Comparing Non-Visual and Visual Guidance Methods for Narrow Field of View Augmented Reality Displays

Alexander Marquardt*, Christina Trepkowski*, Tom David Eibich, Jens Maiero, Ernst Kruijff, and Johannes Schöning

Abstract—Current augmented reality displays still have a very limited field of view compared to the human vision. In order to localize out-of-view objects, researchers have predominantly explored visual guidance approaches to visualize information in the limited (in-view) screen space. Unfortunately, visual conflicts like cluttering or occlusion of information often arise, which can lead to search performance issues and a decreased awareness about the physical environment. In this paper, we compare an innovative non-visual guidance approach based on audio-tactile cues with the state-of-the-art visual guidance technique EyeSee360 for localizing out-of-view objects in augmented reality displays with limited field of view. In our user study, we evaluate both guidance methods in terms of search performance and situation awareness. We show that although audio-tactile guidance is generally slower than the well-performing EyeSee360 in terms of search times, it is on a par regarding the hit rate. Even more so, the audio-tactile method provides a significant improvement in situation awareness compared to the visual approach.

Index Terms—Augmented Reality, view-management, guidance, audio-tactile cues, performance, situation awareness

✦

1 INTRODUCTION

Locating virtual objects in augmented reality (AR) applications can be a challenging task. A major problem of many augmented reality displays is their relatively narrow field of view (FOV) that only covers a small part of human vision. The human visual system has a binocular FOV of about 210° horizontally and 150° vertically [40]. In comparison, the popular Microsoft HoloLens AR display has a FOV of 30°H and 17.5°V. Studies have shown that when the density of information increases, narrow FOV can negatively affect search performance [13]. Furthermore, conflicting visual cues can make it difficult to process and interpret stimuli [47] and can lead to a certain degree of sensory overload as human processing capacities are limited [56].

In this context, view-management techniques are dealing with the layout and appearance of augmentations [9]. However, designing effective view-management systems for narrow FOV displays is still an open issue in research. Depending on the application at hand, view management may need to handle both in-view and out-of-view information adequately. A major problem for narrow FOV displays is typically overlapping information (e.g., labels), where augmentations occlude each other and/or the reference object in the scene [23]. In-view labelling can aggravate the problem [48] as this method tries to place additional labels inside the limited FOV that refer to out-of-view objects. This can lead to visual conflicts that can cause visibility, legibility, depth ordering, scene distortion and object relationship issues [47]. With respect to solving problems related to out-of-view targets, researchers have focused on developing different so-called guidance approaches (see [12] and [79] for an overview). We can roughly differentiate between visual (e.g., [12, 32]) and non-visual guidance methods (e.g., [20, 42, 62]). Most research is directed towards visual methods [79]. Non-visual guidance methods typically look at reducing visual overload or conflicts by minimizing the number of stimuli in the visual sensory channel. To achieve this, sensory substitution - the transfer of information to a different sensory channel - can be used [58]. Sensory substitution is commonly used to overcome the limitations of blocked sensory channels, e.g., for people with visual disabilities [59]. In the context of dense information in narrow FOV displays, sensory substitution is believed to be a fruitful direction to improve user performance [62].

The effectiveness of the most common visual guidance techniques has been compared in a number of object search tasks [12]. Results indicated that EyeSee360 was performing very well against other well-established visual guidance methods [12, 33]. However, there is a lack of understanding how well non-visual guidance compare to their visual counterparts. To address this lack, the aim of this paper is to compare the performance of visual and non-visual guidance methods in the context of head-worn displays with narrow FOV. For this work, we used the currently widely used Microsoft HoloLens (first version with a diagonal FOV of about 35°) as a reference model. In our studies, we look into aspects like search performance, accuracy, cognitive load, and situational awareness (SA) of visual and non-visual guidance methods. We compare visual and non-visual guidance methods and investigate the above-mentioned aspects in three sub-studies. In contrast to previous research that mostly considered optimal laboratory conditions, we study real-world conditions more closely by examining the methods in a simulated real-world environment instead of relying on purely abstract use cases. We assumed that non-visual feedback can have a positive effect on both usability and performance in AR systems with a narrow FOV. We also expected that visual complexity can be reduced by transferring that visual information to another perceptual channel. This approach might be particularly useful in visual complex environments, as it can potentially lead to a reduction of visual workload [75] and, in direct relation, may increase situational awareness (SA).

1.1 Contributions

Through the results of our user studies, we present the following contributions that provide new insights into the effectiveness of non-visual guidance methods in comparison to a state-of-the-art visual technique. We do so in context of guidance in narrow FOV AR displays. The complete study was performed in virtual reality (VR), simulating an AR environment (for details, see Section 4.3). While comparing visual and non-visual guidance methods, we place a strong focus on SA. SA can be a fundamental factor in AR systems and considerations should be given for the usage of AR in real-world conditions [39]. Unfortunately, SA is frequently not taken into account sufficiently when addressing guidance in AR.

- In study part 1 we compared audio-tactile guidance with EyeSee360 during a simple object collection task in terms of general task performance. We showed that audio-tactile guidance could

---

- *Alexander Marquardt, Christina Trepkowski, Tom David Eibich, Jens Maiero, and Ernst Kruijff authors are with Bonn-Rhein-Sieg University of Applied Sciences. E-mail: {alexander.marquardt, christina.trepkowski, tom.eibich, jens.maiero, ernst.kruijff}@h-brs.de.*
- *Ernst Kruijff is also with Simon Fraser University.*
- *Johannes Schöning is with University of Bremen. E-mail: schoening@uni-bremen.de.*

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx.

**The first two authors contributed equally to this work.*

compete and even slightly exceed EyeSee360 regarding the hit rate. However, search times were considerably shorter for EyeSee360.

- In study part 2 we increased the difficulty by adding visual noise and optical flow to the same task as performed in study 1. Furthermore, a small noticeability test was added to have a first indicator for SA. We showed that an increased task difficulty likely does not have an influence on search performance for both guidance methods. However, the noticeability test indicated already a notably higher SA for the audio-tactile mode.
- In study part 3 the task difficulty was increased again by adding a secondary task. The performance of the secondary task was also used to measure SA. We showed that SA was significantly higher with audio-tactile guidance while performance values of the object collection task (search times, hit rate) for both modes were not affected by the secondary task.

Summarizing, it has not been shown yet how non-visual guidance cues can compete with current visual guidance techniques. We do so by discussing performance measurements while solving an object collection task under different degrees of difficulty. Furthermore, we show how to improve SA in case audio-tactile guidance is used for the localization of out-of-view objects in AR.

2 RELATED WORK

Our studies touch upon several fields of research, namely view management, visual and non-visual guidance methods, and situational awareness in AR, which we describe below.

2.1 View Management

Designing and optimizing the layout of information in view management methods have been researched over a longer period of time [9]. Studies so far have mainly focused on label placement for size and position [6, 9], depth-placed ordering [70, 71] and the appearance of labels (e.g. foreground-background issues [29] or the legibility of text [28, 52]). While in recent times some research has been done on view management for wide FOV displays [44, 48], not many researchers have focused on narrow FOV displays yet, except e.g., [13, 76].

2.2 Narrow Field of View

Current-generation AR devices still suffer from a limited FOV. Limiting the FOV typically leads to various problems like perceptual and visuomotor performance decrements for real and virtual environments [8]. Even though most studies that focus on FOV limitations were performed on virtual reality (VR) systems, it can be assumed that insights also apply to AR applications to a certain degree. Another intensively discussed issue is the consistent underestimation of distances for head mounted displays (HMD) with limited FOV in VR scenes [92] and for AR applications [84]. Dense information spaces in narrow FOV have also been shown to affect search performance negatively [13], while a decreased FOV can lead to a significant change in visual scan pattern and head movement, which may in turn also affect search performance [18, 83]. With respect to spatial awareness, it has been shown that FOV restrictions are degrading the abilities of developing spatial knowledge and navigation [1, 90]. Finally, a restricted FOV can result in decreased search performance [3] as well as selection performance [25].

2.3 Visual guidance

With respect to visual guidance, effects like the pop-out effect [36] or attention guiding techniques [77, 89] have found some reasonable application. Less obtrusive methods like subliminal cueing [72] and saliency modulation in AR [88] have also been discussed. Furthermore, head-up displays (HUDs) are also widely used for guidance, e.g., in the aircraft sector, for basic navigation, flight information [5, 66] and pathway guidance [27]. Other common examples for guidance with visual aids are specific pointers to targets like arrows and attention tunnels [81], 3D arrows [33], radars and halos [12] or EyeSee360 [30]. Latter showed superior performance compared to five other visual guidance techniques in different scenarios regarding completion time, usability (SUS score) and workload [12].

2.4 Non-visual Guidance

Non-visual guidance can be implemented in various ways. In terms of vibro-tactile cues, they can be used to direct navigation [53, 85], for 3D selection tasks [2, 60], for supporting pose and motion guidance [7, 61], and visual search tasks [51, 54]. In [62], we reported on different audio-tactile approaches that guide the user in 3D space. The used setup was specifically designed for AR displays with narrow FOV and is inspired by the ring-based tactile guidance systems of Oliveira et al. [20]. Similar head mounted tactile setups have been explored in a two dimensional manner (e.g. Haptic Radar [15], ProximityHat [10]) or as high-resolution tactor grid [42]. Alternative haptic feedback devices exist that can provide directional feedback towards the head, e.g., by using a robot-arm attached to a HMD [91], however their applicability may be limited in AR systems.

Regarding auditory cues, research has looked at supporting visual search [65, 87] and navigation [41]. Studies showed that spatial auditory cues can improve search performance by up to around 25% [64]. Regarding visual search tasks, cross-modal effects have been researched for audio-tactile cues [45, 67] or conflicts between visual and auditory cues [46]. Sonification strategies also use auditory cues to inform or guide the user. It typically modulates sound attributes like pitch and loudness with respect to the presence of the auditory reference [22]. This metaphor can also be found in modern parking car systems, where the distance information is provided through a decreasing time interval between impulse tones [69].

2.5 Situational Awareness

Situational awareness describes the "perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future" [24]. AR technologies found a broad application in improving SA for diverse areas. The AR tool InfoSPOT [37] for example helps facility managers accessing required building information. SA is enhanced by overlaying device information on the view of the real environment. AR is also widely used in the aviation sector for pilots [26], military operations [17, 57], and driver assistance systems in cars [55, 68] to provide the user additional information, e.g., about incoming threads.

