The Influence of Label Design on Search Performance and Noticeability in Wide Field of View Augmented Reality Displays

Ernst Kruijff, Jason Orlosky, Naohiro Kishishita, Christina Trepkowski, and Kiyoshi Kiyokawa,
member IEEE

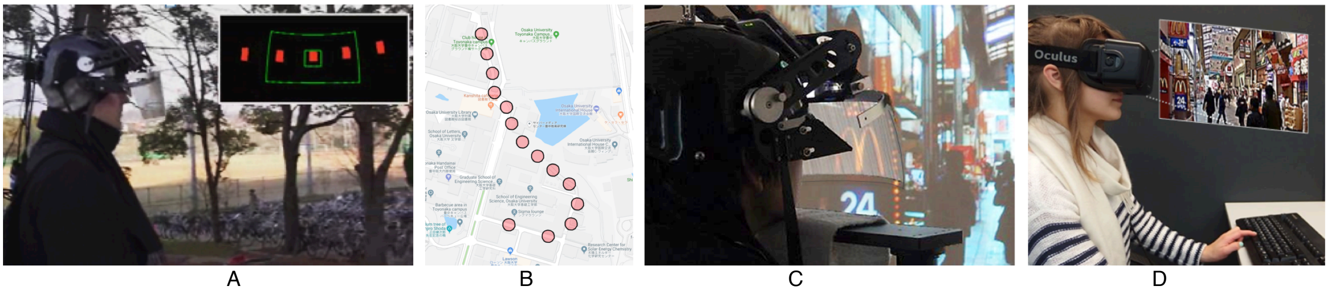

Fig. 1. A) An experiment participant with our wide field of view optical see-through display navigating an outdoor environment in our mobile search task, B) the movement path for the approximately 800 meter-long search task, conducted on a college campus, C) an experiment participant with the optical see-through display evaluating virtual label characteristics on dynamic, projected backgrounds, and D) emulation of optical see-through in virtual reality to evaluate characteristics in video see-through modes.

Abstract— In Augmented Reality (AR), search performance for outdoor tasks is an important metric for evaluating the success of a large number of AR applications. Users must be able to find content quickly, labels and indicators must not be invasive but still clearly noticeable, and the user interface should maximize search performance in a variety of conditions. To address these issues, we have set up a series of experiments to test the influence of virtual characteristics such as color, size, and leader lines on the performance of search tasks and noticeability in both real and simulated environments. We evaluate two primary areas, including 1) the effects of peripheral field of view (FOV) limitations and labeling techniques on target acquisition during outdoor mobile search, and 2) the influence of local characteristics such as color, size, and motion on text labels over dynamic backgrounds. The first experiment showed that limited FOV will severely limit search performance, but that appropriate placement of labels and leaders within the periphery can alleviate this problem without interfering with walking or decreasing user comfort. In the second experiment, we found that different types of motion are more noticeable in optical versus video see-through displays, but that blue coloration is most noticeable in both. Results can aid in designing more effective view management techniques, especially for wider field of view displays.

Index Terms—Augmented Reality, Head Mounted Display, Perception, Peripheral Vision, Visualization.

◆

1 INTRODUCTION

DESIGNING applications that make effective use of available screen space for information dense environments is still a significant challenge in augmented reality (AR) applications. However, only recently have researchers started to take interest in issues related to peripheral vision and AR to expand and overcome limitations of more commonplace displays with a smaller field of view (FOV). There is still a great deal of uncertainty as to what view management characteristics will improve search performance without burdening the user. Most research has focused on managing occlusion or display of sets of labels in central vision, but experimentation and concrete evidence on perception and performance of labels in the peripheral FOV [1] is still largely lacking. The extended space provided by a wider FOV can be used to display more information depicted by augmentations and afford an uncluttered layout thereof. The additional room can also be used to serve other purposes such as showing additional information like warnings or system control elements.

In this paper, we closely look into characteristics such as label placement, type, color preference and motion perception that are important for the design of augmentations displayed in the periphery – and thus, overall view management - for example to improve search performance. We want to delve down into the exact features that are most noticeable for users, and to do so we have designed two experiments to test these effects, each of which employs an optical see-through (OST) wide FOV display, next to an Oculus Rift DK2 that simulates a video see-through (VST) display, as shown in Figure 1.

The first experiment, a real-world outdoor experiment (dynamic search performance), was conducted with the OST display to test some of the most important characteristics related to FOV, and to determine what types of indicators could improve target acquisition without increasing

---

- *Ernst Kruijff is with the Bonn-Rhein-Sieg University of Applied Sciences and Simon Fraser University. Grantham-Allee 20, 53757 Sankt Augustin, Germany. Email: ernst.kruijff@h-brs.de*
- *Jason Orlosky is with Osaka University. 1-32 Machikaneyama, Toyonaka, Osaka 560-0043, Japan. Email: orlosky@lab.ime.cmc.osaka-u.ac.jp*
- *Kishishita Naohiro is with Osaka University*
- *Christina Trepkowski is with the University of Bonn-Rhein-Sieg University of Applied Sciences. Grantham-Allee 20, 53757 Sankt Augustin, Germany. Email: christina.trepkowski@h-brs.de*
- *Kiyoshi Kiyokawa is with Nara Institute of Science and Technology. 8916-5 Takayama-cho, Ikoma, Nara 630-0192 Japan. Email: kiyo@is.naist.jp*

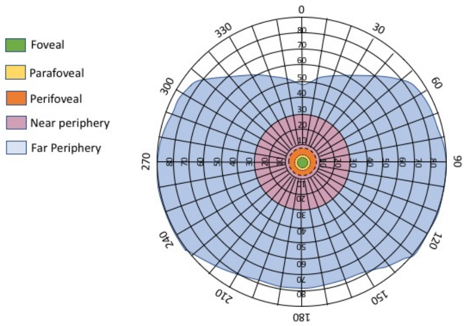

Fig. 2. The different areas in central and peripheral vision. Redrawn from [16].

distraction. Once we found that FOV affected user performance in both seated and mobile search tasks, we then wanted to figure out exactly which label characteristics might affect the user's ability to notice information. To do so, the second experiment (label noticeability) was designed to test which label characteristics were most noticeable and consisted of two sub-studies. It was deliberately performed indoors to control lighting/background issues in an exact manner, as this is almost impossible in an outdoor experiment, since lighting was expected to have a high impact on the results. Noticing a label often drives search behavior by drawing attention towards certain object features [2], meaning improper design or placement can result in missing essential information. Comparing two hardware platforms in controlled environments also allowed us to examine how differences in brightness and optical characteristics affected user perception and preference.

Our contributions in this paper include the experiments with several key findings, more specifically:

- **dynamic search performance**: the dynamic, mobile search experiment focused on determining which types of labels are most effective at improving peripheral search performance (in particular discovery rate) in outdoor AR, showing that decreasing FOV negatively impacts search performance, and
- **label noticeability**: the more controlled experiment with OST and VST displays on dynamic video backgrounds looked into noticeability of label design characteristics, and showed clear choices for all label design categories (color, motion, size), while motion perception (blinking versus circular motion) differed, and color choice was quite uniform between display types.

Results are intended to form the basis for further optimizing label design to improve view management in future work: without understanding the effects of label characteristics, effective view management techniques for wide FOV displays can be hard to design. The results presented here can not only be used to improve spatial awareness and search performance, but can also help define future experiments testing awareness and other aspects, such as safety (distraction) and peripheral motion, in AR.

2 RELATED WORK

Increasing FOV has long been a goal of researchers and display designers alike. Wider FOV displays include prototypes from Nagahara et al. [3], Kiyokawa et al. [4], and Cheng et al. [5], which have given rise to newer, smaller form factors, such as the display developed by Maimone et al. [6]. Furthermore, affordable commercial display solutions such as the Meta 2, offering around 90 degrees of FOV, have become available. Some researchers have also looked at hybrid display solutions to extend FOV, however, they come with constraints. For example, though FoveAR combines see-through glasses with projection-based AR, it will mostly only work indoors [7].

Video see-through versions of the Oculus Rift have also been used in studies on perception of discernibility [8] and widening peripheral vision using fisheye-lenses [9]. Other displays currently in development such as the StarVR promise to provide even 210 degrees. These advancements necessitate further study of the peripheral visual field and influences on information awareness.

2.1 Perception and Peripheral Vision

In real life, we are often dependent on cues in the peripheral visual field. Many aspects of peripheral vision are related to the perception of virtual content and how perception can affect a user's actions. For example, peripheral vision may attract attention to moving objects we may want to avoid, or help us to search and find information [10]. Such behavior may be driven by preattentive processing of specific features that are in contrast to surrounding distractors, where basic features such as color or orientation can take an important role [2], [11]. While searching, object features and context (distractors) can affect search behavior considerably. Search can be defined as the identification whether a target object is present or absent, and if present, where it is located [27]. Basic visual features (like color, size and orientation) are pre-processed before actual attention is placed upon a certain object [12], prior to moving the head so that the object is in central vision [2]. Searching through preattentive objects is made possible in guided search processes. In such processes, the combination of two or more feature processors operating in parallel in the visual field is used [33]. Thus, the features of an object in a preattentive object file can have an effect on driving attention in search behaviour, which can be affected by so-called visual asymmetries [13]: we will reflect on this more closely in the discussion of our results. These characteristics are important to consider since noticing something like a peripheral label will have a direct effect on the user's ability to avoid, deal with, or otherwise interact with the information indicated by the label. We specifically focus on these issues in Experiment 2.

Overall, research suggests that the size of peripheral vision can also affect motion perception [14][15], and scale and distance [16], [17], which can for example be important for navigation abilities [18]. In addition, limiting FOV can result in averse symptoms, including the increase of cognitive load during spatial learning tasks [19]. These issue that can negatively affect the performance of AR applications. Thus, maintaining a wide FOV is of high interest [18].

The sensitivity of the central vision is very different from that of peripheral vision. Due to the distribution of rods and cones in the retina, peripheral vision has a poor resolution [20]. This lower cell density towards the border leads to the degradation in vision of colors, shapes, and text, meaning our perception of or ability to notice virtual content can be very different based on angular position. Throughout this paper, we regard peripheral vision as vision outside the perifovea, being approximately 17 degree radius of central vision [21] (See Fig. 2). In our studies, we cover the periphery to around 100degrees, constrained by technical limitations of current displays. For further information about the different regions in central and peripheral vision, please refer to [10].

Some studies provide evidence for differences in color perception in peripheral vision, pointing towards higher sensitivity of green and brown and slight differences between shapes [22]. In contrast, peripheral vision is still relatively good for motion detection [23]. However, motion types which are most noticeable for semi-transparent elements (such as labels) have yet to be explored. There is also some evidence that light stimuli in the peripheral field can improve distance [24] and spatial scale [12] perception. These studies are somewhat similar to other studies that looked into sparse stimuli in the peripheral field, including [25], that showed that these stimuli can improve situation awareness.

Finally, though exocentric motion has been studied [14], findings are limited to a small range of motions for physical objects, not virtual. It is very likely that the sensitivity of the different areas in the retina will affect the perception, but we do not yet know how.

Because of increasing prevalence of wider FOV displays, perceptual studies will need to focus more extensively on peripheral vision. Still, previous perception studies can aid in our experiment design even though they may focus on central vision. This includes studies on stereo perception such as by Livingston et al. [26], Peterson et al. [27], and Nguyen et al. [28]. On the other hand, a few recent studies have begun to examine peripheral perception. For instance, methods of display and effective FOV have shown to influence on a user's ability to search for targets in wider FOV displays [29].

