

Effective Control of a Car Driver’s Attention for Visual and Acoustic Guidance towards the Direction of Imminent Dangers

Marcus Tönnis*, Gudrun Klinker†

Technische Universität München, Fakultät für Informatik
Boltzmannstraße 3, 85748 Garching bei München, Germany

ABSTRACT

In cars, Augmented Reality is becoming an interesting means to enhance active safety in the driving task. Guiding a driver’s attention to an imminent danger somewhere around the car is a potential application. In a research project with the automotive industry, we are exploring different approaches towards alerting drivers to such dangers. First results were presented last year. We have extended two of these approaches. One uses AR to visualize the source of danger in the driver’s frame of reference while the other one presents information in a bird’s eye schematic map. Our extensions were the incorporation of a real Head-up Display, improved visual perception and acoustic support. Both schemes were evaluated both with and without 3D encoded sound. This paper reports on a user test in which 24 participants provided objective and subjective measurements. The results indicate that the AR-based three-dimensional presentation scheme with and without sound support systematically outperforms the bird’s eye schematic map.

CR Categories: H.1.2 [Models and Principles]: User/Machine Systems—Human Information Processing; H.5.2 [Information Interfaces and Presentation]: User Interfaces—Ergonomics;

Keywords: usability studies and experiments, multimodal (input and) output, (display and) view management, acceptance of MR/AR technology

1 INTRODUCTION

Head-up Display (HUD) technology is becoming increasingly integrated into modern cars. At the moment screen-fixed icons are displayed in the HUD, indicating for example the current speed or some navigational information. Such approaches can reduce the time drivers need to move their eyes away from the road in order to glance at instruments to monitor driving relevant information. Here the HUD passively assists the driver in his driving task.

Upcoming in-car head and eye tracking systems enable Augmented Reality-based visualizations. Virtual objects can become integrated into the environmental world around the car. Essential information can be displayed in the driver’s frame of reference directly in the area, where the task of driving is supervised – the windshield. This concept fits Geiser’s [11] classification of tasks in a car, where the primary task of driving should be mapped to the windshield area [24]. Therefore such visualizations should explicitly be used to actively assist the driver.

Advanced driver assistance systems (ADAS) aim at supporting drivers in their driving task [8]. As driver inattention is a primary cause for up to 78 % of crashes and 65 % of near-crashes [14], driver support systems should help reduce driver inattention. To this end,

an ADAS system needs to monitor and track the car’s environment, exploiting the increasing availability of sensors to detect imminent dangerous situations of traffic and other obstacles in the car’s near distance. Such sensor data must be provided via suitable output channels to catch a driver’s attention and to guide him in that direction. At ISMAR 2005 we reported on such a system [23] and on our first experiments with it. This paper analyzes findings of the work and extends the approach towards guiding drivers without distracting them from the driving task.

A particular problem of imminent danger is the fact that it does not necessarily occur in the driver’s field of view – e.g., from behind. How can Augmented Reality be used in this context? A direct indication of the dangerous location is not possible. Instead, drivers must realize that they should turn their heads towards the corresponding direction. The question arises, whether 3D Augmented Reality technology can outperform classic 2D assistance systems, that for instance can show a map like overview and thus can show the location of the danger.

We investigated, improved and evaluated two visualizations, resulting in one using an animated bird’s eye perspective direction indicator and the other one using an animated 3D arrow. The new bird’s eye perspective scheme presents a car and a 2D arrow pointing in the direction of the imminent danger (Figure 4). It is located in a position in front of the driver’s in-car position and is placed on the HUD. The extended 3D arrow is attached to a pole that appears to be fixed to the car’s front bumper when it is presented in the HUD (Figure 3).

A further extension concerns the use of sound for support in directional guidance. As tracking systems also enable 3D encoded sound, we furthermore generated a warning sound that came out of the direction of the dangerous situation.

An experiment was set up and was executed to determine whether the 3D presentation scheme can outperform the bird’s eye scheme by simultaneously reducing workload. We also conducted the experiment with and without providing 3D encoded sound, to determine whether multiple output channels enhance or deteriorate directional perception and workload. The experiment [15] was conducted with 24 participants, none of which had participated in previous studies on the topic of attentional guidance. Objective as well as subjective measurements were collected and analyzed.

This paper reports on general questions of fast and efficient perception and discusses ambiguities in the perception of sound. The experiment is illustrated and we discuss the results, showing recent advances of our enhanced setup.

2 RELATED WORK

Although egocentric visualization aides have proven superior to exocentric schemes for local guidance tasks in other scenarios, previous studies [23] on guidance of a car driver’s attention resulted in almost better results for well known 2D presentations. The experiment was conducted in a driving simulator which did not have a real Head-up Display. Instead, HUD-based augmentations were shown on the same flat projection wall as the landscape of the simulated

*e-mail: toennis@in.tum.de

†e-mail: klinker@in.tum.de

environment. In case of the perspective rendered 3D arrow, its planar projection and its basic shape, a cone and a cylinder, made it difficult for drivers to perceive the orientation of the arrow. Furthermore, sometimes the experiment revealed a mental translation of the floating arrow from its original location 3 m in front backwards to the position of the driver's head. Except for its virtual appearance about 3 m in front of the driver's head, it had no further fixation on the car, complicating the estimation of the exact position of the arrow. Therefore, it was very difficult for drivers to obtain a good mental picture of the place of the imminent danger towards which the arrow was pointing. As a result, drivers looked too far backwards and generated higher errors in determining the correct direction than with the bird's eye map.

Navigational tasks are related to guidance of attention. Various research on such tasks and wayfinding in virtual environments has been conducted. The work of Chittaro and Burigat [7] concerns the design of the visual part of presentation schemes. They examined aides for enhanced navigation in virtual environments by comparing 3D arrow schemes with 2D schemes. Experiments in a flight simulator resulted in the finding that a majority of the participants found their destination easiest by use of the 3D navigational arrow. The participants also preferred this kind of scheme. The visualization consisted of an arrow defined by a cone and a cylinder. A sphere at the back end of the arrow defined the mounting point for rotations. Since more than one arrow could be seen simultaneously, the sphere specified the origin that was rotationally invariant.

Smith and Hart [22] evaluated various presentation schemes for wayfinding tasks in virtual environments, measuring their cognitive loads. Among other schemes, a 2D graphical compass and a 3D graphical plan were evaluated. The 3D graphical plan was presented as a long ribbon floating some meters above the ground, and going from the start to the destination. The user had to walk underneath the ribbon to get to the destination. From all evaluated schemes, the 2D graphical compass and the 3D graphical plan required the highest mental load. Smith and Hart did not yet evaluate a 3D compass.

