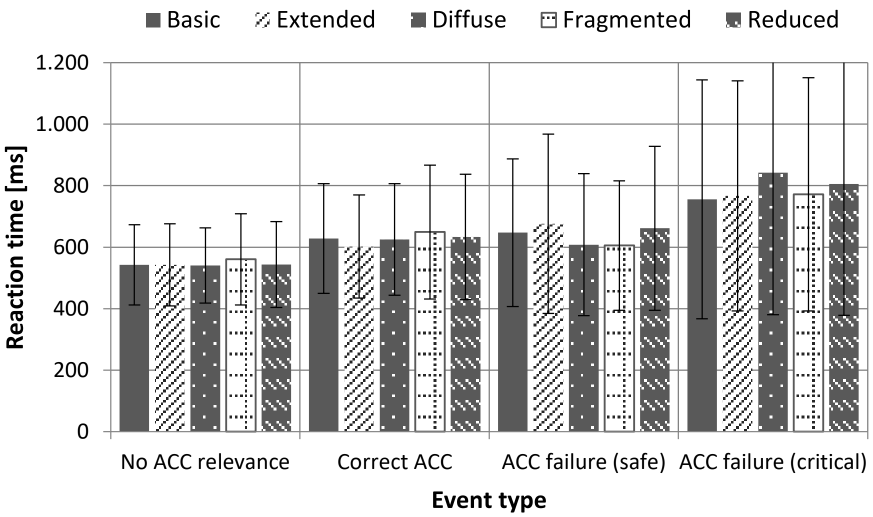

Figure 6. Mean reaction times of the tDRT.

Only slight differences were found between the HMI conditions regarding secondary task performance. No significant main effect of the visualization type could be observed neither for hit rates (F (4, 216) = 0.67, p = .612), nor for false alarms (F (4, 216) = 0.975, p = .422) or reaction times (F (3.57, 185.81) = 0.57, p = .664). Furthermore, there were no significant interactions of the factors event type and visualization type. This leads to the conclusion that there is no evidence for differences in the mental workload due to the type of the evaluated augmented reality visualizations.

Discussion

Limitations in the precision of tracking technology and maps lead to registration errors in augmented reality. Concerning an augmented reality head-up display, this might affect driving performance, usability (Pfannmüller, 2017) and subjective system acceptance (Sadovitch & Wittkämper, 2018). There are several technical approaches discussed in literature to improve tracking and rendering regarding precision and latency (e.g. hybrid tracking, predictive tracking, image displacement). Additionally the problem can be addressed by the use of an error-tolerant HMI, which does not improve the actual but the perceived registration precision.

The approach was evaluated using the example of an augmented reality adaptive cruise control assistance system in a stereoscopic driving simulator. Besides a basic frame to mark a detected vehicle ahead, four different potentially error-tolerant visualization types were tested regarding a counteracting effect on the perception of registration errors (see figure 2).

Subjective data showed that the participants perceived the highest registration precision with the use of an extended and a diffuse HMI, although the actual precision was not better than in the other trials. The extended frame was also very easy to detect and it was rated as very useful. Though subjects did not like its design as much as the fragmented ones for instance. A combination of these visualization types should be considered.

In contrast, the diffuse HMI revealed some more disadvantages. Participants criticized the design as well as the masking effects and a bad detectability. Therefore an optimization of this visualization type should be aimed. In total only very slight differences could be observed in the subjective assessement of error-tolerance. Overall registration precision were not assessed as poor as authors have assumed before. Therefore it should be taken in consideration to increase registration errors for further evaluations. Especially as the occurance of stronger errors is to be expected in real driving conditions.

The reduced visualization type showed the lowest error-tolerance and usefulness aswell. This might be explained by a missing connection to the marked vehicle. Participants remarked that they had even problems to determine the position in depth of the virtual element. For this reason the authors advise not to use too simple and abstract elements in augmented reality that have no direct connection to the environment.

There were no indications found for any differences in driving performance and mental workload due the visualization type. This allows implementing further augmented reality visualizations that are less similar to the basic frame. Even some dynamic patterns should be considered and evaluated additionally to the examined static visualization types.

