

# User-Centered Perspectives for Automotive Augmented Reality

Victor Ng-Thow-Hing<sup>\*1</sup>, Karlin Bark<sup>1</sup>, Lee Beckwith<sup>1</sup>, Cuong Tran<sup>1</sup>, Rishabh Bhandari<sup>2</sup>, Srinath Sridhar<sup>3</sup>

<sup>1</sup>Honda Research Institute USA, Mountain View, CA, USA

<sup>2</sup>Stanford University, Stanford, CA, USA

<sup>3</sup>Max-Planck-Institut für Informatik, Saarbrücken, Germany

## ABSTRACT

Augmented reality (AR) in automobiles has the potential to significantly alter the driver's user experience. Prototypes developed in academia and industry demonstrate a range of applications from advanced driver assist systems to location-based information services. A user-centered process for creating and evaluating designs for AR displays in automobiles helps to explore what collaborative role AR should serve between the technologies of the automobile and the driver. In particular, we consider the nature of this role along three important perspectives: understanding human perception, understanding distraction and understanding human behavior. We argue that AR applications should focus solely on tasks that involve the immediate local driving environment and not secondary task spaces to minimize driver distraction. Consistent depth cues should be supported by the technology to aid proper distance judgement. Driving aids supporting situation awareness should be designed with knowledge of current and future states of road users, while focusing on specific problems. Designs must also take into account behavioral phenomena such as risk compensation, inattentional blindness and an over-reliance on augmented technology in driving decisions.

**Index Terms:** H.5.1 [Multimedia Information Systems]: Artificial, augmented, and virtual realities—; H.5.2 [User Interfaces]: User-centered design—

## 1 INTRODUCTION

Augmented reality (AR) in automobiles can potentially alter the driver's user experience in significant ways. With the emergence of new technologies like head-up displays (HUDs) that are AR capable, designers can now provide visual aids and annotations that alter what the driver focuses on, and how they accomplish the driving task. While this can potentially alleviate cognitive load and create more enjoyment in the driving task, it can also introduce new risks.

In contrast to AR applications on smart phones or tablets, the windshield offers a direct and larger field-of-view of the actual environment. Automobiles can be equipped with a wider variety of powerful sensors and computational power. In addition, automobiles are generally constrained to roadways, which helps to limit the domain of possible contexts applications need to focus on. Driver distraction is a clear danger. The National Highway Traffic Safety Administration (NHTSA) in the United States identifies three types of driver distraction[30]: visual distraction (eyes off road), cognitive distraction (mind off driving), and manual distraction (hands off the wheel). AR can directly affect the first two types.

There are looming societal trends that the future automobile must deal with. In most developed countries, the number of elderly

drivers are growing. Many of these older drivers must choose between a radical change in their lifestyle or the risk of continued driving with impaired visual capabilities[45]. AR can be used to increase saliency of important elements in the driver's view and has the potential for augmenting the driver's situation awareness. These benefits can trickle down to all drivers, regardless of age. Since AR in the car provides a personal display to the driver, information can be customized for the individual, both in terms of content and in visual parameters to match the driver's visual and cognitive capacities. In fact, there is a growing number of advanced driver assistance systems (ADAS) that provide aids to the primary task of driving such as pedestrian detection[16] and blind spot detection[40].

Cars will have access to a vast array of new information about their environment through sensors and internet connectivity. AR can serve as an intermediary for presenting this information as part of an immersive experience, while keeping the driver's eyes on the road and not on a secondary display on the dashboard. New sensors capable of depth measurements provide higher precision information about objects in the environment[24]. AR can provide a natural way to convey spatiotemporal information aligned with the moving positions of objects relative to the ego-centric view of the driver. The potential for greater state understanding that comes with improved intelligence in the car can also be communicated with augmented reality. This can be seen with AR navigation aids that directly display paths in the driver's view[15, 20].

The demonstration of autonomous vehicles on public roads[27] has been seen as a viable alternative to dealing with the dual problems of a growing number of elderly drivers that can no longer safely drive and increased channels of incoming information that can distract the driver. However, to increase acceptance and comfort levels using this technology, the car needs to communicate its decision-making process and perceptual awareness[31] so the driver is ensured the car is operating safely. AR, combined with other modalities such as sound or haptics, can serve as the intermediary for driver-car collaboration. In particular, the car can communicate its own situation awareness and driving plans to the passenger of autonomous vehicles using an AR interface. These human-machine interfaces must be developed simultaneously with technologies for driver automation.

For all these promising areas of application of automotive AR, the stakes are high. Visual perception plays a large role in determining a driver's situation awareness of the environment. *Situation awareness* is the perception of environment elements in time and space. If information is presented incorrectly, either in style or content, situation awareness may be improperly represented or driver distraction can occur, leading to dangerous driving conditions. Massive deployment of these technologies without any design guidelines for application developers can produce an unacceptable number of accidents. Dying is a bad user experience.

AR can be a medium of collaboration between the driver and the technology of the automobile. Technology enables better sensing of a car's immediate environment, access to location-based services,

\*e-mail: vng@honda-ri.com

and new modalities of interaction with a vehicle. AR can provide a way for drivers to interface with these technologies using a visual modality that is integral to the driving task. There is a need to design both the AR display technology and the applications that they will host carefully, considering the behavior and physiological constraints of the driver.

## 1.1 Contributions

This paper describes how we adapted a design framework to incorporate user-centered perspectives in automotive AR solutions. We discuss three important perspectives for successfully designing and implementing augmented reality in automobiles. First, we must understand the visual perception of humans. Here, we are referring to the visual processing that occurs in the brain, prior to the higher state recognition processes such as object recognition or assessing situation awareness. We present our own studies testing the importance of consistency among multiple depth perception cues. Second, we examine issues of driver distraction and how AR can influence it. Finally, we seek to understand how inherent aspects of natural human behavior should influence the design of solutions for driving aids. Specifically, it is important not to inadvertently introduce undesirable behavioral changes that can create dangerous driving habits. Throughout our discussion of these three points, we will illustrate with examples we are currently designing and developing.

## 1.2 Outline

We review the related work in automotive AR solutions in Section 2, followed by a description of our design process in Section 3. The three user-centered perspectives are discussed in detail in Section 4, Section 5, and Section 6. Conclusions are in Section 7.

## 2 RELATED WORK

One advantage an AR-based display has over a basic HUD display that conveys driving data like vehicle speed or other state information is the emphasis on contextual information as it relates to the external environment of the car. Tönnis et al.[42] created AR driving aids for longitudinal and lateral driver assistance that visually conform to the road’s surface. Park et al.[33] found that drivers had faster response times to lane-changing information when it was directly augmented in perspective over the road surface compared to being displayed as 2-D icons. Participants described the AR representation as intuitive, clear and accurate. This information can help disambiguate instructions by using augmented graphics in consistent perspective with environmental elements in the driver’s view. In the literature, these types of HUDs have been described as being contact-analog[35].