Wang [25] also compared schemes for navigational tasks. He describes a dependency in exocentric systems between the viewpoint's distance to the own avatar and global and local control. The larger the distance is, the better global awareness is, but similarly local guidance performance deteriorates. To guarantee good control of the close environment, short distances for presentation schemes are recommended.

Ergonomic aspects of ADAS systems and especially collision warning systems have been evaluated. Experiments on such assistance systems cover visual, as well as acoustic and multi-modal driver feedback. Campbell et al. [6] compared various symbols for ADAS systems, which were intended to warn about frontal and lateral collisions, as well as about lane departures. They found that 2D presentations from the side were optimal for frontal warnings and bird's eye presentations were best suited for lateral collisions. Their study only evaluated 2D presentation schemes where all ranked almost equally.

Sayer [20] requested study participants to develop dead-angle warnings for lane-change assistance systems. Afterwards all participants had to rank those symbols. In every case, bird's eye perspectives ranked best.

Wickens and Seppelt [21] published a survey on 18 experiments that rate visual and auditory signals for various tasks in an in-car environment. The summary states that generally acoustic warnings should be used, except when the optical warning appears in the foveal field of view of the driver.

Lerner [19] offers extensive guidelines for the design of optical, acoustical and tactile warnings. In general, he recommends the use of multi level warning systems: At least one pre-warning, followed by a more immersive one. On the highest level of warnings, redun-

dancy is important to ensure that a warning is definitely realized. This is supported by Bubb [5]. He states that feedback must be given on as many sensory channels as possible.

Gunther [13] explored ways to use sound to support wayfinding in virtual worlds. He states that object detection is supported, but that getting a global overview can be disturbed. Furthermore he points out a problem in distinguishing spatial sounds from positions in front and behind the subject.

3 PRESENTATION CONCEPT

Using his steering wheel and pedals, the driver indirectly interacts with the world outside the car. Most of the time the driver gets feedback from his driving task and from other traffic through the windshield but also gets feedback through other windows. The driver has to care for the 3D relationships between his car and other cars or obstacles. If an imminent danger occurs, for instance an upcoming crash, the driver not only has to become aware of the danger, but also of the location of dangerous situation. It is conceivable to use bird's eye presentations schemes that are already used in systems providing drivers distance information when they park their cars. They show the car and colored bars to indicate shortening distances. Yet, in contrast to a parking scenario, dangerous situations can occur at any speed. The mental load of an exocentric presentation may be too high and time consuming for drivers going at high speeds. They mentally have to transform their view into the exocentric bird's eye perspective, get an overview of spatial relationships in that perspective, transform them back and then determine the specified location with respect to their egocentric position.

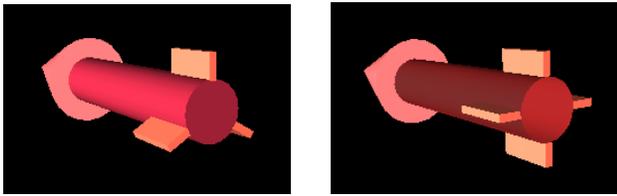
Using an egocentric AR perspective reduces the effort of transforming between several coordinate frames. Unfortunately, a dangerous situation may not be in the driver's field of view. Therefore a location indicator is not useful – the driver's attention needs to be guided. A ribbon metaphor similar to the one of Smith and Hart [22] could be used to guide a driver. The ribbon starts at the position of the driver, first heads straight forward, so that the driver can see it and then turns towards the imminent danger. The eyes just have to follow the course of the ribbon and reach the destination. Such a ribbon cannot be presented above the driver, because it could not reach the dangerous location on the ground. Placing the ribbon like a carpet on the ground would cover a certain area of the environment and therefore is not suitable in ADAS systems. Furthermore it could be misinterpreted as a navigation information. A compass metaphor appears to be more promising for the use in cars like it was used in the previous evaluation. We generated a detailed visual presentation concept with two visualization schemes generated from the compass metaphor. Incorporating the indication of the direction in both visualization schemes maintains comparability for the evaluation and correct analysis. Otherwise it would not exactly be possible to attribute certain influences to the effects on objectively measured variables.

Furthermore we extended the concept by 3D encoded sound to improve importance of warnings and for further directional guidance.

3.1 Visual Presentation Concept

Generally intuitivity and simplicity are important for the design of visualization schemes. Visualization schemes must not be overloaded with details, but use an easy design and they must be easy to understand [19].

Both visualization schemes are animated to appear urgent. Animated icons are perceived better and therefore are more suitable for immediate alerts because they attract the direct attention of the driver [16].



(a) A 3D arrow with three fins - each fin is oriented in another direction
 (b) A 3D arrow with four fins - two opposite fins are aligned along the same spatial axis

Figure 1:

3.1.1 3D Arrow Warning Scheme

The first visualization scheme extends a 3D arrow floating in front of the driver's head, pointing into the direction of the danger. The direction of a simple 3D arrow made out of a cone and a cylinder is difficult to interpret in its direction, when it is pointing directly forward or backward. Attaching fins at the rear-side of the arrow dissolves this ambiguity. After creating mock-ups with three and four equally distributed fins, we selected the three-finned approach. For the three finned approach (see figure 1(a)), two fins (which are at least visible in every rotational state of the 3D arrow) always aim into different directions from the arrow pole and at least one is not aligned along the line of sight. Therefore perceptibility of the direction is increased. For the four finned approach (see figure 1(b)), each two opposite fins are aligned along the same axis and only convey the spatial alignment of this axis. Figure 1 shows both types of 3D arrows and illustrates the alignment of the fins.

A further problem in identifying the exact direction is, that the location of the floating arrow in front of the driver at the distance of the front bumper is unknown and therefore some people translate the arrow to their own head's position. Figure 2 shows this in a sketch. As a result, the imminent danger is searched for in a too far backward orientation. To avoid any ambiguity about the exact location, where the 3D arrow is, we attached a vertical pole to the arrow. From the driver's perspective this pole seems to be mounted at the front of the car. The focal distance of the virtual image in the HUD was placed at the same distance.