A precursor study to the research presented in this paper is presented in section 3.1 to help motivate this work. We previously examined divided attention search tasks, and results showed that search performance drops with an "in-view" labeling technique when compared with "in-situ" labeling, while gradually converging when FOV approaches 100 degrees [30]. The study also revealed that users are more likely to make errors for targets in peripheral vision. However, targets presented to users in these studies were largely similar in color, size, and other characteristics, although some label variation was presented. Finally, several other studies targeting wide FOV have been performed in VR environments [29], [31]–[33].

2.2 View Management and Design

To manage the layout and appearance of augmentations, algorithms have been developed that make up the field typically referred to as view management. Studies often look into optimized label placement for size and position [34], [35] and depth-placed ordering [27], [36]. Several more relevant studies exist on label placement and appearance design [37], [38] in general. A few with particular focus on text, such as that of Gabbard et al., Gatulo et al., and McKee et al. ([39]–[41]), studied the effects of text on certain backgrounds. Such methods have even been designed to take advantage of peripheral vision, like the system proposed by Ishiguro et al. [42].

Making sure information is always visible or never occludes other labels or has proper contrast with a background is important, but these studies only provide us with information about how a limited set of text colors interact with a particular background. Moreover, they are generally conducted in a static setting with non-moving backgrounds and only in central vision. This served as additional motivation for us to explore peripheral label characteristics in more depth, and study how placement throughout a wide FOV display would affect user choices and perception.

We also saw the need to study these effects in dynamic and mobile environments. In contrast with most prior studies and evaluations that are conducted on static images or overlays [39]–[41]. In Experiment 1, we created a search experiment outdoors that would let us test search performance in a real-world mobile task, while in Experiment 2 we specifically chose dynamic background videos to replicate real-world settings that might influence label.

Finally, some work has focused on providing information, or directing attention to off-screen content in narrower FOV displays, including the usage of focus+context techniques [43] and arrows [44]. Most of these techniques try to overcome the limitations associated with a narrow FOV display, especially with respect to search for information [43]. This is somewhat related to providing cues to off-screen content in multi-display systems [45], [46].

To the best of our knowledge, our studies are the first that specifically focus on the design and perception of labels in wide FOV displays. The results of this study will be important for use in designing high visibility labels for both commercial and industrial applications, especially for tasks like aviation or vehicle navigation where high priority directional information is critical.

3 GOALS AND EXPERIMENT SETUP

In this section, we will elaborate on the overall design and detailed conditions in our experiments. We first discuss the results of a preliminary seated, divided attention task that helped motivate our study.

3.1 Building on Divided Attention Tasks

Our first study on view management techniques was conducted on a seated, divided attention task [30]. In this study, we wanted to determine how label placement and FOV would affect performance on a task and conversely understand how a task might affect rate of capture for targets. Participants were seated outdoors and asked to complete a Sudoku (mathematics) puzzle. At the same time, both target and dummy objects were displayed at intervals in both the central field of view and periphery.

Results showed that target discovery rates consistently drop with in-view labelling and increase with in-situ labelling (see Fig. 4. for a discussion) as display angle approaches 100 degrees of FOV. Past this point, the performances of these two view management methods begin to converge, suggesting equivalent discovery rates at approximately 130 degrees of field of view. Results also indicate that users exhibited lower discovery rates for targets appearing in peripheral vision, and that there is little impact of field of view on response time and mental workload. On the other hand, Sudoku solving times were faster with in-situ view methods, so we hypothesized that this tendency might also be observed in mobile tasks. With this in mind, we wanted to follow up with a walking experiment where participants had to perform a similar search task. This would let us study visual behavior when mobile and also determine whether label characteristics or FOV would affect a user's workload for the duration of the stroll.

Fig. 3. Images showing a rear view of the OST display with the tracker used for outdoor label registration (left) and an outdoor test run of the device for the campus navigation task (right).

3.2 Research Goals

By building on the initial experiment, we primarily wanted to answer the following questions in this research.

Experiment 1 – dynamic search performance:

1) What is the effect of FOV on search performance in outdoor environments, and how does the label type affect search performance and mental load during navigation?

Experiment 2 – label noticeability:

2) What label characteristics have the largest effect on user perception with respect to how well it can be noticed, and how are these characteristics affected by FOV, visual background and display types?

Our experiments were designed to address these questions and also to explore the relationships between label characteristics, movement, and a user's surrounding environment. Experiment 2 extends the first experiment by looking more closely into label features that may draw attention towards a certain label. As such, it can lay the foundation to improve search performance, studied in Experiment 1, based on visual feature characteristics.

3.2 Hardware

Hardware for our experiments primarily consisted of the OST display, the VST display, and an Android based tracker for outdoor content registration.

Optical See-through

For the OST display experiments (Experiment 1, and Experiment 2 / sub-study 1), we employed the wide FOV head mounted projective display designed by Nguyen et al. [28], as shown in Figs. 1A and 3. It makes use of a retro-reflective screen placed around the user, with a hyperboloidal mirror and small projectors (3M MPro 110, VGA) attached to the user's head. The projections are reflected back to the eyes from a mirror with a distortion correction algorithm, providing stereo wide FOV (109.5 x 66.6 degrees) and optical see-through capability with the semi-transparent retroreflective screen. The screen is made of thin strips of 3M Scotchlite High Gain Retro-reflective Sheeting 7610 with a 0.35mm interval, attached on a curved acrylic plate. A visual acuity of around 20/200 for observed images is achieved with this configuration. The luminance of the observed white image is 60.2 cd/m2. Further details on the adapted version can be found in [30]. Annotations are displayed in the environment using the GPS sensor and compass of an Android-based smartphone (Samsung Galaxy S II) attached to the back of the display.

In Experiment 1, participants wore the headset and walked along a route on campus as can be seen in Fig. 1B. The Android based smartphone with GPS and compass/gyroscope was used for outdoor tracking. Although the GPS sensor can result in an error of about 3 meters, regions for the outdoor targets were placed far enough from each other so that overlap would not occur and that any error would have a negligible effect on results. The tracker was connected and synchronized with the HMD via the backpack system as can be seen in both images of Fig. 3.

In Experiment 2, the entire setup included the OST HMD, a video projector to display background videos on a uniform wall, and controlled lighting to ensure consistent conditions between trials (Fig. 1C). In the first sub-study, participants were seated 1.5 meters away from the wall showing background videos, while labels were overlaid in stereo at a distance of approximately 1 meter away. This distance was chosen so that the video image, projected at a resolution of 1920 x 1280 pixels, overlapped with the outmost borders of the display's FOV. Participants placed their heads on a chinrest to keep the video centered and to support an ergonomic posture.

Video See-through

For the VST experiment (Experiment 2 / sub-study 2), we utilized an Oculus Rift DK2 that simulated a video see-through background, simulating the same environment as deployed in the OST experiment. This was achieved by rendering a bi-ocular virtual plane that mapped the same video backgrounds to the FOV of the Rift (see Section 5.2). The virtual plane was held stationary in a 3D game-world,

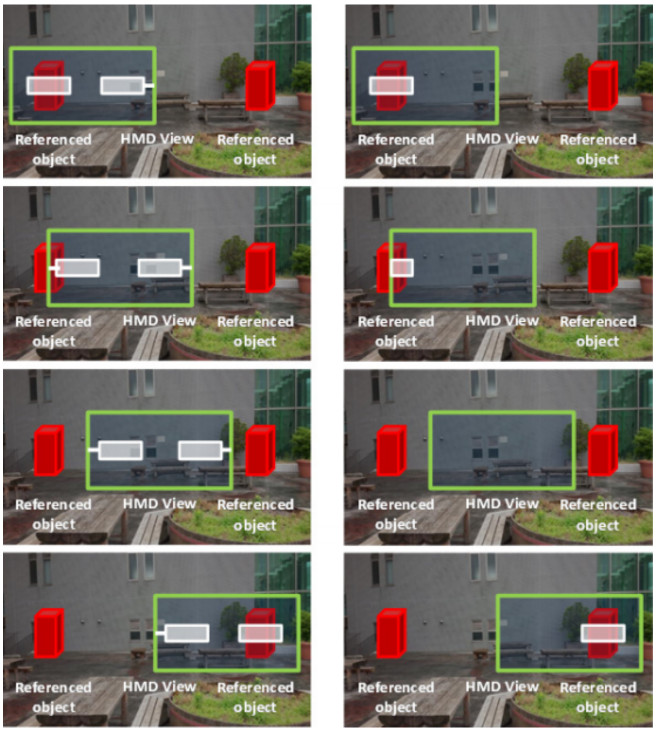

Fig. 4: Schematic views of the two labelling techniques. In in-view labelling (left), all labels appear on the border of the HMD view regardless of the FOV. Overlaps between labels are resolved by minimally shifting their positions. In in-situ labelling (right), labels appear only when the referenced objects are within the HMD view. Label overlaps are not resolved.

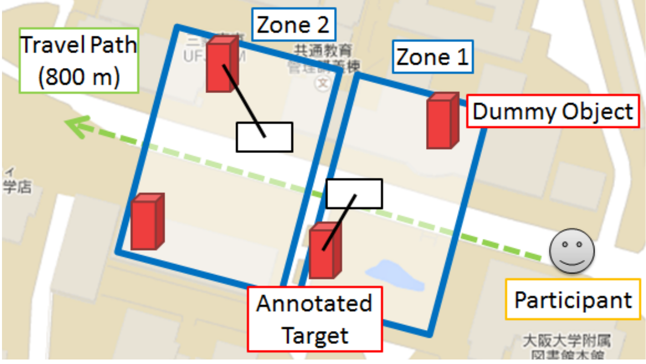

Fig. 5 Image showing two sample zones in which targets appeared along the ~800m path on a college campus (see also Fig. 1B).

and head tracking was engaged so that minor head movements would not induce simulation sickness or fatigue. Display heading was reset at the beginning of each trial to ensure the video was centered to current head orientation. The virtual plane was also positioned to take up the approximate same FOV as the OST display, though label positions were scaled slightly at the horizontal borders to fit the outer edges of the labels positioned at 38 degrees on either side. Vergence was set to approximate the same distance as the projective display (1.5m). Due to the relative low weight of the Rift, users were comfortably seated in a chair, and did not use the chinrest.

Luminance Measurement

In Experiment 2, for each display, general lighting conditions were manually adjusted to be as close as possible, though different areas of the background (with different colors) resulted in different luminance between the displays. The measurements in lumens (lx), along with standard deviation (stdev) are listed in Table 1. Luminance was measured through the viewing windows with a Konica Minolta CS100-A point-luminance meter for both backgrounds and for each label color. Colors for labels were manually adjusted to have perceptually consistent brightness for the pilot study and these color values were preserved throughout experiments, while using CIE Lab (L*a*b* color space) for the two main experiments.

TABLE 1

LUMINANCE VALUES FOR THE DIFFERENT SCREEN ELEMENTS, WITH STANDARD DEVIATION IN PARENTHESIS.