References

Azuma, R.T. (1995). *Predictive tracking for augmented reality*. Dissertation. University of North Carolina at Chapel Hill.

Bengler, K., Götze, M., Pfannmüller, L., & Zaindl, A. (2015). To See or not to See - Innovative Display Technologies as Enablers for Ergonomic Cockpit Concepts. Ergonomic requirements future mobility future functionality. In *electronic displays conference*. Nürnberg, Germany.

Billinghurst, M., Clark, A. & Lee, G. (2015). A Survey of Augmented Reality. *Foundations and Trends in Human–Computer Interaction, 8*, 73–272.

Bortz, J. & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler*. Berlin Heidelberg: Springer-Verlag.

Bruyas, M.-P. & Dumont, L. (2013). Sensitivity of detection response task (DRT) to the driving demand and task difficulty. In *Proceedings of the 7th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design* (S. 64–70). Bolton, NY, USA.

Fuhrmann, A., Hesina, G., Faure, F., & Gervautz, M. (1999). Occlusion in collaborative augmented environments. *Computers & Graphics, 23*, 809–819.

Gish, K. & Staplin, L. (1995). *Human Factors Aspects of Using Head Up Displays in Automobiles: A Review of the Literature* (U.S. Department of Transportation, Eds). Washington D.C.: National Highway Traffic Safety Administration.

Holloway, R.L. (1997). Registration error analysis for augmented reality. *Teleoperators and Virtual Environments, 6*, 413–432.

Horrey, W.J. & Wickens, C.D. (2004). Driving and side task performance: The effects of display clutter, separation, and modality. *Human Factors, 46*, 611–624.

Jachens, A., Zühlsdorf, M., Kern, T.A., & Abel, H. (2017). Augmented Reality Head-up display. System requirements and solutions concerning precise augmentation of information within the road scene. In *ELIV - Electronic in Vehicles* (S. 581–592). Bonn, Deutschland.

Kiefer, R.J. (1998). Quantifying head-up display (HUD) pedestrian detection benefits for older drivers. In *16th International Technical Conference on Experimental Safety Vehicles*. Windsor: NHTSA (S. 428–437).

Kirchner, W.K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology, 55* , 352–358.

MacIntyre, B. & Coelho, E.M. (2000). Adapting to dynamic registration errors using level of error (loe) filtering. Augmented Reality, Proceedings. *IEEE and ACM International Symposium* (ISAR 2000), 85–88.

Pfannmüller, L., Walter, M., & Bengler, K. (2014). Lead me the right way?! The impact of position accuracy of augmented reality navigation arrows in a contact analog head-up display on driving performance, workload and usability.

Pfannmüller, L. (2017). *Anzeigekonzepte für ein kontaktanaloges Head-Up Display*. Dissertation. Technische Universität München.

Sadovitch, V. & Wittkämper, M. (2018). Akzeptanz von Registrierungsfehlern in einem Augmented Reality Head-up-Display. In VDI Wissensforum GmbH (Eds.), *Optische Technologien in der Fahrzeugtechnik* (VDI Berichte 2323, S. 129–144). Düsseldorf.

SAE J3016. (2016). *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*. USA: SAE Society of Automotive Engineers.

Tasaki, T., Moriya, A., Hotta, A., Sasaki, T., & Okumura, H. (2013). Depth Perception Control during Car Vibration by Hidden Images on Monocular Head-Up Display. I*EICE transactions on Information and Systems, 96*, 2850–2856.

Van Krevelen, D.W.F. & Poelman, R. (2010). A survey of augmented reality technologies, applications and limitations. *International Journal of Virtual Reality, 9,* 1-20.

Wegerich, A. (2012). Nutzergerechte Informationsvisualisierung fuer Augmented Reality Anzeigen. Dissertation. Technische universität Berlin.

Zheng, F. (2015). *Spatio-temporal registration in augmented reality*. Dissertation. The University of North Carolina at Chapel Hill, Chapel Hill, NC.