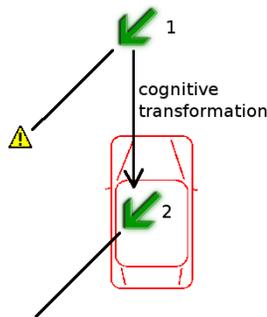


Figure 2: Sketch for the cognitive transformation of the 3D arrow. Arrow 1 shows the real position and arrow 2 the transformed position

AR visualizations must not cover large areas in the windshield, because obstacles in upcoming traffic could be occluded. A minimal volume for indicating various directions of the arrow is reached by rotating the arrow around its midpoint rather than its back end. The arrow is mounted accordingly on the pole. Figure 3 shows the 3D arrow from the driver's point of view.

For enhanced attentional capture and increased spatial percep-

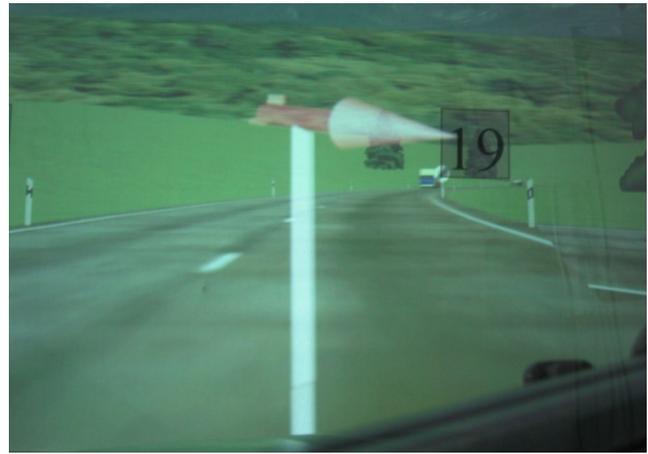


Figure 3: The 3D arrow pointing to the back right

tion of the arrow's direction, a short animated rotation has been implemented. The 3D arrow appears in front of the driver, fixed to the pole. At this moment it is pointing in a direction that is 45° off from the direction of imminent danger, oriented forwards. The arrow immediately starts to rotate horizontally by 45° into the direction of the imminent danger. This animation takes 0.25 s . The rotation around the horizontal axis has been chosen, because it is the most familiar rotation for ground traffic. All other rotations, would have to include vertical rotations and therefore require more spatial interpretation.

3.1.2 Bird's Eye Perspective Warning Scheme

A bird's eye perspective is capable of displaying locations close to a car. To suppress the effect of incomparability of different visualization approaches, as a location indication would manifest w.r.t. to a directional indication, we decided to use the compass metaphor as well. If the first visualization scheme indicates a direction and the bird's eye perspective shows a location, the results of the experiment could not accurately get analyzed. The results can not explicitly show, if a certain benefit of a visualization scheme was generated by the dependency to a frame of reference or by the difference in describing the target. Thus, our second presentation scheme uses a bird's eye sketch of the car with a 2D arrow pointing into the direction of the upcoming danger.

The arrow has been placed in front of the car silhouette. The arrow slightly overlaps with the car to avoid the impression that the arrow might be pointing towards the front of the car – as might be conceivable for indicating, for instance, a malfunctioning front light. For the same reason the silhouette is shown in green, avoiding the impression that existing damages on the car are being indicated. Furthermore, the arrow has been placed in front of the driver's position rather than along the central axis. In principle, the bird's eye presentation scheme is almost a projected visualization of the 3D arrow scheme, seen from above. All these design decisions attribute to the generation of a spatial relationship between the arrow and the car, without indexing the car itself. Figure 4 illustrates the appearance of the bird's eye perspective from the driver's point of view in the HUD.

While the warning scheme is active, it is blinking at a rate of four Hertz, following the recommendation of Lerner [18] to get the driver's attention.

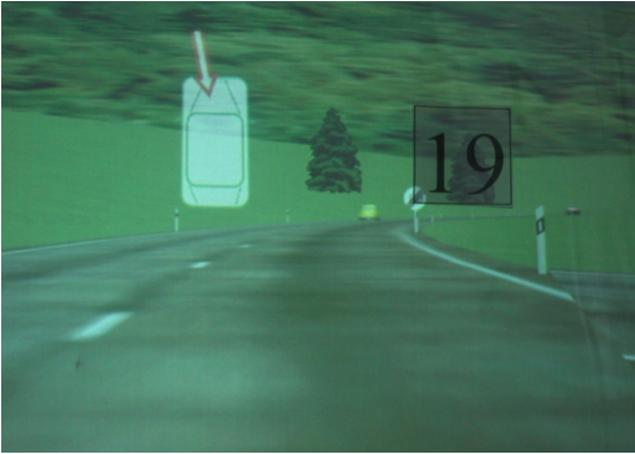


Figure 4: The bird's eye perspective pointing behind the car to the right

3.2 Guidance Using the Acoustic Channel

The auditory channel offers a means for providing additional directionally encoded information. A multi-modal extension with 3D encoded sound supports acoustic guidance. It has been well established that is useful to announce critical information on as many sensory channels as possible [5, 19]. For warnings which require fast reactions, abstract sounds are preferred to the use of speech indicators [16]. 3D encoded warning sounds are well suited to control the attention of car drivers and to give them further hints regarding the position and direction of imminent danger.

It is not clear whether a warning signal is a location identifier or a directional indicator. If sound specified a location, the presentation concept would not be affected, but the experiment could not be executed because two independent variables would influence each other. It would remain unclear, whether the indication of direction of the visual presentation scheme was supported by sound or whether the location indicating property of sound increased the perception. To keep the experiment objectively measurable, this ambiguity of sound has to be discussed. In general, sounds always are hybrids of location and direction indication. On the one hand, a sound originates from a certain direction and therefore indicates this specific direction. For familiar sounds, distance becomes indirectly specified through loudness and frequency of tones. Furthermore sounds are reflected by objects in the environment. Such secondary sounds enable 3D location estimation for the listener. Although a warning sound is not an explicit direction indicator, its use as a directional warning scheme can be considered valid when secondary sound reflections are suppressed.

The auditory channel provides two mechanisms for spatial localization [1]. First, for frequencies below 1500 Hertz localization works through the perception of the differential delay between the left and the right ear. The brain analyzes phase differences of the waves reaching the ear and infers the corresponding direction of the sound. The wave length of the sound has to be longer than the distance of both ears, which causes the restriction to frequencies below 1500 Hz. The second locating mechanism uses perception of intensity for frequencies above 1500 Hz. Shadowing and attenuation effects of the head generate different levels of the waves in the ears, from which the brain estimates the corresponding direction. Only wave lengths that are shorter than the width of the head are attenuated. Frequencies near 1500 Hz are difficult to locate, because both location mechanisms reach physical limits.