	Mall	Landscape	Labels
OST	26.1 (31.8) lx	26.4 (17.2) lx	0.33 (0.09) lx
VST	10.4 (8.8) lx	12.2 (8.5) lx	4.0 (0.33) lx

4 EXPERIMENT 1: INFLUENCE OF LABELING ON OUTDOOR DYNAMIC SEARCH PERFORMANCE

The first experiment was designed to examine label characteristics that might affect search performance in an outdoor mobile task and was conducted with a different set of participants. Resembling the study described in 3.1, this environment was designed to more closely resemble outdoor AR tasks that might be performed in a city setting, such as searching for a restaurant, following evacuation instructions, or general navigation. To simulate such an environment, we set up an ~800 meter outdoor path on a college campus, as shown in Figs. 1B, 5 and 6. Participants were tasked with finding virtual targets while walking through this region, and were instructed to stop and capture targets as they appeared. Through this experiment, we aimed to gather insight as to what type of label characteristics were better suited to outdoor search, but still not invasive for the user. Similar to our previous experiment [30], this experiment followed a divided-attention task setup, though the level of coordination for the main task (walking) was different than the seated gameplay scenario in our previous study, as gameplay required more concentration.

Fig. 6: The image projected onto the display, where black is transparent on the actual display (left), and a participant walking through the outdoor search task (right).

4.1 Setup

To facilitate a dual-attention task, virtual targets (rectangular solids) were overlaid onto a walking path in different locations using the OST HMD as shown in Figs. 5 and 6. These were arranged in 14 zones that were approximately equidistant from each other, as shown in Fig. 1B, where each red circle (approximately 20 meters in diameter) represents one of these zones in the bird's eye view. Each of these zones contained slots for 18 objects, including dummy and target objects, spaced at 20-degree intervals around the user. The perceived locations of the targets (1.2 by 2.4 meters in size) were at approximately 20 meters away above ground (subtending approximately 3.4 by 6.8 degrees), similar to where store signs or billboards would appear in pedestrian environments. A diagram showing a magnified representation of two of these zones and several sample objects are shown in Fig. 5. Once a participant entered a zone, dummy targets were displayed at each of these 18 slots, and he or she was required to find the targets while traversing the course. For 9 out of the 14 zones that were randomly selected, one of the dummies was converted to a real target from between 3 to 8 seconds after the participant entered the zone. Note that the target is indistinguishable from dummies in appearance and was only identifiable by the corresponding referring label. We did not use all zones at a time to prevent the participants from being able to predict subsequent target appearances. Each label measured 10x6 cm, presented at 1 meter away from the user, subtending approximately 5.7 by 3.4 degrees.

Participants started at one end of the path and walked through each of these zones sequentially until reaching the other end of the path. While walking, they used a two second gaze-dwell to "capture" each of the targets along the path, which required keeping the target within a 10-degree selection area in the center of the display. An experimenter accompanied participants at all times to ensure their safety but stood behind so as not to interfere with the search task. This path was traversed 8 times (alternating southbound and northbound) with the different display conditions as described next.

TABLE 2

DISCOVERY RATE, RESPONSE TIMES, WALKING TIME AND SUBJECTIVE EVALUATION (EXPERIMENT 1)

Test	vs.		
Discovery rate	View management	$F_{(1, 120)} = 112.1$, $p < 0.001$	*In-view labeling yielded significantly higher discovery rates than in-situ labeling except at 100° FOV.*
Response time	FOV, view management, and their interaction	$F_{(3, 542)} = 2.432$, $p = 0.06422$ $F_{(1, 542)} = 2.227$, $p = 0.1362$ $F_{(3, 542)} = 0.2671$, $p = 0.8492$	*The response time was relatively constant for different conditions. Participants probably noticed the target immediately, or did not see it at all regardless of target angles.*
Walking time	View management	$F_{(3, 120)} = 13.49$, $p < 0.001$	*In-view labeling yielded significantly longer time because the targets are seen more often and they needed to capture them.*
Subjective evaluation	NASA TLX vs. FOV Ease of noticing vs. FOV Ease of focusing on walking vs. view management	$F_{(3, 120)} = 0.2109$, $p = 0.8887$ $F_{(3, 120)} = 3.496$, $p < 0.05$ $F_{(1, 120)} = 5.619$, $p < 0.05$	*FOV did not significantly impact perceived workload. Higher FOVs lowered ease of noticing especially in in-view labeling. Ease of focusing on walking was lower with in-view labeling but it was not impacted by FOV.*

4.2 Conditions and Data Collection

Our conditions primarily included the position and range of FOV with which target labels were displayed. Although the HMD has a horizontal and vertical field of view of 100 and 50 degrees, respectively, and dummy and target objects were displayed in the full range of the FOV, we included four FOV restrictions for target labels to test the influence of FOV on discovery rate. The study employed a 4 x 2 factorial design, resulting in 8 trials per participant. The order of trials was randomized between participants and was the factorial combination of four FOVs (36° x 20.3°, 54° x 30.4°, 81° x 45.6°, and 100° x 45.6°) and two label techniques (in-view and in-situ, see Fig. 4). These FOVs are the same as our previous experiment (Section 3.1, [30]). This experiment also incorporated the view management scheme from [18], which compared in-view and in-situ labeling techniques for stationary divided attention tasks (see Fig. 4). A total of 16 participants (8 females, mean age = 23.4) participated in the experiment. Before the experiment, each participant was explained the procedure, potential risks (e.g., fatigue and eye strain, via an informed consent form), and that he or she could terminate the experiment at any moment (which did not occur). After completing all of the eight walking tasks, participants answered a survey on subjective ease of noticing labels and ease of focusing on walking for each condition in a five-point Likert scale (1 being very difficult to 5 being very easy). Mental workload was rated for three sub-tasks (target searching, walking, and overall) for each condition using the NASA TLX [47].

Though 9 targets were designed to appear on the path for each condition for each participant, some targets were missing in reality due to data loss, for example by a lost GPS. In the end, the mean number of targets per condition was 8.91 (stdev of 0.33) for a total of 1141 targets throughout the experiment or all participants. Next, analyses of the impact of both FOV and view management are presented and discussed. A two-way ANOVA with Bonferroni correction was used as the primary statistical test. The Likert scale data were analyzed with the aligned rank transform for nonparametric factorial analyses using ANOVA [48].

4.3 Results

Main results are summarized in Table 2. Note that for walking time, the vertical axis starts at 460,000 milliseconds (Fig. 8).

Discovery Rate

Between methods, in-view labeling yielded higher discovery rates than in-situ labeling for FOV smaller than or equal to 81° ($p = 7.67e-10$ for 36°, $p < 0.001$ for 54°, and $p < 0.001$ for 81°) (Fig. 7).

With respect to the main discovery rates for each condition, an interaction between FOV and view management was found ($F_{(3, 120)} = 8.238$, $p < 0.001$) as well as a main effect on view management ($F_{(1, 120)} = 112.1$, $p < 0.001$). With in-view labeling, the discovery rate dropped as FOV increased ($F(3, 60) = 5.338$, $p < 0.01$), and the discovery rate with 100° FOV was significantly lower than those with 36° ($p < 0.01$), 54° ($p < 0.01$) and 81° ($p < 0.05$), respectively. With in-situ labeling, the discovery rate rose as FOV increased ($F(3, 60) = 4.737$, $p < 0.01$), and the discovery rate with 100° FOV was significantly higher than that with 36° ($p < 0.01$).

These results are mostly consistent with those with a sitting task in [30], however, the discovery rates with 100° FOV with in-view labeling were comparatively lower. This suggests that the walking task required more focus on the central visual field. Many participants commented that lighting conditions largely varied during the experiment which may also have impacted the performance. With regards to the mean target discovery rates and target angle for in-view labeling, a main effect on target angle was found ($F(17, 499) = 2.701$, $p = 0.00028$), but the interaction between FOV and target angle was not significant ($F(51, 499) = 0.922$, $p = 0.628$).

Mean discovery rates were lower with larger FOV in the periphery, which is again consistent with [30], but with larger variances, perhaps as the surrounding environment varied more dynamically in the walking task, making augmentations in the periphery less visible.

Finally, a main effect on target angle ($F(17, 498) = 40.34$, $p < 0.001$) as well as an interaction between FOV and target angle ($F(51, 498) = 1.572$, $p < 0.01$) were both significant with in-situ labeling.

Compared to the results in [30], mean discovery rates are higher with larger FOV (specifically 100°), not only for targets at 50° and higher angles but also targets near the central visual field. This may be because targets in the central visual field enter the field of view more often due to horizontal head turning during the walking task.

The mean discovery rates around the central visual field are higher than those in the sitting task [30] with both in-view and in-situ labeling, which is probably because the participants focused more onto the frontal direction in the walking task.

Response Time

Table 2 also shows the mean response time with respect to FOV measured from the moment of the target's appearance to the first moment of its "capture" within the 10-degree selection area in the center of the HMD view. The tasks in which participants failed to find a target were excluded. No main effect for FOV ($F_{(3,542)} = 2.432$, $p = 0.06422$), view management ($F_{(1,542)} = 2.227$, $p = 0.1362$), or their interactions ($F_{(3,542)} = 0.2671$, $p = 0.8492$) were present. This means that the response time was relatively constant for different conditions. In the case of in-view labeling, participants probably noticed the target immediately, or did not see it at all.

Unlike our previous experiment [30], the response time did not change much with in-situ labeling. In the sitting task, participants focused on a puzzle task in the middle and fewer targets were noticed at higher angles, resulting in a shorter mean response time. Unlike focusing on a central screen, participants were likely able to concentrate on walking without keeping their heads completely level, which could explain the higher mean response times.

With respect to target angle, a main effect on target angle was found for in-view labeling ($F_{(17,329)} = 10.93$, $p < 0.001$), but the interaction between FOV and target angle was not significant ($F_{(50,329)} = 0.8599$, $p = 0.7376$). With in-situ labeling, a main effect on target angle was found ($F_{(12,111)} = 5.984$, $p < 0.001$), but the interaction between FOV and target angle was not significant either ($F_{(22,111)} = 0.8724$, $p = 0.6337$). Response time for targets at 30° and higher angles are longer and less stable than those in the sitting task [30], which could explain why FOV had no significant impact.

Walking Time

Only a main effect of walking time on view management was found ($F_{(3,120)} = 13.49$, $p < 0.001$), but neither the main effect on FOV ($F_{(1,120)} = 0.2174$, $p = 0.8842$) or the interaction between FOV and view management ($F_{(3,120)} = 0.9567$, $p = 0.4156$) was significant. The walking time was shorter with in-situ labeling than with in-view labeling (see Fig. 8). This is likely because targets appeared less often within the user's view with in-situ labeling so that participants spend less time to capture them.

Subjective Evaluation

With respect to the mean NASA TLX scores for target searching, walking, and overall subtasks, an interaction between FOV and view management was not significant for all of the three subtasks ($F_{(3,120)} = 0.3763$, $p = 0.7703$ for target searching, $F_{(3,120)} = 0.3448$, $p = 0.7929$ for walking and $F_{(3,120)} = 0.2109$, $p = 0.8887$ for overall). FOV did not significantly impact perceived workload. For ease of noticing labels, main effects on FOV (F(3, 120) = 3.496, $p < 0.05$) and view management (F(1, 120) = 93.333, $p < 0.001$) as well as their interaction (F(3, 120) = 10.486, $p < 0.001$) were all significant. In the case of in-view labeling, ease of noticing labels dropped as FOV increased (by a Friedman test, $X^2_r = 18.41$, $p < 0.001$). There were significant differences between ease of noticing labels at 100° FOV and those at smaller FOVs ($p < 0.001$ for 36°, $p < 0.001$ for 54°, and $p < 0.02$ for 81°). These results are consistent with the results of discovery rates (conditions with higher discovery rates felt easier to notice labels). With in-view labeling, where labels often appear on borders of the view, it is possible to achieve high objective and subjective evaluations at least up to 81° FOV. In the case of in-situ labeling, the main effect of FOV was not significant (by a Friedman test, $X^2_r = 3.338$, $p = 0.342$). The difference between the four FOV conditions was less significant, probably because targets appeared less often within the view compared to in-view labeling. Regarding ease of focusing on walking, only a main effect on view management was found ($F_{(1,120)} = 5.619$, $p < 0.05$). Ease of focusing on walking was relatively stable for different FOVs, probably because walking is a natural everyday activity and little attention was necessary to walk despite different ease of noticing labels.

