The Effect of Narrow Field of View and Information Density on Visual Search Performance in Augmented Reality

Christina Trepkowski*
Bonn-Rhein-Sieg University of Applied Sciences
David Eibich†
Bonn-Rhein-Sieg University of Applied Sciences
Jens Maiero‡
Bonn-Rhein-Sieg University of Applied Sciences
Alexander Marquardt§
Bonn-Rhein-Sieg University of Applied Sciences
Ernst Kruijff¶
Bonn-Rhein-Sieg University of Applied Sciences
Simon Fraser University
Steven Feiner‖
Columbia University

ABSTRACT

Many optical–see-through displays have a relatively narrow field of view. However, a limited field of view can constrain how information can be presented and searched through. To understand these constraints, we present a series of experiments that address the interrelationships between field of view, information density, and search performance. We do so by simulating various fields of view using two approaches: limiting the field of view presented on a Microsoft HoloLens optical–see-through head-worn display and dynamically changing the portion of a large tiled-display wall on which information is presented, for head-tracked users in both cases. Our results indicate a significant effect of information density and field of view on search performance, with potential search performance benefits of using a larger FOV between ca. 7–28%. Furthermore, while grids guided visual search, they did not significantly affect performance.

Index Terms: H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems-Artificial, augmented, and virtual realities

1 INTRODUCTION

Narrow field of view (FOV) optical–see-through displays are commonly used for Augmented Reality (AR). However, their limited FOV implies that only a small part of human vision can be augmented. The human visual system has a binocular FOV exceeding 210° horizontally and 150° vertically [28], while head-worn displays such as the Microsoft HoloLens (ca. 35° diagonal FOV) or the Epson Moverio BT-200 (ca. 23° diagonal FOV) cover only a subset of the human visual field (Fig. 1). Consequently, only a part of the near periphery can be used to display information, in contrast to wide-FOV displays that cover both near and partly far peripheral vision [31]. This can cause visual conflicts; for example, when information density increases, augmentations may need to be compressed inside the small display area afforded by the narrow FOV. This can eventually lead to human information processing problems; for example, in visual search, initial research indicates that narrow FOV can lead to perceptual [32] and cognitive issues [47]. However, due to the prevalence of narrow-FOV displays and the increasing information density in many AR systems [48], a better understanding is useful. In particular, view-management techniques, which deal with the layout and appearance of augmentations, could benefit from a better understanding. Specifically, techniques could be tailored to cause fewer visual conflicts when presenting denser information with a narrow FOV.

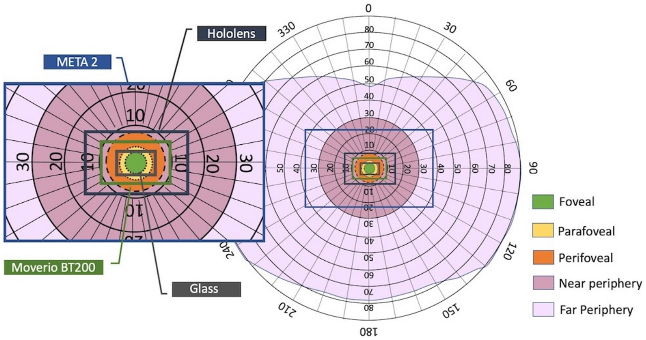

Figure 1: FOV of common narrow-FOV AR eyewear in relation to a medium-FOV display. The FOVs are centered around central vision for direct comparison purposes - e.g., Google Glass is not centered during real-life usage. Visual field image after [31].

Our studies focus on visual search of targets in different information densities. In this paper, we define information density by the overall set size and number of stimuli per area (in our case, based on a grid cell size). We focus on text labels and simple symbols that are characterized by two common features for icons: color and shape. Searching for a target item among similar items (distractors) represents a classical visual search task, and is one of the main tasks in many AR systems [48]. Visual search has been widely studied. For example, research has addressed visual attention, visual processing, and how the different areas in the visual field drive gaze shifts and fixations [37]. The human fovea covers up to 2° of visual angle [26], which should permit someone to simultaneously process and identify two consecutive words in a single eye fixation within most vertical lists, guided by visual attention [40]. Words need to be searched serially when target and distractor features are similar and the target does not pop out [51]. Words differ from symbols, as words have have both visual forms as strings of letters and semantic properties that symbols do not exhibit. After processing a non-target word, search is normally guided by features (but not semantics) in the peripheral visual field that can potentially be preprocessed. Words that have features common with the target word attract the most attention [34]. Preattentive processing normally accelerates visual search [9, 59]. The absence of such features in parts of the near and especially the far peripheral field of narrow-FOV displays means that this information cannot be processed preattentively. As such, we expect a narrower FOV to limit

---

*e-mail: christina.trepkowski@h-brs.de
†e-mail: david.eibich@smail.inf.h-brs.de
‡e-mail: jens.maiero@h-brs.de
§e-mail: alexander.marquardt@h-brs.de
¶e-mail: ernst.kruijff@h-brs.de
‖e-mail: feiner@cs.columbia.edu

visual search performance. Generally, visual search yields faster search times and fewer errors when features are much different between search target and distractor, which is also quite resistant to set size. However, once targets are similar to distractors, search times and errors increase, in direct dependence on set size [52]. On the other hand, increasing eccentricity [50] of a visual target towards the periphery can increase search times and error rates [8], decreasing performance [27].

Even though visual search has been studied widely, the effect of FOV on visual search in AR is unclear. Through our studies, we shed light on this gap in understanding, by studying the relationship between information density, FOV and search task performance. We address search task performance without any additional active guidance—we only deploy a grid that may guide search behavior passively. Targets had to be found among two levels of distractor densities. Search was performed within three levels of FOV, between very small (e.g., Google Glass) and medium (e.g., HoloLens). Such devices stimulate the parafoveal and near peripheral area in different manners (Fig. 1). FOVs were simulated with a Microsoft HoloLens and with a tiled-display wall. In all studies, the search task was accompanied by a secondary task, where users had to react to a tactile stimulus to indicate cognitive load. The secondary task was included to mimic real-world AR usage, which is often characterized by divided attention [30].

2 CONTRIBUTIONS

Through our study results, we make the following contributions:

- FOV and information density can significantly affect search performance of both symbol and text search during divided attention tasks. While the effect of information density was not unexpected, the effect of FOV is an aspect interface designers should consider. Namely, search performance improvements can be achieved between 7–28% by increasing FOV, depending on task and initial FOV, while also significantly reducing the number of overlooked targets.

- Users search differently when processing symbols and text. However, the pattern seems to be more affected by personal preference than by task or condition. We used grids (or bezels) in most of our studies, which likely guided search: search behavior was structured, and users confirmed they found grids helpful for understanding what part of space was searched already. However, grids did not significantly improve performance.

The results indicate the need to design improved view-management techniques for narrow-FOV displays, as these displays are more affected by how effectively information can be conveyed and processed. Results also motivate the development and usage of wider-FOV displays. We extend previous work by providing detailed results about using narrow-FOV, rather than wide-FOV, AR displays for search tasks, which has hardly been studied until now. We are also unaware of any structured analysis of grid usage on visual processing in AR: our results show that grids can be useful for processing more complex information.