Another phenomenon of sound is that certain frequency levels

are attributed to certain directions [17, 1]. Frequencies around 300 Hz and 3000 Hz seem to be in front, while frequencies around 1000 Hz seem to come from the rear.

Therefore we designed a warning sound composed of eight harmonic frequencies, 600, 1000, 1400, 1800, 2200, 2400, 2800, 3200 Hz [12]. The frequencies above 2200 Hz were intensified to satisfy the requirement for high urgency. The use of frequencies below and above 1500 Hz enables both locating mechanisms. This mixture also compensates for unconscious association with certain directions.

The sound has been modeled with a periodicity of 4 Hz [18], to be synchronized with the blinking of the visual warning scheme.

4 EXPERIMENT

We have implemented the presentation schemes and have set up a test environment in a driving simulator. This setup enables measuring how driving performance of car drivers is affected by alerts due to external dangers, e.g. potential collision with a car coming from behind.

4.1 Physical Setup

The stationary driving simulator is shown in figure 6. It consists of a BMW E30 convertible, equipped with a steering wheel, a gas and a breaking pedal to control the motion path of the car during the simulation. The corner speakers of a Dolby 5.1 sound system have been installed in a modified dashboard, respectively on top of the backrest of the backseats.

A rural landscape and simulated traffic scenes are shown on a planar projection wall at a focal distance of 3 meters in front of the car driver. The simulation projects the scenery from a fixed general driver's head position in the car and covers a 50-degree visual field of view [23].

The car is surrounded by 20 evenly spaced, letter-sized sheets of paper. The sheets are labeled with numbers between 1 and 20. They are arranged in random order around the car in height of the driver's head.

4.2 Construction of the Head-Up Display

One major problem is the technical presentation of a 3D shape in an in-car environment, where no additional hardware like Head-mounted Displays (HMD) or polarized glasses are allowed [9, 10].

Approaches in driving simulators which projected HUD visualizations onto the same projection wall as the landscape scenery complicate spatial perception of these shapes' alignment. The 2D image of the 3D object is projected onto the projection wall. The participants have to mentally reconstruct the 3D shape from the 2D image. This works acceptably well for objects that are associated with other objects, such as another car, whose external shape is familiar in the context of roads and the overall environment. Additional objects like a 3D arrow without a fixed attachment to the car are more difficult to interpret since they should not be associated with the physical projection wall.

Spatial decoupling of HUD presentations and virtual scenery is the way out of this issue. Placing virtual objects, where no associate real object such as the projection wall exists, increases volumetric perception of the virtual 3D objects. To achieve this we used the same approach as car manufacturers do with their HUDs [2]. Figure 5 illustrates the general setup.

Our approach requires a larger field of view of the HUD. We placed a second projection wall besides the right side of the car, where it is almost covered by the right A-pillar from the driver's point of view (see figure 6(a)). A flat glass plane in front of the

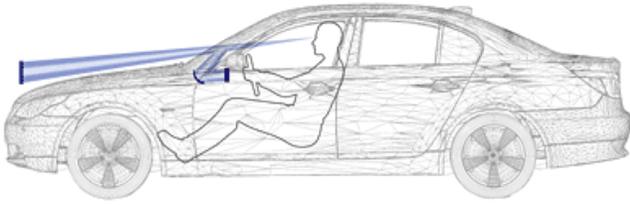


Figure 5: The principle of the HUD of BMW, a bright display behind the dashboard generates the image that is mirrored into the windshield and appears in distance of the front bumper

driver acts as a combiner (see figure 6(b)) and mirrors the augmented objects into the real environment. Added distances between the second smaller projection wall, the combiner and the driver's eyes result in a focal image plane at about 2 meters in front of the driver, right at the distance of the front bumper (see figure 6(c)).

This depth is exactly the depth, at which the 3D arrow can be mounted to the car's front. Thus a imaginable alignment to another object is produced. On the one hand, 3D perception is enhanced by a size relationship to the well known car and on the other hand, the superimposed 3D arrow is decoupled in depth from the projection wall in 3 m distance.

The second projector for the HUD screen was appropriately calibrated to be combined with the projector of the driving simulation that shows the superimposed scenery from the general driver's head position in the car.

4.3 Participants

The experiment was executed with 24 participants none of which had participated in a similar study. 14 participants were male, ten female. The average age was 39.2, the standard deviation 13.3, and the largest group was between the age of 46 and 55 years. The youngest participant was 20 year old, the oldest 60 years. The subjects were paid 15 Euro for their participation.

4.4 Test Procedure

After filling out a demographic questionnaire, the participants had a test drive of about 10 minutes to familiarize themselves with handling a stationary driving simulator. Afterwards they were presented each of the warning schemes pointing forward, sideward and backward, so that they got a first impression of the presentation schemes. In the next step the objective measurements were performed. We used a Within-Subjects design [4] for the experiment. For each of the four presentation schemes, the subjects drove for about 15 minutes on a rural landscape course. The experimenter activated each of the 20 different directional indicators in permuted order. The subjects then had to react to the presented alert by turning their heads in the perceived direction of imminent danger and reading out loud the number of the closest paper sheet. The time from activating an alert to task completion by the subject was recorded. Finally the participants had to fill out a subjective questionnaire.

4.5 Independent Variables

The selected presentation scheme was the independent variable in the Within-Subjects design. All 24 participants were exposed to all four presentation schemes: bird's eye perspective and 3D arrow, both without and with 3D sound. Groups of six participants each started with a different presentation scheme. The follow up schemes also were permuted for each group accordingly. In each

case, 20 alert situations were generated for each scheme. The sequence of alert positions was permuted in each session.

4.6 Objective Dependent Variables

During the tests, *speed* and *lane deviation* of the car were recorded in a protocol file at a frequency of 30 Hz. Deduced speed change and the lane deviation are indicators how much drivers are distracted by the presentation schemes of the alerts, since their interpretation adds a secondary cognitive load on drivers – in addition to the primary driving task. If presentation schemes capture too much of the driver's attention, they are not suitable for use in cars. Those schemes could lead to further imminent dangers such as, for example, the situation of a driver leaving his own lane and going offroad.

Furthermore the protocol recorded discrete tokens of the points in time when a presentation scheme was activated and when the identification of a paper sheet was accomplished. The time differences describe the *response time*. This variable is the most important objective measurement, because in case of a dangerous situation, the driver's reaction time is more important for avoiding a time critical accident than any mistake in keeping the lane or keeping speed.

