HiveFive: Immersion Preserving Attention Guidance in Virtual Reality

Daniel Lange
University of Oldenburg
Oldenburg, Germany
daniel.lange@uol.de

Tim Claudius Stratmann
Uwe Gruenefeld
OFFIS - Institute for IT
first.lastname@offis.de

Susanne Boll
University of Oldenburg
Oldenburg, Germany
susanne.boll@uol.de

Figure 1: Our proposed technique *HiveFive*, in which a swarm visualization is used as a diegetic cue for guiding attention in Virtual Reality (in two different environments - left: forest, right: city). *Best seen in color.*

ABSTRACT

Recent advances in Virtual Reality (VR) technology, such as larger fields of view, have made VR increasingly immersive. However, a larger field of view often results in a user focusing on certain directions and missing relevant content presented elsewhere on the screen. With *HiveFive*, we propose a technique that uses swarm motion to guide user attention in VR. The goal is to seamlessly integrate directional cues into the scene without losing immersiveness. We evaluate *HiveFive* in two studies. First, we compare biological motion (from a prerecorded swarm) with non-biological motion (from an algorithm), finding further evidence that humans can distinguish between these motion types and that, contrary to our hypothesis, non-biological swarm motion results in significantly faster response times. Second, we compare *HiveFive* to four other techniques and show that it not only results in fast response times but also has the smallest negative effect on immersion.

Author Keywords

Attention guidance, virtual reality, immersion, eye-tracking, particle swarms, user studies

CCS Concepts

•**Human-centered computing** → **Virtual reality;** *User centered design; User studies;*

CHI'20, April 25–30, 2020, Honolulu, HI, USA

ISBN 978-1-4503-6708-0/20/04. . . $15.00
DOI: https://doi.org/10.1145/3313831.3376803

INTRODUCTION

In the last decade, Virtual Reality (VR) devices have evolved from expensive and uncomfortable to wear headsets, mostly used in research projects or specialized use cases, to affordable consumer products that are used in everyday life. The technology is defined by Pimentel and Teixeira as "an interactive, immersive experience generated by a computer" [35], in which immersion is understood as the extent to which the used technology provides a living illusion of reality for the human senses [34]. With recent advances, such as improved tracking [30], higher screen resolutions [8] and wider fields of view [36], VR technology has become increasingly immersive.

However, larger fields of view amplify the problem of user missing relevant content presented on the screen, as they are looking in another direction. For example in Cinematic Virtual Reality, one can choose the viewing direction freely and thereby might miss details relevant for the storyline [39]. Similar problems exist in gaming [14] or collaboration [2].

In previous research, different techniques have been proposed to address this problem. For example, techniques that rotate the virtual world to change the direction in which user look [31, 27, 15], techniques that manipulate the environment [45, 18, 9], or techniques that highlight relevant content [1, 26]. However, all these techniques alter the visual properties of objects such as brightness or color, or introduce unnatural features in the virtual environment to draw attention and, therefore, could be perceived as unrealistic and negatively affect the immersion. On the contrary, visual cues can be used that are perceived as part of the scene. For example, a firefly that flies towards a target location of the scene to which the attention should be guided [31]. These visual cues are referred to as diegetic stimuli and have been suggested in Cinematic Virtual Reality

[39]. Diegetic cues have shown to not strongly influence the perceived immersion [40, 39]. Therefore, we think that diegetic cues offer great potential for immersion preserving attention guidance. However, to our knowledge, no diegetic technique exists that is generic enough to be used in a wide range of scenarios. Mostly because a diegetic cue needs to be perceived as a realistic part of the scene.

To address this problem, we investigate in how far swarms can be used to guide attention in VR. Swarms exist in different manifestations in nature (e.g., as swarms of birds, fish or insects) and the general idea of individuals that move as collective can also be applied to particles that are affected by other forces such as wind (e.g., dust, leaves, sparks). Depending on the scene, a specific type of swarm can be selected that is perceived as realistic and therefore, may not negatively affect perceived immersion (e.g., a swarm of bees can be used in a forest environment). Another advantage of swarms is that the individuals of the swarm are in constant motion, which is a strong stimulus in the periphery and therefore, may attract the user's attention. Furthermore, in previous work, it was shown that humans are able to distinguish different types of motion (e.g., human motion [21] or animal motion [28]) and that biological and non-biological motion is processed differently in the brain [12, 11, 20]. Therefore, in a first user study, we compare biological motion of swarms with non-biological motion of swarms to evaluate which motion is perceived faster. Afterward, in a second study, we take the swarm that was perceived quickest and compare it to four other state-of-the-art attention guidance techniques to evaluate which technique results in fastest user responses and is perceived as the most immersive.

Our research contributions are:

- We propose *HiveFive*, a swarm visualization which guides attention in VR and present a comparison of biological and non-biological motion for *HiveFive* in two different environments.
- We evaluate *HiveFive* by comparing it to four existing attention guidance techniques with regard to response times and perceived immersion.

RELATED WORK

Visual Perception and Motion

Visual information enters the human eye in the form of light and is then processed by light-sensitive receptors on the retina [7]. The distribution of these receptors on the retina varies, with most of the cones (color vision receptors) in the center of our vision (referred to as fovea) and more rods (black-and-white vision receptors) outside the center (referred to as periphery) [22]. As a result, humans have sharp and very good color perception in the fovea, while vision in the periphery is less sharp with less good color perception (e.g., yellow can be perceived best and is perceived up to 50° in the periphery) [7, 33]. However, the peripheral vision has a higher temporal resolution than foveal vision, allowing a good perception of motion [32]. Motion is perceived when successive receptors of the retina are stimulated [7]. A special form of motion is biological motion, which originates from biological organisms. Different studies have been conducted that show humans can identify human motion [21] or animal motion [28] presented with dynamic point lights. Furthermore, researchers have shown that biological motion is processed in different brain areas than non-biological motion [12, 11]. Since motion is a strong stimulus in peripheral vision [16] and humans can distinguish between biological and non-biological motion, we investigate if the use of biological motion for guiding attention has a positive effect on response times compared to non-biological motion.

Swarms with Biological and Non-Biological Motion

A swarm is a group of individuals that moves as a collective. In nature, different manifestations of such swarms exist (e.g., flocks of birds, schools of fish or insect swarms). The collective behavior of swarms inspired many researchers and found it's way into different mathematical models and optimization methods (e.g., particle swarm optimization [23]). To simulate swarms in computer-generated environments, Reynolds suggested an algorithm with three simple rules, which each member has to follow: 1) avoid collisions with swarm members, 2) try to match speed with nearby swarm members, 3) try to stay close to other swarm members [38]. To play back swarms with real-world biological motion in a virtual environment, Sinhuber et al. recorded a swarm of insects by tracking the three-dimensional, time-resolved trajectories of individual insects of laboratory insect swarms and making them available as a data set [43]. When investigating the perception of biological motion in swarms, Seiffert et al. found that humans perceive their behavior as biological motion when the whole swarm is visible [42]. To compare biological motion with non-biological motion for guiding attention, we present two different swarms to the user. One swarm representing biological motion based on the recordings by Sinhuber et al. [43] and another swarm representing non-biological motion based on the algorithm by Reynolds [38].

Attention Guidance

We classify previous research focusing on attention guidance in VR into four categories: 1) change of the users' view, 2) manipulation of the environment, 3) highlighting of relevant content and 4) diegetic cues.