3 RELATED WORK

View management is often essential for the design of effective AR systems, and has been studied for long [6]. Researchers have looked into label placement for position and size [4, 6], depth-placed ordering [41, 42], foreground–background issues [23], and the legibility of text [22, 36]. While view management for wide-FOV displays has found some interest [30, 31], there has little work on view management for narrow-FOV displays—even the study reported in [45] made use of a considerably larger FOV than, for example, the HoloLens.

Studies have shown that limiting **FOV** can lead to perceptual and visuomotor performance decrements in both real and virtual environments [5]. Some studies of age-related vision degradation have showed deterioration of visual-processing performance based on deficits in attentional disengagement, but not necessarily FOV degradation itself [14], indicating attention factors have a strong influence. Most studies on FOV have rather been performed with Virtual Reality (VR) or flight simulator systems instead of AR systems. It remains to be seen if and to what extent results apply to AR. In VR, FOV restrictions have been shown to degrade the ability to develop spatial knowledge and navigate [1, 16, 56], and can also result in decreased task performance in searching and locating a target by turning one's head [3]. The FOV in a CAVE simulation can also affect performance in selection tasks, depending on the input method, as an advantage of direct pointing over raycasting is negated with decreasing FOV [19]. In flight-simulator studies, pilots navigated more accurately with a wider FOV of 55° than with a 25° FOV [7]. Furthermore, sequential decrease of the effective FOV can lead to a change in visual scan pattern, an increased range of head movements and decrease in flight control performance [13]. However, a larger FOV has not always been found to be better (e.g., on path-following aviation tasks [54]).

Finally, **visual search** has been studied intensively in perception research [60]. It is affected by the types of features the search target and distractors elicit, which have been discussed in various theories [43]. Specific aspects relevant for AR such as target eccentricity, orientation [10] and depth [38] have been addressed, while search behavior itself has also found some interest in AR, for example by using eye tracking [21]. We distinguish ourselves from previous AR search-task studies in the following manner. First, our study differs from that of Ren et al. [45], which simulated wider FOVs, the smallest being $45^\circ \times 30^\circ$. Users had to find information in one of three charts, after which a link had to be followed to the related object. Results indicated that a constrained FOV resulted in slower task completion than a full FOV. It remains unclear to what extent their results are transferable to other settings, as they used a quite specific environment and information visualization technique. In contrast, we employ a classical search task in an abstract and thus more transferable setting aiming at more generalizable results. The labels in our study are presented without annotated objects. That is, we did not examine the step of relating labels and objects. Second, Kishishita et al. [30] also used wider FOVs (the smallest being $36^\circ \times 20^\circ$) and compared secondary task performance in a search task between in-view and in-situ annotations with different FOVs. They found that performance differences decreased as FOV approached 100° and beyond. Our work differs from theirs, as target search was not part of a secondary task but the main task here. We also use a much higher information density and focus on the effect of different narrow FOVs and information density instead of comparing annotation techniques. In addition, they annotate only the target item, thus partly guiding search, which we do not focus on.

4 RESEARCH QUESTIONS

RQ1: Is there an effect and possible interaction between FOV and information density on search performance?

Based on previous research [9, 59], we assumed that search performance would decrease when limiting the FOV. Limitations in peripheral visual information caused by a narrow FOV would degrade search performance, as some cues normally present with a wide FOV could not be preprocessed to guide search. Moreover, research has shown that reaction time increases with eccentricity in the near peripheral field [10]. We expected that in dependency to information density, more targets would need to be scanned, resulting in longer search times and different search behavior (as expressed by head movement) in high compared to low density conditions in our studies. Furthermore, we expected that

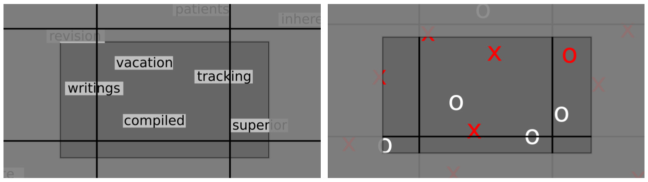

Figure 2: Schematic illustration of text search with search window performed on display wall with bezels forming grid (left) and symbol search with limited FOV on HoloLens with augmented grid lines (right). Items are displayed larger for illustrative purposes.

performance decrements that result from restricting the FOV would be stronger in symbol compared to word search. As shown in Fig. 1, different FOV displays cover different areas of the perifoveal and near periphery. Target symbol features can be processed in the periphery without being fixated [58]. In contrast, text semantic properties cannot be preprocessed in the periphery. Also, in our studies target words showed no orthographic properties to easily distinguish them from distractors, as they had the same length and heterogeneous letter compositions [34]. As such, text search likely would be affected more by set size than symbol search, based on its low search target and distractor feature differences [52]. On the other hand, we expected FOV to have less effect on text search due to lower dependency on peripheral information preprocessing.

RQ2: Is there an effect of a search grid on visual search performance and behavior in AR?

Visual stimuli on the tiled-display wall were displayed with bezels in between screens. The displays formed a grid that we expected could affect the search path as users could scan through the area grid-wise. Previous work [55] has indicated that bezels can help to segment information, although no significant effect on search time was found. However, the actual effect of bezels on user performance is not conclusive, with other studies finding benefits but also disadvantages [46]. We wanted to address if a grid would guide search performance in AR, assuming the grid would allow the participant to scan through potential targets using a more structured scan pattern when experiencing a narrow FOV, as the bezels/grid would serve as "anchor" for the search task. We expected differences to be visible in head motion behavior analysis.

5 SYSTEM AND IMPLEMENTATION

During our experiments, we made use of a tiled-display wall, and the Microsoft HoloLens. The tiled-display wall is a high-resolution curved display, consisting of 35 Full-HD displays (46 inch) connected at an offset of 10° (row-wise) to create the curvature. The display wall covered 160° of the participants visual field, while the display space measured 7 x 3 meters. The experiment was implemented in Unity3D version Unity 4.6.0f3. The experiments for the HoloLens were implemented in Unity3D version Unity 2018.1.0f2 using the Microsoft Mixed Reality Toolkit. Retroreflective markers were mounted on the head-worn devices (HoloLens and bike helmet) to enable head-tracking with an ART tracking system, mounted above the tiled-display wall. Tracking was used for moving the search window. To control the application, two simple single-handed controllers were used. Participants sat on a swivel chair during the experiment.

6 STUDIES

We first performed a text and symbol search on the tiled-display wall (Study 1). Study 2 included the same text and symbol search tasks, but rendered content on the HoloLens (RQ1). Study 3 was performed at the tiled-display wall and assessed the absence of FOV restrictions (RQ1) and the effects of wearing an AR headset on search performance. We assumed that the weight of the AR headset and reflections of the see-through glasses would make search more effortful. Study 4 deployed the HoloLens, and compared grid versus no grid conditions in text and symbol search tasks (RQ2). In our studies, we simulated different FOVs. Simulating an AR environment represents one option to examine the problem of searching visual information in narrow-FOV displays independent of changes in the environment in a controlled laboratory setting that can easily be replicated [2, 29]. The validity of AR simulation has been assessed before [17, 33, 45]. In the future it can be expected that AR headsets will provide wider fields of view and that some current technical problems will be solved. Anticipating these changes, we simulated an AR environment both on the HoloLens and on a tiled-display wall. The wall doesn't have some of the usual problems with head-worn devices, including color separation or blur caused by head motion. Moreover, the wall provides us with the opportunity to perform further experiments in which wider FOVs can be compared. While a VR headset has been used to simulate AR [31], in our case we did not consider its usage as the lower pixel density would make text reading non-comparable to true AR displays.