Clear and intuitive information annotating elements in the driver’s field of view unencumbered by the need to wear sensor apparatus, has tempted many researchers to explore AR use in car navigation[20, 29]. Medenica et al.[28] demonstrated that an AR-based navigation system caused drivers to spend more attention looking ahead on the road than non-AR systems. A hybrid system that smoothly transitioned between egocentric route information and a 2-D overhead map was simulated by Kim and Dey[22] to show its potential to reduce navigation errors and reduce driver distraction for elderly drivers.

The perceptual issues with augmented reality display systems have been identified and categorized. For stereoscopic displays, Drascic and Milgram[11] examined sources of perceptual mismatches with calibration, inter-pupillary distances, luminance and conflicting depth cues of accommodation and vergence. Kruijff et al.[23] expanded the examination of issues to include new mobile device displays. In these mobile displays, differences in the viewer’s and display’s field of view and the viewing offset angle of the real world and the display can effect perception when using

AR applications. Mobile displays also suffer from a lack of depth cues that can lead to underspecified depth of objects in the scene. In Section 4.1, we quantify how inconsistent depth cues can lead to inaccurate estimates of distances for driving applications.

A classification of presentation principles in AR was described by Tönnis and Plecher[43]. The focus of the classification was on different implementational details of AR solutions: temporality, dimensionality, registration, frame of reference, referencing, and mounting. Fröhlich et al.[15] describe a design space focused on different aspects of realistic visualizations (constituents and styles). These categories are related to the technical implementation of the AR solution. In contrast, the goal of our work is to examine AR solutions in the context of how they influence human cognition and how human cognition should play a role in creating the design of a solution. The incorporation of these user-centered perspectives should occur during the early conceptual design phases, before a particular prototype’s presentation and implementation has been finalized.

## 3 DESIGN PROCESS

With these perspectives of the driver in mind, the problem becomes less of how to implement or describe an idea technically, but what is the appropriate form of solution to a driver’s problem.

### 3.1 Understanding the Problem

The process of building up rationale for our designs begins with understanding the problem. Research reports[5] documenting the causes and types of accidents can give important statistics to help prioritize problems areas based on factors such as societal impact and size of casualties. However, they do not provide the personal perspective of the driver. To understand the problem from the driver’s perspective, contextual inquiry and design can be employed as others have done for human-machine interface (HMI) design[17]. We conducted in-car interviews with different demographic groups to understand their daily driving habits, concerns, and how driving is integrated in their daily work/life schedules (Figure 1). Conducting the interview in their own car, allows interviewees to demonstrate directly how they interact with elements in their cars.



Figure 1: In-car interviews conducted during the contextual inquiry process.

### 3.2 Ideation

Once information is gathered from interviews and other research reports, the process of analyzing and organizing this information into prevailing themes begins the creative process of idea generation or *ideation*[21]. To facilitate this process, we employed a structured brainstorming technique called *affinity diagramming*[7]. Observations from interviews as well as facts gathered from other research

reports are written down on notes of paper and organized according to similar themes or affinities (Figure 2). This grouping can be done at several layers to develop a hierarchical structure.

We interviewed 12 people (6 elderly drivers (age over 60) and 6 Generation-Y drivers (age 20-30). Through affinity diagramming we identified themes of safety, traffic, integration with mobile devices and exploration. In the area of safety, elderly drivers had concerns about their personal abilities regarding awareness of their surrounds and navigating intersection scenarios with pedestrians. The Generation-Y drivers mentioned more scenarios regarding the behavior of other drivers such as other people not seeing them. Unintentionally, we discovered individuals in both groups who could be classified as extreme driving enthusiasts. They explicitly avoided all distractions, such as phone use and music in the car, in order to concentrate on perfecting execution of driving maneuvers and developing better situation awareness. Studying these individuals inspired our designers to create solutions to help normal drivers become engaged in the primary task of driving rather than focusing on creating solutions involving distractive secondary tasks. We developed several solution concepts centered on situation awareness, focusing on providing drivers better information to improve their driving decisions rather than explicitly telling drivers what maneuvers to make. The aim of the former strategy is to engage the driver in understanding the relevant situation awareness factors for a particular driving situation. We felt that following the latter strategy would make drivers more reliant on the AR technology and might weaken their inherent driving skills over time.



Figure 2: Affinity diagrams for ideation. Blue notes represent groups collecting commonly-themed observations (in yellow). Pink notes are higher level categories that group themes together.

### 3.3 Sketching and prototyping

From the pool of solution concepts, the most promising are chosen and undergo development from early conceptual sketches, to computer-generated test animations to working prototypes in our driving simulator before eventual field testing in a vehicle. Figure 5 depicts the typical lifecycle of a driving aid. We refine key ideas and choices at the early sketch stages, before investing engineering effort to build higher fidelity prototypes. For example, we used Go-Pro cameras[19] mounted on the windshield to record video footage of 120 degree field-of-view, driver perspectives of various road situations. To simulate the restricted field of view, a white paper circle the same radius as our HUD outgoing lens was placed on the dashboard to visually indicate the HUD field of view via the circle's reflection off the windshield (Figure 3). We could then use video editing software to prototype different augmented graphics

over this footage using the field-of-view guides to limit where we could place augmented graphics. Higher fidelity prototypes eventually involved implementing driving aids to work within our driving simulator which projects road scenery on a 120 degree curved screen while augmented graphics are shown on a see-through HUD display in front (Figure 4).



Figure 3: A white circle was placed on the dashboard so that its reflection could indicate the usable field of view of the HUD.

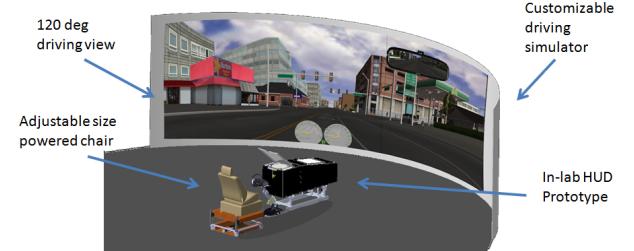


Figure 4: Driving simulator and HUD prototype: Augmented graphics appear on a the see-through windshield HUD display while road scenery is projected on the curved screen behind.