An example of a change of the users' view is forced rotation introduced by Nielsen et al. [31]. The technique rotates the virtual body of the user towards relevant content while the user is free to move the head in any direction. Later, Lin et al. compared a similar technique to an arrow-based approach, finding that users preferred the auto-rotate when lots of head-movement is required [27]. However, the virtual rotation may negatively affect the perceived immersion. Therefore, Gugenheimer et al. [15] developed the SwiVRChair, a chair that turns a user's real body in the direction of relevant content, with users rating the experience as significantly more immersive without suffering from simulator sickness. However, additional hardware is required to use this technique.

Another approach is to manipulate the environment. In recent work, Smith et al. used blurred and non-blurred areas in videos to show that viewers can be directed to regions with little or no spatial blur when the rest of an image is blurred [45]. Based on this technique, Hata et al. [18] used Gauss filters to find thresholds at which blur can no longer be perceived without losing guidance. Later, Grogorick et al. [9] implemented blurring in a virtual environment and tested it in combination with a head-mounted display and a full-dome real-time video projection system. In combination with the head-mounted display, blurring was perceived fastest. However, blurring is an unnatural effect that may affect immersion negatively.

Different techniques have been explored to highlight relevant content in VR. For example techniques that use different shapes to point towards the relevant content such as Arrows [5, 3, 19] or Halos [13], or techniques located at the relevant content such as Subtle Gaze Direction (SGD) [1] or DeadEye [26]. SGD changes the area of an image by modulated luminance and warm/cool flickering while DeadEye tries to draw attention to certain objects by rendering them on one eye only. However, all techniques may also affect immersion negatively.

Recently, different so-called diegetic visual cues have been investigated in the field of cinematic virtual reality. These cues can be perceived as part of a movie by both actors and viewers to attract attention. For example, Nielsen et al. proposed a firefly that appeared in the user's field of view and flew into the target area [31]. The majority of the participants found the firefly helpful to follow the story. Rothe et al. [40] also used diegetic stimuli by moving existing objects in a cinematic VR environment. They found that moving objects can attract attention even without sounds [40]. However, there is no diegetic visual cue that is applicable in a wide range of settings. Therefore, we utilize swarms to guide the user's attention because swarms have different manifestations and therefore, may be perceived as realistic in various settings.

Our technique *HiveFive* is inspired by the diegetic visual cues such as the firefly from Nielsen et al. [31]. To evaluate the performance, we selected several state-of-the-art techniques and compare our technique *HiveFive* against these existing techniques in a second user study.

GENERAL APPROACH

We investigate if it is possible to direct the attention of users in a virtual environment using diegetic stimuli without negatively influencing the feeling of immersion. Since motion is a strong stimulus in peripheral vision [16], our technique is based on swarm motion that attracts attention at the target location. Since past research has shown that there is a difference in human perception of biological [12, 11] and non-biological motion, we investigate whether there is a difference in the response times with which a biological swarm can be perceived. There are indications in the literature that biological motion is perceived as early as ~110 ms of stimulus onset [4], whereby the processing of biological motion occurs unintentionally and incidentally, so probably instead in a bottom-up manner compared to other types of motion [20, 46]. Therefore we suspect that biological swarms perform better than non-biological swarms, because biological motion is better perceived than unstructured non-biological motion as it also stands out between other moving interfering objects [20]. Furthermore, biological swarms are well-perceivable [42]. In a second study, we want to find out in direct comparison how quick *HiveFive* can guide the attention compared to other known techniques and, above all, in how far the techniques have a negative effect on the perceived immersion. We suspect that *HiveFive* is more pleasantly perceived as the other non-diegetic techniques.

STUDY 1: COMPARING SWARM MOTION VARIANTS

Study Design

Our study had two independent variables: swarm (biological motion vs. non-biological motion) and environment (forest vs. city). The environment variable was counterbalanced. The swarms appeared equally distributed over three different angles (15° vs. 30° vs. 45°) with two directions (left vs. right) for each swarm variant twice in random order, resulting in 24 trials per swarm and environment. To mix both types of motion, we randomized all trials within two equally sized blocks of 12 trials. We used quantitative methods to evaluate user performance by using the time to first fixation as our dependent variable. The time to first fixation (TTFF) is a standard eye tracking metric defined as the time between stimulus onset (swarm appears) and the first fixation on the area of interest (swarm).

For this study, we asked: **(RQ1) What effect do biological and non-biological swarm motions have on the response times for attention guidance and in how far does motion in the scene influence these times?**

Depending on our concept and our study design, we formulated the following hypothesis:

H_1 We expect the swarm with biological motion to be perceived faster than the swarm with non-biological motion due to them being processed differently in the human brain [12, 11] and for evolutionary reasons.

H_2 We expect that additional motion in the scene masks the motion of the non-biological swarm and negatively influences its perception.

Procedure

At the beginning of the study, the participants received an introduction to Virtual Reality and the HTC Vive head-mounted display. The eye tracker was calibrated and the participants had a short look around in each environment (forest, city) to get familiar with it. Each of the four combinations of swarm and environment was tested with 24 trials in a separate block. Each block was divided into two parts of 12 trials to recalibrate the eye tracking. At the beginning of each block, participants were standing on a marker and there was a short training trial to get familiar with the task.

In both environments, the task was to follow a sphere moving through the field of vision in the form of a Bernoulli lemniscate (see Figure 2). A lemniscate is a polar curve in the shape of an infinity symbol. We used the lemniscate as a trajectory for the distractor stimulus in the form of a moving sphere in our studies. Once started, the sphere took exactly five seconds to get back to the starting point. We did this for two reasons: 1) participants should move their eyes permanently so that we could test under realistic conditions, 2) by moving the eyes to one side, we could test for higher angles to the other side. For

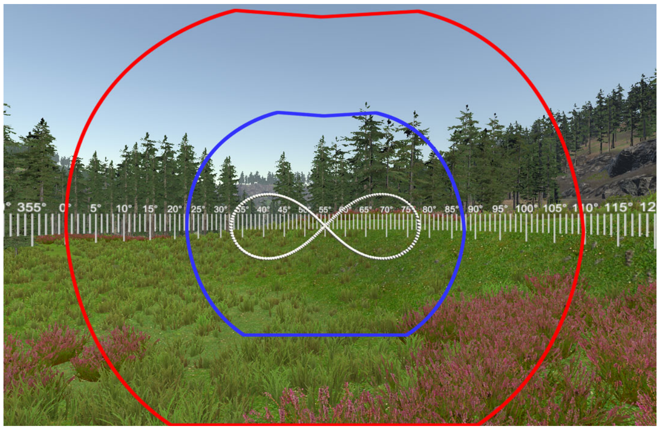

Figure 2: The visible FOV of the HTC Vive is given as 110° (red line) under optimal conditions. However, since the visible range depends on many factors such as the fit of the headset, facial geometry and the distance between eyes and lenses, we have identified an average FOV of 65° (blue line). The lemniscate is located in the centre with a total width of 45° (white line). *Best seen in color.*

example, without moving the eyes, we could present stimuli up to 22.5° to the left or right side. By moving the eyes by 22.5° to one side, we could test up to 45° on the other side.

In each trial, participants were asked to look straight ahead and the sphere was faded in. By clicking a button, the participants could start moving the sphere on their own. Whether the participants were following the sphere was monitored by the study director on the PC who reminded them if necessary. Thereby, we utilized the smooth pursuit eye movement, which is often used in the VR context (e.g., as an input technique [24]), and participants were not able to actively look for the target. After a random time interval of 25±5 seconds, the swarm appeared. Participants were instructed to look at the swarm as soon as possible and were asked to press a button to confirm the perceived stimulus. After the button press or five seconds, the trial was finished and the next trial started with a button press by the participant.