The experimenter wrote down the spoken-out number of each paper sheet in a written protocol for each participant, together with the correct number, i.e. the number that was used for generating the alert. From this data, the *average mistake* and the *error quotient* were computed. The average mistake specifies by how many paper sheets an answer was off in comparison to the correct paper sheet. The error quotient counts the number of wrong answers relative to all exposures. Both variables are indicators for the spatial accuracy of the presentation schemes.

4.7 Subjective Questions

The final questionnaire covered the subjective impression, the participants had during and about the experiment. The subjects had to answer the following four questions, grading them according to the scheme in German schools from 1 (best) to 6 (worst):

- How much did you like a certain variant?
- How well could you deal with it?
- How quickly could you nominate the corresponding paper sheet?
- How exactly could you identify the paper sheet?

These four self-appraisals had to be filled out for all four presentation schemes.

The session ended with a final question aimed at ranking how much participants were pleased with the four presentation schemes with respect to each other.

5 RESULTS AND DISCUSSION

The following section reports on computed results. The tables list mean and standard deviation of the measured values. Furthermore they show the computed significances as α -values for the aggregated results. The significances are presented first, followed by mean and standard deviation. The text furthermore shows minimum and maximum values and discusses the results.



Figure 6: Construction of the Head-Up Display

5.1 Objective Measurements

Objective measurements were gathered from 24 subjects for four times 20 dangerous situations, resulting in 1920 records, 480 for each visualization scheme. Significances were computed for all measurements, given as α -values. The significance niveau is $\alpha = 0.83\% = 0.0083$. We used a one way analysis of variance with Bonferroni-correction [3]. The Bonferroni-correction adjusts the significance niveau to the possible amount of combination of pairs. The four variants of the experiment allow six combinations, therefore dividing a general significance niveau of 0.05 by a value of 6.

5.1.1 Response Time

The response time indicates how long the subject needed from the first appearance of a presentation scheme until they nominated a paper sheet. The aggregated results for the reaction time can be seen in table 1. The 3D arrow presentation scheme with sound achieved the best results with a mean of 2.66 s (std.dev 0.66 s), followed by the same variant without sound (mean 2.93 s, std.dev 0.94 s). The next best reactions with a mean of 3.32 s (std.dev 0.93 s) was the bird's eye presentation scheme with sound. The slowest variant was the bird's eye presentation scheme without sound (mean 3.77 s, std.dev 1.46 s). The results are significant, except for the pairs that only differ in having sound or not.

α -Values		no sound		with sound	
		3D ar-row	Bird's eye	3D ar-row	Bird's eye
no sound	3D arrow		0.001	0.049	0.006
	Bird's eye			0.000	0.068
with sound	3D arrow				0.000
	Bird's eye				
Mean [s]		2.93	3.77	2.66	3.32
Std.dev [s]		0.94	1.46	0.66	0.93

Table 1: Response Time

The longest reaction time was 7.89 s in the bird's eye perspective presentation scheme. The shortest reaction time was 1.65 s under the 3D arrow with sound.

For the response time, the 3D arrow scheme is superior to the bird's eye scheme and each time the variant with sound is superior to that without sound. Thus sound serves its purpose to increase

attentional capture. Especially the amount of very slow reaction times is reduced dramatically by use of acoustic signals.

The 3D perception time of the arrow is reduced by decoupling it from the projection wall. Thus a monoscopic Augmented Reality presentation like the 3D arrow, that is placed in free space in front of the driver is easier to perceive in its 3D shape than one that is projected onto a projection wall. This fact must be taken into account for further driving simulator experiments.

5.1.2 Average Mistake

The average mistake indicates how many paper sheets a participant's answer was off from the correct paper sheet. The content of table 2 indicates: The average mistake for the 3D arrow is minimally lower than the average error for the 3D arrow scheme with sound (0.71 to 0.73, with std.dev 0.34 without sound to 0.32 with sound). The third best result (0.83, std.dev 0.33) occurred by the bird's eye scheme with sound, while the worst results were computed from the pure bird's eye presentation (0.94, std.dev 0.68). None of the results is significant.

α -Values		no sound		with sound	
		3D ar-row	Bird's eye	3D ar-row	Bird's eye
no sound	3D arrow		0.113	0.709	0.129
	Bird's eye			0.184	0.183
with sound	3D arrow				0.396
	Bird's eye				
Mean []		0.71	0.94	0.73	0.83
Std.dev []		0.34	0.68	0.32	0.33

Table 2: Average Mistake

The largest average mistake occurred in the pure bird's eye scheme with a value of 3.8. The lowest average mistake was 0.2 measured in the 3D arrow experiment without sound.

Transformation into the bird's eye perspective and back into the egocentric perspective seems to increase angular errors. Future ADAS systems should thus strive for keeping subjects within their own frame of reference. Furthermore the use of a pole to fix the 3D arrow on a certain position prohibits the potential mental translation of the driver's position in the car. The results show that sound generally does not increase accuracy, but in case of the bird's eye scheme, sound can reduce large mistakes in interpreting directions.

5.1.3 Error Quotient

The error quotient shows how often the participants failed to name the indicated paper sheet. Again, the 3D arrow without sound performed best with an error quotient of 0.48 (std.dev 0.13), closely followed by the same variant with sound (error quotient 0.52, std.dev 0.14). Both bird's eye presentation schemes (with and without sound) achieved almost similar values of 0.58 for the error quotient (std.dev 0.14).

There are no significances in the paired error quotient comparison, as table 3 shows.

α -Values		no sound		with sound	
		3D ar-row	Bird's eye	3D ar-row	Bird's eye
no sound	3D arrow		0.019	0.223	0.013
	Bird's eye			0.124	0.810
with sound	3D arrow				0.098
	Bird's eye				
Mean []		0.48	0.58	0.52	0.58
Std.dev []		0.13	0.14	0.14	0.14

Table 3: Error Quotient

The maximum error quotient with a value of 0.85 occurred in the pure bird's eye presentation, while the minimum value of 0.20 again was generated from the pure 3D arrow without sound. The pure 3D arrow presentation scheme is the only one with a majority of correctly identified paper sheets. Here again, the use of the same frame of reference for the indication of the direction proves to be more useful than the exocentric perspective. Furthermore 3D perception is highly increased by placing the warning icons, especially the 3D arrow, onto a separate spatially decoupled display.

5.1.4 Mean Lane Departure

Lane departures can occur in both lateral directions. To illustrate this, we use negative values for departures to the left and positive values for departures to the right side. These values indicate a departure from the perfect trajectory in the own lane.