To measure SA, various techniques can be used (see [80] for an overview). SAGAT - a freeze probe technique - is one of the most common approaches to validate SA. On the other hand, measuring task-depending characteristics of the operators performance is the probably simplest way to examine the impact of SA. Performance measures are non-intrusive as they are produced through the natural flow of the task and is used to indirectly measure SA [80].

3 RESEARCH QUESTIONS

The user study reported in this paper compares guidance performance of non-visual to visual cues under three different degrees of difficulty. In each study part users used guidance cues to identify a target among distractor objects. Task difficulty (from now on referred to as "task load") can be typically modulated by adding noise or including a secondary task next to the main task [19]. During our experiment, task load is increased by adding visual noise, namely through a dynamic environment, and a secondary task. While the background was kept static and neutral in the first study part, the second and third study part were set in a vivid virtual city environment causing rich visual background noise and optical flow. In order to increase task load again in the third study part, users further had to perform a secondary task next to the guided search task. This allows us to examine the user's SA more closely [24].

These studies addressed our research questions, formulated as follows:

RQ_1: How well do non-visual guidance methods perform compared to visual guidance methods for a search task on different levels of task load (inferred by a static/dynamic environment and secondary task)?

H_1: We hypothesize that EyeSee360 will outperform audio-tactile guidance in low task load (static environment) conditions. On the other hand we expect in the high task load (dynamic environment and secondary task) conditions a higher performance for the audio-tactile method because of the reduced workload and less visual clutter compared to EyeSee360.

RQ_2: Is there an effect of guidance method on situation awareness when a secondary task is included?

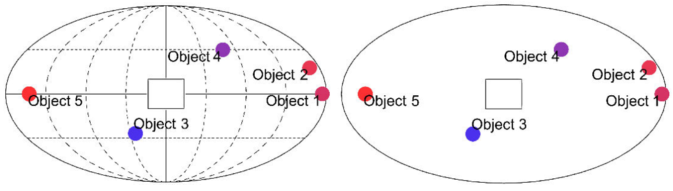
H_2: We hypothesize that the usage of EyeSee360 contributes to a lower SA compared to audio-tactile guidance. We expect this behaviour because of a higher mental workload due to a higher density of visual information compressed inside a small FOV.

4 USER STUDY

For visual guidance we used the EyeSee360 technique [30]. EyeSee360 was created for visualizing out-of-view objects in 360° around the user, depending on the user's orientation, and was improved over time in terms of reducing visual clutter and mental workload [31, 34]. Following, we will describe both methods in more detail. To provide non-visual cues, we used a modified version of our audio-tactile guidance interface reported in [62] and encoded latitude by audio and depth by vibration cues. Previously, we tested this cue combination against other non-visual audio-tactile feedback encodings and showed that it provides a superior performance regarding guidance accuracy and search time [62].

4.1 EyeSee360

The original EyeSee360 technique (see Figure 1a) maps the 3D space to a 2D ellipse with a smaller rectangle in the central point. Colored dots (called proxies) are positioned in this 2D map inside the inner rectangle to indicate target locations of objects in the 3D space inside the user's FOV, while out-of-view objects are displayed inside the ellipse but outside the rectangle (see Figure 1). The inner rectangle is sized so as not to occlude the user's focus. The horizontal line corresponds to the eye level of the user, the distance to this line indicates elevation level of the target. A proxy above this line indicates that the target is above eye level, a proxy beneath the line means that the target is located below. Distance to the vertical line indicates the longitudinal position of the target, which can be on the left or right side of the user. To illustrate the distance of the object, the proxy can take a color of a gradient from red (target is close) to blue (target is far away), as can be seen in Figure 1. This heatmap-inspired coding is intended to make the interpretation of distances as intuitive as possible. The original version of EyeSee360 included helplines in addition to the horizontal and vertical line (Figure 1a). However, we decided to use the improved variant with zero helplines only (Figure 1b) as it has been shown to cause less distraction and resulted in a better search performance compared to the variant with helplines [34].



(a) All helplines.

(b) No helplines.

Fig. 1: Out-of-view visualization with EyeSee360. **(a)** shows the initial method presented in [30]. **(b)** shows an improved variant of this method without helplines [34].

4.2 Audio-tactile guidance

Audio-tactile guidance encodes spatial information on longitude, latitude and depth to guide the user to a position in the 3D space. Here, we briefly describe these encodings, for more detail, please refer to [62], as we basically replicated the methods reported therein. In the aforementioned paper, we investigated different approaches of non-visual guidance in terms of performance, accuracy and information localization. These metaphors are partially adapted from Oliveira et al. [20]. For the purpose of this paper, we used the best-performing metaphor as reported in [62].

The user is informed about the relative position of the target on the **longitude** by the position of the vibrating tactor in the vibro-tactile setup (see Figure 2 upper, system description in Section 4.3). If the target angular position was located horizontally between two tactor positions, both motors vibrated. Motor intensity of both motors was set in relation to the angular distance of the target. This was done to achieve an interpolation effect to indicate that a target lied in between the physical motor setup, similar to the phantom effect described in [38]. Once activated, the corresponding motors were running at a frequency from about 50 hz up to 200 hz, depending on the current angular distance of the target. These values were chosen as we previously showed that this feedback was clearly perceptible without being considered as disturbing [62]. If the head is turned towards the indicated direction (Figure 2 lower), the vibration "wanders" with the head rotation until the feedback at the center of the forehead informs the user that the target is located directly in front of his view direction. In case the target angle temporarily lies above 90°or below -90°of the current head rotation, the corresponding outermost vibration motor keeps vibrating until the user rotates the head closer to the target direction.

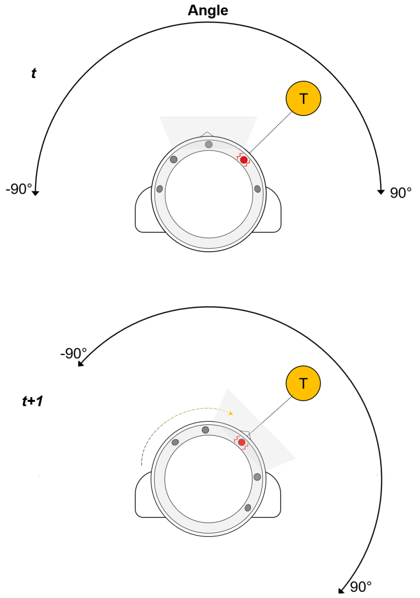


Fig. 2: Longitudinal encoding (top view). Initially (*time t*), the tactor position indicates the target direction. At time *t+1* the vibration feedback "wanders" with head rotation.

The **latitude** was provided by auditory feedback that used a modulating function with a quadratic growth, as this function has been demonstrated to be the best working one in conjunction with latitudinal encoding [20]. The modulating function adjusts the pitch and the volume of the sound source depending on the difference between the user viewing angle and the target elevation level as shown in Figure 3. If the viewing angle is far from target elevation, the auditory feedback is low for both volume and pitch, starting from about 300 Hz. As soon as the viewing angle gets closer to target elevation, pitch and volume increased to indicate the rapprochement on the latitudinal plane. Pitch and volume is the highest at about 1300 Hz if the viewing angle corresponds right to target elevation level to inform the user that the correct elevation angle is spotted. The mid-range spectrum from 300 $\sim$ 1300 Hz was chosen as the human auditory system is particularly attuned in this range and frequency discrimination works sufficiently good. Higher frequencies however can sometimes be perceived as annoying or even painful over time [14].

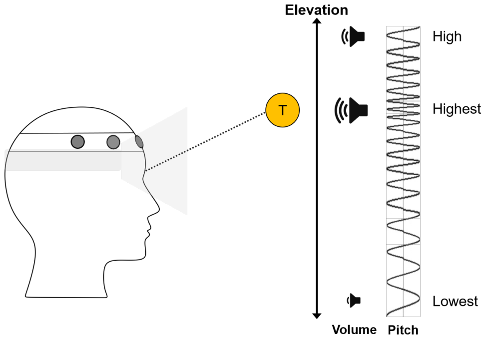


Fig. 3: Latitudinal encoding by auditory cues. Pitch and volume gets adjusted by the viewing angle of the user.

To provide information about target **depth** (distance), we used the implementation for target localization *with absolute depth feedback* [62]. This method uses the currently selected motor from longitudinal encoding (see Figure 2) and applies a variable on/off pattern for activating the vibration motors - hereafter referred to as *pulse feedback* (see Figure 4). Pulse Feedback is inspired by commonly used car parking metaphors that encode distance information through a decreasing time interval between impulse tones [69], but in a vibro-tactile manner. This makes pulse feedback easy to understand for most people since it is a commonly used real world metaphor. One pulse is described as the time the motor is turned on and off again for a specific interval. These on/off times have always the same length and are set to periods from 100ms up to 500ms. A long pulse of 500ms would indicate that the target lies very far away from the user while a very short pulse of 100ms length signals that the target is positioned right in front of the user. The maximum pulse speed of 100ms was chosen to comply with the physical restrictions of the vibration motors, e.g., overcoming motor inertia and braking time without provoking interferences [60].

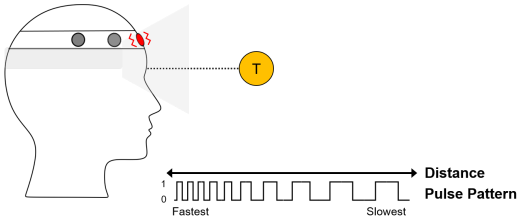
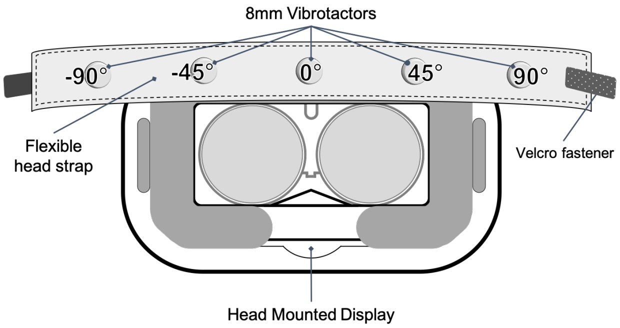


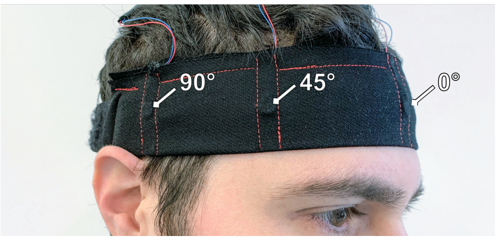
Fig. 4: Depth encoding by pulse feedback. Pulse duration gets adjusted by target depth.