After each block participants were asked to fill out a questionnaire about simulator sickness (VRSQ) [25] and a demographic questionnaire at the end of the experiment. Overall, each participant took approximately 60 minutes to finish the experiment.

Implementation

We implemented our virtual environment and the swarms in the 3D game engine Unity[1] together with an HTC Vive[2] and an aGlass DKII eye tracker [3] of the company 7invensun. The Unity project was set up using SteamVR and the eye tracker was used with the aGlass DKII 2.1.0.5 SDK. The manufacturer indicates the accuracy of the tracker with < 0.5°. Our Unity project is available as an Open Source project on GitHub[4].

Environments

Two environments were chosen to investigate the influence of additional non-biological motion on the perception of the swarms. To create the environments we used the free "Windridge City" package from the Unity Asset Store as it provides both an artificial city environment and a natural forest environment [5]. In the city environment, the participants were placed on a busy street to create permanent non-biological motion in the environment with moving cars, which were animated to cross the user's field of view. A wind simulation also generated motion in the few visible trees. In the forest environment, the participants were placed in a forest with mountains visible in the background. The wind simulation was deactivated in this environment to prevent any motion from grass and trees.

Lemniscate

The implemented Lemniscate with a degree scale can be seen in Figure 2. The width was 45° visual angle, the height was 15° visual angle. We chose the width because the lenses of the HTC Vive create a blurry image in the outer peripheral areas and we wanted to give the participant the possibility to see the sphere sharply at any time. We chose the height because we wanted to focus on the center horizontal area, as this is the area that is most often scanned by the eyes during consumption of VR applications [44].

Shared parameters of the swarms

While designing HiveFive, we took a lot of inspiration from bees, as bees appear in swarms and would fit well within any outdoor environment. At the beginning, we defined the basic parameters that should apply to both swarms: A swarm should be large enough to be well perceived in peripheral vision without covering too much of the background and the motion must also be perceivable in the outer areas of peripheral vision. To generate well visible bees, we have limited ourselves to the color yellow, since some bee species have yellow color components and yellow is also best perceived in peripheral vision. In order to be able to examine only the motion, we haven't used additional stimuli like 3D bee models with moving wings. Our bees consist of a sphere with a diameter of 5mm. Since an average worker bee is 5 - 7mm [6] in size and the background should remain clearly visible, we have decided to use the lower limit. To avoid further graphical influences we used a unicolor shader without reflections and switched off the received and cast shadows for the bees. In order to find out the right speed for bee motion, the right amount of bees and spread for the entire swarm, we conducted a pilot study with three people. A first clue was provided by McKee et al. [29], who found that an object with a size of about 20° at 40° eccentricity must move at more than 30°/second to be perceived as moving, which means that with increasing eccentricity, stimuli must be relatively large and fast. However, since our swarms consist of several small units that move freely and chaotically in all directions within their boundaries, other values seem to fit. Since our swarms move in 3D space, we give the speed in the

[1] http://www.unity3d.com, last retrieved January 7, 2020
[2] http://www.vive.com, last retrieved January 7, 2020
[3] http://www.aglass.com, last retrieved January 7, 2020
[4] https://github.com/danllng/hivefive
[5] https://assetstore.unity.com/, last retrieved January 7, 2020

unit m/second. Derived from the speed at which the recorded midges moved, we found a velocity of 0.35m/second with a spread of 5° visual angle in each direction most pleasant. Too fast movements seem too restless and stressful, too little motion could no longer be perceived well in the outer areas of peripheral vision. We chose 10 bees to avoid cluttering the background too much since too few bees were not well perceivable and too many bees overlaid the background too much. Since our technique is inspired by bees (beehive) and has a spread of 5° in each direction, we call it HiveFive. Due to the shared basic parameters, we ensure that the two swarm variants differ only in the way the particles move. Without prior knowledge of swarms, the two variants are otherwise visually indistinguishable.

Swarm with Biological Motion

To implement a biological swarm in VR, we have used the data set published by Sinhuber et al., which consists of tracked recordings of trajectories of midge swarms [43]. To our knowledge, it is the most accurate tracking data available. A total of 19 data sets with average swarm sizes of 14 to 94 midges were published. The average recording times are between 162 and 1028 frames. Since the midges were tracked automatically, the system partly lost midges and found them again later, so that there are data sets with small swarm sizes but several hundred observations. Each data point in the dataset contains an observation number, coordinates, current speed, tracked time (in milliseconds) and acceleration. For our first study, we tried to find a subset of the dataset that tracks 10 midges continuously over an identical period of 5 seconds, with a spread not greater than 5° visual angle in each direction with normalized vectors. Since a swarm contracts and spreads, its spreading should not fall below 5° visual angle and should not exceed 10°. We chose the data set "Ob4" with to have a number of midges that are tracked continuously over a period of 5 seconds (time the stimulus was presented to the participant). Due to the left out midges, gaps in the swarm can occur, but this is unproblematic, since [21] showed that data of biological movement patterns can be removed without affecting perception. In order to achieve a spreading of 5° in each direction and match the difference in unit measures between the recordings of the data set and our environment, we have reduced the scale of the biological swarm to 0.08% of the original size. To reach the speed of 0.35m/second, we have used 1.8 times the original speed as a counteraction to the rescaling and to restore the original speed.

Swarm with Non-biological Motion

To create a corresponding non-biological swarm, we used the Reynolds flocking algorithm [38] with the ruleset as described in the related work part. We have set the speed to 0.35 and the spread to 0.1 (5° radius).

Participants

We recruited 20 volunteer participants (7 female), aged between 23 and 60 years (M=31.50, SD=11.68). None suffered from color vision impairments, 10 had normal vision, and 10 had corrected-to-normal vision. We asked the participants to rate their experience with Virtual Reality on a 5 point Likert scale. The participants stated that they had very limited experience (Md=1, IQR=1).

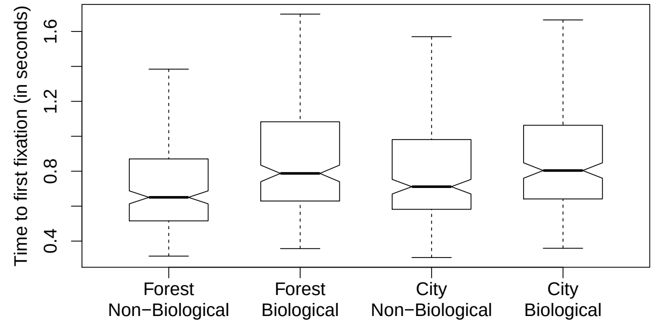

Figure 3: Boxplot of times to first fixation for both swarm motion types for each environment (forest, city).

Results

Time to first fixation

We consider the effect of swarm and environment on TTFF. The median times to first fixation for the city environment are: swarm with non-biological motion=0.71s (IQR=0.40s) and swarm with biological motion=0.80s (IQR=0.42s), and for the forest environment are: swarm with non-biological motion=0.65s (IQR=0.35s) and swarm with biological motion=0.79s (IQR=0.45s). The times are compared in Figure 3.

A Shapiro-Wilk-Test showed that our data is not normally distributed ($p < 0.001$), and thereafter we used Wilcoxon Signed-rank tests to check for significant differences in TTFF between conditions. In the city setting, the swarm with non-biological motion led to a significantly lower TTFF than the swarm with biological motion (W=10685, Z=-2.28, r=0.11, p=0.023). In the forest setting, with less environmental motion, the effect was even stronger, the swarm with non-biological motion led to a significantly lower TTFF than the swarm with biological motion (W=8446, Z=-4.54, r=0.21, p<0.001).