The 3D arrow with sound has the smallest mean value (-0.02 m, std.dev 0.27 m). It is followed by the 3D arrow scheme with sound which has a mean value of -0.04 m (std.dev 0.26 m). The third best presentation scheme is the bird's eye perspective with sound (mean value -0.07 m, std.dev 0.22 m). The worst presentation scheme is the bird's eye presentation without sound with a average lane deviation of -0.13 m (std.dev 0.21 m). The results are not significant, as table 4 illustrates.

α -Values		no sound		with sound	
		3D ar-row	Bird's eye	3D ar-row	Bird's eye
no sound	3D arrow		0.023	0.767	0.297
	Bird's eye			0.031	0.094
with sound	3D arrow				0.156
	Bird's eye				
Mean [m]		-0.04	-0.13	-0.02	-0.07
Std.dev [m]		0.26	0.21	0.27	0.22

Table 4: Mean Lane Departure

It is remarkable, that 19 from 24 participants (76%) using the bird's eye presentation without sound left the perfect trajectory to the left, departing towards the opposite lane. All other variants had more evenly distributed values.

As in the previous experiment, lane deviation is lower for the 3D arrow. Executing the task of getting spatial orientation is easier in the driver's personal frame of reference.

5.1.5 Standard Deviation of Lane Departure

To get a proportion for the lane deviation, we computed the standard deviation of the lane departure values. The base values were taken out of the interval, a presentation scheme was activated until the paper sheet's number was spoken out.

The pure 3D arrow presentation gave the best results for the standard deviation of the lane departure with a value of 0.17 meters (std.dev 0.06). The 3D arrow scheme with sound is minimally worse with a value of 0.18 m (std.dev 0.17 m). Both bird's eye presentations almost got the same lane departure values, with a standard deviation of the lane departure of 0.20 m, but the standard deviation of the variant with sound is greater (std.dev 0.21 m), as the meter value of the variant without sound (std.dev 0.08 m). Table 5 shows that the α -Values are too high for significances.

α -Values		no sound		with sound	
		3D ar-row	Bird's eye	3D ar-row	Bird's eye
no sound	3D arrow		0.081	0.819	0.403
	Bird's eye			0.488	0.991
with sound	3D arrow				0.141
	Bird's eye				
Mean [m]		0.17	0.20	0.18	0.20
Std.dev [m]		0.06	0.08	0.17	0.21

Table 5: Standard Deviation of Lane Departure

The largest variation was 1.14 m on the test drive with the bird's eye presentation scheme with sound. The straightest road course with a variation of only 0.04 m was produced under the 3D arrow with sound.

As for the total lane departure, the standard deviation during detection of the paper sheet is minimized by use of the 3D arrow.

5.1.6 Speed Variation

Speed variations indicate problems in dealing with a presentation scheme. The least speed variations were generated by the sound supported 3D arrow presentation scheme with a value of -0.61 km/h (std.dev 1.62 km/h). The variant without sound gave a value of -0.69 km/h (std.dev 1.76 km/h). The bird's eye perspective with sound reaches the third position with a value of -0.81 km/h but has a high standard deviation of 3.04 km/h. The worst results are generated by the pure bird's eye perspective (-1.11 km/h, std.dev 1.90 km/h). No value is significant here, as table 6 shows.

5.2 Subjective Measurements

The personal opinions were given according to the German school grade system, ranging from 1 (best) to 6 (worst).

For the question *How much did you like a certain variant?*, the 3D arrow without sound reached the first position with a mark of

α -Values		no sound		with sound	
		3D ar-row	Bird's eye	3D ar-row	Bird's eye
no sound	3D arrow		0.210	0.622	0.799
	Bird's eye			0.092	0.538
with sound	3D arrow				0.629
	Bird's eye				
Mean [km/h]		-0.69	-1.11	-0.61	-0.81
Std.dev [km/h]		1.76	1.90	1.62	3.04

Table 6: Speed Variation

2.71 and a standard deviation of 1.20. It is followed by the bird's eye perspective with a rank of 2.92 (std.dev 1.14). The third place is reached by the 3D arrow with sound (3.04, std.dev 1.15). The bird's eye perspective with sound came to the fourth place with a mark of 3.17 (std.dev 1.43). All results for this question are not significant (see table 7) because they exceed the significance niveau of 0.0083.

Most participants did not like the sound, they said the sound was annoying. This was due to the experimental nature of the study, where the participants were exposed to the sounds very often. Since its activation will only happen in critical situations, such annoyance should not be taken into account.

α -Values		no sound		with sound	
		3D ar-row	Bird's eye	3D ar-row	Bird's eye
no sound	3D arrow		0.564	0.188	0.293
	Bird's eye			0.781	0.388
with sound	3D arrow				0.722
	Bird's eye				
Mean []		2.71	2.92	3.04	3.17
Std.dev []		1.20	1.14	1.15	1.43

Table 7: Question: *How much did you like a certain variant?*

The question *How well could you deal with it?* brought the 3D arrow with sound to place number one with a mark of 2.67, standard deviation 1.27. It is followed by the same 3D arrow, but without sound (2.87, std.dev 0.99). With a larger offset, the bird's eye perspective follows with a mark of 3.29 (std.dev 1.23). The fourth place goes to the bird's eye perspective without sound (3.33, std.dev 1.17). No result is significant (see table 8).

Here the presentation schemes that use the 3D arrow are better than the bird's eye schemes. In all cases the variant with sound performed better than the corresponding one without sound.

α -Values		no sound		with sound	
		3D ar-row	Bird's eye	3D ar-row	Bird's eye
no sound	3D arrow		0.141	0.364	0.233
	Bird's eye			0.088	0.857
with sound	3D arrow				0.044
	Bird's eye				
Mean []		2.87	3.33	2.67	3.29
Std.dev []		0.99	1.17	1.27	1.23

Table 8: Question: *How well could you deal with it?*

The question *How quickly could you nominate the corresponding paper sheet?* resulted in the 3D arrow with sound on rank position number one with a mark of 2.50 (std.dev 1.18). The second best result got the 3D arrow without sound (2.79, 1.14). Place three in the ranking of this question is taken by the bird's eye presentation scheme with sound and a mark of 3.21 (std.dev 1.14). It is followed by the same scheme without sound and a mark of 3.63 with a std.dev of 1.01. Only one comparison is significant: The pair given by the bird's eye presentation and the 3D arrow presentation with sound, see table 9 for details.

The 3D arrow presentation ranked better than the bird's eye ones and every time the variants with sound were ranked better than the ones without.