4.3 System and Implementation

In this work, we compare visual and audio-tactile guidance for AR applications. However, for the user studies we used virtual environments to ensure the same preconditions (e.g. lightning, visual and auditory noise) and to allow an overall comparability between the various study parts [74]. As such, we follow a similar approach as reported in [39]. For this, the FOV of the Microsoft HoloLens as current state-of-the-art AR headset is simulated in VR. This was achieved by placing a virtual display of about 35°(diagonal) size 3cm in front of the user's eyes. This display used a semitransparent glass-like material in order to gain the impression of using an actual AR device (compare to [77]). Virtual augmentations are only visible for the user inside that simulated AR FOV, as can be seen on Figure 6 and Figure 9a and 9b. To provide tactile cues, we created an extension that is usable in combination with various AR/VR HMDs. In contrast to our previous system [62], it consists of a headband which is made out of stretchable, comfortable-to-wear cotton instead of a solution integrated in the headset. In this headband, 5 vibrotactors are placed along the temples and the forehead in 45° intervals (see scheme in Figure 5a. We used Precision Microdrives 8mm vibration motors (2mm type, model number 308-107). These motors were placed into sewn pockets, so both sides of the motors are protected by fabric, as can be seen in Figure 5b. By this design, we avoid direct skin contact and uncomfortable pressure against the forehead while still maintaining clearly noticeable vibration feedback, even when it is worn below a HMD.



(a) Scheme of the vibro-tactile interface in combination with a HMD.



(b) Wearing the interface. Vibrotactors are hidden in the sewn pockets.

Fig. 5: Custom made head strap attached with 5 vibrotactors placed in 45°intervals. The interface can be worn below a HMD.

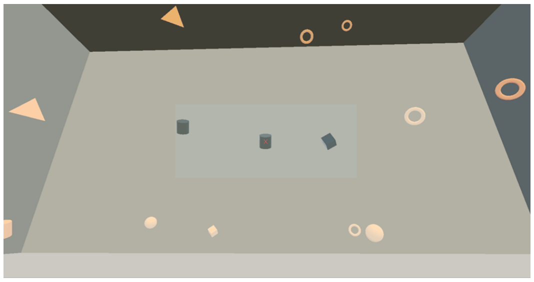
The system was implemented using Unity 2019.2. We used the HTC VIVE Pro VR headset including VIVE Pro controller as VR platform. Participants performed the study in a laboratory room in seated position on a rotating chair which is adjusted to a comfortable position beforehand. Spatial audio was enabled by the Steam Audio Spatializer plugin, using the integrated earphones of the HMD. The vibrotactors were controlled by a Raspberry Pi 3 Model B+ running a Python-based version of Open Sound Control to communicate with the Unity App.

To model the visual noise conditions (see details in the next section), for study part 2 and 3, virtual pedestrians were created with a random appearance by the UMA 2 package (Unity Multipurpose Avatar). For the car traffic, from a pool of 8 different looking cars, models were distributed in the scene. To simulate simple crowd and car traffic movement, NavMesh Agent behaviour managed a continuous movement over random predefined paths in the scene. Furthermore, ambient city sound effects are used to enhance immersion and to create additional auditory noise.

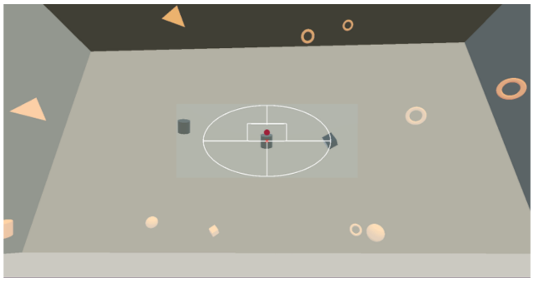
4.4 Study Design

Both guidance methods were compared in VR in three study parts to examine how they perform under different levels of task load in a fully controlled environment.

In each study part and trial the user had to identify a target among distractor objects that could not be differentiated by their appearance or position alone. All objects took a random shape of one of five primitives (sphere, cylinder, cube, pyramid, ring) with equal size. The primitives ultimately represented locations and objects within an urban environment. Therefore they were coloured in various shades of colours that are supposed to appear predominantly in urban surroundings. For this we analyzed a static image of the city scene in study part 2 and 3 and extracted four independent color clusters (see Figure 7a and 7b) by using k-means clustering. For the experiments, the primitives of the scene were randomly given one of the color of the resulting clusters (see Figure 7c).



(a) Study part 1 using audio-tactile guidance.



(b) Study part 1 using EyeSee360.

Fig. 6: Guidance methods in comparison without any visual distractors in study part 1. The center region shows the simulated AR display with the according guidance method. Out-of-view objects were visualized in semi-transparent orange color and were not visible outside the simulated AR display during the experiment.

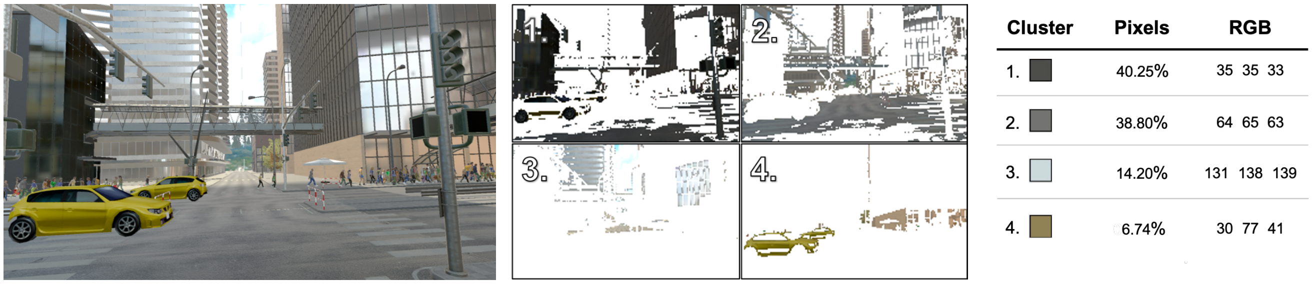
For each trial, 40 objects (1 target and 39 distractors) were distributed in the scene. To spatially distribute the targets around the user and to prevent them from overlapping each other, a virtual spherical grid is placed on the user´s position, similar to [62]. The spherical grid contains rows and columns, describing the angular distances to the user. We used three rows as elevation angles (on 0° respectively eye level, $22.5°, 45°$) and ten columns along 180° (when looking straight ahead, the user was facing 90°). Initially, the objects are all placed in the center of each used row/column combination. Afterwards, each object was given a minor random horizontal and vertical offset and set to a random distance of 15-30 meters to the user to create different depth levels. We did not include further initial target elevation angles below 0° due to physical limitations [20, 62] and to prevent that an object would be occluded by the ground level. During the studies only the distributed primitives without the spherical grid were visible to the user. This setup also ensured that items were not occluded by each other or by any objects of the environment. Although there was no city environment in study part 1, we kept the depth range the same between studies for comparability reasons. The user was sitting on a swivel chair throughout the study and was not supposed to stand up or walk. To search the objects the user rotated the head (±90° left/right and up to 45° up) or turned the body on the stool. The user was always shown a crosshair in both guidance modes in the center of the display to select the virtual target. To select a target, the user had to orient the head towards the target object and place the crosshair over it, and could then press the trigger of the controller for confirmation.

In *study part 1* there was no background noise and no secondary task. It was set in a 3D space with uniformly gray floor, walls and ceiling (see Figure 6). A light source was also included in the scene to create an impression of depth. In *study part 2 and 3* the target had to be found under conditions with background noise. Objects (distractors and target) were generated in the same way as in study part 1, but were located in a busy city environment. In the VR setting the user was standing in front of a broad street and was facing the opposite roadside. To create areas with increased optical flow at the 0° elevation level, cars were moving fast from left to right on the street and vice versa, while people walked around the pedestrian walkway in front of the user. To mimic real-world conditions, targets and distractors could not be occluded by buildings, but could be partially (and briefly be) occluded by pedestrians and cars. Horizontal optical flow between 22.5° and 45° was realized by recurrent wind gusts that transported small (visible) particles.

We added further distractors and minor optical flow for study 2 and 3 in the form of flying birds. The birds appeared in irregular intervals (every 12-17 seconds in study part 2 and 15-20 seconds in study part 3) on a path between the target elevation levels on 11.25° and 33.75°. The chosen path of the bird was depended on the current target elevation. If the target elevation was set on 0°, the bird was flying on 11.25°. If the target was on 45°, the bird was flying on 33.75°. Finally, if the target was placed on 22.5°, one of both path's was chosen randomly. We did this to ensure the user always had the possibility to notice the bird during the search task. One bird was always present at the same time, visible for about 12 seconds. It followed a sinusoidal trajectory around the user from left-to-right or right-to-left on the selected elevation (see Figure 9c). If the user selected a possible target object while a bird was already flying in the scene, the bird adapted its elevation according to the new target object. Next to creating additional optical flow to the scene, the flying birds address two issues related to SA, namely noticeability and performance in a dual task condition. We focused on general perception (noticeability) in the first half of study part 2. For this, we let the birds fly more frequent and at a closer distance to the user at about 12 meters to make them clearly recognisable. Participants started every study part either with the visual or with the audio-tactile guidance method. Afterwards, they repeated the same object collection task with the other method. To receive an impression about the general perception of the environment, we asked every user after finishing the first mode of study part 2 which movable object was noticed in the scene. Prior to the study, participants were not explicitly advised to pay attention to their environment. The general perception was achieved in case the user indicated that he noticed the bird. With regards to measuring dual task performance we also included a secondary task next to the object collection task in study part 3. Here, we let the birds fly less frequent (every 15-20 seconds) and further away at about 20 meters to make them less obvious, yet still well visible for the user. The performance of the secondary task was primarily measured by the number of correctly detected birds during the regular object collection task with each of the two guidance methods. We also measured how often and long the bird was visible in the total FOV of the HMD (not the simulated AR FOV).