In total, the swarm with non-biological motion (Md=0.69s, IQR=0.38s) led to a significantly lower TTFF than the swarm with biological swarm motion (Md=0.80s, IQR=0.43s) averaged over both settings (W=38188, Z=-4.81, r=0.16, p<0.001).

VR Sickness

The Virtual Reality Sickness Questionnaire (VRSQ) [25] confirmed that the city environment with more motion in the scene negatively affected VR sickness. Table 1 shows the individual scores for each environment including the sub-scores for oculomotor (fatigue, headache, eyestrain, and difficulty focusing) and disorientation (vertigo, dizziness, and blurred vision).

Discussion

Biological vs. Non-biological Swarm Motion

As the times to first fixation show, there is a significant difference between the reaction times of biological and non-biological swarms. E. Hiris [20] found that biological motion is easier to detect than unstructured non-biological motion. It seems to be related to the underlying form of biological

Table 1: The Virtual Reality Sickness Questionnaire (VRSQ) confirmed that the city setting with more motion in the scene negatively affected VR sickness.

Setting	Oculomotor	Disorientation	VRSQ Score
City	47.92	41.00	44.46
Forest	44.17	38.33	41.25

motion. Seiffert et al. [42] identify biological swarm motion as biological motion because it is linked to the motion of biological entities. However, a swarm does not have an underlying form and is unstructured, since there are no fixed connections between the components of a swarm and the entities can move freely. Seiffert et al. found that humans perceive swarm motion as biological motion when the whole swarm is visible. Since our compared swarms are both unstructured and the entities of the non-biological swarm move according to a mathematical model and no biological motion is underlying, we predicted a difference between the reaction times. However, contrary to our hypothesis H_1, the response times for swarms with biological motion were slower than for swarms with non-biological motion. Therefore, we cannot accept our hypothesis H_1.

Influence of the Environment

The additional non-biological motion in the city environment caused worse reaction times, but only for the non-biological swarm. The performance of the biological swarm hardly changed, so that we can accept our hypothesis H_2.

Motion Parameter of HiveFive

Since non-biological swarm motion is perceived significantly faster than biological swarm motion, we now use the non-biological motion according to Reynolds as the basis for our method. This has some advantages as it is easier to implement and configure than the biological variant.

STUDY 2: EVALUATION OF HIVEFIVE

Study Design

To evaluate the performance of different attention guidance techniques, we conducted a within-subjects controlled laboratory study in Virtual Reality with the HTC Vive. Our study had one independent variable *Technique* (*HiveFive* vs. Arrow vs. Blurring vs. DeadEye vs. Subtle Gaze Direction). All techniques were counterbalanced using a Latin square design. The techniques appeared equally distributed over three different angles (15° vs. 30° vs 45°) with two directions (left vs. right) each two times in a randomized order, resulting in a block of 12 trials per *Technique*. We used quantitative methods to evaluate user performance, taking TTFF, object selection accuracy, perceived presence measured with the IGROUP Presence Questionnaire (IPQ) [37, 41], perceived workload measured with the raw NASA task load index (Raw-TLX) [17], and subjective Likert-items as our dependent variables. Different to the first study, TTFF was measured as the time between stimulus onset (technique appears) and the first fixation on the area of interest (apple). We have changed this metric because this study is not about which method is seen the fastest, but whether the method successfully helps to solve the problem to find the right target.

For this study, we asked: **(RQ2) Which technique guides users attention the fastest while preserving the perceived immersion in Virtual Reality?**

H_3 We expect no other technique to outperform *HiveFive* with regard to response time.

H_4 We expect HiveFive to have the least negative impact on immersion.

Procedure

At the beginning of the study, the participants received an introduction to Virtual Reality and the HTC Vive head-mounted display. The process of the eye tracker calibration was shown and the participants had a short look around the environment to get familiar with it. At the beginning of each block, there was a short training run, so that the participants understand the task.

The task aimed to find the randomized target apple in an apple tree, filled with decorative apples. The target apple could only be found with the help of the techniques because otherwise, it had no visible features. First of all the participants were asked to look straight ahead and follow a sphere moving on the path of a Bernoulli lemniscate with their eyes, with identical parameters to those of the first study. There was a start marker on the ground to ensure that all participants were standing in the same place. After alignment, the participants looked straight ahead and the sphere was faded in. By clicking a button, the participants could start moving the sphere on their own. Whether the participants were following the sphere was monitored by the study director on the PC and reminded them if necessary. After a period of random 25±5 seconds, the technique was activated and was supposed to draw the participant's attention to the target apple. The participants were instructed to instantly look at the apple as soon as they perceived the stimulus from the technique and press a button. At this point, the participant was allowed to move his head. As soon as the button was pressed, the stimulus disappeared and a selection tool appeared to select the apple that the participants thought was meant. As the selection tool, we used the SteamVR laser attached to the VR controller. After selecting an apple, the next trial was started. If no apple was selected within 5 seconds, the trial was aborted. If the technique was not noticed within 5 seconds, the trial was also aborted and the next trial started. The selection time was limited to 5 seconds, so that participants could not select apples randomly, and trials do not consume to much time. In pre-tests, the selection time was lower than 2 seconds on average. Since participants had to follow a sphere that was moving within their view, we introduced a button press to identify false positives in which participants did not perceive the stimulus. After each block, the participants were asked to fill out an immersive questionnaire (IPQ), a Raw-TLX questionnaire, a subjective Likert scale questionnaire and an additional demographic questionnaire at the end of the experiment. Overall, each participant took approximately 80 minutes to finish the experiment.

(a) Apple tree without techniques (b) Arrow (c) Blurring

(d) DeadEye (e) HiveFive (f) Subtle Gaze Direction

Figure 4: The apple tree used in the second study as overview (a) and the compared techniques (b-f). *Best seen in color.*

Search task

The study consisted of a search task that took place in the virtual forest setting from the first study. To create a realistic, immersive scenario, the wind simulation was reactivated, so the trees could move, and a large apple tree was added. The tree was filled with decorative apples that served as target objects. Some apples were arranged in the central field of view within the dimensions of the Bernoulli lemniscate in 10 columns and 3 rows, thus the distances between the centers of the apples corresponded to 5° visual angle. Since the immersion feeling has to be measured, the apples were slightly misplaced in order not to appear too unnatural (see Figure 4a). The apples had a diameter of 2° visual angle.

Arrow

Arrow is an obvious stimulus that has already been investigated for VR purposes by Lin et al. [27]. Unlike described there, the arrow does not point in the direction of the target object but is located directly at the edge of the target. The arrow is a three-dimensional arrow as used in [3] for virtual environments. The model was also taken from the Unity Asset Store. The arrow has a length of 5° visual angle and has a yellow color since yellow is the color that can also be perceived in peripheral vision in the outer areas of the periphery (see Figure 4b). The tip of the arrow is located directly on the apple and circles around it, while the back of the arrow points, orthogonally to the apple, in the direction of the participant's gaze.

Blurring

Blurring was implemented as described by Hata et al. [18] and used in its strongest form ($\sigma = 5$). The target was placed centrally in an area with a diameter of 5° visual angle and not blurred (see Figure 4c). Due to the high degree of blurring, the stimulus is no longer subtle but was the only one that could be perceived in our study design with constantly moving eyes and a moving background, since the background is also permanently blurred by the eye movements.