It is during the sketching process, where we work out the mechanics and appearance of a design, that we take into consideration the three user perspectives. For example, when designing the left-turn aid, one might have been tempted to draw the left-turn path the car should take, but we realized during our interviews that the main concern was deciding the timing of the left turn. One interviewee said, "It's difficult for me to judge how far away the oncoming traffic is, how fast they are going, and if it is safe to proceed". Further low-fidelity computer animations confirmed that adding a green turning trajectory did not accentuate the danger of an oncoming vehicle. In fact, it diverted attention away from the other car. Instead of telling drivers what to do, our solution (Figure 14) provides additional cues to the driver to enhance their situational awareness, but the decision to turn is left up to them.

### 3.4 Evaluation

Design techniques such as contextual inquiry, ideation and prototyping can be employed to identify needs, conceive solutions and implement them at various levels of fidelity[4]. This is important because our current understanding of the neurological processes mapping human perception to human behavior is imperfect. Only through observational evidence and evaluation of solution ideas can we validate if AR designs are useful.

Since experiencing AR in automobiles is not common, many diverse solutions must be examined in order to converge to a common set of design patterns that have established utility. In order to do so

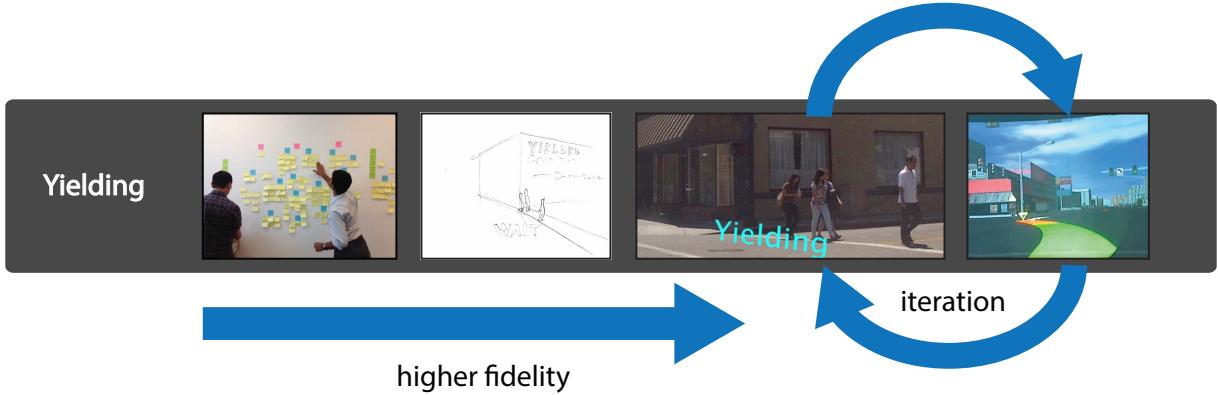


Figure 5: Typical design lifecycle of AR driving applications (e.g., yielding right-of-way): From user research, the ideation process creates solution concepts which are refined from initial sketches to a series of higher fidelity prototypes.

in a timely fashion, early evaluation should be done prior to expensive investment of resources implementing actual working prototypes. Designers may attempt to anticipate how users will react, but human behavior is often unpredictable. These early evaluations can help identify whether the designs are effective and allow for rapid iteration to help designers come up with a compelling solution.

In one example that illustrates the usefulness of early evaluation, a driving aid designed to deter drivers from prematurely cutting off a pedestrian's right-of-way was evaluated to observe both initial visceral reactions and effectiveness. Early concepts for the aid focused on visible textual barriers that increased the spatial footprint of the pedestrian, which we hoped would make the driver leave a bigger margin of safety around them.

For low-fidelity prototyping, we initially transitioned from hand-drawn sketches (Figure 5) to using pre-rendered graphics over pre-recorded driving scenery (Figure 6 and Figure 14). Registration of 3-D graphic elements with moving elements from video was done using a process called matchmoving [10]. These animations were presented to drivers for user acceptance testing. We received initial visceral reactions and visual response using eye-tracking equipment[39].

For the yielding aid, positive reactions were recorded from an elderly user:

"The visuals are helpful and confirm when it is safe to proceed... This is a very good confirmation and a second opportunity to be cautious and yield. I prefer this... it reminds me when my wife accompanies me when driving... she is usually vigilant and confirms such risks too... I like this!"

Other useful suggestions on color, placement and motion of the aid were noted by users.

Although qualitative observations provide insightful feedback, evaluations that can capture the user's behavior or intended actions are also a necessity. A challenge is to develop low-fidelity prototypes that allow user enactments for AR with correct driving mechanics implemented. One method that our group has employed for these enactments is to use short, focused clips of real life driving footage and animation overlays to estimate how a user would act in different situations. Users can be asked to indicate when they would perform different actions using push buttons or gaming input devices. We focus on identifying one or two basic parameters to measure, which keeps the study short and simple.

In the case of the yielding example, the main intent of the yield text was to provide a virtual barrier that would encourage drivers



Figure 6: Text is used as a barrier to protect pedestrians.

to wait a longer period of time before driving through the crossing. A short, two minute study was conducted to observe whether or not drivers presented with the yield text would indeed wait longer. With a single question in mind, a controlled study with numerous measures was not necessary. Instead, participants were shown a short, ten second video clip that depicted a scenario with pedestrians (Figure 6) and participants were asked to indicate the time at which they would proceed past the pedestrians after the pedestrians finished crossing the street. Half the participants were shown video clips with the driving aid, while the other half were shown the original recorded footage. Surprisingly, we found that the driving aid did not have the intended effect. In fact, we found that participants who viewed the video without the driving aid tended to wait a longer period of time ( 0.5 s) before proceeding than those who had the driving aid. This evaluation, which only took one day to conduct, provided valuable data highlighting the ineffectiveness of the design. Although further follow up studies are needed to determine precisely why this was the case, early detection of this behav-

ior allows the team to iterate and improve upon the design prior to conducting a larger, controlled study.

Once a solution has been refined, one can then identify and examine design variables, perhaps using proposed classification schemes like described in [43]. Since detailed studies require a large investment of time, recruitment of subjects and evaluation of collected data, they should be done only after a candidate idea has been sufficiently iterated and refined to the point of being testable. Controlled studies can be employed to identify which design variables affect driving performance. Physiological measures such as eye-tracking and skin conductance sensors, in addition to driving metrics recorded from driving simulation sessions, can be used to quantitatively evaluate a driver's performance. Eye-tracking in particular can help confirm if a driver's visual perceptions are being influenced by AR elements applied to the driver's field of view.