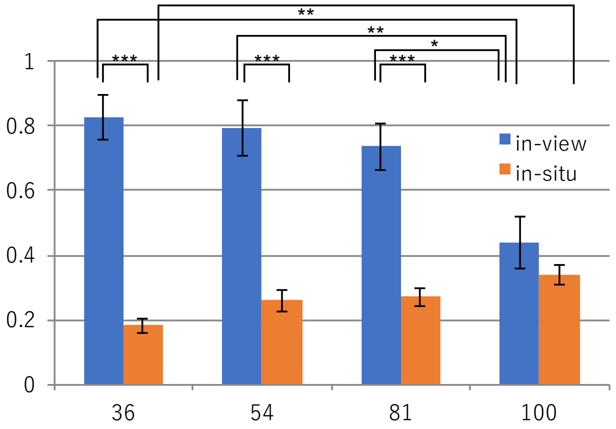

Fig. 7: Target discovery rates with regard to FOV and view management in Experiment 1. (*: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$)

5 EXPERIMENT 2: INFLUENCE OF FUNDAMENTAL LABEL CHARACTERISTICS ON NOTICEABILITY

In Experiment 2, two studies were conducted to determine the most noticeable label characteristics. The general task was to view labels on a moving background recorded from an outdoor scene and then to select the most noticeable out of a series of selectable characteristics. In contrast to Experiment 1, this experiment did not concern a divided-attention task.

5.1 Pilot Study: Initial Test of Label Characteristics for Optical See-through

Before conducting the main study in Experiment 2, we conducted a pilot study with our OST display to get a better idea of appropriate design variables and background conditions. In particular, we wanted to find characteristics that most affect noticeability in the OST display, including color, size, and motion (as shown in Fig. 9.) and leader lines. These characteristics would then help define a primary experiment to test the most influential characteristics in both OST and VST displays

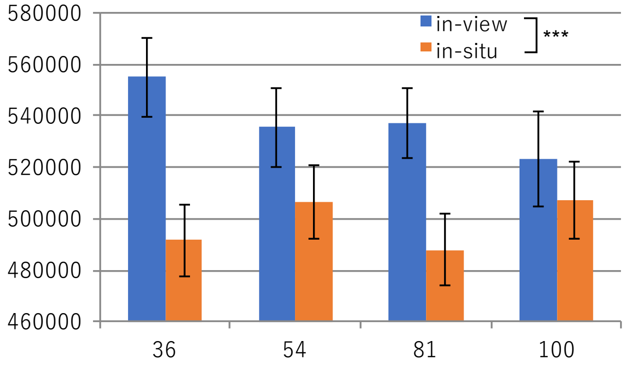

Fig. 8: Walking time (in milliseconds) with regard to FOV and view management in Experiment 1. (***: $p < 0.001$)

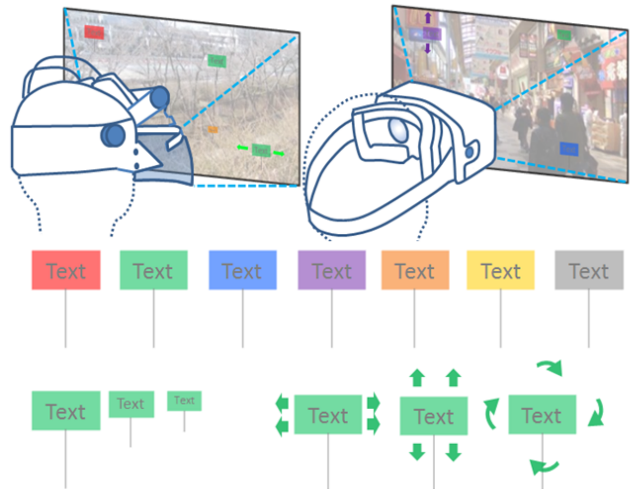

Fig. 9: Diagrams of the two display setups (top) showing examples of several different label positions, colors, sizes, and motions (bottom) overlaid onto each background (Experiment 2). Note that only one label was displayed at a time during actual trials. The left background depicts an unstructured natural environment, whereas the right shows a more structured, human-made environment (Fig. 10).

To test these characteristics, we overlaid labels onto the user's field of view, and allowed him or her to select the option that he or she thought was most noticeable. Labels were placed over three different backgrounds (landscape, street, mall) and in six positions in the peripheral visual field, including 38, 26 and 13 degrees to the left or right of screen center. Four kinds of label design parameters (color, motion, size and leader line) were included. Participants were allowed to select the option they felt was most noticeable. For example, if a particular trial tested color, the participant would select one of seven colors. The study employed a within-subjects design, resulting in 72 trials (3 backgrounds x 6 positions x 4 characteristics) per participant. Nine participants (1 female, mean age = 25.78) participated in the pilot study, performed in controlled lighting conditions.

Pilot study results are summarized since detailed findings are discussed in the primary experiments. In a nutshell, for the primary experiments we reduced the number of backgrounds, increased the number of peripheral label positions, removed leader lines, and increased the number of colors for selection. More specifically, the number of backgrounds for the main study in Experiment 2 could be reduced to two, as the street background (a street scene with some nature) did not reveal any significant results. Adding further rows of label positions allowed us to analyze results with respect to distance from the center of the visual field. Because the user's heads were stationary, leader lines were removed. Finally, after further literature review, blinking was added as an additional motion cue since it could potentially incite a higher level of visual change to which the peripheral visual field can be receptive [20][49].

5.2 Setup

To compare characteristics, we had participants evaluate a series of labels distributed throughout the visual field overlaid onto one of two video backgrounds. A visual representation of this task as well as the different label options are shown in Fig. 9. The labels were displayed in each of our HMDs (the OST and the VST), and participants had to select the most noticeable label among a set of labels with different colors, motions or sizes.

Fig. 10: Frames from each of the background videos used in the experiments, including the *landscape* (left, unstructured) and the *mall* (right, structured).

The experiment consisted of cycling through each label design variable for a particular condition and selecting a single option. For example, a trial testing color could consist of the following process: 1) a label appears at a particular position (e.g. top left) over a particular background, 2) the user cycles through all available colors by a keyboard, 3) the user selects the most noticeable color. We asked participants to look straight ahead at the crosshair at the center of the projected image, without looking directly at the label. Participants were suggested to finish selection within one cycle of the video duration (15 sec), but noted the experiment was not performance driven.

Labels were augmented over different background videos, as shown in Fig. 10, which were 15 seconds in length and repeated to provide enough time to select the most noticeable color / size / motion from a list.

Motivated by a recent study that employed dynamic moving backgrounds to study content placement preferences in real time [50], we chose representative video backgrounds including different structures, colors and dynamicity. These backgrounds were selected to mimic different real-life settings, as different backgrounds may affect the noticeability of labels due to perceptual foreground-background conflicts [1] and reflect similarity (mall) or difference (landscape) between (preattentively processed) label and background features (distractors) [2]. It is important to note the videos were recorded at a stationary position, hence, did convey object motion but no self-motion cues (like associated with walking). The videos were recorded on a tripod, capturing natural motion in the scene such as pedestrians and leaves on trees. The street background used in the pilot was removed for the main experiment. Labels always appeared in the same set of angular positions relative to the generated video backgrounds in both HMDs to maintain consistent foreground-background effects.

5.3 Conditions and Data Collection

Aside from the differences in display configurations, the conditions in both the OST and VST trials were the same. Informed by related work showing differences in the peripheral visual field on these aspects, we included the following label designs (Fig. 9):

Color: red, magenta, cyan, green, yellow, orange and gray
Motion: no motion, horizontal motion (~4 degrees left / right, 2Hz), vertical motion (~4 degrees up / down, 2Hz), circular motion (around a diameter of ~4 degrees , 2Hz), and blinking (2Hz).

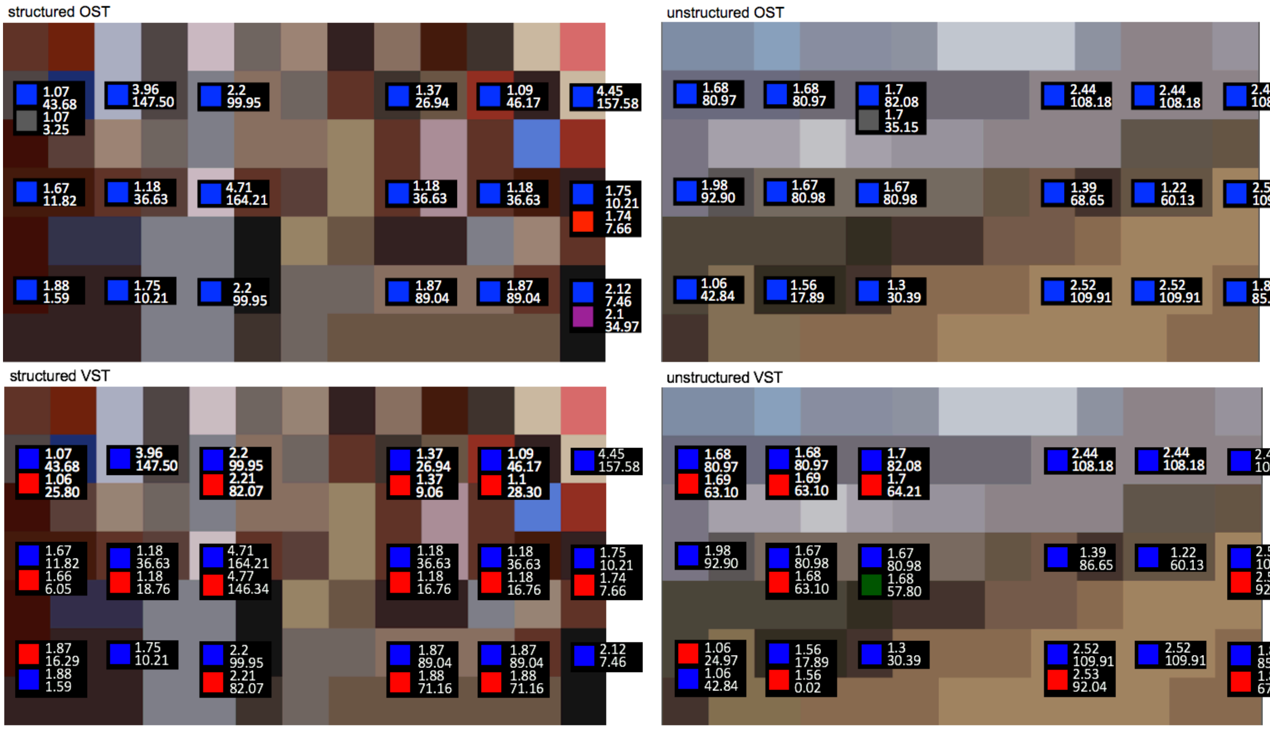

Fig. 11: Color choice analysis in structured (left) and unstructured (right) environments, for both the optical see-through (top) and video see-through (bottom) systems. The colored box in every label in the graph depicts the chosen color in relation to the center of the label in the test environment, the provided statistics are the colour (first value) and brightness contrast (second value) of every chosen color.