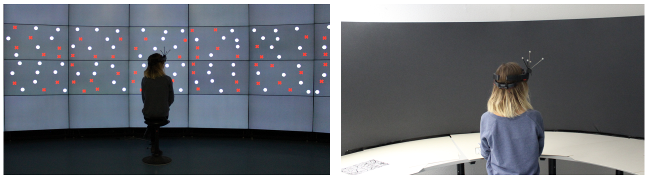

Figure 3: Tiled-wall display showing conjunction search without FOV restriction (left) and physical setup for the HoloLens studies (right).

6.1 Apparatus and stimuli

The experimental setup concerned a visual search task, where either a target text or symbol had to be found. Details about the search tasks—which were very similar—can be found in Sections 6.2 and 6.3. In each trial, visual stimuli were composed of one target among many distractors. In our experiment, we performed search tasks without any active guidance. Locations of labels were randomly determined. To avoid visual clutter and occlusion, a minimal distance between the labels and the screen bezels (or grid lines, in the HoloLens) was set. This also avoided having the text be split by screen bezels. The target position was balanced across central and peripheral areas 0–3. Areas refer to display or grid columns consisting of two displays/grid cells (Fig. 6).

We made use of the tiled-display wall in study 1 and partly in study 3. During studies, we only made use of the two middle rows of the tiled-display wall, a grid of 2 x 7 displays (Fig. 3). In the HoloLens studies, we mimicked this setup by using a 2 x 7 visual grid. Displays were at eye height. The participant was seated 1.3 meters from the center of the tiled-display wall, facing the middle row at the start of every trial. To improve comparability between the display wall and the HoloLens setup and to better simulate AR device-wearing conditions, participants also wore the turned-off HoloLens while performing most tasks at the tiled-display wall. To avoid misunderstanding, we refer to the HoloLens in the tiled-display wall studies as the "head-worn display" (HWD), while referring to "the HoloLens" only in studies where content was rendered on the HoloLens. The HoloLens was not activated, because syncing would have resulted in latency that would affect search performance. As such, we made use of external head tracking to select the target in studies 1 and 3. In study 2 and in study 4, visual stimuli were presented on the HoloLens. Visual stimuli were arranged in grid cells (2 x 7 grid, as at the tiled-display wall) along a cylindrical curved plane (radius 1.3 m) in front of the participant, to create a roughly equidistant arrangement of visual stimuli.

In study 1, on the tiled-display wall we put a mask on the array of visual stimuli so that targets and distractors were visible only in a rectangular search window. In study 1, the search window was used to simulate the same three FOVs as in study 2. In study 3, we used no FOV search window in one condition to support a search process without FOV restrictions. The search window was calculated to be the exact same size on the physical display surface as would be achieved by a view frustum in AR glasses at a constant distance (1.3m). The search window was display-aligned (with the displays at the tiled-display wall) and left constant in size to guarantee the same number of text or symbols would be visible within the search window for a specific FOV condition. The position of the search windows depended on the head orientation, so participants could move the search window to scan the whole search area by turning their head. Outside the search window on the tiled-display wall, a uniformly lit, homogeneous background in a neutral mid grey color without labels was shown. The background inside the search window was slightly darker to make the user aware of the borders (see Fig. 2). Study 3, task 1 deployed the same setup as study 1, yet without search window. In HoloLens tasks 2 and 3 of study 3 and in study 2, the neutral background was not displayed by screens but made out of a curved cardboard in mid grey, very similar to the tiled-display wall (see Fig. 3). Thus, in study 2 and tasks 2 and 3 of study 3, labels and grid lines were generated by the HoloLens, whereas the background was physical. Participants held a controller in each hand. The dominant hand was used to perform the visual search task, the other one for the secondary task.

6.2 Text Search

In the text label search task, targets and distractors were composed of English words. The font was Segoe UI, the standard font used in the HoloLens, in black with a white background. The label size was 10.5cm x 2.8cm on the wall and 7.5cm x 1.9 cm on the HoloLens. It was determined in a pilot test with five users to ensure good legibility also for words on displays in more peripheral regions. The word data pool for the textual search was generated by the English Lexicon Project word list generator and met the following conditions. Words consisted of eight letters (average in English language) and occur rather frequently with a HAL frequency between 15,000 and 20,000 per million words (average 10,000) to facilitate memorizing. HAL frequency is recommended to be used as word frequency measure (Balota et al., 2007). We also ensured that there were no words in the data pool that were too similar to each other. Target words were chosen randomly from the generated data pool of words with the restriction of always being different.

6.3 Symbol Search

In the symbol search task, target and distractors were defined through a conjunction of feature values that better reflects visual search in everyday life than single feature search. An example of conjunction search is a "green car" while a single feature search would be a "car." When working with picture labels (e.g., map icons that are associated with categories such as restaurants or tourism) a single feature often is not sufficient for categorization, so frequently two features are combined. A common combination of features includes color and shape (e.g., a red color and a cross shape stands for a medical building). In our studies, the target symbol differed from its distractors in either color or shape and had one feature in common with each distractor. The symbols that were used were red and white "X"s and red and white "O"s (similar to [53, 61]). That is, if the target was a white "X", then distractors were red "X"s and white "O"s, as each of them had one feature in common with the target . The target symbol was randomly determined. Colors and shapes were chosen for their low confusability as well as their symmetricality and similar brightness values. Symbol size was 9 cm^2 (display wall) and 6.6 cm^2 (HoloLens).

6.4 Design and Procedure

Participants were recruited via university mailing list and received a voucher for 10 euros as a reward for participation. Studies 1 and 2 took around an hour each. The same participants participated in studies 1 and 2. Between studies 1 and 2, users took a break that lasted at least one week. Study 3 took around 20 minutes and study 4 around 25 minutes.

6.4.1 Studies 1 and 2

Design and procedure were the same in studies 1 and 2. Each comprised two tasks: text and symbol search. Both studies employed a 3 x 2 factorial within-subjects design, being the factorial combination of three search window sizes (*very small*, simulating, e.g., Google Glass, but binocular instead of monocular: 14°; *small*, simulating the Epson Moverio BT-200, binocular: 23°; and *medium*, based on HoloLens, binocular: 35° [15]) and two label densities. Display degrees are defined by their diagonal FOV. In our studies, information density is defined by set size and the number of labels per m^2. The grid size of the wall-display was 0,59m^2, while the grid size of the HoloLens was 0,29m^2. The grid size for the HoloLens was adjusted to match the physical setup of our laboratory setup while favoring a similar field of regard (160°) as the tiled-wall display. Set sizes were chosen to create the same density/m^2 in both displays. In the low density condition on the tiled-display wall set size per grid cell (display) was 12 while being 6 per grid-cell on the HoloLens. In the high density condition set size per display/grid cell was 24 on the tiled-wall and 12 on the HoloLens. As a result there was approximately one label per 0,0245 /m^2 in the highest, and one label per 0,049 /m^2 in the lowest density on both displays. As we used 14 displays/grid cells, depending on study and condition users had to search through either 84 (low density in the HoloLens study), 168 (low density on the tiled-display wall, high density on the HoloLens) or 336 potential targets (high density on the tiled-display wall).