Driving simulations are an effective tool for controlled studies. Using driving simulations, we can test driving behavior as well as model dangerous situations that are not feasible with real driving. Many of the perceptual and behavior processes in driving are not conscious. Driving simulation is a very important tool to observe driver behavior while using the AR-based driving aids. Furthermore, the basic application can be tested for usability before additional research and development effort is made investing on sensors and computer recognition algorithms for detecting elements in the car's environment. Nevertheless, additional engineering effort is required to implement the aids to the stage that can work correctly with driving data generated from a driving simulator.

In the next three sections, we examine further several user perspectives that must be taken into account when conceiving AR applications for the car.

#### 4 UNDERSTANDING HUMAN PERCEPTION

In this section, we discuss several perceptual cues and processes that are important to consider when adding synthetic imagery to the visual field in AR automotive applications.

##### 4.1 Depth Perception

Proper handling of depth perception is important in automotive AR for the driver to gauge distances of augmented objects with respect to real objects in the scene. In Tönnis et al.[44], a study using an egocentric view of a 3-D arrow to guide the driver's attention produced slower reaction times than an indirect bird's eye view locating the area of immediate danger. The results may seem counterintuitive as one would think the arrow in the egocentric view would be more direct at localizing the source of attention. However, the 3-D arrow was rendered as a projection on a 2-D display, which the authors speculated might have been the reason for the slower reaction times as observers were missing important stereoscopic depth cues. Current techniques to display AR utilize a see-thru display by reflecting the computer graphics imagery off a windshield or see-thru surface mounted near the windshield. However, there are significant challenges in displaying images properly to the driver with correct and consistent depth cues.

For a 2-D display, monocular cues such as size, occlusion and perspective can be modeled directly with the standard computer graphics pipeline. However, with the introduction of optical see-through displays[2], the eyes must be able to make sense of an entire combined scene of synthetic and real elements at the same time. *Accommodation* is an important depth cue where the muscles in the eye actively change its optical power to change focus at different distances to maintain a clear image. *Vergence* is the simultaneous inward rotation of the eyes towards each other to maintain a single binocular image when viewing an object. If an image is created directly on the windshield, as our eyes converge to points beyond the windshield into the environment, a distracting double image of the windshield visual elements will appear. The fact that the eye

can only focus on one distance at a time presents a problem when displaying AR imagery on see-thru displays.

Displays that generate images directly from the windshield, such as by exciting embedded transparent phosphors in the glass via laser projection [46], will cause the eye to focus directly on the surface of the windshield, making the entire scene beyond the car's windshield out of focus. Furthermore, the head position must be tracked to align the virtual image with the real objects they are meant to augment in order to properly simulate motion parallax. Although head-tracking technology has become more robust with face recognition algorithms, this still requires additional hardware and computer vision processing which can add to latency in the proper rendering of graphics.

Methods that use optical combiner-HUDs can generate imagery beyond the windshield. Most HUDs in cars today, including so-

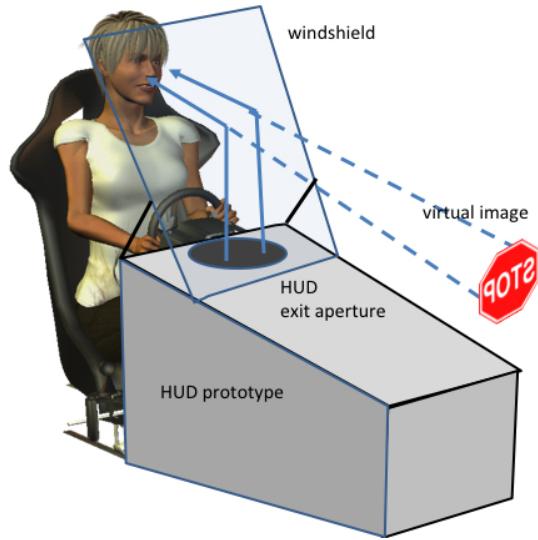


Figure 7: Diagram of HUD prototype with virtual sign generated. Sign distance depicted is much shorter than actual distances used in the study for space reasons.

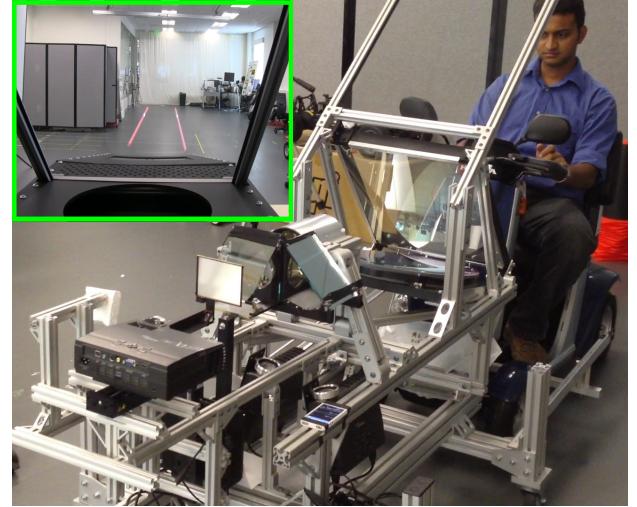


Figure 8: Actual prototype shown without cover, displaying optics and actuated projector displays. Inset: Virtual guiding lanes aligned with ground plane are generated with the prototype from the driver's point of view.