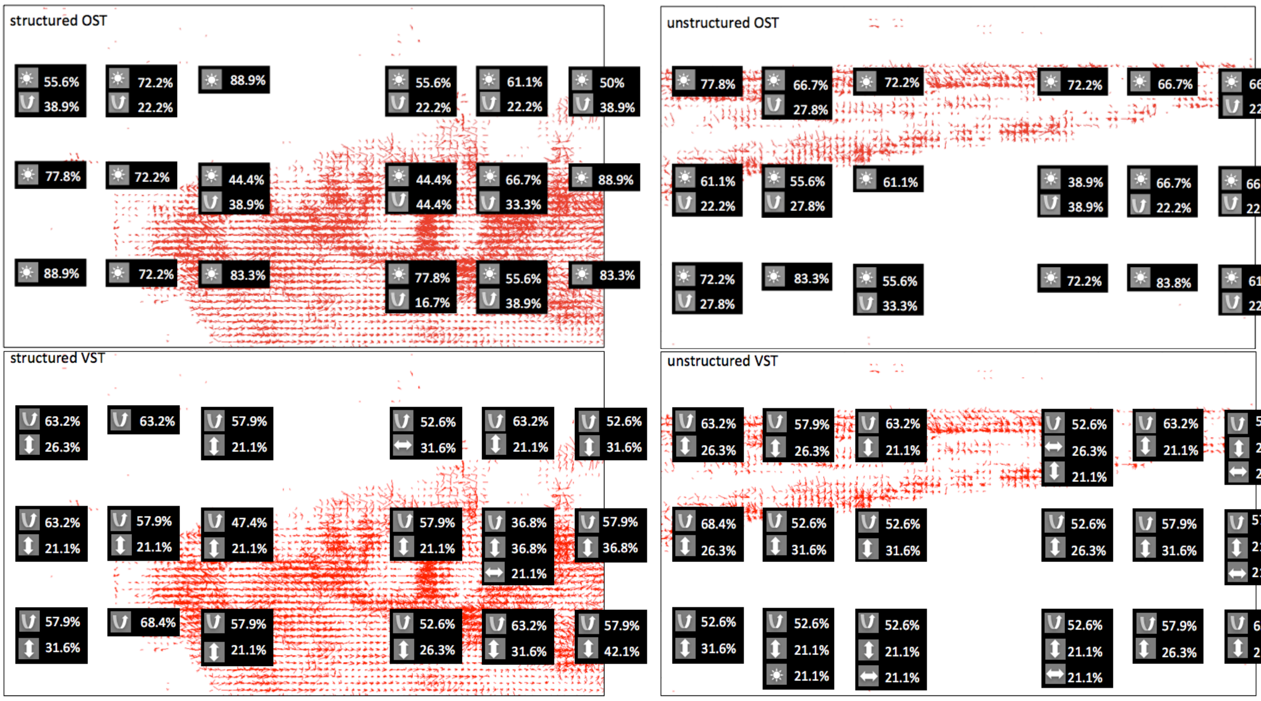

Fig. 12: Motion choice analysis in unstructured and structured video backgrounds, for both the optical see-through and video see-through systems. The gray box in every label depicts the center of the label in the test environment. Icons depicting a horizontal arrow represent horizontal motion, the vertical arrow is vertical motion, the sun is blinking, the bend arrow is circular motion.

Size: small (5 x 2.5 degrees), medium (10 x 5.0 degrees), and large (15 x 7.5 degrees of FOV, respectively). Participants selected the smallest size still easy to notice.

The OST study was performed as a within-subjects study, with 18 participants (3 females, mean age = 23.12, age range = 19 - 30) and employed a 2 x 18 x 3 factorial design, resulting in 108 trials per participant. The order of all trials was randomized between participants and was a factorial combination of two background types (landscape, mall), 18 different positions (horizontal FOV 38, 26, and 13 degrees, both left and right, vertical FOV 20, 0 and -20 degrees, Table 3) and three kinds of label aspects (color, motion, size). Labels were rendered opaque but appeared semi-transparent due to display optics. Participants could cycle through all label aspects variables freely. 19 participants participated (4 females, mean age = 29.32, age range = 19 - 62) in the VST study, which was just as in the OST study a 2 x 18 x 3 factorial design with 108 trials per participant. VST labels were rendered as opaque. After both experiments, participants answered several general questions with regards to user comfort and simulator sickness. Prior to both experiments, participants were screened on color blindness using the Ishihara color test. Furthermore, all participants signed an informed consent form in which they were informed about the content of the experiment and collection and storage of data.

TABLE 3

DIVISION OF REGIONS BY HORIZONTAL (ROW) AND VERTICAL (COLUMN) ANGLE USED FOR ANALYSIS OF DIFFERENT AREAS OF THE PERIPHERY.

	-38°	-26°	-13°	0°	13°	26°	38°
+20°	3	2	1		1	2	3
0°	2	1				1	2
-20°	3	2	1		1	2	3

5.4 Results

Using the 1944 trials of the OST experiment, and the 2052 trials for the VST experiment, we analyzed display characteristics with respect to their position in the FOV, investigated differences between displays, and explored possible interactions. Significant results are summarized in Table 4. Details with respect to each characteristic are described below. Repeated measures ANOVA, multinomial logistic regression and Friedman-test with Wilcoxon signed-rank post-hoc tests were used as statistical tests.

Color Preference

Analysis revealed a general color preference in the OST ($F(1.65, 28.02) = 128.51$, $p < .001$, $\eta^2 = 0.88$) and VST condition ($F(1.45, 26.04) = 51.08$, $p < .001$, $\eta^2 = 0.74$) as blue was the most preferred color, and was chosen significantly more often than the other colors. In the OST condition blue was chosen in 77% of all cases, other colors in less than 10%. In the VST condition blue accounted for 67% of the total color choices, red for 22.7% and other colors for less than 5%. Multinomial logistic regression was then performed to analyze

how color choice was affected by background, region, vertical position and laterality. Only outcome variables that had at least a 20% representation at one position were considered. As for the OST, for color analysis blue, red, gray and purple were considered with blue as the reference category as it was chosen most often. Background and vertical position affected color choice in the OST experiment. For the VST, colors blue, red and green were considered in the analysis with blue as reference category, as also here it was chosen most often. Only the hemifield showed an effect on color choice (Table 3).

To more closely address the potential effect of background on choice of color, we performed a foreground-background color contrast analysis, assessing both color and brightness contrast. Contrast ratio was determined by (L1 + 0.05) / (L2 + 0.05), where L1 is the relative luminance of the lighter of the colors, and L2 is the relative luminance of the darker of the colors. The .05 value used is based on Typical Viewing Flare (IEC-4WD [51]). Contrast ratios can range from 1 to 21, while a ratio of 3:1 is the minimum level recommended by ISO-9241-3 for standard text and vision. Color brightness gives a perceived brightness for a color and was determined by the following formula: ((Red value X 299) + (Green value X 587) + (Blue value X 114)) / 1000 (W3C, [52]).

Based on the pixelated image abstraction tool provided by [53], we produced images consisting of 7 rows and 13 columns grid using a 24 color palette, to assess the effect of label color preference per relevant cell in the grid: each cell provided us with the most relevant color for the respective location in the background to assess the foreground-background contrast analysis. Color brightness at all background positions except of one was greater than that of the blue label which had also the lowest color brightness of all labels. As a result, blue showed the highest dark on light brightness difference (BG color brightness – label color brightness) at almost all positions (see Fig. 11) of both backgrounds, followed by red. Regarding absolute differences, some label colors showed higher values at some positions than blue but this difference was always negative then, that is, there was a higher light on dark contrast for these labels sometimes. Color contrasts between the label and the background were very similar between labels at all positions and differences were negligible.

Some artefacts in the lower right corner of OST unstructured and in the center of the structured VST conditions can be noticed, but can hardly be traced back to variants based on background color. In general, it can be noticed that contrasts were not always high, and did not differ much in between colors, with yellow producing the highest color contrast. Overall, the choice of the blue color overlapped with the choice of most pleasing color gained from the post-experiment questionnaire, as blue was preferred by 77.8% by the OST participants and 50% of the VST participants, who also noted a higher preference for green (27.8%).

Motion Selection

Analysis revealed that display device affected the choice of label motion significantly though we originally hypothesized that background would also be a deciding factor. In general, blinking labels were preferred in the OST condition, whereas circular motion was preferred in the VST condition.

Multinomial logistic regression was performed to analyze how the outcome variable motion choice was affected by background, hemifield, vertical position and region. Again, only variables with at least a 20% representation at one or more positions were considered in the regression model. As for the OST display, for the motion analysis only circular and blinking outcomes were included. Regions did not affect motion choice and there were also no main effects of background, hemifield or vertical position.

For the VST all motion types were included in the analysis with circular motion as a reference category since it was dominant. Motion choice was only affected by the region here (Table 3).

We also assessed the potential interplay between label motion type choice and background motion. To do so, we produced optical flow density images by graphically accumulating the optical flow in the 15-second video sequence using MATLAB. Fig. 12 depicts the various backgrounds and the motion type choices, again with choices under 20% omitted. Interestingly, while the images show some of the variations, for example central versus peripheral (as noted in the statistics), there is still no clear dependency of motion choice in comparison to background motion. Based on motion patterns in the background, mostly horizontal with a slight diagonal in both mall and landscape environments, the choice for circular and blinking could be explained, as these motion types are quite in contrast with the prevalent motions in the backgrounds. However, also outside the visual motion areas, these particular types were chosen – hence, it can be assumed that the choice was not necessarily dependent on the motions in the background. Furthermore, background motion alone cannot explain why blinking is preferred in OST conditions and circular motion in the VST case.

Minimum Size

Size preferences were mostly as expected, with increased selection of larger labels towards the peripheral visual borders. A Friedman test was used to analyze size choices for each experiment, and background, region, laterality and vertical position were within-factors (see Table 3 for statistics). A Wilcoxon signed-rank test with Sidak correction was conducted for post-hoc analysis. In the OST experiment, background, region and hemifield did not influence size choice, whereas the vertical position did. Post-hoc tests confirmed a significant choice of larger labels at +20° compared to 0, but no differences between 0° and -20° or -20° and +20°. In contrast, in the VST experiment, regions affected size choice, as larger labels were chosen more often in the peripheral than in the more central field of view. Similar to OST, larger labels were selected more often at +20° than at 0° in the VST experiment. As in the OST experiment, there was no significant difference between hemifields.

6 DISCUSSION

The results revealed several interesting tendencies, which we summarize below with respect to each of our research questions. Overall, while other studies reported on FOV effects on scale [16], distance [12], and velocity [54] estimations, these issues did not directly affect our experiments, as labels were at the same disparity plane throughout all experiments. However, once the labels will be directly linked to real-world objects, it is likely that in particular depth estimation will become a factor. For example, we assume that a wider FOV will afford a more precise label matching to real world objects, by supporting more accurate judgments upon their distance and location (relative size, based on [16] [12]). However, at current we cannot make exact assumptions based on our study results. Furthermore, previous work has shown improvements in situation awareness [25]. It will be worthwhile to study situation awareness once we transfer our acquired understanding of label appearance and search affects into the real-world domain, to address to what extent FOV will affect the acquisition, storage and recall of label information.