4.5 Procedure

Participants were recruited via a university mailing list and received a 10 Euro voucher as a reward for participation. We employed a 2x3 within-subject design to examine the effect of factors guidance feedback (visual versus audio-tactile) and task load (no noise, noise, noise and secondary task) on search time performance and hits/errors. Study parts 1 to 3 were always completed in ascending order as difficulty increased from study part 1 to 3. It was intended users got used to the guidance feedback when first noise and secondly an additional task was added to the search task. Users had to complete 30 training trials in total, ten before performing each study part. The task during the training trials was identical to the task of the performance session, except that the user had no time limit to find the targets in order to understand how the guidance methods work. Within each study part, guidance feedback was tested block-wise: First all trials with one guidance method were completed, then all trials with the other guidance method followed: (Mode A →Mode B) or vice versa (Mode B→Mode A). Therefore, for three study parts, there are 2^3 possible feedback orders to perform the complete experiment that were balanced across participants. At the beginning of each study part a fixation point was shown to ensure the correct starting position of the user. As soon as the trial started, the guidance feedback was provided depending on the current condition to inform the user about the target location. The user could select the target by placing the cursor on an object and pressing the trigger to finish the trial. A red "x" was used as cursor, placed in the center of the simulated AR FOV (see Figure 6 and Figure 9a and 9b). This shape and color was used to be clearly visible in both visual and non-visual



Cluster	Pixels	RGB
1. ■	40.25%	35 35 33
2. ■	38.80%	64 65 63
3. ■	14.20%	131 138 139
4. ■	6.74%	30 77 41

(a) Static input image.

(b) Resulting cluster partitions.

(c) Color clusters values.

Fig. 7: A static image (a) of the environment used for study part 2 and 3 is taken as input for color analysis. Four cluster partitions (b) are extracted from the input and the resulting values (c) are used to color the target and distractor objects for the user study.

guidance mode. After confirming the selected target by pressing the trigger, the next trial started automatically, making it a continuous object collection task. The procedure for study part 1 and 2 was the same, which only differed with respect to the background.
In addition to the object collection task, the user was supposed to do a secondary task in study part 3. A bird in either red, black or blue color appeared in the scene, flying from one side of the street around the user to the other side (shown in Figure 9c). The secondary task was about to react as quickly as possible by pressing a button as soon as the bird was spotted. The bird had to be visible inside the user's total HMD FOV while the button was pressed to be counted as hit. This enabled the user to select a target in the search task and to indicate the discovery of the bird at the same time. The three main study parts took 30 minutes (10 minutes for each part - 5 minutes with each guidance method). Including introduction, training and filling out the questionnaire, the whole study took 45 minutes.

5 RESULTS

16 users (4 females), aged between 19 and 60 years ($M = 29.1, SD = 9.2$) took part in our study. The majority of participants played video games daily (50%) or weekly (31.3%) and indicated that the gaming console and the computer were their most often used mediums (37.5 % each) followed by the smartphone (18.8%). Regarding the experience with AR glasses 43.8% stated that they were using them sometimes.

A 2x3 repeated measures ANOVA was used to analyze the effect of task load (no noise, with noise, with noise and secondary task) and mode (EyeSee360, audio-tactile) on hit rate (hits/trials), each absolute and signed row, column total errors and errors per trial, trial duration and total number of trials. Greenhouse-Geisser correction was applied when necessary. Row error was computed as the difference between the row of the chosen object and the row of the actual target on the spherical grid (see Section 4.4). Column error was computed analogously. We further used Pearson correlation to analyze the association of the target distance and performance measures. We assumed that identifying and selecting targets that are far away and thus look smaller, could be more difficult and made a separate analysis accordingly.

5.1 Performance, noise and guidance mode

In the following, please note that when we refer to task load, this relates to the load inherent to the task itself, while we refer to workload as the cognitive demand on the user side. The factor task load did not affect hit rate as neither background noise nor the secondary task did lead to performance decrements here ($p = .103$). Hit rate was consistently high with mean values ranging from 0.93 to 0.96. Guidance mode, on the other hand, significantly affected hit rate. Even though both modes facilitated hit rates above 0.9, mean hit rate of the audio-tactile guidance turned out be a significant 3% higher compared to the EyeSee360 technique, $F(1,15) = 8.45, p = .011, \eta_p^2 = .36$ (see Table 1). Trial duration was further affected by both, mode ($F(1,15) = 84.72, p < .001, \eta_p^2 = .85$) and the task ($F(1.1,17.1) = 6.3, p = .019, \eta_p^2 = .296$) and marginally by their interaction ($F(1.2,18) = 4.01, p = .054, \eta_p^2 = .211$).

Table 1: Mean values and standard errors of the hit rate which has been affected by the interface but not by task load.

Interface	Hit rate
EyeSee360	0.93 (0.02)*
audio-tactile	0.96 (0.01)*
Task load	
No noise	0.93 (0.02)
Noise	0.94 (0.01)
Noise + 2nd task	0.96 (0.01)

* p <.05

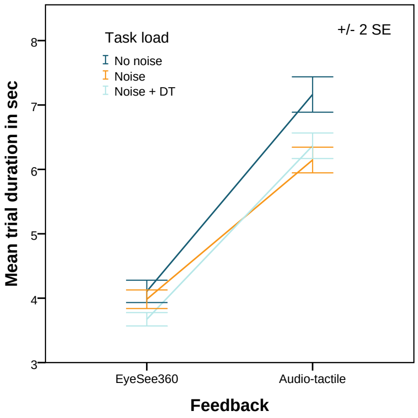


Fig. 8: Trial duration in seconds by task load and guidance method. Trial duration was significantly longer with the audio-tactile mode in each task at p level <.001.

Main effects analysis showed that in each study part, trial duration was longer with the audio-tactile mode than with EyeSee360 ($p < .001$). Furthermore we found a trend that trial duration decreased slightly in the EyeSee360 condition from study part 1 (no noise) to 3 (noise and dual task) ($p = .062$), while with the audio-tactile guidance trial duration decreased from part 1 to 2 (noise) ($p = .038$) (see Figure 8). The figure also shows that several values deviate upwards. We assume that these outliers can result from different factors: 1) Selection difficulties, 2) target locations where the background color was more similar to the target and 3) targets which were particularly close to distractors. As we did not log these variables, it is not possible to clearly trace it back. However, some of these aspects will be considered in the upcoming discussion.

5.2 Effect of target distance

We computed the euclidean distance from the user's viewpoint to the target position for each trial. Following, we analyzed the correlation between the distance, trial duration and hits in both guidance conditions. Only in the EyeSee360 condition there was a sig-


(a) Study part 2 and 3 using EyeSee360.
(b) Study part 2 and 3 using audio-tactile guidance.
(c) Bird for SA measures used in study part 2 and 3.

Fig. 9: Busy city environment that is used for study part 2 and 3 to create visual noise and optical flow. The same object collection task is required to solve with EyeSee360 (a) and audio-tactile guidance (b). To measure SA in study part 3, the user has to react to a bird flying through the scene as secondary task (c). The dotted line visualizes an exemplary route of the bird. Note that out-of-view objects were visualized in semi-transparent orange color and were not visible outside the simulated AR display during the experiment.

nificant positive correlation between the distance and trial duration ($r(3662) = 0.135, p < .001$), and a negative correlation between distance and hits ($r(3662) = -0.047, p = .005$). That is, when using EyeSee360, the further away the target was, the longer the participants took and the fewer hits they made. Correlations were not significant in the audio-tactile condition. We also categorized data by near and far target distance and included distance as two-level factor in the ANOVA model. As targets were placed in a random depth between 15 to 30 meters, targets below 22.5 meters are classified as near distance, everything above as far distance targets. The repeated measures analysis shows a significant influence of distance (near/far) on trial duration, $F(1,15) = 31.7, p < .001, \eta_p^2 = .848$. Users generally needed a little more time when the target was located in the far area ($M = 6s, SE = 0.51$) compared to the near area ($M = 5.5s, SE = 0.46$). There was also a marginally significant interaction of guidance and distance on trial duration, $F(1,15) = 3.48, p = .082, \eta_p^2 = .188$: Main effects analysis showed that only with EyeSee360 users needed more time for distant compared to near targets ($far : M = 4.8s(SE = 0.48), near : M = 4s(SE = 0.32), p = .001$). In the audio-tactile condition performance was similar for near ($M = 7.1s, SE = 0.61$) and far targets ($M = 7.3s, SE = 0.59$). Regarding hit rate, distance and the guidance method showed a marginally significant interaction effect, $F(1,15) = 4.05, p = .062, \eta_p^2 = .213$. Main effects analysis revealed that when comparing the performance between guidance methods for near and far distance targets separately, EyeSee360 and the audio-tactile technique differed only at the far target level: Hit rate was significantly higher with audio-tactile guidance ($M = 0.96, SE = .011$) than with EyeSee360 ($M = 0.92, SE = .017$), $p = .006$. That is, in case of far targets the audio-tactile guidance performed 4.2% better than EyeSee360. At the near distance level hit rates were also high for both feedback modes (EyeSee, M = 0.94, SE = 0.02, audio-tactile, M = 0.96, SE = 0.01) but did not differ significantly ($M = 0.92, SE = .017, p = .138$). When comparing the guidance methods at both distance levels, EyeSee360 had shorter search times at each level ($p < .001$): At the far target level, users needed 4.8s on average ($SE = 0.48$) and were 34% faster than with the audio-tactile mode ($7.3s, SE = 0.59$). At the near target level the EyeSee360 mode ($4s, SE = 0.32$) showed a 44% shorter mean search time than audio-tactile guidance ($7.1s, SE = 0.61$).

5.3 Secondary task

After having finished the first block of trials with one mode in study part 2, users were asked which moving elements in the scene they had noticed. Users were not previously advised to pay special attention to the background. As the question could only be asked one time, a t-test for independent samples had to be performed to compare two groups of participants as half of them started with EyeSee360 and the other half with the audio-tactile mode. In the audio-tactile group 7 out of 8 users noticed birds in the background ($M = 0.87, SD = 0.35$), but only 2 of 8 with EyeSee360 ($M = 0,25, SD = 0.46), t(14) = 3.04, p = .009$. To further analyze how the mode affected the detection of the bird in the background we conducted t-tests for dependent variables in study part 3, where the secondary task was to press a button when the bird was noticed. In case the assumption of normality distribution was not met, the Wilcoxon signed rank test was used as non parametric analysis. Mean values and standard errors are summarized in Table 2.

Table 2: Mean values and standard errors of bird performance measures for both guidance methods in study part 3.

Mode	Total time in FOV in s	Number of FOV entries	Number of correct detections	Misses
EyeSee360	2.5 (0.8)***	1.2 (0.1)*	13.7 (3.1)**	3.1 (3.2)**
Audio-tactile	1.8 (0.6)***	1.1 (0.1)*	16.4 (1.4)**	0.6 (1.1)**

* p <.05, ** p <.01, *** p <.001.