α -Values		no sound		with sound	
		3D ar-row	Bird's eye	3D ar-row	Bird's eye
no sound	3D arrow		0.010	0.032	0.195
	Bird's eye			0.002	0.015
with sound	3D arrow				0.026
	Bird's eye				
Mean []		2.79	3.63	2.50	3.21
Std.dev []		1.14	1.01	1.18	1.14

Table 9: Question: *How quickly could you nominate the corresponding paper sheet?*

The fourth question from the questionnaire was: *How exactly could you identify the paper sheet?* The participants thought, that they were best in identifying the paper sheets by getting the direction from the 3D arrow with sound. Here they ranked it in an average to 2.70 (std.dev 0.95). The second best-ranked presentation scheme was the 3D arrow without sound (2.83, std.dev 0.87). This is followed by the bird's eye perspective with an average mark of 3.17 (std.dev 0.87). α -Values never are significant - see table 10.

α -Values		no sound		with sound	
		3D ar-row	Bird's eye	3D ar-row	Bird's eye
no sound	3D arrow		0.027	0.185	0.224
	Bird's eye			0.014	0.056
with sound	3D arrow				0.094
	Bird's eye				
Mean []		2.83	3.42	2.70	3.17
Std.dev []		0.87	0.88	0.95	0.87

Table 10: Question: *How exactly could you identify the paper sheet?*

Finally, the participants had to rank the presentation concepts directly. They had to specify an order, listing which presentation scheme they thought was best, second best, second worst and worst. The best placement was given to the 3D arrow presentation with sound (2.13, std.dev 1.15), followed by the pure 3D arrow averaged to 2.33 (std.dev 1.09). On position three, the bird's eye presentation scheme with sound was placed with 2.67 (std.dev 0.96), while the bird's eye presentation without sound earned the fourth rank with 2.88 (std.dev 1.19). Table 11 shows that the α -values are above the critical threshold of 0,0083, therefore the results are not significant.

α -Values		no sound		with sound	
		3D ar-row	Bird's eye	3D ar-row	Bird's eye
no sound	3D arrow		0.163	0.504	0.421
	Bird's eye			0.122	0.468
with sound	3D arrow				0.120
	Bird's eye				
Mean []		2.33	2.88	2.13	2.67
Std.dev []		1.09	1.19	1.15	0.96

Table 11: Question to rank the presentation schemes w.r.t. pleasure

Summarizing all subjective measurements, the subjective order of the participants preferences states as follows:

1. 3D arrow with sound
2. pure 3D arrow
3. bird's eye presentation with sound
4. pure bird's eye presentation

The 3D arrow presentation is generally preferred, similarly the variant with sound.

5.3 Comparison to Results of the Previous Experiment

Even though comparison between the previous experiment and the newly reported experiment is only valid to a limited degree, this section tries to relate both experiments one to another. Limitations to a perfect comparison are different participants with different age in the experiments and the use of a constructed HUD for presentation of the optical presentation schemes in the second experiment. In addition, the previous experiment did not take the difference of indicating a direction or a location into account. Only the driving simulator, the landscape and the test procedure were kept unchanged. These factors are dependent to the results of the experiment and generate dependent aberrations in the resulting values.

Five values of the objective measurements can be compared to the previous experiment. The most crucial value to be measured for alerting systems is the response time, which people need to react on a certain trigger. All other measurable values, lane deviation, speed variation, error quotient and average mistake are less relevant and only indicate the driver's distraction, respectively the accuracy. Comparing the reaction time between both experiments, the 3D arrow presentation with sound of the current experiment get's on first place, followed by the same variant without sound. The next best is the bird's eye presentation scheme with sound, followed by the variant without sound, both from the current experiment. The almost similar variant from the previous experiment came to the second last place with almost the same reaction time. The worst reaction time was measured for the 3D arrow of the previous experiment. With respect to lane deviation, the 3D arrow of the current experiment was also best. For all other measurements, speed variation, error quotient and average mistake, the bird's eye presentation of the previous experiment held its superiority. But all three results are placed in the interval spun between the two types of the previous experiment, shrinking down the distance between the different variants.

There were four subjective questions in the previous experiment, asking, which presentation scheme was *liked most*, could *deal with best*, could *fastest name* and *exactly name the paper sheet*. In all four cases, the previous experiment's bird's eye view resulted best with marks around a value of 2 (1 is best, 6 is worst), followed by both current experiments 3D arrow presentation schemes with

and without sound with marks around 2.8. The next best are the current experiment's bird's eye presentation schemes with marks in the area of 3.2. The worst results were attributed to the 3D arrow of the previous experiment, pointing to a mark of 4.

This comparison should really be read with caution, because the previous experiment did not clearly deal with the topic of pure directional indication and also brought in a direct location marker (bird's eye scheme), with which to deal is easier, but impossible to realize for the case of a egocentric presentation, when the target is not in the field of view of the driver.

6 CONCLUSION AND FUTURE WORK

In this paper, we have reported on our continuing studies pertaining to the guidance of driver attention. We have extended the approach to use AR in cars to assist in the driving task. This effort embarks in an important new line of research, analyzing the use of AR-based presentation metaphors in situations where users have to divide their attention between several spatially-based tasks. While managing the car and spatial relationships in the environment, the driver might not observe an imminent danger. Guiding a driver's attention towards the direction of such danger while remaining in the driver's frame of reference is superior to a presentation in another frame of reference. Primarily detection times are significantly reduced, while other factors such as speed and lane deviation are better than in an exocentric warning scheme. In time critical situations, the reaction time is the most important factor for safety, so this is the most relevant achievement. Reducing the reaction time by concurrently reducing driver distraction is the main goal to achieve when testing spatial alerting systems. Furthermore, multi-modal warning presentations which use sound in combination with a visual scheme prove to be a further source of improvement towards providing fast attentional capture and towards minimizing complete misinterpretations of visual presentations.

Our findings are based on two major extensions. First, the use of a monoscopic HUD to present visual information that is registered in 3D. Second, increased perceptibility of the position and orientation of the 3D arrow.

The HUD which has been integrated into the driving simulator presents a monoscopic projection of the 3D arrow. Surely a stereoscopic HUD would further increase 3D perception significantly, but up to now there seems to be no acceptable way of introducing such a setup in a real car. Our approach instead, using the set up of available HUDs, has an extended field of view.

Another problem was identified as being the design of the 3D arrow. Generally it is recommended to use a shape as basic as possible, but spatial perception requires some more details. Extending the 3D arrow with fins as well as a short animation enhanced the perception of the orientation. Cognitive translations can be avoided by attaching the arrow to a pole that connects the 3D arrow to the car. In our experiment, the pole was very thick. Future evaluations should decrease the thickness of the pole, so that less of the driver's view is blocked.