6.1 Effect of FOV on Search Performance in a Real-World Task

Results showed that search performance drops (in-view labelling) or increases (in-situ labelling) smoothly up to 100 degrees of FOV. In-view labeling yielded higher search

TABLE 4

PRIMARY CHOICES FOR COLOR, MOTION, AND SIZE IN EACH OF THE HMDS (EXPERIMENT 2, STUDY 1 AND 2).

Test		Primary choices for OST		Primary choices for VST	
color	vs.		blue		blue
	Background	$\chi2$ (3) = 8.83, p = .03	Preferring red over blue is less likely on the mall than on the landscape background (Odds ratio = 0.25, $p < .05$)	none	
	Position	Vertical: $\chi2$ (6) = 12.80, p = .046	Gray chosen more likely than blue at +20° compared to 0° (Odds ratio = 2.28, $p < .05$).	Hemifield: $\chi2$ (2) = 6.33, p = .002	Preferring red over blue is more likely on the left (29.6%) than on the right (18.8%), Odds ratio = 2.2, $p < .001$.
	While some variations over the field of view could be noticed, overall, blue was found to be significantly better noticeable in both display types				
motion		blinking		circular	
	Screen Region			$\chi2$ (6) = 2.65, p = .015.	It was more likely to prefer horizontal (Odds ratio = 2.36, p = 0.32) and blinking (Odds ratio = 5.16, p = 0.27) over circular motion in region 1 than in region 3.
	Display types significantly affected the choice of label motion type, where blinking scored best for OST, and circular for VST				
size		large	medium in center, large toward border	large	medium in center, large toward border
	Screen Region	none		$\chi2$(2) = 19.51, p < .001	1 vs. 2, Z = -2.76, p = .006, 1 vs. 3, Z =-3.28, p = 0.001 Larger icons chosen more often in outer than inner periphery
	Vertical Position	X^2(2)= 6.61, p = .037	Larger Icons at +20 vs. 0, Z = -2.27, p = .023	$\chi2$(2) = 8.22, p = .016	Larger Icons at +20 vs. 0 Z = -2.65, p = .008
	Label size preference was significantly affected by label position, with preference for larger labels towards the periphery				

performance with lower ease of focusing on walking.

FOV had a significant impact on target discovery rates, which were significantly lower or higher at 100° FOV compared to smaller FOVs with in-view and in-situ labeling, respectively. With in-view labeling, all the labels always appear in the HMD FOV at their original positions or on the border of the FOV, so the labels are easier to find with the smaller FOV. With in-situ labeling, all the labels stay at their original positions and only those within the HMD FOV will be seen, thus the labels are easier to find with the wider FOV. At 100° FOV, discovery rates with in-view and in-situ labeling were similar, suggesting a higher FOV may not be necessary in mobile situations. These results are in line with what we found in our previous divided-attention study [30].

With respect to preattentive processing [2], it remains to be seen what exact effect the in-view labelling method has. Clearly, it conveys additional information in the periphery, however, how this information is processed in the different zones in the periphery requires further study. With respect to our label features and search performance, there is some relationship with work performed in the field of visual asymmetries [13]. Mainly, the work on asymmetries looks into the difference between features of a search object among distractors. The most efficient of searches are those in which the target is defined by a single basic feature (e.g., color or shape) and in which the distractors are homogeneous. The least efficient searches are those in which targets and distractors share the same basic features. We made use of higher contrast labels (“features”, Fig. 1A). While the OST brightness is lower than our VST displays, we did not note any problems with users having issues separating labels from the background (“distractors”). Also, due to the color difference (red versus white), the in-view label was well distinguishable from the other labels. We will provide further discussion on visual asymmetries in section 6.3.

FOV also had a significant impact on a perceived task difficulty on the HMD. On the other hand, FOV had little impact on response time and perceived workload. Another key finding in Experiment 1 is that view management had an even more significant impact on task performance than FOV. To better exploit a wide FOV, it is very important to design an appropriate view management policy for the given task. A variety of visualization techniques for indicating off-screen content have been proposed [43]–[46], but more specific techniques for AR with wide-view HMDs should be explored.

In conclusion, we have provided valuable new insights into the understanding of the effects of different FOVs on search performance in mobile outdoor AR, showing mostly coherent results with a previous study with a stationary divided attention task [30] and search task studies with varying FOVs reported in [32].

6.2 Effect of Label Type and FOV on Mental Load

In experiment 1, FOV had little impact on perceived workload and focus on the task in the real environment. However, a more complicated task may be more sensitive to FOV. On the other hand, FOV had a significant impact on ease of noticing information on the HMD. Displaying information at high angles beyond approximately 40° to the left or right in the periphery can lead to stronger perceived task difficulty. As such, it will be interesting to validate even wider FOV displays (larger than 110 degrees) to see if the expanded space can still be used without increasing the (perceived) mental load drastically. Once lighter and wider FOV displays become available, we will target this issue in further studies.

6.3 Effect of FOV and Background on Label Noticeability

Results showed that with respect to label color, blue yielded the highest preference quite uniformly through the different display locations, while circular motion (VST) and blinking (OST) were chosen most frequently. As expected, increasing size of labels towards the periphery yielded best results.

Experiment 2 looked into issues that overlap with related work on preattentive processing: normally, a visual stimulus is divided into objects preattentively, and holds local features such as color or size [2]. With respect to feature/distractor characteristics in visual asymmetries research, our study shows some resemblance as we asked participants about noticeability of a label (features) against a dynamic background (distractors). However, it should be clearly stated again that while users attended the object, no direct focus was placed on the object. This is in contrast to visual search tasks in visual asymmetries research in which the user looks at the features directly. While work on asymmetries mainly regards cues at the same disparity plane, some studies have also been performed in 3D space, including [55]. In our case, a label (with varying features) had to be noticed while being overlaid over a background (distractor), which was at a different disparity plane. Research is ongoing on what basic features are important for visual search. Among others, while color, size and orientation are somewhat agreed upon, shape is not completely understood [2], [11]. These features overlap with our label appearance features color and size [56], but also to some extent to motion [57]. With respect to visual asymmetries in our backgrounds, we had both overlap and clear distinction between label features and distractors. While the background hardly shared features with our labels in our landscape scene, it did in the mall scene (Fig. 10).

In our study, with respect to color, users clearly chose blue as the most noticeable. This choice also coincided with their subjective preference of color (which we assessed after the experiments), yet contradicts some previous research suggesting that green and brown are better for identifying peripheral content [22]. Interestingly, we noted that green and gray were more often chosen in the center in the VST condition than expected. Anatomically, blue coincides with the sensitivity of cones in the peripheral visual field to blue light, which further explains this choice [20]. With regards to contrast, blue yielded the highest color brightness (dark on light) contrast at all positions (followed by red). If we compare these results to research in legibility, early studies have found that increasing the brightness difference between the color of lettering and background enhances legibility, while a high brightness difference improves legibility particularly in dark on light contrast conditions [58]. Furthermore, studies have found an advantage of dark on light for reading and character recognition in printed media [59], [60]. Hence, the high dark on light contrast of label colors blue and red at first sight may give some insight into the preference pattern. However, we could not find any direct interaction between color choice and background, as discussed here after, which may contradict this. Though blue was the most chosen color, we also noted some effects of vertical degree and hemifields on color choice. However, these variances did not necessarily correspond to color patterns in the background. As we used green as main color for the size and motion conditions [22], it remains to be seen if a change to blue will affect user choice for size and motion. Surprisingly, we did not always find a clear effect of background on choices, even though the cross-hair was on disparity plane of the video image: only color was affected by background in the OST experiment.

With respect to motion, circular motion (VST) and blinking (OST) were chosen most frequently. This may be due to the level of visual change they depict, to which the peripheral visual field is receptive [10], [20]. FOV did not have any significant effect on motion type choice, which is in line with previous research. While the spatial determinants of velocity discrimination follow the change in resolution found with eccentricity, peripheral temporal sensitivity is nearly equal to foveal temporal sensitivity [23]. However, in our experiment we only checked for individual labels – in situations where denser label sets are used, label motion may need to be adjusted accordingly, or even avoided. Our results extend findings in [61], which looked into the effectiveness of certain "popout" cues (like motion, flashing, but also color and shape) at different angles in the FOV, to find an item among distractors. The study only used diagonal motion, whereas we used horizontal, vertical and circular motion. The study showed that motion, flashing and luminance performed better in the periphery, while color and shape work well closer to central vision. Important to note is the study also showed that motion effects are highly accurate even at wide angles and at subtle levels. With regards to view management, it would be interesting to address different levels of drawing attention in relation to FOV angle by also using variations of popout cues. Furthermore, another venue of future work is if certain cues (like blinking) may also provide outside the normal display FOV at low resolution, for example by using a sparse peripheral display [25].

Finally, with respect to label size, we showed that while in the central visual field users still selected smaller minimum sizes, choice for larger sizes increases when moving into the periphery. This result was not unexpected: it is supported by the density of cells in the human eye, as smaller items are more difficult to see the further they are moved towards the border of our vision [20], while our study revealed how this increase progresses for virtual elements.

The lack of effect or irregularities of the background somewhat contradicts perception studies as previous work on text legibility has shown effects [39], [41]. As stated before, only color was affected by background in the OST experiment. In general, this is in contradiction to research results on visual search affected by asymmetries, as it has been shown that asymmetries in color search are dependent on the relationship between the stimulus colors and the background color [56]. It was noted that most models confirm objects can be preattentively be segmented from the background [62], while the background color can affect the perception of the target and distractors. When a background is more complex, it tends to take more time to process (search) for a target [62]. Furthermore, with respect to background motion, our local analysis did not show any significances, which again is in contradiction to what has been found in search asymmetry research: for example, previous research indicates that local rather than global properties of flow fields are of importance for searching a stationary target [57]. Some of the irregularities we found can potentially be attributed to upper and lower visual field differences, with a lower visual field advantage occurring for motion, global processing and coordinate spatial judgments, while upper visual field advantages occur for visual search, local processing and categorical judgments [63]. However, more research is required to fully address differences, as results are not conclusive with respect to the issues studied in our experiments. Furthermore, while varying differences between disparity planes (and associated vergence and accommodation issues) may have an effect on foreground-background issues in noticeability tasks [64], we cannot assess this at current time.

In conclusion, we were able show clear results for color (blue), motion (circular and blinking) and minimum size requirements (larger towards border) that can aid in designing more effective view management systems for AR, especially those that deploy wider FOV displays. Such optimizations can potentially further improve search performance (experiment 1), an area for future work.

6.4 Label Design Preference Differences between OST and VST

The biggest and most surprising difference between the two displays was the choice of motion. We initially hypothesized that blinking content would be most noticeable in all cases, especially considering blinking is used for many types of warning lights like strobes or rear bicycle lights. Though blinking was selected most often in OST, circular motion was selected in VST by a significant margin.

In hindsight, we realized that circular motion is also often found on the rotating lights of emergency vehicles and many road warning signs. The real question is why the difference between selected motions was so distinct between OST and VST. One explanation could be that blinking was more noticeable in the OST display because the label brightness is added onto the background whereas it is replaced in the VST display. This has very interesting implications for motion perception for transparent versus opaque content. It is possible that the motion or structure of the background also influenced perception of movement, since the transparency of the OST actually mixes virtual content with a background that has content moving in very different ways. Another explanation could be that translational motion was less noticeable in the OST display because the label was slightly blurred due to different accommodation distances, whereas VST displays generally have higher visual quality [65].