Users noticed the bird 28% faster when using the audio-tactile guidance mode than with EyeSee360. The total time the bird spent in the total HMD FOV until it was noticed was significantly lower in the audio-tactile condition ($t(15) = 5.28, p < .001$). Also, the time from the last FOV entry of the bird until it was found was significantly lower for the audio-tactile mode ($audio-tactile : M = 1.67, SE = 0.51, EyeSee360 : M = 2.16, SE = 0.17, t(15) = 4.64, p < .001$) as well as the average number of FOV entries of the bird per trial, $t(15) = 2.5, p = .024$. In addition, the mean number of detected birds was higher in the audio-tactile condition than with EyeSee360 , $Z = 3.06, p = .002$. The overall error (misses and false-detections) was significantly higher for EyeSee360 ($M = 3.38, SE = 3.24$) than with the audio-tactile mode ($M = 0.81, SE = 1.22$), which could be attributed to the misses. Their number was higher in the visual condition than in the audio-tactile one ($Z = 3.21, p = .001$) while the number of false-detections did not differ significantly between conditions. Mean values and standard errors are displayed in Table 2. We further analyzed potential correlations between the performance of the search task and the performance of the secondary task, namely time until the bird was found. Regarding the audio-tactile guidance method, there was no correlation between primary and secondary task performance measures. When being guided by EyeSee360 a higher hit rate in the search task was associated with a faster detection of the bird after it entered the FOV ($r = -.656, p = .006$), the total time the bird spent in HMD FOV until it was noticed ($r = -.536, p = .032$) and bird detections ($r = .745, p = .001$).

Table 3: Significant differences between questionnaire ratings about distractors and task performance for study part 3.

	EyeSee360	Audio-tactile
Feeling disturbed by moving objects	4.9 (3.2)	4.1 (2.8)*
Fast secondary task performance	6.2 (2.3)	7.4 (2.3)*
Precise secondary task performance	5.8 (2.5)	7.1 (2.6)**
Concentration on secondary task	5.9 (2.4)	7.8 (2.4)*
Ease of judging the vertical position	9.4 (0.9)*	8.1 (2.2)

* p <.05, ** p <.01

5.4 Questionnaire ratings

With regards to **cognitive measures**, a 2 x 3 repeated measures ANOVA was used to analyze the effect of task load and guidance method on workload through overall (raw) NASA TLX rating scores and on subscales. The overall NASA TLX score ranged from 0 to 100, ratings on a subscale from 1 to 21. Task load showed a significant effect on the overall NASA TLX score, $F(2,30) = 12.11, p < .001, \eta_p^2 = .447$. It was significantly lower in study part 1 compared to part 2 ($p = .001$) and to part 3 ($p = .001$). Regarding the analysis of subscales, task load affected mental demand ($F(1.43, 217.67) = 6.5, p = .011, \eta_p^2 = .3$), marginally physical demand ($F(2,30) = 3.18, p = .056, \eta_p^2 = .175$) and performance ($F(2,30) = 6.19, p = .006, \eta_p^2 = .292$). Post-hoc comparisons revealed significantly higher ratings for mental demand for study part 3 compared to part 1 ($p = .019$) and higher ratings on the performance subscale in study part 1 than in 3 ($p = .012$). The effort subscale was not affected by neither task load nor guidance method. In contrast, frustration subscale was affected by both, task load ($F(2,30) = 6.18, p = .006, \eta_p^2 = .292$) and guidance method ($F(1,15) = 4.34, p = .055, \eta_p^2 = .23$). Frustration was higher in study part 3 compared to part 1 ($p = .033$) (see Figure 10) and higher with EyeSee360 ($M = 7.4, SE = 1.1$) than with the audio-tactile interface ($M = 5.8, SE = 1.1$) across all study parts. There was no interaction effect between study part and mode, however mean values and standard errors by both factors are displayed in Table 4.

We further compared usability ratings regarding **distractors and task performance** factors between EyeSee360 and the audio-tactile mode, see Table 3. In study part 3 but not in study part 2, users felt more disturbed by moving objects while performing the search task with EyeSee360 than with the audio-tactile guidance, $t(15) = -2.36, p = .032$. They further indicated they thought they had performed the secondary task faster ($t(15) = 2.40, p = .03$) and more precisely ($t(15) = 3.47, p = .003$) with the audio-tactile mode. Participants were also better able to concentrate on the secondary task ($t(15) = 3.21, p = .006$) and on the main task with audio-tactile guidance ($t(15) = 2.3, p = .036$). However, judging the vertical position was perceived to be easier with EyeSee360 ($t(15) = -2.44, p = .028$). Other usability ratings as the ease of performing the task, ease of learning, performing the main task fast and precisely, judging the horizontal position and distance, fatigue did not differ significantly between guidance modes.

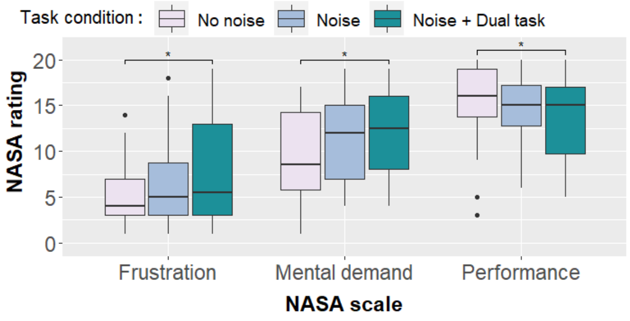


Fig. 10: NASA TLX scores across both guidance modes for the frustration, mental demand and performance subscale show significant differences between task load conditions, * = p <.05.

Regarding the **overall usability of the system**, users provided high ratings, indicating that they coped well with the task and the setup (see Table 5). Finally, users were asked post hoc which of the two guidance methods they would **prefer** using in VR/AR technologies and which method is potentially better to pay more attention on the surroundings on a 7-point scale (1 = audio-tactile, 7 = EyeSee360). With regard to the first point, user ratings ($M = 3.13 SD = 1.49$) indicated a slight tendency for the usage of audio-tactile feedback for the purpose of guidance in AR. On the latter point, ratings ($M = 2.56, SD = 1.77$) show a clear trend that users have the feeling to be more aware of their surroundings when using audio-tactile cues.

Table 4: Mean values and standard deviations by study part and guidance mode for NASA TLX subscales frustration, mental demand and performance.

Study part	Scale	Audio-tactile	EyeSee
1	F	4.8 (3.8)	6.3 (3.9)
	MD	9.3 (5.2)	10.4 (4.5)
	P	15.8 (3.1)	15.3 (5.1)
2	F	5.8 (5.2)	7.4 (4.6)
	MD	11.4 (5.3)	11.6 (4.9)
	P	14.2 (4.5)	14.9 (3.4)
3	F	6.7 (5.8)	8.5 (5.7)
	MD	11.9 (4.8)	12.5 (4.6)
	P	13.4 (4.4)	13.3 (4.8)

F = Frustration, MD = Mental demand, P = Performance

Table 5: Mean level of agreement with comfort and usability statements for the overall system on 11-point Likert items and standard deviations.

Statement	Mean Rating (SD)
Easy to detect targets	7,38 (2,47)
Sitting comfort	8,50 (2,24)
Interface (HMD+head strap) comfort	8,88 (1,76)
Task was easy to understand	10,13 (0,78)
Concentration on task	9,38 (1,17)
Easy to recognize targets	8,75 (1,52)
Improvement over time	9,25 (2,05)
Fun of use	9,88 (1,27)

6 DISCUSSION

RQ$_1$: *How well do non-visual guidance methods perform compared to visual guidance methods in a search task on different levels of task load?*

In H_1 we expected that the performance of the guidance method would be related to the degree of subjective *workload* as experienced by the user. The hypothesis implied that EyeSee360 would outperform audio-tactile guidance on a low level of visual task load, but performance differences would be decreasing as soon as the task load would increase. The performance of the audio-tactile guidance on the other hand was expected not to decrease in the high task load conditions. The self-evaluated mental workload generally increased with a higher task load during the experiment as intended, whereas users indicated that they did not put more effort into solving the task. Interestingly, there was no significant difference of the mental workload ratings across the guidance modes, which is in contradiction to our first hypothesis H_1. This was assumed since visual guidance methods usually compress a high level of information in a limited FOV (compare [30,33]). However, this outcome has probably been reduced by recent improvements of the EyeSee360 method [31].

Surprisingly, our results show that the **task load** did not have an effect on task performance of both methods. Regarding hit rate, the audio-tactile guidance was on a par with EyeSee360 across study parts, which also indicates a comparable performance to other visual guidance techniques [32]. The overall hit rate of the audio-tactile mode was 3% better then EyeSee360 which was a small but significant difference in mean values (EyeSee360: 0.93% vs. audio-tactile 0.96%). However, the search duration per trial was considerably longer for audio-tactile guidance in comparison to using EyeSee360. This may be explained by the fact that audio-tactile cues used for guidance are relatively difficult to interpret and therefore require additional training until it can be used at a higher speeds (see [62]). In comparison, the $focus + context$ approach used in EyeSee360 allows it to locate out-of-view objects directly and mostly intuitive (e.g., the proxy encodes already in which direction the user has to move their head to locate the object) [30,32]. Another possible explanation regarding the consistently good guidance performance of EyeSee360 in the noise conditions could be explained by the observations that participants were partially able to fade out the background and focus mainly on the projection plane of the EyeSee360 interface (also see [16]). Therefore the increased noise level did not affect the visual guidance method as much as expected and users were able to concentrate on the search task in a straightforward manner.

However, using EyeSee360 led to a significantly higher **frustration** level compared to the audio-tactile technique. We assume that this effect was partially caused by the occasional selection issues. During the selection phase, users sometimes needed several attempts to place the crosshair reliably on the target object, required for selection. This happened mostly if the object was placed at a high distance. We assume this problem arises as a part of stereoscopic depth disparities [16, 47, 50] in which users need to focus on different focal planes: EyeSee360 in the foreground for target guidance and object selection in the background. This problem might be aggravated if both planes have a large distance to each other - as is the case the target object is placed far away from the user and thus appears smaller. In this context we found out that users generally took longer if targets were placed in higher distances with both methods. However, the tendency to an interaction between distance and guidance on hit rate shows that slightly fewer hits were made in EyeSee360 condition when targets were distant. This is comparable to [35], which also reports about a reduced selection accuracy of EyeSee360 compared to other visual methods. However, improving target selection, e.g., integrating a combination of head- and eye-based approach [49] or using novel interaction devices like 3D pen [73], might lead to a higher accuracy and overall hit rate for both visual and non-visual guidance.