DeadEye

DeadEye was implemented as described by Krekhov et al. [26] by rendering a target on one eye only. In our configuration, we rendered the target on the right eye, not on the left (see Figure 4d). Since according to Krekhov the consideration of the dominant eye brings only slight performance improvements, all participants were tested with the same configuration.

Subtle Gaze Direction

Subtle Gaze Direction was implemented as described by Bailey et al. [1], by placing a two-dimensional round shape directly on the target. The pixels were each brightened or darkened by 9.5%, with a frequency of 10 Hz. Since all other techniques are visible techniques, where the participants can be sure that they have seen them, the luminance modulation was not deactivated after a saccade towards the cue as described by Bailey et al., but remained until the button push (see Figure 4f).

Participants

We recruited 20 volunteer participants (10 female), aged between 24 and 60 years (M=32.00, SD=10.55). None suffered from color vision impairments, 14 had normal vision, and 6 had corrected-to-normal vision. We asked the participants to rate their experience with Virtual Reality on a 5 point Likert-scale. The participants stated that they had limited experience (Md=2, IQR=1.25).

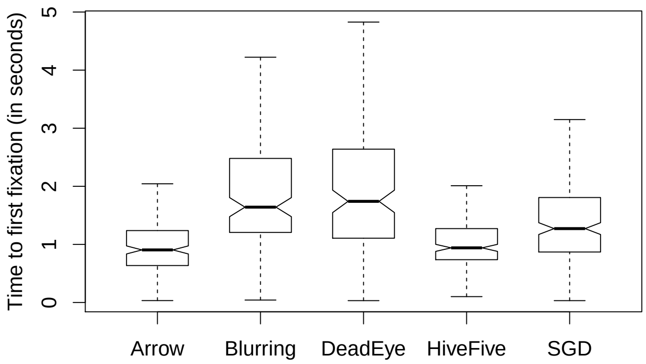

Figure 5: Boxplot of times to first fixation for all techniques.

Results

Number of Perceived Stimuli

We consider trials with a recorded TTFF, a trial in which the participant perceived the stimuli. The total number of perceived stimuli for the different techniques in descending order are: SGD=213/240 (88.8%), *HiveFive*=194/240 (80.8%), Arrow=189/240 (78.8%), DeadEye=157/240 (65.4%), and Blurring=150/240 (62.5%).

A Shapiro-Wilk-Test showed that our data is not normally distributed ($p < 0.001$), and thereafter we ran a Friedman test that revealed a significant effect of *Technique* on the number of perceived stimuli ($\chi^2(4)=35.11$, $p<0.001$, N=20). A posthoc test using Wilcoxon Signed-rank with Bonferroni-Holm correction showed significant differences between some of the conditions (see Table 2). Here, we can conclude *SGD, HiveFive, Arrow > DeadEye, Blurring* for the number of perceived stimuli.

Table 2: Pairwise comparisons of techniques with significant results for the number of perceived stimuli (r: > 0.1 small, > 0.3 medium, and > 0.5 large effect).

Comparison			W	Z	p	r
SGD	vs.	DeadEye	210	-3.96	<0.001	0.63
SGD	vs.	Blurring	11	-3.41	<0.001	0.54
HiveFive	vs.	DeadEye	134	2.87	0.003	0.45
HiveFive	vs.	Blurring	152	2.38	0.015	0.38
Arrow	vs.	DeadEye	146	2.70	0.005	0.43
Arrow	vs.	Blurring	158	2.57	0.009	0.41

Time to First Fixation

The median times to first fixation in ascending order are (trials with no recorded TTFF excluded): Arrow=0.91s (IQR=0.60s), *HiveFive*=0.94s (IQR=0.53s), SGD=1.27s (IQR=0.94s), Blurring=1.64s (IQR=1.27s), and DeadEye=1.74s (IQR=1.53s). The times are compared in Figure 5.

A Shapiro-Wilk-Test showed that our data is not normally distributed ($p < 0.001$), and thereafter we ran a Friedman test that revealed a significant effect of *Technique* on TTFF ($\chi^2(4)=58.11$, $p<0.001$, N=20). A posthoc test using Wilcoxon Signed-rank with Bonferroni-Holm correction showed significant differences between some of the conditions (see Table 3). Here, we can conclude *Arrow, HiveFive < SGD < Blurring, DeadEye* for the TTFF.

Table 3: Pairwise comparisons of techniques with significant results for the TTFF (r: > 0.1 small, > 0.3 medium, and > 0.5 large effect).

Comparison			W	Z	p	r
Arrow	vs.	Blurring	1076	-7.12	<0.001	0.33
Arrow	vs.	DeadEye	859	-7.65	<0.001	0.35
Arrow	vs.	SGD	3280	-6.13	<0.001	0.28
HiveFive	vs.	Blurring	827	-7.41	<0.001	0.34
HiveFive	vs.	DeadEye	992	-7.46	<0.001	0.34
HiveFive	vs.	SGD	3626	-5.91	<0.001	0.27
SGD	vs.	Blurring	5974	3.79	<0.001	0.17
SGD	vs.	DeadEye	6975	4.05	<0.001	0.19

Object Selection Accuracy

For the object selection accuracy, we consider all trials in which the participant perceived the stimuli. The percentage of correctly selected objects in descending order are: SGD=98.1% (209/213), *HiveFive*=97.9% (190/194), Arrow=97.4% (184/189), DeadEye=84.1% (132/157), and Blurring=77.3% (116/150).

Presence

The IGROUP Presence Questionnaire (IPQ) rates presence in four subscales. Table 5 shows the presence ratings per subscale per technique.

Task Load

HiveFive induced the lowest mental, physical and temporal demand, as well as the lowest effort and frustration compared to all other techniques as measured by the Raw-TLX (see Table 4). Arrow closely followed by *HiveFive* led to the best perceived performance. Blurring and DeadEye induced the highest mental demand, effort, and frustration while performing worse than all other techniques.

Subjective Measures

After each condition, we asked the participants to answer two questions with 5-point Likert-scale items (1=strongly disagree, 5=strongly agree). The results are shown in Figure 6. Participants stated that they were not distracted by *HiveFive* (Md=1, IQR=0.25), SGD (Md=1.5, IQR=1.25), and Arrow (Md=2, IQR=2), while they were neutral for Blurring (Md=3, IQR=2) and distracted by DeadEye (Md=4, IQR=1.25). Further, participants stated that *HiveFive* (Md=5, IQR=1) and SGD (Md=4, IQR=1.25) were well embedded in the environment, while they disagreed for Blurring (Md=2.5, IQR=2.25), DeadEye (Md=2, IQR=3), and Arrow (Md=1, IQR=3).

Discussion

Number of Perceived Stimuli

In our final evaluation study of *HiveFive*, we found several interesting results. The attention guidance techniques SGD, Arrow and our approach *HiveFive* dominate significantly over DeadEye and Blurring in the number of perceived stimuli (see Table 2). This effect can be explained by the fact that SGD, HiveFive, and Arrow are methods that are placed at or on

Table 4: Raw-TLX ratings for all techniques (values range from 0 (very low) to 20 (very high)).