We addressed how these factors affected our dependent variables: search task performance (measured by search time), hits/errors (target found/not found), times the target was inside the search window without being noticed, visual search behavior (search path from head tracking data) and objective and self-reported workload (TDT performance, NASA-TLX).

Each factor combination was repeated four times, resulting in 24 trials per search task, randomized across conditions. As participants performed text and symbol search, studies 1 and 2 consisted of 48 trials each. Search task order was balanced across participants in both studies. Participants performed study 2 after finishing study 1. We calculated training effects to address any potential confounds.

Each trial started with the participant facing straight ahead. The target was shown at the beginning of each trial in the center location of the display system (tiled-display wall study 1) or overlaid on the grey wall (HoloLens study 2). Once it was memorized, the participant pressed the confirmation button on the controller. Then the search array and search window became visible and time measurement started. On the tiled-display wall, once the target was found and visible within the search window, the participant pressed the confirmation button again. In the HoloLens study, the target had to be gazed at using the mouse pointer shown at the center of the display. Label search time was logged and a hit was recorded when the target was within the search window or overlaid by the mouse pointer; otherwise an error was logged. The participant then reoriented towards the starting position to be shown a new target and to initiate the next trial with a new search array. Participants were told to conduct the task as quickly and accurately as possible. While working on the visual search task, participants simultaneously performed a secondary tactile detection task (TDT) by pressing a button on the controller held in the non-dominant hand. They were instructed that the visual search task should be the highest priority, followed by the TDT. Logfiles recorded the aforementioned hits or errors, search behavior (head movement), and reaction times to the secondary task, described in Section 6.4.3. In studies 1 and 2, each time after having finished a search task a (raw) NASA-TLX questionnaire [25] had to be completed by the participants to assess the subjective workload. Also, participants rated their level of agreement on 11-point Likert items (1 = "fully disagree" to 11 = "fully agree") related to task easiness, understanding, concentration, comfort, usability, and perceptual issues at the end of each study.

6.4.2 Studies 3 and 4

Study 3 was the symbol search task at the tiled-wall display and employed a 2 x 2 factorial within-subjects design to examine how a full FOV (full view versus 35° (HoloLens standard) and wearing a head mounted device (with/without HWD) affect search performance. In the full FOV condition, visual stimuli were displayed without a mask on all 14 screens (see Fig. 3). In the condition without a HWD, participants wore a lightweight helmet that was only necessary for head tracking. Conditions with and without HWD were tested blockwise, with the order balanced across participants. In each block, FOV conditions were repeated four times in randomized order resulting in 8 trials per block, that is, 16 trials in total. Finally, participants rated their level of agreement on 11-point Likert items addressing the self-evaluation of main and secondary task performance and device wearing comfort for different conditions.

Study 4 comprised the same text and symbol search task as in Studies 1 and 2 (see measures and search task procedure in Sections 6.3 and 6.4.1), It was performed on the HoloLens with the standard HoloLens FOV (35°). Visual stimuli were set at medium density (12 symbols per display). Dependent variables were search task performance, visual search behavior and self-reported workload. Employing a within-subjects design, dependent measures were compared between treatment conditions with and without grid-lines that were displayed by the HoloLens. Each condition was repeated eight times, resulting in 16 trials for each the text and symbol search task. The order of search tasks was balanced across participants. After having finished both tasks, participants provided open comments (anonymously on the PC) on their search behavior and secondary task performance in the text and symbol search.

6.4.3 Secondary task: Tactile Detection Task

The TDT is a typical detection response task that is commonly used as secondary task and shows sensitivity to cognitive workload [12, 39]. Attention demands are measured in terms of reaction times and the hit rate [18]. Tactile stimuli were transmitted via a vibration motor that was mounted on a wrist band on the left arm. Participants should react on a vibration stimulus by pressing a button on a controller held in the non-dominant hand. The vibration stimulus was triggered every 3–5 seconds with a duration of 1 second. The activation was interrupted when a reaction was carried out to provide the participant feedback to the response. The mental demand is determined as a result of the reaction times and the hit rate, in which only a response time within 200–2000 milliseconds was counted as a hit [18]. Reactions faster than 200 ms were labeled as cheats and responses slower than 2000 ms as misses. The metric hit rate is the number of hits divided by the number of stimuli [18]. The TDT method was simultaneously applied to the visual search task under different conditions.

Study		Display Wall		HoloLens	
		Text	Symbol	Text	Symbol
FOV	very small	64.4	21.5	32.2	10.9
	small	51.5	16.8	23.3	9.7
		20%	22%	28%	11%*
	medium	42.8	10.6	21.6	8.6
		17%	37%**	7%	11%
Density	low	41.5	14.2	20.4	8.7
	high	65	17.4	34	11
		-57%**	-23%**	-67%***	-26%**

Table 1: Median search times in sec in studies 1 and 2 by task, FOV and density (m^2/label). The change in performance in % compared to the next smaller field of view or density condition is shown underneath medians. P-values refer to Wilcoxon signed-rank pairwise comparisons, *= $p < .05$, **= $p < .01$, ***= $p < .001$.

7 RESULTS

The Aligned Rank Transform (ART, [57]) procedure was used to analyze the effect of factors on search time and times the target was overlooked (inside the search window without being noticed) in all studies. In study 1, we also analyzed TDT performance using ART. Repeated measures ANOVA was not used due to violations of the normality assumption. Wilcoxon signed-ranks tests were applied for post-hoc pairwise comparisons and to compare questionnaire ratings. Pearson's correlation coefficient r was used to measure effect size. We only report on the salient results.

Regarding search time, median values were computed instead of means as they are not affected by extreme values. Extreme search time values in our studies mainly resulted from starting the search in the wrong direction and/or overlooked targets that were already considered as separate dependent measure. The number of hits approximated 100% in all studies, showing that without time constraints participants searched until the target was found instead of skipping trials.

7.1 Results of Study 1

7.1.1 Performance data

Sixteen users participated in this study (3 female, mean age = 27.44, SD = 9.26). Half wore glasses or contact lenses, the other half had normal vision. Half of the users had never used AR-glasses before (50%). Some of the participants played video-games daily (44%) and primarily online on the PC (50%).

In both search tasks, there was no interaction effect between FOV and density on performance. In the **text** search task search time was affected by label density (F(1,363) = 15.42, $p < .001$) as participants needed more time at high than at low density (see Table 7.1.1). There was only a small tendency to an effect of the FOV (F(2,363) = 2.55, $p = .08$). The number of times the target was overlooked was not affected by label density or FOV.

When searching for **symbols** search time was affected by density (F(1,363) = 7.63, $p < .01$) and also the FOV (F(2,363) = 17.4, $p < .001$). It took longer to find the target at high compared to low density and longer with the very small or small FOV compared to the medium FOV (Z = -3.21, $p = .001$, r = 0.80; Z = -2.69, $p < .01$, r = 0.67). The number of times the target was overlooked was higher at high density (F(1,363) = 9.13, $p < .01$) and also affected by the FOV ((F(2,363) = 10.29, $p < .001$). The target was missed more often with the smallest FOV compared to the medium FOV (Z = - 2.51, $p < .05$, r = 0.63).