Until problems with occlusion, focal depth, and color reproduction are completely solved in OST and VST displays, content designers should take motion into account when building warning labels or content designed to grab the user's attention. Moreover, in the future motion preference may actually serve as a method to determine whether a particular AR display is correctly reproducing content perceptually. This idea deserves further experimentation.

7 CONCLUSION AND FUTURE WORK

In this article, we presented the results of two experiments that explore the perception, in particular search and noticeability, of virtual label characteristics relative to wide FOV display types.

In Experiment 1, we confirmed the effect of FOV on search performance using in-view and in-situ labeling, extending the results presented in our previous study [30]. In Experiment 2, we found that perceptual differences between OST and VST displays differed with respect to label motion, though color choice, primarily blue, was largely the same. Minimum noticeable size of labels increased linearly with distance from central vision for both displays, which should be taken into account in future iterations of view management algorithms. Though background is often shown to affect readability, our results suggest that the influence of background may not be as strong for noticeability, which is somewhat in contrast to findings from research in visual asymmetries.

While our experiments reveal new information about peripheral perception and search behavior, we also need to narrow down new research directions that will cover for some of the limitations of our experiments. First, it would be beneficial to compare a greater variety of devices such as Pinlight [6] or light field displays to cover differences in optical quality and associated aberrations more effectively. We did not fully explore the near peripheral visual area or perifovea as they affect narrow and medium FOV displays in particular, yet for comparative studies these areas will be highly relevant. While some studies on depth judgment differences between wide and narrow FOV have been performed [66], more closely comparing perceptual differences between displays of varying FOV and focal depth will be important. Furthermore, we revealed significant differences, in particular motion choice, which strongly warrant additional studies on peripheral perception in OST versus VST devices. Motion is important in particular since egocentric motion perception functions as an early warning system for hazards in the human visual system. If circular or horizontal motion happen to be particularly noticeable (or distracting) when a pedestrian is crossing the street, adding labels could endanger the user's wellbeing. As such, our results further motivate studies on safety of in-situ AR. This may even have legal implications, as an advertiser that places a moving ad in the individual's periphery may be held liable for causing an accident. At the same time, factory work could benefit greatly from displaying more noticeable warnings that could help prevent injury.

To more closely address attention issues, it will be necessary to analyze eye motion using eye-tracking hardware. While in Experiment 2, we specifically instructed participants not to look at labels directly, involuntary eye fixations at the label may have occurred that could have affected results. Even more so, in Experiment 1, eye tracking could have been useful to analyze search patterns. Overall, eye tracking analysis can be seen in cohesion with calibration procedures relevant for fine-tuning optical characteristics to optimize view management [67]. Eye tracking also has a unique connection to user attention and focus. Though some exceptions exist, it is almost always the case that gaze point indicates attention. Taking advantage of this fact, further studies on how label design affect attention can be carried out, which can also help us understand how virtual information can guide (or misguide) the user's attention in everyday life, and how this may improve search performance. Further experiments should also be looking more closely in divided attention tasks, similar to [30]. Some research has already been reported on workload in divided attention tasks in car head-up displays [68], [69]. Yet, as results were contradicting, workload should be further addressed in relation to different view management styles. Another important aspect of this attention could possibly be interaction, i.e., labels that function as widgets or icons with which the user has to physically control. Though some of our in-situ targets were selectable with gaze, users largely did not have to manipulate them in any way. Targets that are graspable, that have interactive buttons, or that move with the user will likely show very different tendencies in terms of the user's performance of a main task. Other types of search tasks could be considered.

With regards to the video backgrounds, to replicate even more realistic conditions, the next step is to validate the results of experiment 2 against a real outdoor validation, where the high dynamic range (HDR) of foreground-background and lighting conditions are expected to have a stronger effect. Such outdoor experiments should also more closely consider direct mapping of labels to real-world objects, as to address potential effects of FOV on, in particular, distance [12] and scale [16] perception, which may also have an impact on situation awareness [25]. Even though in our current experiment we did not find clear foreground-background interactions on sub-regions of the background, the overall background sometimes had an effect. Post-experiment analysis on HDR environment data is another significant venue for research, which has been shown to be challenging yet important [70]. Finally, as users only fixated on a single depth defined by the video background, an interesting issue is also the effect of different disparity planes. One more research direction is the testing of differing focal depth (and associated vergence and accommodation issues) and its influence on label design.

Overall, our studies form a solid basis for extending research in peripheral label design and can contribute to design guidelines for designing more effective view management systems for wide FOV AR systems, while the results can also encourage new studies to compare results in narrower FOV displays.

ACKNOWLEDGEMENT

This work was partially funded through the Deutsche Forschungsgemeinschaft (DFG, grants Kr 4531/1-1 and 4531/2-1), partially by the United States Department of the Navy, Office of Naval Research Grant #N62909-18-1-2036, and the JSPS Grant-in-aid for Scientific Research #15H02738. We also would like to thank all the participants for their time and effort while participating in the experiments. Finally, we kindly thank Katharina Stollenwerk for producing the optical flow density figures in MATLAB.

REFERENCES

[1] E. Kruijff, J. E. Swan, and S. Feiner, "Perceptual issues in augmented reality revisited," in *IEEE International Symposium on Mixed and Augmented Reality*, 2010, no. 13–16 Oct. 2010, pp. 3–12.

[2] J. Wolfe and S. Bennett, "Preattentive object files: shapeless bundles of basic features.," *Vision Res.*, vol. 37, no. 1, pp. 25–43, Jan. 1997.

[3] H. Nagahara, Y. Yagi, and M. Yachida, "Wide field of view head mounted display for tele-presence with an omnidirectional image sensor," in *Conference on Computer Vision and Pattern Recognition (Workshop)*, 2003, vol. 7, p. 86.

[4] K. Kiyokawa, "A wide field-of-view head mounted projective display using hyperbolic half-silvered mirrors," in *IEEE and ACM International Symposium on Mixed and Augmented Reality*, 2007, pp. 1–4.

[5] D. Cheng, Y. Wang, H. Hua, and J. Sasian, "Design of a wide-angle, lightweight head-mounted display using free-form optics tiling.," *Opt. Lett.*, vol. 36, no. 11, pp. 2098–100, Jun. 2011.

[6] A. Maimone, D. Lanman, K. Rathinavel, K. Keller, D. Luebke, and H. Fuchs, "Pinlight displays," *ACM Trans. Graph.*, vol. 33, no. 4, pp. 1–11, Jul. 2014.

[7] H. Benko, E. Ofek, F. Zheng, and A. D. Wilson, “FoveAR: Combining an Optically See-Through Near-Eye Display with Projector-Based Spatial Augmented Reality,” in *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology - UIST '15*, 2015, pp. 129–135.

[8] W. Steptoe, S. Julier, and A. Steed, “Presence and discernability in conventional and non-photorealistic immersive augmented reality,” in *IEEE International Symposium on Mixed and Augmented Reality*, 2014, pp. 213–218.

[9] J. Orlosky, Q. Wu, K. Kiyokawa, H. Takemura, and C. Nitschke, “Fisheye vision: peripheral spatial compression for improved field of view in head mounted displays,” in *ACM symposium on Spatial User Interaction*, 2014, pp. 54–61.

[10] H. Strasburger, I. Rentschler, and M. Jüttner, “Peripheral vision and pattern recognition: a review.,” *J. Vis.*, vol. 11, no. 5, p. 13, Jan. 2011.

[11] J. Wolfe and J. DiMase, “Do Intersections Serve as Basic Features in Visual Search?,” *Perception*, vol. 32, no. 6, pp. 645–656, Jun. 2003.

[12] J. Jones, D. Krum, and M. Bolas, “Vertical Field-of-View Extension and Walking Characteristics in Head-Worn Virtual Environments,” *ACM Trans. Appl. Percept.*, vol. 14, no. 2, pp. 1–17, Oct. 2016.

[13] J. Wolfe, “Asymmetries in visual search: an introduction.,” *Percept. Psychophys.*, vol. 63, no. 3, pp. 381–9, Apr. 2001.

[14] T. Brandt, J. Dichgans, and E. Koenig, “Differential effects of central versus peripheral vision on egocentric and exocentric motion perception,” *Exp. Brain Res.*, vol. 16, no. 5, Mar. 1973.

[15] P. Pretto, M. Ogier, H. Bülthoff, and J.-P. Bresciani, “Influence of the size of the field of view on motion perception,” *Comput. Graph.*, vol. 33, no. 2, pp. 139–146, Apr. 2009.

[16] J. Jones, E. Swan, and M. Bolas, “Peripheral Stimulation and its Effect on Perceived Spatial Scale in Virtual Environments,” *IEEE Trans. Vis. Comput. Graph.*, vol. 19, no. 4, pp. 701–710, Apr. 2013.

[17] J. Jones, E. Swan, G. Singh, and S. Ellis, “Peripheral visual information and its effect on distance judgments in virtual and augmented environments,” in *Proceedings of the ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization - APGV '11*, 2011, p. 29.

[18] P. Alfano and G. Michel, “Restricting the Field of View: Perceptual and Performance Effects,” *Percept. Mot. Skills*, vol. 70, no. 1, pp. 35–45, Feb. 1990.

[19] E. Barhorst-Cates, K. Rand, and S. Creem-Regehr, “The Effects of Restricted Peripheral Field-of-View on Spatial Learning while Navigating.,” *PLoS One*, vol. 11, no. 10, p. e0163785, 2016.

[20] G. Osterberg, *Topography of the Layer of Rods and Cones in the Human Retina*. 1935.

[21] B. Wandell, *Foundations of vision*. 1995.

[22] C. Lou *et al.*, “Object recognition test in peripheral vision: a study on the influence of object color, pattern and shape,” in *Proceedings of the 2012 international conference on Brain Informatics*, 2012, vol. 7670, pp. 18–26.

[23] S. McKee and K. Nakayama, “The detection of motion in the peripheral visual field.,” *Vision Res.*, vol. 24, no. 1, pp. 25–32, Jan. 1984.

[24] B. Li, J. Walker, and S. A. Kuhl, “The Effects of Peripheral Vision and Light Stimulation on Distance Judgments Through HMDs,” *ACM Trans. Appl. Percept.*, vol. 15, no. 2, pp. 1–14, Apr. 2018.

[25] R. Xiao and H. Benko, “Augmenting the Field-of-View of Head-Mounted Displays with Sparse Peripheral Displays,” in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*, 2016, pp. 1221–1232.

[26] M. Livingston, Z. Ai, and J. Decker, “A user study towards understanding stereo perception in head-worn augmented reality displays.,” in *IEEE International Symposium on Mixed and Augmented Reality*, 2009, pp. 53–56.

[27] 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, pp. 169–176.

[28] D. Nguyen, T. Mashita, K. Kiyokawa, and H. Takemura, “Subjective Image Quality Assessment of a Wide-view Head Mounted Projective Display with a Semi-transparent Retro-reflective Screen,” in *International Conference on Artificial Reality and Telexistence*, 2011.

[29] N. Kishishita, J. Orlosky, T. Mashita, K. Kiyokawa, and H. Takemura, “Investigation on the peripheral visual field for information display with real and virtual wide field-of-view see-through HMDs,” in *IEEE Symposium on 3D User Interfaces*, 2013, pp. 143–144.