Technique	Mental Demand Median (IQR)	Physical Demand Median (IQR)	Temporal Demand Median (IQR)	Performance Median (IQR)	Effort Median (IQR)	Frustration Median (IQR)	Task Load Score %
Arrow	4.0 (4.50)	2.5 (2.25)	2.0 (3.50)	**19.0 (3.00)**	2.5 (2.00)	1.0 (2.25)	28.83
Blurring	8.0 (8.00)	3.0 (6.25)	5.5 (6.50)	9.0 (9.00)	8.5 (6.50)	10.0 (10.50)	39.33
DeadEye	11.5 (8.25)	4.5 (4.50)	5.0 (6.50)	7.0 (6.50)	12.0 (10.25)	10.5 (10.00)	45.83
HiveFive	**2.0 (2.25)**	**2.0 (2.00)**	**2.0 (2.75)**	18.0 (2.25)	**2.0 (2.00)**	**0.0 (2.00)**	**25.50**
SGD	3.5 (3.25)	2.0 (3.50)	3.5 (4.50)	16.5 (4.00)	4.0 (4.00)	2.0 (3.00)	31.92

Table 5: IPQ Presence ratings for all techniques (values range 0 (very low) to 6 (very high)).

Technique	General Presence	Spatial Presence	Involve-ment	Exp. Realism
Arrow	4.75	4.44	3.40	2.61
Blurring	4.40	4.22	3.23	2.35
DeadEye	4.45	4.44	3.16	2.70
HiveFive	**5.05**	**4.65**	**3.65**	**3.30**
SGD	4.70	4.45	3.31	2.88

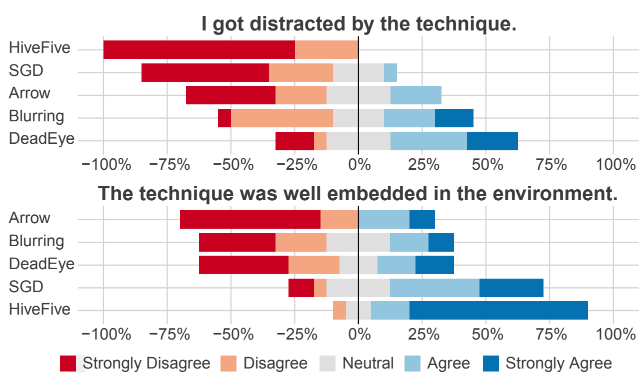

Figure 6: Results from the 5-point Likert-scale questionnaire.

the target while blurring changes the environment but not the target [18]. Furthermore, DeadEye seems to create an effect that is difficult to detect with constantly moving eyes.

Time to First Fixation

We found significant results regarding TTFF, with Arrow and *HiveFive* being perceived the fastest, closely followed by SGD and Blurring, while DeadEye being perceived the slowest (see Table 3). The fact that Arrow and *HiveFive* were perceived the fastest is not surprising since they are the most obvious stimuli, but the significant distance to SGD is remarkable, since flickering is also a strong stimulus in the peripheral vision [1, 16]. We suspect that the participants' concentration on the sphere, combined with the permanent eye movement and motion within the scene, have made perception more difficult, which Grogorick et al. also assumed [10]. Although *HiveFive* was very fast, Arrow was perceived minimally faster, which is why we cannot accept hypothesis H_3.

Object Selection Accuracy

Whether the technique successfully helped to find the correct apple is shown by the object selection accuracy. SGD was the most reliable, followed by *HiveFive*, Arrow and finally DeadEye and Blurring. The fact that *HiveFive* leads to mistakes despite the obvious stimulus could be because the swarm is constantly in motion. This movement touches other apples in the area of the target apple due to its spread, so that in rare cases participants could not correctly assess which apple was meant.

Presence

The results of the IPQ shows that the participants experienced both the highest sense of presence and the highest degree of realism with *HiveFive*, which allows us to accept our hypothesis H_4 (see Table 5). As expected, the diegetic stimuli of *HiveFive* allows us to guide attention while maintaining the sense of presence and thus immersion. In contrast to Arrow, SGD, Blurring and DeadEye, which generate attention through artificial stimuli while manipulating the scene, *HiveFive* fits well into the environment, and therefore, has a positive effect on the perceived immersion.

Task Load

The results of the Raw-TLX rating reflect the previous results (see Table 4). *HiveFive* is ahead in all ratings, except for the performance, where the participants rated themselves slightly better with Arrow. Overall, the task load is lowest for HiveFive. Here it is particularly interesting that the load seems to be lower compared to Arrow, which is a further indication that diegetic stimuli are perceived as well integrated into the scene and suggest that they offer a natural way to guide attention.

Subjective Measures

A large majority of the participants shared the opinion that *HiveFive* fits best into the environment, which suggests that *HiveFive* was perceived as a diegetic stimulus (see Figure 6). SGD also achieved good results due to the unobtrusive stimulus. Blurring and DeadEye did not perform so well and Arrow did perform the least well in the environment. While very few people felt distracted by *HiveFive*, DeadEye and Blurring distracted them the most. A Shapiro-Wilk-Test showed that our data is not normally distributed ($p < 0.001$), and thereafter we ran a Friedman test that revealed a significant effect of *distraction* (Likert-item) on technique ($\chi^2(4)=41.84$, $p<0.001$, $N=20$). A posthoc test using Wilcoxon Signed-rank with Bonferroni-Holm correction showed that HiveFive is less distracting than SGD ($p=.02$), Arrow, Blurring, and DeadEye ($p<.001$). This means that with *HiveFive*, we have developed a technique that attracts attention very fast but without being perceived as unnatural or distracting in the environment.

GENERAL DISCUSSION

Adaptability to the Scenario

HiveFive is a technique that must be carefully adapted to the environment. In our experiments, we have tried to create a diegetic stimulus that is useful in many different environments. Since our technology is strongly inspired by bees, we have decided to use fitting environments. However, the technology must always be adapted to the scenario and the environment to preserve immersion. A swarm of bees in an office does not make as much sense as dust or midges flying in front of a light. The results of our work show, however, that motion stimuli that fit well into the environment of a VR scene are perceived by the users as more pleasant while not losing their effectiveness.

Finding the Right Parameters

HiveFive's noticeability can probably be improved by adjusting the parameters. Changes in the color, speed, and spread of the particles are possible. One can also increase the number of particles and use animated models, such as 3D models of bees. An animated flapping of the wings could even add a reasonable flickering to the stimulus. However, since *HiveFive* is based on Reynolds' flocking algorithm, the speed of motion and the distance between the particles have to be fine-tuned to create a reasonable swarm effect [38]. With unfavorable parameters, it could occur that the particles circle around a point and thus form a too even structure, which may lead to less effective guiding of attention.

Transferability of the Results

With cinematic VR or gaming as the main application fields of our technique, there is a high chance of having motion in the scene. Our first study showed two interesting things. First, we saw that additional motion influences the performance of the non-biological swarm. Second, it appears as though the biological motion is less affected by this and that even with more motion in the background the perception of the biological swarm stays more or less the same. Suggesting that the biological motion could be a good alternative for scenes with a lot of motion.

Limitations

Since *HiveFive* is using a visual diegetic stimulus, it must be carefully adapted to the context of the used environment. However, this also means that there will be environments in which the method makes no sense. For example, flies in an office environment may make sense, but it is not common for them to be in swarms and users might feel irritated by a swarm visualization in this context. Although we have not had any negative experiences with our visualization, we can imagine that the swarm movement could have a negative effect on some people if they have phobias about animals that move in swarms.