To sum up, as can be seen in Table 7.1.1, reducing the FOV from medium (35°) to small (23°) or increasing label density resulted in longer search times in the text and symbol search tasks. A further reduction in the FOV to very small (12°) did not affect search time significantly. The FOV affected times the target was overlooked only in the symbol search task: targets were overlooked more often when the FOV was restricted from medium (35°) to very small (12°). Label density did not affect the number of overlooked targets. As studies 1 and 2 were always performed in the same order, we examined possible training effects in each search task in study 1. The trial number was correlated with performance measures (using Spearman's rho) for all trials and for trials by condition but no significant correlation was found.

For the TDT, the hit rate (see 6.4.3) was computed for each trial as measure for secondary task performance. ART analysis

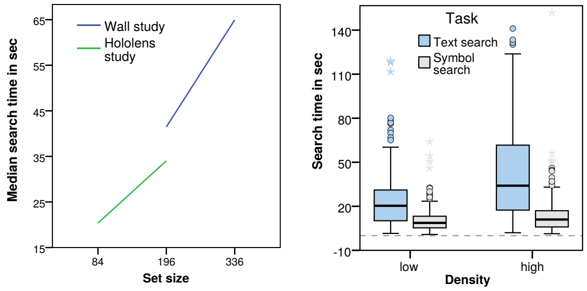

Figure 4: Left: Median search time in the text search task by set size for studies 1 and 2. Right: Boxplots of search time by density in m^2 per label for text and symbol search tasks performed on the HoloLens in study 2. The scattered distribution is very similar in all tasks in Studies 1 and 2.

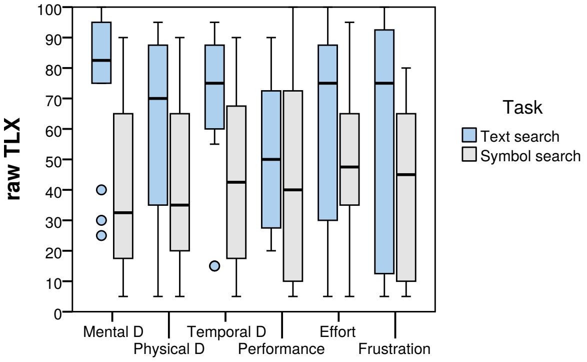

Figure 5: Study 1: Boxplots of raw NASA-TLX scores of each subscale (0-100) for the text and symbol search task (D = demand).

showed no effect of density or the FOV on the hit rate and also no interaction. Mean hit rate was generally between 0.77 and 0.76 (SD between 0.21 and 0.23). Based on these results, we kept the TDT in studies 2 and 3 to keep consistency (divided attention task) between studies, but did not further analyse the TDT results. For more information, refer to the discussion section.

7.1.2 Questionnaire ratings

Raw NASA-TLX ratings were computed for each subscale and compared between the text and symbol search task. Subjective workload was rated significantly higher for the text search on each subscale except of performance (see Fig. 5). Statements concerning task easiness, understanding, concentration, usability, and perceptual issues were rated positive in general. However, wearing the HoloLens over time was rather uncomfortable (Median = 5). Comparing ratings between text and symbol search showed that it was slightly easier to remember target symbols than words ($Z = -2.75$, $p < .01$, $r = 0.69$) and also easier to maintain concentration during the symbol search compared to the text search task ($Z = -3.32$, $p = .001$, $r = 0.83$). The text search was rated as easier with a medium and small FOV compared to the very small one ($Z = -2.21$, $p < .05$, $r = 0.55$; $Z = -2.71$, $p < .01$, $r = 0.68$). In the symbol search, the medium or small FOV conditions also got higher ratings than the very small FOV ($Z = -3.47$, $p = .001$, $r = 0.87$; $Z = -3.54$, $p < .001$, $r = 0.89$). Easiness ratings for the medium and small FOV condition showed no significant differences between each other in study 1.

7.1.3 Search behavior

To analyse head movement behavior, we plotted head movements against the display/grid cells, and searched for dominant behavior in all performed trials. Figure 6 shows representative samples. Users actively used the display grid to search through the stimuli. This occurred both display-wise (hence, display after display) or scanning, in which either horizontal (row) or vertical (column) head movements were made that were not constrained by search within a single grid cell. Not all users used either search pattern: some mixed the patterns, which sometimes occurred at random, sometimes a clear switch occurred between two modes (especially scan). Also, we could not find a clear effect of information density on head motion. Of particular interest is also the change in behavior between symbol and text search. Text search had a predominantly display-wise search mode, whereas for symbol search the behavior varied much more (see Table 4). Within the displays, search patterns were both row and column-wise. Finally, while we expected FOV could potentially affect search pattern, we could not detect any clear differences.

Statement	Median level of agreement (IQR) Study 1: Wall	 Study 2: HoloLens
Ease of remembering the target word	8.5 (2)	8.0 (3)
Ease of remembering the target symbol	10 (1)	11 (2.5)
Sitting comfort	8 (2)	9 (3.5)
HoloLens wearing comfort	5 (3.5)	6 (5)
Ease of understanding the task	11 (2.5)	11 (1)
Concentration during text search	6 (5.5)	8 (5)
Concentration during symbol search	10 (2.75)	10.5 (2)
Legibility of labels on peripheral displays	8 (4.5)	9 (5.)
Ease of using the interface	11 (2)	10.5 (2)
Fun	9 (4.75)	8.5 (3)
General usability	9 (2)	8.5 (3.75)

Table 2: Median level of agreement and interquartile ranges (IQR) of statements on task understanding, execution and usability rated on a 11-point Likert scale for studies 1 and 2.

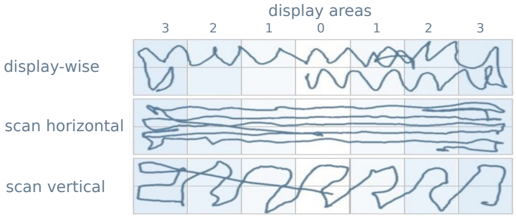

Figure 6: Exemplary head movement paths of typical search patterns

7.2 Results of Study 2

7.2.1 Performance data

In study 2, the same participants took part as in study 1. In the **text** search task search time was affected by label density (F(1,363) = 42.15, $p < .001$) and FOV (F(2,363) = 6.5, $p < .01$). Participants needed more time at high than at low density and more time with the very small FOV compared to the medium FOV ($Z = -2.43$, $p < .05$, $r = 0.61$). The number of times the target was overlooked was not affected by label density or the FOV.

When searching for **symbols** search time was also affected by density (F(1,363) = 8.5, $p < .01$) as participants were faster at low compared to high density. Boxplots of search time by density for both search tasks show a less scattered distribution of values in the symbol search task (see Fig. 4). The FOV also affected search time (F(2,363) = 4.92, $p < .01$) as targets were found faster with a medium or small FOV compared to the very small FOV ($Z = -2.17$, $p < .05$, $r = 0.54$; $Z = -2.02$, $p < .05$, $r = 0.51$). Table 7.1.1 provides an overview of median search times for different conditions and significant differences. Furthermore, in the symbol search task the target was more often overlooked at high compared to low density (F(1,363) = 12.98, $p < .001$) whereas the FOV had no effect.

7.2.2 Questionnaire ratings

As in study 1, NASA-TLX ratings were significantly higher in the text search task than in the symbol search task on all subscales except for the performance subscale ($Z = -0.52$, $p = .6$). In both search tasks, median task easiness was highest for the medium FOV, followed by the small and very small FOV (see Table 3). In contrast to study 1, ratings of the medium and small FOV differed significantly from each other. Other Likert items were rated similarly positive as in study 1 (see Table 3).