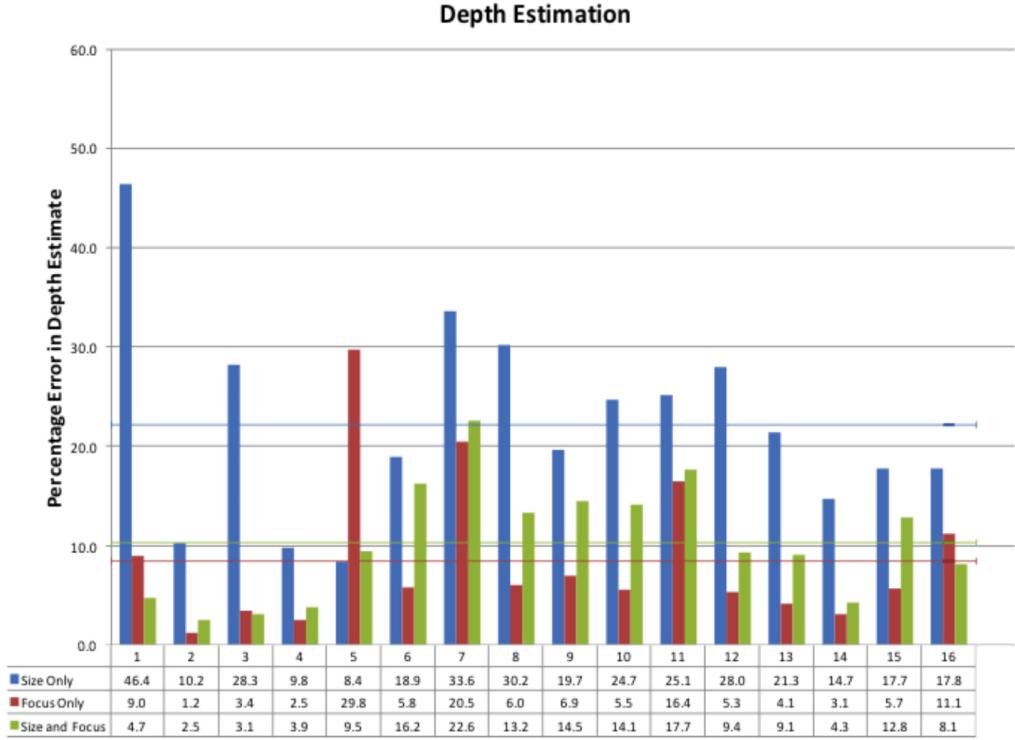


Figure 9: Percentage errors for estimating distances of an AR street sign using different depth cues for 16 subjects

lutions that directly reflect a smartphone or tablet display off the windshield, result in an image that is in a fixed focal plane a relatively close distance beyond the windshield, not more than three meters[3]. Consequently, drivers will not be able to correctly gauge correct depth perception because the depth cues on the HUD display will not match the depth of the objects in the environment they are meant to augment, which are typically beyond 5 meters. Displays must be built that have dynamic focal planes that can be adjusted to arbitrary distances beyond 5 meters to several hundred meters beyond the car, providing enough time for a driver to see potential hazards and safely react to them.

We built a prototype HUD display (Figures 7 and 8) that has actuated optical elements that enable the focal plane to be dynamically adjusted from 5m to infinity. When testing with a group of 16 participants (8 males and 8 females), aged 19-78 ( $\mu = 42$ ,  $\sigma = 20$ ), there is a clear difference in depth estimation when people are asked to judge the distance of a virtual stop sign drawn at different distances (from 7m-20m) using a fixed focal plane at 7m compared to one that can adjust its focal plane to match the targeted distance (Figure 9). Subjects were given different depth cue conditions from size only, focus only (accommodation) and size+focus. It is clear that size cues alone had the largest distance estimation percentage errors among the three conditions. The mean percentage error in depth perception by participants with size-based cues is 22% compared to 9.5% with focal cues ( $p < 0.01$ ). The condition for size + focus cues showed no significant difference from the focal cue only condition. Interestingly, for the size-only condition, participants estimated the depth to be close to the fixed focal length that was set for that condition (7 meters). This suggests that the depth cue for accommodation can dominate potentially weaker cues like size, reinforcing the need to express accurate depth cues for AR elements. See-thru display systems must have the ability to adjust the focal planes at a large range of distances beyond 5m to properly imple-

ment driving aid augmented reality applications. Incorrect depth perception can lead to incorrect driver decisions.

## 4.2 Field of view

Although humans are capable of over 180 degrees of horizontal field of view[23], only the center 2 degrees of the fovea has highest visual acuity [14]. The visual acuity contributes to the driver's useful field of view (UFOV) which is defined as the visual area over which information can be extracted at a brief glance without head movements[18]. Due to the limited size of the UFOV, drivers need to turn their heads to attend to different parts of the environment. From a driver's viewpoint in a moving vehicle, the most stable parts of an image are further away, which coincides with where drivers should be directing their gaze to have time to react to changes in road conditions. Highlighting parts of the view corresponding to the periphery of a windshield would correspond to objects only a few meters in front of a vehicle. For a fast moving vehicle, augmentations at the periphery of a windshield would give drivers little time to react to them before they leave the driver's field of view. For these reasons, it may not be necessary to build an AR display with a field of view that covers the entire windshield. This is fortuitous, as it is technically difficult to build a full windshield display with optical-combiner HUD technology due to constraints placed on the optics caused by available dashboard space. In addition, the field of view of the HUD display will be fixed in size as determined by the optical design (our prototype's FOV is 20 degrees). These restrictions on the field of view imply that most AR applications should concentrate on the forward view for displaying augmentations. For creating situation awareness of entities beyond the field of view, other strategies may need to be employed. Using low-fidelity prototyping methods as seen in Figure 3, we can design solutions that take this restricted field of view into account. In Figure 10, a cautionary callout for making the driver aware of

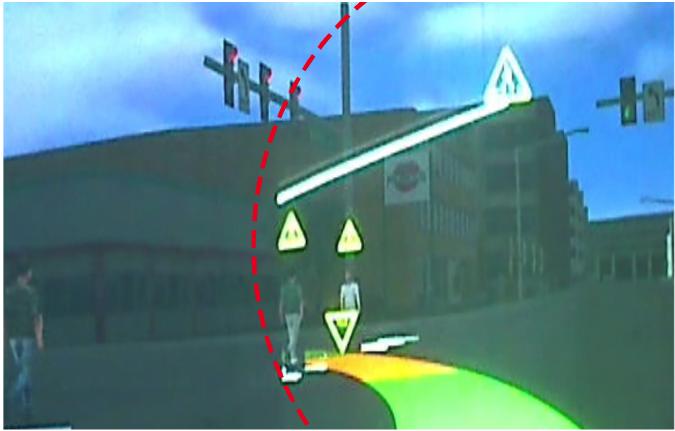


Figure 10: The pedestrian on the extreme left is out of the HUD's field of view. The field of view boundary is denoted by a dashed red curve (added to diagram for clarity). Note: Augmented graphics are on a see-through display while city and pedestrians are from a driving simulator projected on a screen in front of the display. The warning sign call-out with the white line helps to direct the driver's attention to the pedestrian out of the field of view.

pedestrians at intersection crossings was created to direct the gaze of drivers at pedestrians which may exist outside of their Ufov.

## 5 UNDERSTANDING DISTRACTION

AR can potentially add to visual and cognitive distraction. If an AR application offers interactivity with the user (e.g., with gesture-based input), it contributes to manual distraction as well [30]. Any kind of cognitive distraction directly influences visual distraction. Research has shown that even in the absence of manual distraction when using a hands-free phone, driver interference effects are predicted more on the cognitive component of conversations while driving[32]. This can be explained by examination of the human attention system.