FUTURE WORK

We think that *HiveFive* offers great potential for future work of attention guidance in complex interactive VR worlds. In our controlled experiments, the technology was positioned somewhere in the field of view. For our future work, we will investigate how the technique can be positioned at the edge of the visible area and if this supports head movements and attention shift towards the appropriate direction. We will investigate, based on the visual phenomenon of smooth pursuit, if HiveFive can be adapted to direct the user's attention not only within the visible area but also in the non-visible area of the immersive environment. For this purpose, individual particles or the whole swarm could fly from a visible area into a direction outside the field of view.

CONCLUSION

In this paper, we investigated the difference in the perception of biological and non-biological swarm motion in an immersive environment. We found that humans immersed into a Virtual Reality can distinguish between both forms of motion and that the quality of perception of non-biological motion deteriorates when additional non-biological motion takes place in an environment. In addition, we developed a new diegetic visualization technique for attention guidance in VR and compared it with four other state-of-the-art techniques. Although *HiveFive* did not produce the fastest response times, it is one of the fastest techniques without negatively affecting the user's sense of presence and thus maintaining immersion. We expect that *HiveFive* can be further improved in future work and adapted to make users aware of targets outside of the field of view.

REFERENCES

[1] Reynold Bailey, Ann McNamara, Nisha Sudarsanam, and Cindy Grimm. 2009. Subtle gaze direction. *ACM Transactions on Graphics* 28, 4 (Aug. 2009), 1–14. DOI: http://dx.doi.org/10.1145/1559755.1559757

[2] Cullen Brown, Ghanshyam Bhutra, Mohamed Suhail, Qinghong Xu, and Eric D. Ragan. 2017. Coordinating attention and cooperation in multi-user virtual reality narratives. In *2017 IEEE Virtual Reality (VR)*. 377–378. DOI:http://dx.doi.org/10.1109/VR.2017.7892334

[3] Stefano Burigat and Luca Chittaro. 2007. Navigation in 3D virtual environments: Effects of user experience and location-pointing navigation aids. *International Journal of Human-Computer Studies* 65, 11 (Nov. 2007), 945–958. DOI: http://dx.doi.org/10.1016/j.ijhcs.2007.07.003

[4] George Buzzell, Laura Chubb, Ashley S. Safford, James C. Thompson, and Craig G. McDonald. 2013. Speed of Human Biological Form and Motion Processing. *PLoS ONE* 8, 7 (July 2013), e69396. DOI: http://dx.doi.org/10.1371/journal.pone.0069396

[5] Luca Chittaro and Stefano Burigat. 2004. 3D location-pointing as a navigation aid in Virtual Environments. In *Proceedings of the working conference on Advanced visual interfaces - AVI '04*. ACM Press, Gallipoli, Italy, 267. DOI: http://dx.doi.org/10.1145/989863.989910

[6] Manuel A. Giannoni-Guzmán, Arian Avalos, Jaime Marrero Perez, Eduardo J. Otero Loperena, Mehmet Kayım, Jose Alejandro Medina, Steve E. Massey, Meral Kence, Aykut Kence, Tugrul Giray, and José L. Agosto-Rivera. 2014. Measuring individual locomotor rhythms in honey bees, paper wasps and other similar-sized insects. *The Journal of Experimental Biology* 217, 8 (April 2014), 1307. DOI: http://dx.doi.org/10.1242/jeb.096180

[7] E. Bruce Goldstein and James R. Brockmole. 2017. *Sensation and perception* (10th edition ed.). Cengage Learning, Boston. OCLC: 952665360.

[8] Fangwang Gou, Haiwei Chen, Ming-Chun Li, Seok-Lyul Lee, and Shin-Tson Wu. 2017. Submillisecond-response liquid crystal for high-resolution virtual reality displays. *Opt. Express* 25, 7 (Apr 2017), 7984–7997. DOI: http://dx.doi.org/10.1364/OE.25.007984

[9] Steve Grogorick, Georgia Albuquerque, and Marcus Maqnor. 2018. Gaze Guidance in Immersive Environments. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, Tuebingen/Reutlingen, Germany, 563–564. DOI: http://dx.doi.org/10.1109/VR.2018.8446215

[10] Steve Grogorick, Michael Stengel, Elmar Eisemann, and Marcus Magnor. 2017. Subtle gaze guidance for immersive environments. In *Proceedings of the ACM Symposium on Applied Perception - SAP '17*. ACM Press, Cottbus, Germany, 1–7. DOI: http://dx.doi.org/10.1145/3119881.3119890

[11] Emily D. Grossman, Lorella Battelli, and Alvaro Pascual-Leone. 2005. Repetitive TMS over posterior STS disrupts perception of biological motion. *Vision Research* 45, 22 (Oct. 2005), 2847–2853. DOI: http://dx.doi.org/10.1016/j.visres.2005.05.027

[12] Emily D. Grossman, M. Donnelly, R. Price, D. Pickens, V. Morgan, G. Neighbor, and Randolph Blake. 2000. Brain Areas Involved in Perception of Biological Motion. *Journal of Cognitive Neuroscience* 12, 5 (Sept. 2000), 711–720. DOI: http://dx.doi.org/10.1162/089892900562417

[13] Uwe Gruenefeld, Abdallah El Ali, Susanne Boll, and Wilko Heuten. 2018. Beyond Halo and Wedge: Visualizing Out-of-view Objects on Head-mounted Virtual and Augmented Reality Devices. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '18)*. ACM, New York, NY, USA, Article 40, 11 pages. DOI: http://dx.doi.org/10.1145/3229434.3229438

[14] Uwe Gruenefeld, Ilja Koethe, Daniel Lange, Sebastian Weiß, and Wilko Heuten. 2019. Comparing Techniques for Visualizing Moving Out-of-View Objects in Head-mounted Virtual Reality. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 742–746. DOI: http://dx.doi.org/10.1109/VR.2019.8797725

[15] Jan Gugenheimer, Dennis Wolf, Gabriel Haas, Sebastian Krebs, and Enrico Rukzio. 2016. SwiVRChair: A Motorized Swivel Chair to Nudge Users' Orientation for 360 Degree Storytelling in Virtual Reality. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*. ACM Press, Santa Clara, California, USA, 1996–2000. DOI: http://dx.doi.org/10.1145/2858036.2858040

[16] Carl Gutwin, Andy Cockburn, and Ashley Coveney. 2017. Peripheral Popout: The Influence of Visual Angle and Stimulus Intensity on Popout Effects. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*. ACM Press, Denver, Colorado, USA, 208–219. DOI: http://dx.doi.org/10.1145/3025453.3025984

[17] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Advances in Psychology*. Vol. 52. Elsevier, 139–183. DOI:http://dx.doi.org/10.1016/S0166-4115(08)62386-9

[18] Hajime Hata, Hideki Koike, and Yoichi Sato. 2016. Visual Guidance with Unnoticed Blur Effect. In *Proceedings of the International Working Conference on Advanced Visual Interfaces - AVI '16*. ACM Press, Bari, Italy, 28–35. DOI: http://dx.doi.org/10.1145/2909132.2909254

[19] Niels Henze, Benjamin Poppinga, and Susanne Boll. 2010. Experiments in the wild: public evaluation of off-screen visualizations in the Android market. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction Extending Boundaries - NordiCHI '10*. ACM Press, Reykjavik, Iceland, 675. DOI:http://dx.doi.org/10.1145/1868914.1869002

[20] Eric Hiris. 2007. Detection of biological and nonbiological motion. *Journal of Vision* 7, 12 (Sept. 2007), 4. DOI:http://dx.doi.org/10.1167/7.12.4

[21] Gunnar Johansson. 1973. Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics* 14, 2 (June 1973), 201–211. DOI: http://dx.doi.org/10.3758/BF03212378

[22] James Kalat. 2015. *Biological psychology*. Nelson Education.