Another source of frustration is potentially **visual clutter**. In this context, sensory overload is a relevant topic as visual guidance methods usually compress information into a relatively small FOV (see [23,47]). Even though EyeSee360 was optimized to somehow reduce mental workload and visual clutter [31, 34], these techniques might still suffer from a limited FOV. By transcoding visual information into audio-tactile cues we potentially reduce the visual complexity and the number of distractors within the FOV. This allows the AR system to use the free display-space for any other non-guidance related further information. In this regard, it would also be interesting to investigate the search behaviour between visual and non-visual guidance in context of a limited FOV (compare [18, 83]). Also information density could be an additional factor which might affects search performance [13]. Generally, further studies are required to find out whether search behaviour and performance differ considerably between visual and non-visual guidance methods. Also, while it makes sense to use wider FOV to reduce cognitive load, previous studies have only dealt with relatively low information density so far [8]. In addition, considerations should be given to how these factors might affect real AR environments compared to a highly controlled simulated AR environment in VR. While it can be assumed that results can be applied to AR systems up to a certain degree, simulated AR still has clear challenges related to the fidelity of the real world component in the system. For example, physical conditions like relative brightness and contrast between real and virtual objects or the level of opacity of the virtual objects might have an additional impact on the user performance [74].

Finally, **attention mechanisms** likely play an important role. Human attention is primarily attracted to visually salient stimuli. Visual selective attention allows human perception only to focus on a small area of the visual field at a given moment [93]. However, multisensory integration and crossmodal attention have a large impact on how we perceive the world, potentially enhancing selection attention in AR tasks. Providing information over multiple sensory channels has the potential to enable sensory stimulus integration. For this, attention mechanisms are used to process and coordinate multiple stimuli across sensory modalities, which also affects the way of managing resources [21]. However, multiple stimuli require a correct synchronization, otherwise sensory integration does not take place as stimuli could be interpreted independently [82]. This topic should be more closely addressed in further studies comparing visual and non-visual guidance methods.

In conclusion, with respect to H_1 we can state that although the audio-tactile guidance is slower, it is able to provide a similar and even slightly better hit rate compared to a well established visual guidance method like EyeSee360. That is, when fast search times are not prioritized, the audio-tactile method allows precise guidance while freeing up the visual channel for other non-guidance information. The audio-tactile feedback can also be interesting for visually impaired people like [41, 78], since the same information is substituted to another sensory channel [59] without degradation of hit rate performance. For this purpose also depth cues can be particularly helpful. Regarding to common depth judgement issues in VR/AR (see [84, 92]) the presented tactile depth cues might be supportive for a more accurate depth estimation.

RQ$_2$: *Is there an effect of guidance method on situation awareness when a secondary task is included?*
As expected EyeSee360 performed reasonably well in terms of an abstract object collection task. But regarding SA, an effect was noticeable as soon as the **task load** increased from study part 2 to 3. Audio-tactile guidance performed significantly better with regards to general perception (study part 2, noticeability) and SA performance measures (study part 3, secondary task performance). This outcome confirms our second hypothesis H_2 that a higher SA is achieved by using audio-tactile guidance. This, however, is not related to a higher workload when using EyeSee360 as initially supposed. With respect to the general perception, it was easy for most users (7 of 8) to notice the bird if audio-tactile guidance was used in study part 2. In contrast, only 2 of 8 users were able to notice the bird while solving the collection task with EyeSee360. This performance difference may be attributed to the **focal disparity**, in which users tend to focus on the AR-plane to primarily follow the visual guidance cues while blurring out the background (compare [16]). By this behaviour, small details and objects are simply being overlooked by the user.

Concerning the performance measures, a significant difference between both methods became apparent. Almost all SA measures were significantly better with audio-tactile guidance than with EyeSee360 in the secondary task. Subjective questionnaire ratings also showed that users felt they could perceive their surroundings significantly better using audio-tactile guidance. This indicates a higher SA in terms of environmental perception using the audio-tactile interface and can probably be attributed to the fact that the user did not have to deal with visually related issues (clutter, occlusion, selection issues). Therefore users were able to handle the main and secondary task in more balanced manner compared to the visual mode. In addition, frustration and workload also showed a significant difference for EyeSee360 in study part 3. Users were possibly more stressed when solving the main and secondary task at the same time. Since **human capabilities** for processing information are limited, there might not be enough capacity to solve a secondary task sufficiently while using a visual guidance method [63]. This could be due to the fact that that users tend to allocate their resources to higher priority-task components as soon as the arousal increases. Even though participants were briefed that both target search and secondary task were equally important to solve, some users might have prioritized the target search subconsciously since it was a continuous task over the whole user study. Furthermore, the level of immersion might be higher in a simulated environment if a visualization method is used for the search task compared to a non-visual approach. This possibly results in a trade-off between the degree of immersion and situation awareness, like reported in [39]. Generally, the usage of AR can cause distraction from the real world since it requires intensive concentration [4]. For that reason, reducing visual stimuli in the visible area of the AR device and substituting them into other modalities in a more intuitive way seems like a reasonable attempt to increase SA during the use of this technology. However, this approach is highly dependent on the current user task and further considerations have to be taken in case it is still required to display additional visual information inside the FOV.

Finally, we suspected a possible correlation between the main and the secondary task, namely that users tend to focus more on one task while neglecting the other. However, a statistical relationship between those two variables was not ascertainable in case of audio-tactile guidance. For visual guidance, however, the study revealed a quite contrary effect. It turned out that if users performed the object collection task well, they also showed a reasonable performance in the secondary task. In conclusion, with respect to H_2 we can state that a higher SA can be achieved using audio-tactile guidance. However, this result could not directly associated with a higher mental workload, but due to other factors that need further exploration.

7 CONCLUSION

In this paper we compared EyeSee360, a state-of-the-art visual guidance method, with a non-visual guidance approach using audio-tactile stimuli. Doing so, we addressed head-mounted displays with narrow FOV. The main focus was on measuring performance, accuracy, cognitive load and SA during a object selection task in simulated AR. We used a vibration headband that consists of five vibration motors to create vibro-tactile feedback along the forehead and temples. By providing audio-tactile cues, it is possible to guide the user in the 3D space on the longitudinal, latitudinal, and depth plane. In particular, it can restrict negative effects like visual clutter or occlusion compared to common visual guidance approaches. As a result, we showed that users are more aware of their environment with audio-tactile in comparison with EyeSee360 with 16.5% more correctly detected background targets and 28% faster detection times, which implies a higher SA when using audio-tactile methods. However, during the main task, audio-tactile guidance performed slower than EyeSee360 regarding search times. As such, the choice of technique is context dependent - for example, if target search time is prioritized, EyeSee360 is preferable. Usage contexts that require improved SA and limited visual clutter can benefit from audio-tactile guidance. Despite the fact that the FOV in AR devices will increase in the future, it is still a challenging goal to build displays that could cover the entire human visual field [43]. Even if next generation devices included a larger FOV, they would still be limited. Therefore, problems associated with visual methods like cluttering, occlusion and a potentially high workload likely remain, especially with higher information density. Here, audio-tactile cues can help in order to address and improve these problems by substituting some of the visual information.

Future work includes an improvement of the physical setup. By extending the headband with more vibration motors (similar to [42]), a higher resolution and thus an increasing accuracy would be possible. This would allow us to investigate multiple target guidance in more complex situations more closely. Using a different motor driver technology like linear resonant actuators or piezoelectrics would also be reasonable in terms of usable bandwidth and acceleration characteristics to improve accuracy and performance. Another interesting venue of research is the combination of methods to assess characteristics like performance and SA in more detail. In this constellation, audio-tactile cues could be responsible for the target point guidance while a visual method like EyeSee360 could be used to increase SA, e.g., warning of incoming objects similar to [39]. However, it still needs to be investigated how visual metaphors work best in combination with audio-tactile guidance without distracting or overloading the user with information. In this context it would also be worthwhile to consider the integration of transition signals to attract the users attention as soon as certain information enters the FOV. Finally, eye tracking techniques can be used to enhance the effectiveness of both visual and non-visual guidance metaphors, support object selection [11], and could be used an additional indicator for situation awareness [86].

ACKNOWLEDGMENTS

We would like to thank Uwe Gruenefeld for providing the source code and useful comments for the EyeSee360 technique. This work was partially supported by the Deutsche Forschungsgemeinschaft (grant KR 4521/2-1).

REFERENCES

[1] P. Alfano and G. Michel. Restricting the field of view: Perceptual and performance effects. *Perceptual and motor skills*, 70(1):35–45, 1990.

[2] O. Ariza, G. Bruder, N. Katzakis, and F. Steinicke. Analysis of proximity-based multimodal feedback for 3d selection in immersive virtual environments. In *Proceedings of IEEE Virtual Reality (VR)*, p. (accepted), 2018.

[3] K. Arthur and F. Brooks Jr. *Effects of field of view on performance with head-mounted displays*. PhD thesis, University of North Carolina at Chapel Hill, 2000.

[4] J. W. Ayers, E. C. Leas, M. Dredze, J.-P. Allem, J. G. Grabowski, and L. Hill. Pokémon GO—A New Distraction for Drivers and Pedestrians. *JAMA Internal Medicine*, 176(12):1865–1866, 12 2016.

[5] R. Azuma. A survey of augmented reality. *Presence: Teleoperators & Virtual Environments*, 6(4):355–385, 1997.

[6] R. Azuma and C. Furmanski. Evaluating label placement for augmented reality view management. In *Proceedings of the 2Nd IEEE/ACM International Symposium on Mixed and Augmented Reality*, ISMAR '03, pp. 66–. IEEE Computer Society, Washington, DC, USA, 2003.

[7] K. Bark, E. Hyman, F. Tan, E. Cha, S. Jax, L. Buxbaum, and K. Kuchenbecker. Effects of vibrotactile feedback on human learning of arm motions. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 23(1):51–63, 2014.

[8] J. Baumeister, S. Ssin, N. ElSayed, J. Dorrian, D. Webb, J. Walsh, T. Simon, A. Irlitti, R. Smith, M. Kohler, et al. Cognitive cost of using augmented reality displays. *IEEE transactions on visualization and computer graphics*, 23(11):2378–2388, 2017.