In contrast to the previous experiment, we now can recommend the 3D arrow for use as a warning indicator specifying the direction of an imminent danger. Reaction times are significantly shorter and other objective measurements are not worse.

It remains to be examined to what extent a tethered viewpoint, showing a 3D car and the relative direction from a lifted backward position could be used as a presentation scheme for informing a driver about certain dangers.

Further experiments will evaluate the effects of mounting the warning presentation scheme glance-fixed, so that it always remains near the driver's viewing direction. Drivers could then follow the 3D arrow until they look directly at the imminent danger. The 3D arrow will then point away from the driver. Our hypothesis is, that

the results will provide further indications towards the superior usability of the 3D arrow. The bird's eye presentation scheme deteriorates because the driver also has to take the rotation of his own head into account to determine the direction.

ACKNOWLEDGEMENTS

We would like to thank Uli Koch for his enduring work in this project. Uli Koch created various mock-ups of alternatives, executed and analyzed the experiments. Furthermore we would like to thank Dr. Martin Gründl, a psychologist from the Universität Regensburg, for many valuable discussions with respect to location-direction ambiguity of 3D sounds. Finally we would like to thank Christian Lange, a colleague from the neighbouring chair for ergonomics of the mechanical engineering faculty, who provided much valuable insights regarding ergonomic issues in cars and the analysis of the collected data.

REFERENCES

- [1] Jens Blauert. *Communication Acoustics*. Springer, 2005.
- [2] BMW. Head-Up Display Homepage. <http://www.bmw.co.za/products/hud/default.asp>.
- [3] C. E. Bonferroni. Teoria statistica delle classi e calcolo delle probabilità, 1936. Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze 8, pages 3-62.
- [4] Jürgen Bortz. *Statistik für Human- und Sozialwissenschaftler*. Springer Verlag, 2004.
- [5] H. Bubb. Systemergonomische Gestaltung. In *Ergonomie (Schmidtke, H., Editor)*, Hanser Verlag, München, 1993.
- [6] J.L. Campbell, D.H. Hofmeister, R. J. Kiefer, D. J. Selke, P. Green, and J.B. Richman. Comprehension testing of active safety symbols. In *SAE International*, 2004.
- [7] Luca Chittaro and Stefano Burigat. 3D Location-pointing as a Navigation Aid in Virtual Environments. In *AVI '04: Proceedings of the working conference on Advanced Visual Interfaces*, pages 267-274, New York, NY, USA, 2004. ACM Press.
- [8] C. Christen. Intelligent Safety Systems - Sustainability Report in Economics and Mobility, 2002. Adam Opel AG.
- [9] DIN Deutsches Institut für Normung. *Road vehicles - Ergonomic aspects of transport information and control systems - Dialogue management principles and compliance procedures*, Mar. 2003. Ref.No.: DIN EN ISO 15005:2003.
- [10] DIN Deutsches Institut für Normung. *Road vehicles - Ergonomic aspects of transport information and control systems - Specifications and compliance procedures for in-vehicle visual presentation*, Oct. 2003. Ref.No.: DIN EN ISO 15008:2003-10.
- [11] G. Geiser. *Man Machine Interaction in Vehicles*. ATZ 87, 1985. pages 74 - 77.
- [12] P. Green, W. Levison, G. Paelke, and C. Serafin. Suggested Human Factors Design Guidelines for Driver Information Systems. UMTRI-93-21, FHWA-RD-94-087, 1994.
- [13] Ryan T. C. Gunther. The use of 3d sound as a navigational aid in virtual environments. Master's thesis, University of Waterloo, Ontario, Canada, 1997.
- [14] Sheila G. Klauer, Vicki L. Neale, Thomas A. Dingus, David Ramsey, and Jeremy Sudweeks. Driver inattention: a contributing factor to crashes and near-crashes. In *Proc. of HFES 49th Annual Meeting*, pages 1922-1926, 2005.
- [15] Uli Koch. Vergleichende Untersuchung zu verschiedenen augmentierten Anzeige Konzepten für Warngaben von Fahrerassistenzsystemen. Experimentelle Semesterarbeit, TU München, 2006.
- [16] D. Lee, C. Carney, S. Casey, and J. Campbell. In-Vehicle Display Icons and Other Information Elements Volume I: Guidelines. FHWA-RD-03-065, 2004.
- [17] Alexander Lerch. Unterlagen zum Tutorium in den ersten Semestern des Tonmeisterstudiums, Chap. 5.5: Richtungshören bei einer Schallquelle. <http://www.ton.hdk-berlin.de/tutorium/html/>.
- [18] N. Lerner, D. Dekker, G. Steinberg, and R. Huey. Inappropriate Alarm Rates and Driver Annoyance. DTNH22-91-C-07004, 1996.
- [19] N. D. Lerner. Preliminary Human Factors Guidelines for Crash Avoidance Warning Devices. NHTSA Project No. DTNH22-91-C-07004, 1996.
- [20] Tina Brunetti Sayer. The Development and Evaluation of Icons for Side Obstacle warning systems. In *Proceedings of 46th Human Factors and Ergonomics Society (HFES)*, 2002.
- [21] B. Seppelt and C. D. Wickens. Interference With Driving or In-Vehicle Task Information: The Effects of Auditory Versus Visual Delivery. AHFD-02-18/GM-02-3, 2002.
- [22] Shamus Smith and Jonathan Hart. Evaluating Distributed Cognitive Resources for Wayfinding in a Desktop Virtual Environment. In *Proceedings of the 1st IEEE Symposium on 3D User Interfaces (3D UI)*, March 2006.
- [23] M. Tönnis, C. Sandor, C. Lange, G. Klinker, and H. Bubb. Experimental Evaluation of an Augmented Reality Visualization for Directing a Car Driver's Attention. In *Proceedings of ISMAR*, October 2005.
- [24] Marcus Tönnis, Verena Broy, and Gudrun Klinker. A Survey of Challenges Related to the Design of 3D User Interfaces for Car Drivers. In *Proceedings of the 1st IEEE Symposium on 3D User Interfaces (3D UI)*, March 2006.
- [25] W. Wang. *Human Navigation Performance Using 6 Degree of Freedom Dynamic Viewpoint Tethering in Virtual Environments*. PhD thesis, Univ. of Toronto, Mechanical and Industrial Engineering, 2004.