	Wall		HoloLens	
	Text search	Symbol search	Text search	Symbol search
very small	4 (4.75)*	4 (4.5)**	4 (3)**	6 (5.5)**
small	7 (3.75)**	8.5 (3.5)***	6 (1.75)**	8 (2.75)**
medium	7.5 (5)	11 (1)	9 (2.75)**	10 (1.75)**

Table 3: Median task easiness and interquartile ranges (IQR) for different FOVs in the text and symbol search task on a 11-point Likert scale for studies 1 and 2. P-values refer to the comparison with the next smaller FOV using Wilcoxon, very small is compared to medium. *= $p < .05$, **= $p < .01$), ***= $p < .001$.

7.2.3 Search behavior

Search behavior in the HoloLens showed different patterns in comparison to the tiled-display wall study (see Fig. 4). Text search was predominantly display (grid-)based, with few mixed modes, and little preference for scanning horizontally or vertically. For symbol search, most users used vertical scanning, followed by switched scan patterns. Generally, the modes were far more dispersed than on the tiled-display wall.

7.3 Results of Studies 3 and 4

7.3.1 Performance data

Twelve people participated in study 3 (3 female, mean age = 35.42, SD = 14.18). Time differences between the condition without and with HWD (wearing a helmet versus turned-off HoloLens) were not significant, whereas participants found the target symbol much faster with a full FOV (Median = 4.36, IQR = 3.6) than with the search window (Median = 10.64, IQR = 12.08), F(1,369) = 86.03, $p < .001$. That is, the median search time was improved by 59% with a full FOV. The number of times the target was overlooked could only be compared between HWD and helmet in trials with the restricted FOV. Participants did not overlook more targets with the HWD. Another 12 people performed study 4 (3 female, mean age = 34.25, SD = 14.18). The grid that was displayed by the HoloLens did not influence search time, either in the text or in the symbol search task, or the number of overlooked targets.

7.3.2 Questionnaire ratings

In study 3, participants' self-rated performance was generally high (Median > 8) and matched objective measures as it was higher for the full FOV condition compared to the restricted FOV ($Z = -2.84$, $p < .01$, $r = 0.82$). Wearing the helmet was judged as more comfortable than wearing the HoloLens ($Z = -1.97$, $p < .05$, $r = 0.57$). Participants thought they performed equally well on the secondary task in different conditions.

In study 4, almost all participants (11/12) reported that the grid affected their search behavior. Searching grid by grid was , as it was easier to remember which part of the search area was already scanned. The majority (8/11) stated performing the secondary task was easier during symbol search, while the others perceived it as equally easy in both tasks. All participants stated the secondary task performance was not affected by the grid.

	Condition	Display	Scan-hor	Scan-ver	Scan-switch	Mixed
Study 1:	Text	81.3%	6.3%	0%	6.3%	6.3%
Display wall	Symbol	37.5%	6.3%	0%	12.5%	43.75%
Study 2:	Text	56.25%	0%	12.50%	6.25%	25.00%
HoloLens	Symbol	6.25%	12.5%	56.25%	25.00%	6.25%
Study 3:	Full FOV	0%	91.67%	0%	8.3%	0%
Display Wall	limited FOV	50%	16.67%	0%	12.5%	20.83%
Study 4: HoloLens	Grid-Symbol	0.0%	58.3%	8.3%	16.7%	16.7%
	No Grid-Symbol	0.0%	50.0%	8.3%	25.0%	16.7%
	Grid-Text	8.3%	8.3%	33.3%	16.7%	33.3%
	No Grid-Text	8.3%	8.3%	25.0%	25.0%	33.3%

Table 4: Percentages of users performing search in a specific pattern ("hor" = horizontal, "ver" = vertical).

7.3.3 Search behavior

The grid versus no-grid search tasks in study 4 revealed that users hardly changed their search pattern between grid and no-grid conditions. Rather, they seemed to stick to a certain pattern they may have liked or believed effective. In contrast to the tiled-display wall study, the grid did not result in a clear grid-wise search pattern, though scanning patterns seemed to follow the grid-structure. These findings did not correspond to the self-reported grid-wise search behavior in the post-hoc questionnaire. Interestingly, when comparing the occurrence of search patterns between corresponding conditions across studies, the percentages of users performing specific patterns differed quite a lot (Table 4). This also provides evidence the usage of a specific search pattern depended rather on user preferences than conditions. In contrast, head-movement data from study 3 showed that almost all participants used the horizontal search pattern in the full FOV condition (92%), whereas with a restricted FOV search behavior was again more diverse.

8 DISCUSSION

RQ1: Is there an effect and possible interaction between FOV and information density on search performance?

Our results confirmed text and symbol search were significantly affected by **information density** in the tiled-display wall and HoloLens studies (studies 1 and 2). Relative performance difference between symbol and text search, and set size effects resemble what is frequently found in perception research [37, 58]. While set size differed between studies 1 and 2, densities were comparable between low tiled-display and high HoloLens density conditions regarding area taken up per label. We note that our highest set size is likely higher than the usual density displayed in many AR systems. However, many systems (e.g., phones) often need to compress many labels into the display area, showing similar densities as as our search tasks, defined by distance between labels and overall number of labels seen on the screen [32]. Regarding Fig. 4, note that the set size functions are quite similar between the tiled-display wall and the HoloLens. Set-size search time slopes are typically linear for conjunction searches [58], which gives us the possibility to compare the slopes. However, search times were lower with the HoloLens (at a comparible density). Furthermore, the lower slope seems to indicate that the HoloLens is somewhat more resistant to information density. This is somewhat surprising, as the HoloLens tends to have more technical limitations than the tiled-display (e.g., rendering latency that can be experienced as smearing). We also used the mouse pointer instead of the search window (tiled-display wall) to select, which may have been slightly slower. Further research is required to explain the slightly better slope for the HoloLens. As we did no direct statistical comparison, it would be interesting to aim at perfectly matching setups and to finally compare performance measures directly. With respect to the nature of the search tasks, as we only displayed labels on a cylindrical, single depth, it will be interesting to see how search performs with labels at different depths, and to what extent object-label mapping—the step after finding the label—is affected by visual clutter. Extended understanding of depth and object-label mapping could provide further input to view-management techniques.

Regarding **FOV**, we expected text to be less affected by a smaller FOV than symbols, as text search would be potentially less affected by the lack of peripheral cues. As it turned out, the effect of FOV was significant in symbol search tasks and text (HoloLens). While the effect of FOV was not significant in the wall study, there was a tendency toward an effect. Also, targets were missed significantly more often with the smallest compared to the medium FOV. Generally, as we showed in the search times in Table 7.1.1, increasing FOV can significantly improve performance: even when comparing medium to small FOVs, performance improvements of around 20% could often be achieved. In comparison, we showed search without FOV restrictions was around 59% faster than medium FOV, showing the relevance of developing wide-FOV displays. In both wall and HoloLens studies, text was unaffected by density and FOV with respect to overlooked targets. However, symbol search was affected in both cases, although at the HoloLens only in dependency with density. Search performance improvements by increasing FOVs extends previous findings [45] showing similar results for wider FOVS. However, it remains to be seen to what extent text search can be improved by increasing FOV, as previous work has indicated that the far peripheral visual field may not be useful for preprocessing text features [44].