[9] B. Bell, S. Feiner, and T. Höllerer. View management for virtual and augmented reality. In *Proceedings of the 14th Annual ACM Symposium on User Interface Software and Technology*, UIST '01, pp. 101–110. ACM, New York, NY, USA, 2001.

[10] M. Berning, F. Braun, T. Riedel, and M. Beigl. Proximityhat: A head-worn system for subtle sensory augmentation with tactile stimulation. In *Proceedings of the 2015 ACM International Symposium on Wearable Computers*, ISWC '15, pp. 31–38. ACM, New York, NY, USA, 2015.

[11] J. Blattgerste, P. Renner, and T. Pfeiffer. Advantages of eye-gaze over head-gaze-based selection in virtual and augmented reality under varying field of views. In *Proceedings of the Workshop on Communication by Gaze Interaction*, pp. 1–9, 2018.

[12] F. Bork, C. Schnelzer, U. Eck, and N. Navab. Towards efficient visual guidance in limited field-of-view head-mounted displays. *IEEE Transactions on Visualization and Computer Graphics*, 24(11):2983–2992, 2018.

[13] C. Trepkowski, and T. Eibich and J. Maiero and A. Marquardt and E. Kruijff and S. Feiner. The effect of narrow field of view and information density on visual search performance in augmented reality. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 2019.

[14] A. Case and A. Day. *Designing with Sound: Fundamentals for Products and Services*. O'Reilly Media, 2018.

[15] A. Cassinelli, C. Reynolds, and M. Ishikawa. Augmenting spatial awareness with haptic radar. In *2006 10th IEEE International Symposium on Wearable Computers*, pp. 61–64, Oct 2006.

[16] P. Chakravarthy, D. Dunn, K. Akşit, and H. Fuchs. Focusar: Auto-focus augmented reality eyeglasses for both real and virtual. *IEEE transactions on visualization and computer graphics*, PP, 09 2018.

[17] M. Chmielewski, K. Sapiejewski, and M. Sobolewski. Application of augmented reality, mobile devices, and sensors for a combat entity quantitative assessment supporting decisions and situational awareness development. *Applied Sciences*, 9(21):4577, 2019.

[18] J. Covelli, J. Rolland, M. Proctor, J. Kincaid, and P. Hancock. Field of view effects on pilot performance in flight. *The International Journal of Aviation Psychology*, 20(2):197–219, 2010.

[19] D. Damos. *Multiple task performance*. CRC Press, 1991.

[20] V. de Jesus Oliveira, L. Brayda, L. Nedel, and A. Maciel. Designing a vibrotactile head-mounted display for spatial awareness in 3d spaces. *IEEE Transactions on Visualization and Computer Graphics*, 23(4):1409–1417, April 2017.

[21] J. Driver and C. Spence. Crossmodal attention. *Current opinion in neurobiology*, 8(2):245–253, 1998.

[22] G. Dubus and R. Bresin. A systematic review of mapping strategies for the sonification of physical quantities. *PloS one*, 8(12):e82491, 2013.

[23] S. Ellis and B. Menges. Localization of virtual objects in the near visual field. *Human Factors*, 40(3):415–431, 1998.

[24] M. Endsley and D. Garland. *Situation Awareness Analysis and Measurement*. CRC Press, 2000.

[25] B. Ens, D. Ahlström, and P. Irani. Moving ahead with peephole pointing: Modelling object selection with head-worn display field of view limitations. In *Proceedings of the 2016 Symposium on Spatial User Interaction*, pp. 107–110. ACM, 2016.

[26] D. Foyle, A. Andre, and B. L. Hooey. Situation awareness in an augmented reality cockpit: Design, viewpoints and cognitive glue. In *Proceedings of the 11th International Conference on Human Computer Interaction*, vol. 1, pp. 3–9, 2005.

[27] G. French and T. Schnell. Terrain awareness & pathway guidance for head-up displays (tapguide); a simulator study of pilot performance. In *Digital Avionics Systems Conference, 2003. DASC'03. The 22nd*, vol. 2, pp. 9–C. IEEE, 2003.

[28] J. Gabbard, E. Swan, II, and D. Hix. The effects of text drawing styles, background textures, and natural lighting on text legibility in outdoor augmented reality. *Presence: Teleoper. Virtual Environ.*, 15(1):16–32, Feb. 2006.

[29] R. Grasset, T. Langlotz, D. Kalkofen, M. Tatzgern, and D. Schmalstieg. Image-driven view management for augmented reality browsers. In *Proceedings of the 2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, ISMAR '12, pp. 177–186. IEEE Computer Society, Washington, DC, USA, 2012.

[30] U. Gruenefeld, D. Ennenga, A. E. Ali, W. Heuten, and S. Boll. Eyesee360: designing a visualization technique for out-of-view objects in head-mounted augmented reality. In *Proceedings of the 5th Symposium on Spatial User Interaction*, pp. 109–118. ACM, 2017.

[31] U. Gruenefeld, D. Hsiao, and W. Heuten. Eyeseex: Visualization of out-of-view objects on small field-of-view augmented and virtual reality devices. In *PerDis '18*, 2018.

[32] U. Gruenefeld, I. Köthe, D. Lange, S. Weiss, and W. Heuten. Comparing techniques for visualizing moving out-of-view objects in head-mounted virtual reality. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 03 2019.

[33] U. Gruenefeld, D. Lange, L. Hammer, S. Boll, and W. Heuten. Flyingarrow: Pointing towards out-of-view objects on augmented reality devices. In *Proceedings of the 7th ACM International Symposium on Pervasive Displays*, p. 20. ACM, 2018.

[34] U. Gruenefeld, L. Prädel, and W. Heuten. Improving search time performance for locating out-of-view objects. *Mensch und Computer 2019-Tagungsband*, 2019.

[35] U. Gruenefeld, L. Prädel, and W. Heuten. Locating nearby physical objects in augmented reality. In *Proceedings of the 18th International Conference on Mobile and Ubiquitous Multimedia*, pp. 1–10, 2019.

[36] C. Gutwin, A. Cockburn, and A. Coveney. Peripheral popout: The influence of visual angle and stimulus intensity on popout effects. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pp. 208–219. ACM, New York, NY, USA, 2017.

[37] J. Irizarry, M. Gheisari, G. Williams, and B. Walker. Infospot: A mobile augmented reality method for accessing building information through a situation awareness approach. vol. 33, pp. 11–23. Elsevier, 2013.

[38] A. Israr and I. Poupyrev. Tactile brush: Drawing on skin with a tactile grid display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pp. 2019–2028. ACM, New York, NY, USA, 2011.

[39] J. Jung, H. Lee, J. Choi, A. Nanda, U. Gruenefeld, T. Stratmann, and W. Heuten. Ensuring safety in augmented reality from trade-off between immersion and situation awareness. In *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 10 2018.

[40] J. Kalat. *Biological psychology*. Nelson Education, 2015.

[41] B. Katz, S. Kammoun, G. Parseihian, O. Gutierrez, A. Brilhault, M. Auvray, P. Truillet, M. Denis, S. Thorpe, and C. Jouffrais. Navig: Augmented reality guidance system for the visually impaired: Combining object localization, gnss, and spatial audio. *Virtual Reality*, 16:17, 01 2012.

[42] O. Kaul and M. Rohs. Haptichead: 3d guidance and target acquisition through a vibrotactile grid. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA '16, pp. 2533–2539. ACM, New York, NY, USA, 2016.

[43] K. Kim, M. Billinghurst, G. Bruder, H. Duh, and G. Welch. Revisiting trends in augmented reality research: A review of the 2nd decade of ismar (2008–2017). *IEEE Transactions on Visualization and Computer Graphics*, PP:1–1, 09 2018.

[44] N. Kishishita, K. Kiyokawa, J. Orlosky, T. Mashita, H. Takemura, and E. Kruijff. Analysing the effects of a wide field of view augmented reality display on search performance in divided attention tasks. In *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 177–186, Sept 2014.

[45] A. Klapetek, M. Ngo, and C. Spence. Does crossmodal correspondence modulate the facilitatory effect of auditory cues on visual search? *Attention, Perception, & Psychophysics*, 74(6):1154–1167, 2012.

[46] T. Koelewijn, A. Bronkhorst, and J. Theeuwes. Competition between auditory and visual spatial cues during visual task performance. *Experimental Brain Research*, 195(4):593–602, Jun 2009.

[47] E. Kruijff, E. S. II, and S. Feiner. Perceptual issues in augmented reality revisited. In *In Proceedings of the International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 3–12. IEEE Computer Society, 2010.

[48] E. Kruijff, J. Orlosky, N. Kishishita, C. Trepkowski, and K. Kiyokawa. The influence of label design on search performance and noticeability in wide field of view augmented reality displays. *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–1, 2018.

[49] M. Kytö, B. Ens, T. Piumsomboon, G. Lee, and M. Billinghurst. Pinpointing: Precise head-and eye-based target selection for augmented reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–14, 2018.

[50] M. Kytö, A. Mäkinen, T. Tossavainen, and P. Oittinen. Stereoscopic depth perception in video see-through augmented reality within action space. *Journal of Electronic Imaging*, 23(1):1 – 11, 2014.

[51] V. Lehtinen, A. Oulasvirta, A. Salovaara, and P. Nurmi. Dynamic tactile guidance for visual search tasks. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology*, UIST '12, pp. 445–452. ACM, New York, NY, USA, 2012.

[52] A. Leykin and M. Tuceryan. Automatic determination of text readability over textured backgrounds for augmented reality systems. In *Third IEEE and ACM International Symposium on Mixed and Augmented Reality*, pp. 224–230, Nov 2004.

[53] R. Lindeman, R. Page, Y. Yanagida, and J. Sibert. Towards full-body haptic feedback: the design and deployment of a spatialized vibrotactile feedback system. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST 2004*, pp. 146–149, 2004.

[54] R. Lindeman, Y. Yanagida, J. Sibert, and R. Lavine. Effective vibrotactile cueing in a visual search task. In *Proc. of Interact 2003*, pp. 89–96, 2003.

[55] P. Lindemann, T. Lee, and G. Rigoll. Supporting driver situation awareness for autonomous urban driving with an augmented-reality windshield display. In *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 358–363, Oct 2018.

[56] Z. Lipowski. Sensory and information inputs overload: Behavioral effects. *Comprehensive Psychiatry*, 16(3):199 – 221, 1975.