### 5.1 Attention system

Inattention during driving accounts for 78% of crashes and 65% of near crashes [36]. The human visual system has an attention system where the gaze is first attracted to regions of high saliency. In visual design, searching by color has been shown to have the fastest time compared to size, brightness, or geometric shape [6]. The sensitivity of the cues are primed by the current task a person is doing, known as *selective visual attention* [9]. Once the eye has fixated on the visual region, the higher visual acuity regions process this visual information to infer greater contextual information, such as object recognition, followed by cognitive processing to create situation awareness that can incorporate the temporal behavior of all objects in the scene. Secondary tasks that are not related to the immediate driving environment such as phone conversations can cause the attention system to suppress relevant cues important for the primary task of driving. Similarly, putting augmented reality elements related to secondary task information such as calendar items or music album covers can dangerously redirect the attention system away from important driving cues.

Instead, augmented reality can be used to provide markings that can adapt to changing contextual situations to better regulate traffic behavior. In unaugmented driving, civil engineers already do this with lane markings, and road signs, which have proven to be very useful for indicating static aspects of traffic. Augmented reality can be used to provide dynamic markings, that can adapt to changing contextual situations. In Figure 10, a virtual crosswalk

can be placed on a roadway for a pedestrian crossing a road with the pattern moving in the direction of the pedestrian's walk, increasing the saliency of the region as well as projecting into the future the intended pathway of the pedestrian.

One important element of an attention system is the suppression of elements surrounding the focus of attention[9]. The introduction of physical, digital billboards with brighter self-illumination and animation has been shown to draw the attention of drivers more than regular billboards[12]. If the visuals of the advertisements are more salient than surrounding cues, there is the danger that relevant driving hazards could be suppressed. *Inattentional blindness* [26] can occur where the driver stays fixated on some visual items, ignoring unexpected, but important changes elsewhere in the visual field. On the other hand, increasing saliency can be used beneficially to highlight potential hazards on the road like pedestrians and potholes. Using cues such as motion or strong colors of sufficient contrast to the background scenery will not only make things more distinguishable, they offer an important way of informing the driver that these elements are synthetic versus real.

Care must be taken in choosing the cues used to highlight hazards. Schall et al.[38] found that static cues for hazards actually had longer reaction times than using no cues. However, more dynamic cues did reduce reaction times compared to the static cue conditions in their studies. Schall et al.[37] further explored the use of AR cues for elderly drivers and found they improved detection of hazardous objects of low visibility, while not interfering with the detection of nonhazardous objects. This shows potential promise of AR technology being used to compensate for deficiencies in vision as people age.

Further usability studies would be needed to design the proper balance between increasing desirable saliency and avoiding unwanted distraction. Perhaps using visual highlights that fade quickly over time can avoid inattentional blindness. Experiments employing eye-tracking measurement technology[39] can help designers evaluate the eye-gaze behavior of their AR visual artifacts. We utilized this technology when designing a left turn driving aid that helps driver estimate the oncoming vehicle speed by projecting a path in front of the car corresponding to the 3 second projection of the car (Figure 14). One concern regarding the design of the projected path was that drivers would fixate on the projected path, distracting them from observing the road scenery. Eye gaze behavior was tested on different types of visual stylings of the projected path. Though initial qualitative feedback indicated that a chevron style design was more pleasing to users, eye gaze recordings also indicated that users tended to fixate a little more on the chevron styled path with a greater number of rapid eye movements. Therefore, a less eye catching but effective solid red path was chosen. In a pilot study to test the effectiveness of the projected red path (Figure 11), we found that users fixated on the movement of oncoming vehicles and rarely fixated on the projected path. In post-study questionnaires, participants also indicated that the projected path was clearly visible, indicating that they were able to use their peripheral vision to notice the driving aid, but their primary attention visual attention was spent on other objects in their view.

### 5.2 Cognitive Dissonance

Cognitive dissonance describes the discomfort one feels when maintaining two different conflicting beliefs. This phenomena affects the visual perception of physical environments[1]. In the context of AR, if the graphic elements are not registered correctly, or if they exist in two different coordinate systems (2-D image plane vs. 3-D real world), the brain may seek to resolve between these two spaces and the driver may become distracted or even misinterpret the visual scene. Popular AR applications currently overlay 2-D image labels on video images of the world. However, when transitioning to a see-thru display in a car, one cannot focus simul-



Figure 11: Sample snapshot of eye gaze measurement for left turn driving aid. The blue circle indicates where the participant's eyes are focused. Participants did not fixate on the projected path.

taneously on elements of the real world and their corresponding 2-D annotations because they may exist at different focal depths. These effects may become even more disconcerting when stationary 2-D text is placed against a moving 3-D background. Studies [34] have shown that layering labels at different depths rather than in the same visual plane can improve visual search times due to improved clarity and legibility. In Sridhar and Ng-Thow-Hing[41], an AR application places address labels directly in the same 3-D space as the buildings they describe (see Figure 12). Labels move with the buildings so that position and context stay consistent with the changing scenery.



Figure 12: Address labels are displayed in the same 3-D space as the physical road.

## 6 UNDERSTANDING HUMAN BEHAVIOR

So far, the first two perspectives dealt with the experience of the driver through relatively short timeframes when AR systems are activated in specific driver contexts. There have been studies examining the behavioral changes that can be induced through long-term usage of these driving aids. In de Winter[8], continuous haptic guidance in the gas pedal or steering wheel, while reducing mental workload, may inhibit retention of robust driving skills. The author concludes that supplementary information should not be provided continuously, but only as needed. In Norman[31], a strong case is made for the need to continually involve the driver in interactions

with the vehicle, so that a collaborative partnership is established between the driver and the technology of the car. This creates a greater sense of trust and feeling of control when interacting with the car's sensors and driving control systems. Lee and See[25] explains how a breakdown of trust between technology and people can result in misuse and disuse of technology automation. In the case of AR-based driving aids, drivers will either turn off or ignore the AR-based information (disuse) or incorrectly rely on this information, with potential disastrous outcomes if a critical assumption of the technology's accuracy is violated.

### 6.1 Collaboration between human and machine



Figure 13: The safety grid identifies location of vehicles surrounding a car.

AR can serve as one important part of a multi-modal interface that can facilitate the collaboration between a driver and the technology capabilities of the automobile. Lee and See[25] describe how the development of trust can be influenced by the visual display of the interface to a technology, with trust increasing with increasing realism. This may lead to an advantage of AR-based over indirect map-based displays for navigation because drivers can unambiguously see elements in their real view annotated directly with the information the aid provides. Visuals can also be used to manipulate visceral reactions through color, motion or metaphor that can manipulate the level of trust of the driver towards the AR technology in the car.

Visual displays often are the intermediary for interpreting data that would otherwise not be easily interpretable by a driver. For example, sensors in the car that can detect vehicles in the surrounding environment can use AR aids to show the neighboring vehicles in a visual manner easy to understand. In Figure 13, we designed a visual aid to be placed in the sky on a plane parallel to the ground. Detected cars in the car's immediate environment can have corresponding grid squares highlighted. The transparency of the grid square indicates proximity to the vehicle with fully opaque squares implying the vehicle is directly adjacent to the driver. As grid squares become more opaque or transparent, the driver gets a sense if cars are approaching or moving away.