[30] N. Kishishita, K. Kiyokawa, E. Kruijff, J. Orlosky, T. Mashita, and H. Takemura, “Analysing the effects of a wide field of view augmented reality display on search performance in divided attention tasks,” in *International Symposium on Mixed and Augmented Reality*, 2014, vol. 85, no. 6, pp. 177–186.

[31] B. Watson, N. Walker, and L. Hodges, “A User Study Evaluating Level of Detail Degradation in the Periphery of Head-Mounted Displays,” *in Proceedings of Framework for Interactive Virtual Environments (FIVE) Conference*, 1995. [Online]. Available: https://smartech.gatech.edu/bitstream/handle/1853/3575/95-31.pdf. [Accessed: 17-Mar-2014].

[32] K. Arthur, “Effects of Field of View on Performance with Head-Mounted Displays.”

[33] D. Ren, T. Goldschwendt, Y. Chang, and T. Höllerer, “Evaluating Wide-Field-of-View Augmented Reality with Mixed Reality Simulation,” in *IEEE Virtual Reality*, 2016.

[34] B. Bell, S. Feiner, and T. Höllerer, “View Management for Virtual and Augmented Reality,” in *ACM symposium on User Interface Software and Technology*, 2001, pp. 101–110.

[35] R. Azuma and C. Furmanski, “Evaluating Label Placement for Augmented Reality View Management,” in *IEEE/ACM international Symposium on Mixed and Augmented Reality*, 2003.

[36] S. Peterson, M. Axholt, and S. Ellis, “Label segregation by remapping stereoscopic depth in far-field augmented reality,” in *IEEE/ACM International Symposium on Mixed and Augmented Reality*, 2008, pp. 143–152.

[37] K. Uratani, T. Machida, K. Kiyokawa, and H. Takemura, “A study of depth visualization techniques for virtual annotations in augmented reality,” in *IEEE Virtual Reality*, 2005, pp. 295–296.

[38] R. Grasset, T. Langlotz, D. Kalkofen, M. Tatzgern, and D. Schmalstieg, “Image-Driven View Management for Augmented Reality Browsers,” in *IEEE International Symposium on Mixed and Augmented Reality*, 2012, pp. 177–186.

[39] A. Leykin and M. Tuceryan, “Automatic Determination of Text Readability over Textured Backgrounds for Augmented Reality Systems,” *IEEE/ACM International Symposium on Mixed and Augmented Reality*. 2004.

[40] M. Gattullo, A. E. Uva, M. Fiorentino, and G. Monno, “Effect of Text Outline and Contrast Polarity on AR Text Readability in Industrial Lighting,” *Vis. Comput. Graph. IEEE Trans.*, vol. 21, no. 5, pp. 638–651, 2015.

[41] J. Gabbard, I. I. Edward Swan, and D. Hix, “The effects of text drawing styles, background textures, and natural lighting on text legibility in outdoor augmented reality,” *Presence Teleoperators Virtual Environ.*, vol. 15, no. 1, pp. 16–32, 2006.

[42] Y. Ishiguro and J. Rekimoto, “Peripheral vision annotation,” in *Proceedings of the 2nd Augmented Human International Conference on - AH '11*, 2011, pp. 1–5.

[43] H. Jo, S. Hwang, H. Park, and J. Ryu, “Aroundplot: Focus+context interface for off-screen objects in 3D environments,” *Comput. Graph.*, vol. 35, no. 4, pp. 841–853, Aug. 2011.

[44] T. Schinke, N. Henze, and S. Boll, “Visualization of off-screen objects in mobile augmented reality,” in *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services - MobileHCI '10*, 2010, p. 313.

[45] P. Baudisch and R. Rosenholtz, “Halo: a technique for visualizing off-screen objects,” in *Proceedings of the conference on Human factors in computing systems - CHI '03*, 2003, p. 481.

[46] T. Siu and V. Herskovic, “SidebARs: improving awareness of off-screen elements in mobile augmented reality,” in *Proceedings of the 2013 Chilean Conference on Human - Computer Interaction - ChileCHI '13*, 2013, pp. 36–41.

[47] S. Hart and L. Staveland, “Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research,” in *Human Mental Workload*, P. A. a Hancock and N. Meshkati, Eds. Amsterdam: North Holland Press, 1988.

[48] 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 2011 annual conference on Human factors in computing systems - CHI '11*, 2011, p. 143.

[49] N. Milburn and H. Mertens, “Evaluation of a Range of Target

Blink Amplitude for Attention-Getting Value in a Simulated Air Traffic Control Display.” 1997.

[50] J. Orlosky, K. Kiyokawa, and H. Takemura, “Managing mobile text in head mounted displays: studies on visual preference and text placement,” *ACM Mob. Comput. Commun. Rev.*, vol. 18, no. 2, pp. 20–31, 2014.

[51] “IEC/4WD 61966-2-2: Colour Measurement and Management in Multimedia Systems and Equipment - Part 2.1 : Default Colour Space - sRGB.” .

[52] C. Ridpath and W. Chisholm, “Techniques For Accessibility Evaluation And Repair Tools,” 2000. [Online]. Available: https://www.w3.org/TR/AERT. [Accessed: 01-Mar-2016].

[53] T. Gerstner, D. DeCarlo, M. Alexa, A. Finkelstein, Y. Gingold, and A. Nealen, “Pixelated image abstraction with integrated user constraints,” *Comput. Graph.*, vol. 37, no. 5, pp. 333–347, Aug. 2013.

[54] N. Nilsson, S. Serafin, and R. Nordahl, “Establishing the Range of Perceptually Natural Visual Walking Speeds for Virtual Walking-In-Place Locomotion,” *IEEE Trans. Vis. Comput. Graph.*, vol. 20, no. 4, pp. 569–578, Apr. 2014.

[55] F. Previc and J. Blume, “Visual search asymmetries in three-dimensional space,” *Vision Res.*, vol. 33, no. 18, pp. 2697–2704, Dec. 1993.

[56] R. Rosenholtz, A. Nagy, and N. Bell, “The effect of background color on asymmetries in color search,” *J. Vis.*, vol. 4, no. 3, p. 9, Mar. 2004.

[57] C. Royden, J. Wolfe, and N. Klempen, “Visual search asymmetries in motion and optic flow fields,” *Percept. Psychophys.*, vol. 63, no. 3, pp. 436–444, Apr. 2001.

[58] M. Mclean, “Brightness Contrast, Color Contrast, and Legibility,” *Hum. Factors J. Hum. Factors Ergon. Soc.*, vol. 7, no. 6, pp. 521–527, Dec. 1965.

[59] D. Paterson and M. A. Tinker, “Studies of typographical factors influencing speed of reading.,” *J. Appl. Psychol.*, vol. 15, no. 3, pp. 241–247, 1931.

[60] F. Sumner, “Influence of color on legibility of copy.,” *J. Appl. Psychol.*, vol. 16, no. 2, pp. 201–204, 1932.

[61] C. Gutwin, A. Cockburn, and A. Coveney, “Peripheral Popout: The Influence of Visual Angle and Stimulus on Popout Effects,” in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI ’17*, 2017, pp. 208–219.

[62] J. Wolfe, A. Oliva, T. Horowitz, S. Butcher, and A. Bompas, “Segmentation of objects from backgrounds in visual search tasks,” *Vision Res.*, vol. 42, no. 28, pp. 2985–3004, Dec. 2002.

[63] N. Thomas and L. Elias, “Upper and lower visual field differences in perceptual asymmetries.,” *Brain Res.*, vol. 1387, pp. 108–15, Apr. 2011.

[64] D. Hoffman, A. Girshick, K. Akeley, and M. Banks, “Vergence-accommodation conflicts hinder visual performance and cause visual fatigue.,” *J. Vis.*, vol. 8, no. 3, p. 33.1-30, Jan. 2008.

[65] J. Rolland, R. Holloway, and H. Fuchs, “Comparison of optical and video see-through, head-mounted displays,” in *Photonics for Industrial Applications*, 1995, pp. 293–307.

[66] J. Jones, E. Suma, D. Krum, and M. Bolas, “Comparability of Narrow and Wide Field-Of-View Head-Mounted Displays for Medium-Field Distance Judgments,” in *ACM Symposium in Applied Perception*, 2012.

[67] K. Moser, Y. Itoh, K. Oshima, J. E. Swan, G. Klinker, and C. Sandor, “Subjective Evaluation of a Semi-Automatic Optical See-Through Head-Mounted Display Calibration Technique.,” *IEEE Trans. Vis. Comput. Graph.*, vol. 21, no. 4, pp. 491–500, Apr. 2015.

[68] W. Horrey, C. Wickens, and A. Alexander, “The Effects of Head-up Display Clutter and In-Vehicle Display Separation on Concurrent Driving Performance,” in *Proceedings of the Human Factors and Ergonomics Society*, 2003.

[69] M. Toennis, C. Lange, and G. Klinker, “Visual Longitudinal and Lateral Driving Assistance in the Head-Up Display of Cars,” in *IEEE and ACM International Symposium on Mixed and Augmented Reality*, 2007, pp. 1–4.

[70] C. Sandor, A. Cunningham, A. Dey, and V.-V. Mattila, “An Augmented Reality X-Ray system based on visual saliency,” in *IEEE International Symposium on Mixed and Augmented Reality*, 2010, pp. 27–36.

Ernst Kruijff is interim Professor for Computer Graphics and Interactive Systems at the Institute of Visual Computing, Bonn-Rhein-Sieg University of Applied Sciences, Germany, and Adjunct Professor at Simon Fraser University, Canada. He has published around 70 articles at main IEEE and ACM conferences, and is co-author of the standard reference in the field of 3D user interfaces (LaViola et al, 2017). His current research interests include the usage of multisensory 3D interfaces to overcome perceptual and cognitive issues associated with interacting with complex information sources. Ernst is a member of IEEE and ACM.

Jason Orlosky is a professor at Osaka University, adjunct faculty at Augusta University, and a technology consultant. He earned his BS in Computer Engineering from Georgia Tech in 2006 and his PhD in Information Systems Engineering from Osaka University in 2016. He has been awarded best paper at Intelligent User Interfaces in 2013 and Augmented Human in 2016, was a Japan Society for the Promotion of Science Research Fellow from 2015 - 2016, and has most recently received the Kasami Award from Osaka University. His current research includes improving the safety and usability of intelligent interfaces and augmenting human vision. He has been a member of both IEEE and ACM.

Naohiro Kishishita has received B.S. from School of Engineering Science, Osaka University in 2012, and M.S. from Graduate School of Information Science and Technology, Osaka University in 2014. He has worked on wide field-of-view optical see-through head mounted displays and their evaluations. Since 2014, he has worked for Fujitsu Corporation as a system engineer.

Christina Trepkowski. 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 has contributed to various publications at ACM and IEEE conferences and is currently doing research on multisensory interfaces and virtual reality therapy.

Kiyoshi Kiyokawa is a Professor at Nara Institute of Science and Technology, since 2017. He received his M.S. and Ph.D. degrees in information systems from Nara Institute of Science and Technology in 1996 and 1998, respectively. He worked for Communications Research Laboratory from 1999 to 2002. He was a visiting researcher at Human Interface Technology Laboratory of University of Washington from 2001 to 2002. He was an Associate Professor at Cybermedia Center, Osaka University from 2002 to 2017. His research interests include virtual reality, augmented reality, human augmentation, 3D user interfaces, CSCW, and context awareness. He is a member of IEEE and ACM, and a board member of the Virtual Reality Society of Japan (VRSJ). He is currently a Chair of the Special Interest Group on Mixed Reality of the VRSJ.