[57] M. Livingston, L. Rosenblum, D. Brown, G. Schmidt, S. Julier, Y. Baillot, E. Swan, Z. Ai, and P. Maassel. Military applications of augmented reality. In *Handbook of augmented reality*, pp. 671–706. Springer, 2011.

[58] J. Loomis, R. Klatzky, and N. Giudice. Sensory substitution of vision: Importance of perceptual and cognitive processing. In *Assistive technology for blindness and low vision*, pp. 179–210. CRC Press, 2018.

[59] S. Maidenbaum, S. Abboud, and A. Amedi. Sensory substitution: Closing the gap between basic research and widespread practical visual rehabilitation. *Neuroscience & Biobehavioral Reviews*, 41, 01 2013.

[60] A. Marquardt, E. Kruijff, C. Trepkowski, J. Maiero, A. Schwandt, A. Hinkenjann, W. Stuerzlinger, and J. Schoening. Audio-tactile feedback for enhancing 3d manipulation. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, VRST '18. ACM, 2018.

[61] A. Marquardt, J. Maiero, E. Kruijff, C. Trepkowski, A. Schwandt, A. Hinkenjann, J. Schöning, and W. Stuerzlinger. Tactile hand motion and pose guidance for 3d interaction. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*, VRST '18, 2018.

[62] A. Marquardt, C. Trepkowski, T. Eibich, J. Maiero, and E. Kruijff. Non-visual cues for view management in narrow field of view augmented reality displays. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 190–201, Oct 2019.

[63] G. Matthews and I. Margetts. Self-report arousal and divided attention: A study of performance operating characteristics. *Human Performance*, 4(2):107–125, 1991.

[64] J. McIntire, P. Havig, S. Watamaniuk, and R. Gilkey. Visual search performance with 3-d auditory cues: Effects of motion, target location, and practice. *Human Factors*, 52(1):41–53, 2010.

[65] Mudd, S., and McCormick, E. The use of auditory cues in a visual search task. *Journal of Applied Psychology*, 44(3):184–188, 1960.

[66] R. Newman. *Head-up displays: Designing the way ahead*. Routledge, 2017.

[67] M. Ngo and C. Spence. Auditory, tactile, and multisensory cues facilitate search for dynamic visual stimuli. *Attention, Perception, & Psychophysics*, 72(6):1654–1665, Aug 2010.

[68] B. Park, C. Yoon, J. Lee, and K. Kim. Augmented reality based on driving

situation awareness in vehicle. In *2015 17th International Conference on Advanced Communication Technology (ICACT)*, pp. 593–595, July 2015.

[69] G. Parseihian, C. Gondre, M. Aramaki, S. Ystad, and R. Kronland-Martinet. Comparison and evaluation of sonification strategies for guidance tasks. *IEEE Transactions on Multimedia*, 18(4):674–686, April 2016.

[70] S. Peterson, M. Axholt, and S. Ellis. Label segregation by remapping stereoscopic depth in far-field augmented reality. In *Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*, ISMAR '08, pp. 143–152. IEEE Computer Society, Washington, DC, USA, 2008.

[71] S. Peterson, M. Axholt, and S. Ellis. Managing visual clutter: A generalized technique for label segregation using stereoscopic disparity. In *IEEE Virtual Reality Conference 2008 (VR 2008), 8-12 March 2008, Reno, Nevada, USA, Proceedings*, pp. 169–176. IEEE Computer Society, 2008.

[72] B. Pfleging, N. Henze, A. Schmidt, D. Rau, and B. Reitschuster. Influence of subliminal cueing on visual search tasks. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '13, pp. 1269–1274. ACM, 2013.

[73] D. Pham and W. Stuerzlinger. Is the pen mightier than the controller? a comparison of input devices for selection in virtual and augmented reality. In *25th ACM Symposium on Virtual Reality Software and Technology*, VRST '19. Association for Computing Machinery, New York, NY, USA, 2019.

[74] E. Ragan, C. Wilkes, D. A. Bowman, and T. Hollerer. Simulation of augmented reality systems in purely virtual environments. In *2009 IEEE Virtual Reality Conference*, pp. 287–288, 2009.

[75] A. Raj, J. Beach, E. Stuart, and L. Vassiliades. *Multimodal and multisensory displays for perceptual tasks*. 01 2015.

[76] D. Ren, T. Goldschwendt, Y. Chang, and T. Höllerer. Evaluating wide-field-of-view augmented reality with mixed reality simulation. In *Virtual Reality (VR), 2016 IEEE*, pp. 93–102. IEEE, 2016.

[77] P. Renner and T. Pfeiffer. Attention guiding techniques using peripheral vision and eye tracking for feedback in augmented-reality-based assistance systems. In *2017 IEEE Symposium on 3D User Interfaces (3DUI)*, pp. 186–194, March 2017.

[78] F. Ribeiro, D. Florêncio, P. Chou, and Z. Zhang. Auditory augmented reality: Object sonification for the visually impaired. In *2012 IEEE 14th International Workshop on Multimedia Signal Processing (MMSP)*, pp. 319–324, Sep. 2012.

[79] S. Rothe, D. Buschek, and H. Hußmann. Guidance in cinematic virtual reality-taxonomy, research status and challenges. *Multimodal Technologies and Interaction*, 3(1):19, 2019.

[80] P. Salmon, N. Stanton, G. Walker, and D. Green. Situation awareness measurement: A review of applicability for c4i environments. *Applied ergonomics*, 37(2):225–238, 2006.

[81] B. Schwerdtfeger and G. Klinker. Supporting order picking with augmented reality. In *Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*, ISMAR '08, pp. 91–94. IEEE Computer Society, Washington, DC, USA, 2008.

[82] C. Spence and S. Squire. Multisensory integration: maintaining the perception of synchrony. *Current Biology*, 13(13):R519–R521, 2003.

[83] L. Stark, K. Ezumi, T. Nguyen, R. Paul, G. Tharp, and H. Yamashita. Visual search in virtual environments. In *Human vision, visual processing, and digital display III*, vol. 1666, pp. 577–589. International Society for Optics and Photonics, 1992.

[84] E. Swan, A. Jones, E. Kolstad, M. Livingston, and H. S. Smallman. Egocentric depth judgments in optical, see-through augmented reality. *IEEE transactions on visualization and computer graphics*, 13(3):429–442, 2007.

[85] H. Uchiyama, M. Covington, and W. Potter. Vibrotactile glove guidance for semi-autonomous wheelchair operations. In *Proceedings of the 46th Annual Southeast Regional Conference on XX*, pp. 336–339. ACM, 2008.

[86] K. van de Merwe, H. Dijk, and R. Zon. Eye movements as an indicator of situation awareness in a flight simulator experiment. *The International Journal of Aviation Psychology*, 22:78–95, 01 2012.

[87] E. Van der Burg, C. Olivers, A. Bronkhorst, and J. Theeuwes. Pip and pop: nonspatial auditory signals improve spatial visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 34(5):1053, 2008.

[88] E. Veas, E. Mendez, S. Feiner, and D. Schmalstieg. Directing attention and influencing memory with visual saliency modulation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pp. 1471–1480. ACM, 2011.

[89] M. Ward, A. Barde, P. Russell, M. Billinghurst, and W. Helton. Visual cues to reorient attention from head mounted displays. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 60:1574–1578, 09 2016. doi: 10.1177/1541931213601363

[90] M. Wells, M. Venturino, and R. Osgood. The effect of field-of-view size on performance at a simple simulated air-to-air mission. In *Helmet-Mounted Displays*, vol. 1116, pp. 126–138. International Society for Optics and Photonics, 1989.

[91] A. Wilberz, D. Leschtschow, C. Trepkowski, J. Maiero, E. Kruijff, and B. Riecke. Facehaptics: Robot arm based versatile facial haptics for immersive environments. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, p. 1–14. Association for Computing Machinery, New York, NY, USA, 2020. doi: 10.1145/3313831.3376481

[92] B. Wu, T. Ooi, and Z. He. Perceiving distance accurately by a directional process of integrating ground information. *Nature*, 428:73–7, 04 2004.

[93] L. Zhang and W. Lin. *Selective visual attention: computational models and applications*. John Wiley & Sons, 2013.

Alexander Marquardt Alexander Marquardt obtained his B.Sc and M.Sc in Computer Science at the Bonn-Rhein-Sieg University of Applied Sciences (BRSU). As a Ph.D. student at University of Bremen and BRSU, his research interests are focused on the design and development of experimental multisensory user interfaces for narrow field of view head-worn devices.

Christina Trepkowski received a Master of Science in Psychology from the Heinrich-Heine-University of Düsseldorf, Germany in 2017. Alongside her studies she started working as a research associate at the Bonn-Rhein-Sieg University of Applied Sciences in 2014 where she has been active as researcher ever since. She is currently doing her Ph.D. on the influence of multisensory cues on situation awareness in information-rich augmented reality environments.

Tom David Eibich received his Bachelor in Computer Science from the Bonn-Rhein-Sieg University in 2017. He is currently doing the Masters programm and his research interests are augmented reality and interactive environments.

Jens Maiero received the Diploma in mathematics from the Stuttgart Technology University of Applied Sciences, Stuttgart, Germany, in 2006, the M.Sc. degree in Computer Science from the Bonn-Rhein-Sieg University of Applied Sciences (BRSU), Sankt Augustin, Germany, in 2009, and is currently finalizing his Ph.D. degree in Computer Science at Brunel University, London, U.K., and BRSU.

Ernst Kruijff is professor for human computer interaction at the Institute of Visual Computing, Bonn-Rhein-Sieg University of Applied Sciences. He is also adjunct professor at SFU-SIAT in Canada. For over two decades, his research has focused at the human-factors driven analysis, design and validation of multisensory 3D user interfaces. His research looks predominantly at the usage of audio-tactile feedback methods to enhance interaction and perception within the frame of AR view management, VR navigation and hybrid 2D/3D mobile systems.

Johannes Schöning is a Lichtenberg Professor and Professor of Human-Computer Interaction (HCI) at the University of Bremen in Germany. In addition, he is the co-director of the Bremen Spatial Cognition Center (BSCC) and member of the TZI (Technologie-Zentrum Informatik und Informationstechnik) and Minds, Media, Machines (MMM). MMM is an interdisciplinary network of researchers at University Bremen, Germany. His research interests lies at the intersection between HCI, geographic information science and ubiquitous interface technologies. He investigates how people interact with digital spatial information and creates new methods and novel interfaces to help people to do so.