### 6.2 Risk Compensation

Risk compensation describes the phenomena when the presence of extra safety information causes drivers to engage in riskier behavior because they may feel more confident about their surroundings [31]. When multiple vehicles have advanced AR systems, the proper estimation of situation awareness may be inhibited if drivers deliberately change their behavior in response to this information. When we left the safety grid driving aid continuously on (Figure 13), we noticed that test drivers began engaging in more lane changes. One solution to prevent this behavior may be to only display the grid when imminent danger is present.

Another method to counteract this effect is actually making things appear more dangerous than they really are, to help prevent complacency[31]. We designed a left turn aid that helps driver estimate the oncoming vehicle speed by projecting a path in front of the



Figure 14: Oncoming cars show their projected path three seconds in the future to help drivers gauge their speed for making left turns.

car corresponding to the 3 second projection of the car (Figure 14). The rationale is that as a driver makes a left turn at an intersection, the oncoming car increases its occupied area on the road using the projection in a manner proportional to its speed and danger level.

### 6.3 Situation Awareness

A driver's behavior is directly a result of a decision-making process. *Situation awareness* describes how operators (in this case drivers) maintain state and future state information of elements in their environment. Proper situation awareness is critical for maintaining safe driving behavior. In Endsly[13], situation awareness is viewed from an individual's perspective and comprises three levels: I) Perception of elements in the environment, II) Comprehension of their meaning and III) Projection of future system states. In Section 5.1, we can understand how AR's manipulation of saliency can help achieve level I. However, AR can also assist in the cognitive tasks to achieve level II and III, by allowing the computer to infer this information.

As driving is a spatiotemporal task, one way to do this is to develop systems that interpret numerical meter results directly into their spatiotemporal representations using carefully designed visuals. In Tönnis et al.[42], an AR braking bar indicates the distance in front of the car where it will eventually stop if the brake is fully depressed. There is no need to estimate this from the velocity shown in the speedometer.

The braking bar exclusively uses the state information from the driver's car. AR can also enhance the situation awareness of other road vehicles and pedestrians surrounding the driver. In our left turn driving aid (Figure 14), we automatically draw the projected path of the other oncoming cars several seconds in the future based on their current velocity. By articulating what is going on in the external world with other cars, we are augmenting a driver's ability to understand information in the world and help the driver's own decision process rather than directly instructing the driver what to do.

We have found that the best strategy for building driving aids for achieving better situation awareness is to focus on specific problems where driving accidents are prevalent. For the left turn aid, we were motivated by the fact that 22 percent of pre-crash events occur when a vehicle is making a left turn [5]. In contrast, the safety grid (Figure 13) was not designed with a specific situation in mind. As a result, in initial versions of the design, the driver was unsure how to use the safety grid or was tempted to leave it on all the time, as a dangerous substitute for self-gazing to build situation awareness. We are subsequently redesigning the aid to focus on specific highway problems involving neighboring cars.

## 7 CONCLUSIONS

In this paper, we have described three important perspectives centered on the driver: understanding driver perception, driver distraction and driver behavior. We believe that these considerations are

just as important as the technical components needed to implement effective AR solutions in the car. Indeed, the design process needs to consider these human characteristics when conceiving solutions for augmented reality in the vehicle. We also describe a design process that can be used to help create and evaluate driving solutions employing augmented reality. We intend to continue to pursue this process for the current and future driving aids we will design.

In the two different technological efforts toward driver automation and driver enhancements, AR has the potential to contribute to both endeavors. In driver enhancement, providing better situation awareness and saliency to driving hazards or other important elements on the road can be aided with AR. For autonomous driving, AR can serve an important role of communicating to the driver and building trust in the car's decisions, confirming it perceives objects and the rationale for its decisions[31]. For example, AR can show where a planned lane change will occur and what triggers in the environment (e.g., a slow car) instigated the action. This allows the driver to understand the car's state and reduce the anxiety of not knowing what actions a car make take next. In addition, AR can be used to convey the degree of uncertainty a car has about perceived elements on the road to allow the driver to decide if manual intervention is necessary. If this is not done, a polished rendering of an AR element may falsely lead the driver to build an incorrect assumption of situation awareness.

Several major questions not addressed here must be considered. How do multiple driving aids interact together if used at the same time? If people suddenly were to stop using AR after prolonged use, how will their subsequent driving behavior be affected in a non-AR equipped car? Will driving aids improve a driver's performance in a non-AR car? Will native driving skills deteriorate as drivers become overly dependent on AR-based information for maintaining situation awareness? The successful application of AR for automobiles will require well-motivated solutions that carefully consider the driver's mental capabilities, dynamic conditions encountered while driving, and technology to implement these solutions.

## ACKNOWLEDGEMENTS

The authors wish to thank Tom Zamojdo and Chris Grabowski for their stimulating conversations, sage advice, and assistance in building our HUD prototypes.

## REFERENCES

- [1] E. Balceris and D. Dunning. Cognitive dissonance and the perception of natural environments. *Psychological Science*, pages 917–921, 2007.
- [2] O. Bimber and R. Raskar. *Spatial Augmented Reality: Merging Real and Virtual Worlds*. A K Peters, 2005.
- [3] BMW. Sbt e60 - head-up display. Technical report, BMW AG - TIS, 2005.
- [4] B. Buxton. *Sketching User Experiences*. Morgan Kaufmann, 2007.
- [5] E. Choi. Crash factors in intersection-related crashes: An on-scene perspective. Technical report, NHTSA, 2010.
- [6] R. Christ. *Research for evaluating visual display codes: An emphasis on colour coding*, pages 209–228. John Wiley & Sons, 1984.
- [7] R. Curedale. *design methods 1: 200 ways to apply design thinking*. Design Community College Inc., 2012.
- [8] J. de Winter. Preparing drivers for dangerous situations: A critical reflection on continuous shared control. *Systems, Man, and Cybernetics (SMC)*, pages 1050–1056, 2011.
- [9] R. Desimone and J. Duncan. Neural mechanisms of selective visual attention. *Annual Reviews Neuroscience*, 18:193–222, 1995.
- [10] T. Dobbert. *Matchmoving: The Invisible Art of Camera Tracking*. Sybex, 2nd edition, 2013.
- [11] D. Drascic and P. Milgram. Perceptual issues in augmented reality. In *Proc. SPIE: Stereoscopic Displays and Virtual Reality Systems III*, volume 2653, pages 123–134, 1996.