With respect to **cognitive measures**, interestingly we could not find any significant difference between different task conditions from our TDT measures in study 1. However, NASA-TLX subjective measures showed text search being far more demanding than conjunction search: generally, conjunction search was found to be much easier. This difference was expected due the uniformity between target and distractors of text in contrast to symbols, which was also represented in search time. Thereby, it should be said that task duration does not necessarily result in higher mental load ratings in NASA-TLX [24]. Based on the inconsistent findings, we made use of the TDT in the other studies to maintain divided attention tasks, but refrained from deriving any direct assumptions from the TDT data themselves. Generally, studies 2, 3, and 4 produced very similar results to study 1 with respect to cognitive load, indicating text was more demanding than symbol search.

Head-movement behavior differed per task, user, and display medium. While we found that users made use of the grid cells as guidance to structure their search, their behavior differed. While in studies 1 and 2, users adopted different search patterns for text and symbol search, we could not replicate this result in study 4. However, the difference may have been caused by switching between grid and no-grid conditions. Currently, it seems search pattern is somewhat based on personal preference instead of clearly set by task characteristics. In part, generally used patterns may not be usable. For example, previous work has shown that the layout of words in lines or columns within the search display constrains the order in which participants look at words. However, we could not detect such behavior, likely because of our randomized, instead of structured, presentation of labels [11, 35, 40]. What will be of direct interest to more closely assess search behavior is eye movement. Eye and head movements are complementary during visual search [49], where cycles of visual processing (fixation) and data acquisition (gaze shifts) occur [20]. As such, it will be interesting to assess the relationship between head movement and saccades during eye–head coordination using an eye tracker (e.g., to measure potential timing differences between head and eye movements based on different FOVs). Moreover, fixation and gaze shifts are also synced with body movement [20]. As such, it will be interesting to contrast our results with a divided-attention task where people move through an environment and also receive nonsearch task related visual cues, similar to [31].

RQ2: Is there an effect of a search grid on visual search performance and behavior in AR?

In tiled-display systems, bezels have not been found to significantly affect search performance, though they can potentially support structured segmentation [55]. Head-motion behavior analysis in both tiled-display wall and HoloLens studies indicated that users seemed to actively use the features of the bezels, by either searching within a grid cell (display), or using bezels as guidance to scan through the visual stimuli. When comparing grid versus no grid in the HoloLens, the grid cells (or absence thereof) did not affect search performance significantly, in line with previous findings. However, users found them useful to direct search. Strangely enough, we could not detect any change in head-movement behavior between search with and without grid. Perhaps, users explored patterns during grid conditions and reused them in no-grid situations. This issue warrants further research. It would also be interesting to follow up the role of the grid in tasks where labels do not reside at a constant depth, but at different depths, referring to objects in the scene. Here, the potential of 2D or 3D grids can be further assessed.

9 CONCLUSIONS AND OUTLOOK

We presented the results of a series of studies addressing visual search tasks in various narrow FOVs. Visual search—independent of FOV—of text labels and symbols generally adhered to previous work, finding similar effects. However, our results provide new insights into the effect of FOV on search performance. To our knowledge, this is the first study to systematically explore narrow-FOV AR search tasks. We showed that increasing FOV significantly affects search performance, with performance increases of 7–28% over the next smaller FOV. We also illustrated how a grid can subjectively aid search. The impact of our results cannot only drive the choice of an adequate FOV based on application specifications, but can also inform the design of novel view-management methods that can potentially deal better with higher densities in narrow-FOV displays. Such improvements may, for example, take the shape of hierarchical information representations, or perhaps the reduction of visual stimuli in favor of other sensory channels. We expect that, in particular, the way view management deals with visual clutter of labels at different depths will affect visual search [32]—comparing different methods could be of high value.

More work is needed to better understand other factors that likely affect search performance in narrow-FOV displays. First, head movement and eye movement should be studied together, as we expect that different FOVs will have different effects not only on head movement, but also on saccades. Further, as we noted in the discussion, we did not vary label depth or ask labels to be matched to objects. Both factors likely will affect visual processing, and warrant further research. Subsequent studies should also address the effect of search patterns on performance, as people used different strategies that could potentially affect search time. Another factor of interest is to assess search performance in real-world conditions, to determine the possible effect of the background on performance. Finally, while we studied narrow FOV, a further comparison to medium-FOV displays such as the DAQRI Smart Glasses or the Meta 2 will be useful to address to what extent further performance improvements for visual search could be achieved.

ACKNOWLEDGEMENTS

This work was partially supported by the Deutsche Forschungsgemeinschaft (grants KR 4521/2-1) and the National Science Foundation (grant IIS-1514429).

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] J. Arthur, L. Prinzel, K. Shelton, L. Kramer, S. Williams, R. Bailey, and R. Norman. Design and testing of an unlimited field-of-regard synthetic vision head-worn display for commercial aircraft surface operations. In *Enhanced and Synthetic Vision 2007*, volume 6559, page 65590E. International Society for Optics and Photonics, 2007.

[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] 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, pages 66–, Washington, DC, USA, 2003. IEEE Computer Society.

[5] J. Baumeister, S. Ssin, N. ElSayed, J. Dorrian, D. Webb, J. Walsh, T. Simon, A. Irlitti, R. Smith, M. Kohler, and B. Thomas. Cognitive cost of using augmented reality displays. *IEEE Transactions on Visualization and Computer Graphics*, 23(11):2378–2388, Nov 2017.

[6] 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, pages 101–110, New York, NY, USA, 2001. ACM.

[7] M. Brickner and D. Foyle. Field of view effects on a simulated flight task with head-down and head-up sensor imagery displays. In *Proceedings of the Human Factors Society Annual Meeting*, volume 34, pages 1567–1571. SAGE Publications Sage CA: Los Angeles, CA, 1990.

[8] M. Carrasco, D. Evert, I. Chang, and S. Katz. The eccentricity effect: Target eccentricity affects performance on conjunction searches. *Perception & Psychophysics*, 57(8):1241–1261, Nov 1995.

[9] M. Carrasco, B. McElree, K. Denisova, and A. Giordano. Speed of visual processing increases with eccentricity. *Nature Neuroscience*, 6:899–899, 2003.

[10] M. Carrasco, T. McLean, S. Katz, and K. Frieder. Feature asymmetries in visual search: Effects of display duration, target eccentricity, orientation and spatial frequency. *Vision Research*, 38(3):347 – 374, 1998.

[11] T. Colombi and T. Baccino. Le rôle de la mise en page et de la structure syntaxique dans la sélection des liens hypertextuels. *Le travail humain*, 66(1):45–64, 2003.

[12] A. Conti, C. Dlugosch, R. Vilimek, A. Keinath, and K. Bengler. An assessment of cognitive workload using detection response tasks. *Advances in human factors and ergonomics series. Advances in human aspects of road and rail transportation*, pages 735–743, 2012.

[13] 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.

[14] J. D Cosman, M. Lees, J. Lee, M. Rizzo, and S. Vecera. Impaired attentional disengagement in older adults with useful field of view decline. *The journals of gerontology. Series B, Psychological sciences and social sciences*, 67:405–12, 11 2011.

[15] S. Delabrida, A. Loureiro, T. D'Angelo, R. Oliveira, B. Thomas, E. Carvalho, and M. Billinghurst. A low cost optical see-through hmd - do-it-yourself. *2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*, pages 252–257, 2016.