[23] James Kennedy and Russel C. Eberhart. 1995. Particle swarm optimization. In *Proceedings of ICNN'95 - International Conference on Neural Networks*, Vol. 4. IEEE, Perth, WA, Australia, 1942–1948. DOI: http://dx.doi.org/10.1109/ICNN.1995.488968

[24] Mohamed Khamis, Carl Oechsner, Florian Alt, and Andreas Bulling. 2018. VRpursuits: interaction in virtual reality using smooth pursuit eye movements. In *Proceedings of the 2018 International Conference on Advanced Visual Interfaces - AVI '18*. ACM Press, Castiglione della Pescaia, Grosseto, Italy, 1–8. DOI: http://dx.doi.org/10.1145/3206505.3206522

[25] Hyun K. Kim, Jaehyun Park, Yeongcheol Choi, and Mungyeong Choe. 2018. Virtual reality sickness questionnaire (VRSQ): Motion sickness measurement index in a virtual reality environment. *Applied Ergonomics* 69 (2018), 66 – 73. DOI: http://dx.doi.org/https://doi.org/10.1016/j.apergo.2017.12.016

[26] Andrey Krekhov, Sebastian Cmentowski, Andre Waschk, and Jens Krüger. 2019. Deadeye Visualization Revisited: Investigation of Preattentiveness and Applicability in Virtual Environments. *arXiv:1907.04702 [cs]* (July 2019). http://arxiv.org/abs/1907.04702 arXiv: 1907.04702.

[27] Yen-Chen Lin, Yung-Ju Chang, Hou-Ning Hu, Hsien-Tzu Cheng, Chi-Wen Huang, and Min Sun. 2017. Tell Me Where to Look: Investigating Ways for Assisting Focus in 360° Video. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*. ACM Press, Denver, Colorado, USA, 2535–2545. DOI: http://dx.doi.org/10.1145/3025453.3025757

[28] George Mather and Sophie West. 1993. Recognition of Animal Locomotion from Dynamic Point-Light Displays. *Perception* 22, 7 (July 1993), 759–766. DOI: http://dx.doi.org/10.1068/p220759

[29] Suzanne P. Mckee and Ken Nakayama. 1984. The detection of motion in the peripheral visual field. *Vision Research* 24, 1 (Jan. 1984), 25–32. DOI: http://dx.doi.org/10.1016/0042-6989(84)90140-8

[30] Diederick C. Niehorster, Li Li, and Markus Lappe. 2017. The Accuracy and Precision of Position and Orientation Tracking in the HTC Vive Virtual Reality System for Scientific Research. *i-Perception* 8, 3 (2017), 2041669517708205. DOI: http://dx.doi.org/10.1177/2041669517708205

[31] Lasse T. Nielsen, Matias B. Møller, Sune D. Hartmeyer, Troels C. M. Ljung, Niels C. Nilsson, Rolf Nordahl, and Stefania Serafin. 2016. Missing the point: an exploration of how to guide users' attention during cinematic virtual reality. In *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology - VRST '16*. ACM Press, Munich, Germany, 229–232. DOI: http://dx.doi.org/10.1145/2993369.2993405

[32] Thomas E. Ogden and Robert F. Miller. 1966. Studies of the optic nerve of the rhesus monkey: Nerve fiber spectrum and physiological properties. *Vision Research* 6, 9-10 (Oct. 1966), 485–IN2. DOI: http://dx.doi.org/10.1016/0042-6989(66)90001-0

[33] Gustav Osterberg. 1935. *Topography of the layer of rods and cones in the human retina*. A. Busck. https://books.google.de/books?id=DeDrSAAACAAJ

[34] Randy Pausch, Dennis Proffitt, and George Williams. 1997. Quantifying immersion in virtual reality. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques - SIGGRAPH '97*. ACM Press, Not Known, 13–18. DOI: http://dx.doi.org/10.1145/258734.258744

[35] Kan Pimentel and Kevin Teixeira. 1994. *Virtual Reality: Through the New Looking Glass* (2nd ed.). McGraw-Hill, Inc., New York, NY, USA.

[36] Ismo Rakkolainen, Matthew Turk, and Tobias Höllerer. 2016. A Compact, wide-FOV Optical Design for Head-mounted Displays. In *Proceedings of the 22Nd ACM Conference on Virtual Reality Software and Technology (VRST '16)*. ACM, New York, NY, USA, 293–294. DOI: http://dx.doi.org/10.1145/2993369.2996322

[37] Holger Regenbrecht and Thomas Schubert. 2002. Real and Illusory Interactions Enhance Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments* 11, 4 (Aug. 2002), 425–434. DOI: http://dx.doi.org/10.1162/105474602760204318

[38] Craig W. Reynolds. 1987. Flocks, herds and schools: A distributed behavioral model. *ACM SIGGRAPH Computer Graphics* 21, 4 (Aug. 1987), 25–34. DOI: http://dx.doi.org/10.1145/37402.37406

[39] Sylvia Rothe, Daniel Buschek, and Heinrich Hußmann. 2019. Guidance in Cinematic Virtual Reality-Taxonomy, Research Status and Challenges. *Multimodal Technologies and Interaction* 3, 1 (2019). DOI: http://dx.doi.org/10.3390/mti3010019

[40] Sylvia Rothe and Heinrich Hußmann. 2018. Guiding the Viewer in Cinematic Virtual Reality by Diegetic Cues. In *Augmented Reality, Virtual Reality, and Computer Graphics*, Lucio Tommaso De Paolis and Patrick Bourdot (Eds.). Vol. 10850. Springer International Publishing, Cham, 101–117. DOI: http://dx.doi.org/10.1007/978-3-319-95270-3_7

[41] Thomas Schubert, Frank Friedmann, and Holger Regenbrecht. 2001. The Experience of Presence: Factor Analytic Insights. *Presence: Teleoperators and Virtual Environments* 10, 3 (June 2001), 266–281. DOI: http://dx.doi.org/10.1162/105474601300343603

[42] Adriane E. Seiffert, Sean T. Hayes, Caroline E. Harriott, and Julie A. Adams. 2015. Motion perception of biological swarms. (2015), 6.

[43] Michael Sinhuber, Kasper van der Vaart, Rui Ni, James G. Puckett, Douglas H. Kelley, and Nicholas T. Ouellette. 2019. Three-dimensional time-resolved trajectories from laboratory insect swarms. *Scientific Data* 6 (March 2019), 190036. https://doi.org/10.1038/sdata.2019.36

[44] Vincent Sitzmann, Ana Serrano, Amy Pavel, Maneesh Agrawala, Diego Gutierrez, Belen Masia, and Gordon Wetzstein. 2018. Saliency in VR: How Do People Explore Virtual Environments? *IEEE Transactions on Visualization and Computer Graphics* 24, 4 (April 2018), 1633–1642. DOI: http://dx.doi.org/10.1109/TVCG.2018.2793599

[45] Wayne S. Smith and Yoav Tadmor. 2013. Nonblurred regions show priority for gaze direction over spatial blur. *Quarterly Journal of Experimental Psychology* 66, 5 (May 2013), 927–945. DOI: http://dx.doi.org/10.1080/17470218.2012.722659

[46] Ian M. Thornton and Quoc C. Vuong. 2004. Incidental Processing of Biological Motion. *Current Biology* 14, 12 (June 2004), 1084–1089. DOI: http://dx.doi.org/10.1016/j.cub.2004.06.025