- [12] T. Dukic, C. Ahstrom, C. Patten, C. Kettwich, and K. Kircher. Effects of electronic billboards on driver distraction. *Traffic Injury Prevention*, 2012.
- [13] M. Endsly. Towards a theory of situation awareness in dynamic systems. *Human Factors*, 37(1):32–64, 1995.
- [14] M. Fairchild. *Color Appearance Models*. Addison, Wesley, & Longman, 1998.
- [15] P. Fröhlich, R. Schatz, P. Leitner, M. Baldauf, and S. Mantler. Augmenting the driver’s view with realtime safety-related information. In *Proceedings of the 1st Augmented Human International Conference*, pages 1–11, 2010.
- [16] D. Gavrila and S. Munder. Multi-cue pedestrian detection and tracking from a moving vehicle. *International Journal of Computer Vision*, 73(1):41–59, 2007.
- [17] A. Gellatly, C. Hansen, M. Highstrom, and J. Weiss. Journey: General motors’ move to incorporate contextual design into its next generation of automotive hmi designs. In *Proceedings of the Second International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI 2010)*, pages 156–161, 2010.
- [18] K. Goode, K. Ball, M. Sloane, D. Roenker, D. Roth, R. Myers, and C. Owsley. Useful field of view and other neurocognitive indicators of crash risk in older adults. *Journal of Clinical Psychology in Medical Settings*, 5(4):425–440, 1988.
- [19] GoPro. Gopro hero3 camera. <http://gopro.com>, 2012.
- [20] D. Harkin, W. Cartwright, and M. Black. Decomposing the map: using head-up display for vehicle navigation. In *Proceedings of the 22nd International Cartographic Conference (ICC 2005)*, 2005.
- [21] B. Jonson. Design ideation: the conceptual sketch in the digital age. *Design Studies*, 26(6):613–624, 2005.
- [22] S. Kim and A. Dey. Simulated augmented reality windshield display as a cognitive mapping aid for elder driver navigation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 133–142, 2009.
- [23] E. Kruijff, J. Swan II, and S. Feiner. Perceptual issues in augmented reality revisited. In *Proceedings of the 9th IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 3–12, 2010.
- [24] J. Langheim, A. Buchanan, U. Lages, and M. Wahl. Carsense-new environment sensing for advanced driver assistance systems. In *Proceedings of the IEEE Intelligent Vehicle Symposium*, pages 89–94, 2001.
- [25] J. D. Lee and K. A. See. Trust in automation: designing for appropriate reliance. *Human Factors*, 46(1):50–80, 2004.
- [26] A. Mack. Inattentional blindness looking without seeing. *Current Directions in Psychological Science*, pages 180–184, 2003.
- [27] J. Markoff. Google cars drive themselves, in traffic. *New York Times*, 2010.
- [28] Z. Medenica, A. L. Kun, T. Paek, and O. Palinko. Augmented reality vs. street views: a driving simulator study comparing two emerging navigation aids. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*, pages 265–274, 2011.
- [29] W. Narzt, G. Pomberger, A. Ferscha, D. Kolb, R. Müller, J. Wieghardt, H. Hörtner, and C. Lindinger. Augmented reality navigation systems. *Universal Access in the Information Society*, 4(3):177–187, 2006.
- [30] NHTSA. Distraction. <http://www.nhtsa.gov/Research/>
- [31] D. Norman. *The Design of Future Things*. Basic Books, 2007.
- [32] L. Nunes and M. A. Recarte. Cognitive demands of hands-free-phone conversation while driving. *Transportation Research Part F: Traffic Psychology and Behavior*, 5(2):133–144, 2002.
- [33] K. S. Park, I. H. Cho, G. B. Hong, T. J. Nam, J. Y. Park, S. I. Cho, and I. H. Joo. *Disposition of Information Entities and Adequate Level of Information Presentation in an In-Car Augmented Reality Navigation System*, volume 4558, pages 1098–1108. Springer Berlin Heidelberg, 2007.
- [34] S. Peterson, M. Axholt, and S. Ellis. Objective and subjective assessment of stereoscopically separated labels in augmented reality. *Computers & Graphics*, pages 23–33, 2009.
- [35] T. Poitschke, M. Ablassmeier, G. Rigoll, S. Bardins, S. Kohlbecher, and E. Schneider. Contact-analog information representation in an automotive head-up display. In *Proceedings of the 2008 symposium on Eye tracking research & applications*, pages 119–122, 2008.
- [36] P. Salmon, N. Stanton, and K. Young. Situation awareness on the road: review, theoretical and methodological issues, and future directions. *Theoretical Issues in Ergonomics Science*, 13(4):472–492, 2012.
- [37] M. Schall Jr., M. Rusch, J. Lee, J. Dawson, G. Thomas, N. Aksan, and M. Rizzo. Augmented reality cues and elderly driver hazard perception. *Human Factors*, 55(3):643–658, 2013.
- [38] M. Schall Jr., M. Rusch, J. Lee, S. Vecera, and M. Rizzo. Attraction without distraction: Effects of augmented reality cues on driver hazard perception. *Journal of Vision*, 10(7):236, 2010.
- [39] S. I. (SMI). Smi eye tracking glasses. <http://www.smivision.com>, 2012.
- [40] M. Sotelo and J. Barriga. Blind spot detection using vision for automotive applications. *J. Zhejiang University Science A*, 9(10):1369–1372, 2008.
- [41] S. Sridhar and V. Ng-Thow-Hing. Generation of virtual display surfaces for in-vehicle contextual augmented reality. In *Proceedings of the 11th IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 317–318, 2012.
- [42] M. Tönnis, C. Lange, and G. Klinker. Visual longitudinal and lateral driving assistance in the head-up display of cars. In *In Proceedings of the 6th International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 91–94, 2007.
- [43] M. Tönnis and D. A. Plecher. Presentation Principles in Augmented Reality - Classification and Categorization Guidelines. techreport, Technische Universität München, 2011.
- [44] M. Tönnis, C. Sandor, C. Lange, and H. Bubb. Experimental evaluation of an augmented reality visualization for directing a car driver’s attention. In *Proceedings of the 4th IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR ’05)*, pages 56–59, 2005.
- [45] J. Wood and R. Troutbeck. Elderly drivers and simulated visual impairment. *Optometry & Vision Science*, 72(2), 1995.
- [46] W. Wu, F. Blaicher, J. Yang, T. Seder, and D. Cui. A prototype of landmark-based car navigation using a full-windshield head-up display system. In *Proceedings of the 2009 workshop on Ambient media computing*, pages 21–28, 2009.