[16] H. Dolezal. *Living in a world transformed: Perceptual and performatory adaptation to visual distortion*. Academic Press, 1982.

[17] S. R. Ellis, F. Breant, B. Manges, R. Jacoby, and B. D. Adelstein. Factors influencing operator interaction with virtual objects viewed via head-mounted see-through displays: viewing conditions and rendering latency. In *Virtual Reality Annual International Symposium, 1997., IEEE 1997*, pages 138–145. IEEE, 1997.

[18] J. Engström. The tactile detection task as a method for assessing drivers' cognitive load. *Performance Methods for Assessing Driver Distraction: The Quest for Improved Road Safety*, pages 90–103, 2010.

[19] 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*, pages 107–110. ACM, 2016.

[20] Y. Fang, R. Nakashima, K. Matsumiya, I. Kuriki, and S. Shioiri. Eye-head coordination for visual cognitive processing. *PLOS ONE*, 10(3):1–17, 03 2015.

[21] K. Fujii, J. Totz, and G.-Z. Yang. Visual search behaviour and analysis of augmented visualisation for minimally invasive surgery. In C. A. Linte, J. T. Moore, E. C. S. Chen, and D. R. Holmes, editors, *Augmented Environments for Computer-Assisted Interventions*, pages 24–35, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.

[22] J. Gabbard, J. 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.

[23] 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, pages 177–186, Washington, DC, USA, 2012. IEEE Computer Society.

[24] S. Haga, H. Shinoda, and M. Kokubun. Effects of task difficulty and time-on-task on mental workload. *Japanese Psychological Research*, 44(3):134–143.

[25] S. G. Hart and L. E. Stavenland. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock and N. Meshkati, editors, *Human Mental Workload*, chapter 7, pages 139–183. Elsevier, 1988.

[26] M. Jacob and S. Hochstein. Graded recognition as a function of the number of target fixations. *Vision research*, 50(1):107–117, 2010.

[27] M. Juettner and I. Rentschler. Scale-invariant superiority of foveal vision in perceptual categorization. *European Journal of Neuroscience*, 12(1), 2008.

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

[29] K. Keller and H. Fuchs. Simulation-based design and rapid prototyping of a parallax-free, orthoscopic video see-through head-mounted display. In *Proceedings of the 4th IEEE/ACM International Symposium on Mixed and Augmented Reality*, pages 28–31. IEEE Computer Society, 2005.

[30] 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 *Mixed and Augmented Reality (ISMAR), 2014 IEEE International Symposium on*, pages 177–186. IEEE, 2014.

[31] 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*, pages 1–1, 2018.

[32] E. Kruijff, J. Swan II, and S. Feiner. Perceptual issues in augmented reality revisited. In *ISMAR*, pages 3–12. IEEE Computer Society, 2010.

[33] C. Lee, S. Bonebrake, D. A. Bowman, and T. Höllerer. The role of latency in the validity of ar simulation. In *Virtual Reality Conference (VR), 2010 IEEE*, pages 11–18. IEEE, 2010.

[34] L. Léger, J.-F. Rouet, C. Ros, and N. Vibert. Orthographic versus semantic matching in visual search for words within lists. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 66(1):32, 2012.

[35] L. Léger, C. Tijus, and T. Baccino. Spatial layout and words detection. *Le travail humain*, 69(4):349–377, 2006.

[36] 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*, pages 224–230, Nov 2004.

[37] B. McElree and M. Carrasco. The temporal dynamics of visual search: evidence for parallel processing in feature and conjunction searches. *Journal of experimental psychology. Human perception and performance*, 25(6):1517–39, 1999.

[38] E. McSorley and J. Findlay. Visual search in depth. *Vision Research*, 41(25):3487 – 3496, 2001.

[39] N. Merat and A. H. Jamson. The effect of stimulus modality on signal

detection: Implications for assessing the safety of in-vehicle technology. *Human Factors*, 50(1):145–158, 2008.

[40] H. Ojanpää, R. Näsänen, and I. Kojo. Eye movements in the visual search of word lists. *Vision Research*, 42(12):1499–1512, 2002.

[41] 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, pages 143–152, Washington, DC, USA, 2008. IEEE Computer Society.

[42] 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*, pages 169–176. IEEE Computer Society, 2008.

[43] P. Quinlan. Visual feature integration theory: Past, present, and future. *Psychological bulletin*, 129:643–73, 10 2003.

[44] E. Rantanen and J. Goldberg. The effect of mental workload on the visual field size and shape. *Ergonomics*, 42(6):816–834, 1999.

[45] 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*, pages 93–102. IEEE, 2016.

[46] G. Robertson, M. Czerwinski, P. Baudisch, B. Meyers, D. Robbins, G. Smith, and D. Tan. The large-display user experience. *IEEE Computer Graphics and Applications*, 25(4):44–51, July 2005.

[47] E. Sabelman and R. Lam. The real-life dangers of augmented reality. *IEEE Spectrum*, 52(7):48–53, 2015.

[48] D. Schmalstieg and T. Höllerer. *Augmented Reality - Principles and Practice*. Addison-Wesley Professional, 2016.

[49] G. J. F. Solman, T. Foulsham, and A. Kingstone. Eye and head movements are complementary in visual selection. 4(1):160569, 2017.

[50] C. Staugaard, A. Petersen, and S. Vangkilde. Eccentricity effects in vision and attention. *Neuropsychologia*, 92:69 – 78, 2016. Spatial and non-spatial aspects of visual attention: interactive cognitive mechanisms and neural underpinnings.

[51] T. Thornton and D. Gilden. Parallel and serial processes in visual search. *Psychological review*, 114(1):71, 2007.

[52] A. Treisman. The perception of features and objects. In A. Baddeley and L. Weiskrantz, editors, *Attention: Selection, Awareness and Control*, pages 5–35. Clarendon Press, 1993.

[53] A. M. Treisman and G. Gelade. A feature-integration theory of attention. *Cognitive psychology*, 12(1):97–136, 1980.

[54] J. Turner. Evaluation of an eye-slaved area-of-interest display for tactical combat simulation. In *The 6th interservice/industry training equipment conference and exhibition*, pages 75–86, 1984.

[55] J. Wallace, D. Vogel, and E. Lank. Effect of bezel presence and width on visual search. In *Proceedings of The International Symposium on Pervasive Displays*, PerDis '14, pages 118:118–118:123, New York, NY, USA, 2014. ACM.

[56] 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*, volume 1116, pages 126–138. International Society for Optics and Photonics, 1989.

[57] J. Wobbrock, L. Findlater, D. Gergle, and J. Higgins. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 143–146, New York, NY, USA, 2011. ACM.

[58] J. Wolfe. Visual search. In H. Pashler, editor, *Attention*, pages 13–73. Taylor & Francis, 1998.

[59] J. Wolfe. Guidance of visual search by preattentive information. In L. Itti, G. Rees, and J. Tsotsos, editors, *Neurobiology of Attention*, pages 101–104. 12 2005.

[60] J. Wolfe and T. Horowitz. What attributes guide the deployment of visual attention and how do they do it? *Nature reviews. Neuroscience*, 5:495–501, 07 2004.

[61] J. M. Wolfe, K. R. Cave, and S. L. Franzel. Guided search: an alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human perception and performance*, 15(3):419, 1989.