WarpVision: Using Spatial Curvature to Guide Attention in Virtual Reality

Jérôme Kudnick, Martin Weier, Colin Groth, Biying Fu, Robin Horst

Fig. 1: A VR scene of a bistro, showing the warped space created by *WarpVision* in order to guide the visual attention to the menu board (highlighted by the red cross). The strength of the effect depends on eccentricity in the visual field determined by eye tracking (the current gaze position is represented by the eye symbol on the left).

Abstract— With the advent of consumer-targeted, low-cost virtual reality devices and facile authoring technologies, the development and design of experiences in virtual reality are also becoming more accessible to non-expert authors. However, the inherent freedom of exploration in these virtual spaces presents a significant challenge for designers seeking to guide user attention toward points and objects of interest. This paper proposes the new technique *WarpVision*, which utilizes spatial curvature to subtly guide the user's attention in virtual reality. *WarpVision* distorts an area around the point of interest, thus changing the size, form, and location of all objects and the space around them. In this way, the user's attention can be guided even when the point of interest is not in the immediate field of vision. *WarpVision* is evaluated in a user study based on a within-subjects design, comparing it to the state-of-the-art technique *Deadeye*. Participants completed visual search tasks across two virtual environments being supported with *WarpVision* at four different intensities. Results show that *WarpVision* significantly reduces search times compared to *Deadeye*. While both techniques introduce comparable levels of immersion disruption, *WarpVision* has a lower reported impact on the user's well-being.

Index Terms—Gaze guidance, subtle cues, immersion preserving, screen-space filtering, shader, spatial distortion, user navigation, spatial awareness, eye tracking, extended reality, user study.

1 Introduction

Virtual reality (VR) is experiencing increasing adoption, driven by the release of new commercial head-mounted displays. Common applications include immersive gaming and panoramic 360° video content, where the user is free to dynamically choose their viewport within the virtual environment. This flexibility enhances immersion and personalizes experiences, but also introduces the risk of users overlooking important content due to suboptimal viewing directions. For instance, a user may miss a critical in-game event or fail to notice a key narrative element in a 360° video simply because their gaze was directed elsewhere. In industrial contexts such as remote maintenance or virtual training, the consequences of such misalignment are more severe. Failure to identify relevant components or equipment in time can lead to inefficiencies, safety risks, and miscommunication. In these cases, subtle guidance mechanisms that direct the user's attention to relevant areas without disrupting immersion are highly beneficial. Techniques that serve this purpose are referred to as *gaze guidance techniques*.

Gaze guidance techniques aim to direct visual attention to specific points or objects within a scene. Various approaches based on visual, auditory, and haptic cues have been developed, with visual techniques being the most prevalent due to their ease of integration. These are broadly categorized into direct cues—which act on the target itself using visual enhancements such as outlines or brightness changes—and indirect cues—which include symbolic elements like arrows or markers placed in the periphery [22]. A core challenge in gaze guidance lies in balancing subtlety and effectiveness. Techniques that are highly reliable in capturing attention often do so at the cost of immersion, whereas subtle cues may go unnoticed and lose their guiding function. Additionally, methods that rely on the 3D structure of the scene (e.g. spatial indicators) suffer from limited generalizability and are sensitive to scene complexity.

This paper contributes to the current state of research by introducing *WarpVision*, a novel screen-space gaze guidance technique that locally warps the visual environment to direct user attention in VR (see Figure 1). By applying a geometric distortion dependent on gaze and eccentricity around a point of interest (POI), *WarpVision* guides attention even when the POI lies outside the user's immediate field of view. The distortion dynamically adapts based on gaze: regions in

- *Jérôme Kudnick is with RheinMain University of Applied Sciences. E-mail: jerome.kudnick@hs-rm.de*
- *Martin Weier is with RheinMain University of Applied Sciences. E-mail: martin.weier@hs-rm.de*
- *Colin Groth is with Max Planck Institute for Informatics. E-mail: cgroth@mpi-inf.mpg.de*
- *Biying Fu is with RheinMain University of Applied Sciences. E-mail: biying.fu@hs-rm.de*
- *Robin Horst is with RheinMain University of Applied Sciences. E-mail: robin.horst@hs-rm.de*

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

focus remain undistorted, preserving visual fidelity. Operating purely in screen-space, *WarpVision* requires no scene geometry or semantic knowledge, enabling broad applicability across diverse VR content. A user study in artificial and realistic VR environments demonstrates that *WarpVision* outperforms existing subtle guidance techniques in accuracy and reliability while maintaining immersion.

2 Related Work

In the context of VR guidance, several techniques in the visual, haptic, and auditory domain have been developed. We focus on visual cues that provide robust, versatile, and accurate guidance. In addition, elaborate haptic methods and spatial audio are not readily available in consumer-level HMDs. Previous work in the field of visual gaze guidance can be divided into three categories: Overt, subtle, and diegetic techniques.

Overt Techniques Overt techniques describe methods that use cues that can be directly perceived and identified by the user. Direct recognition is characterized by the high and easy visibility (e.g. through high contrasts) of these methods. Probably the most known way to specify a direction and allow guidance to POIs outside the current field of view is to use directional arrows [12, 24, 27]. In the work by Lin et al. [12], for example, such arrows are used directly and visibly in 360° videos to direct the view to the POI. The results showed a high level of efficiency in supporting users when viewing 360° videos and keeping the area to be viewed in focus. Similarly to arrows, Gruenefeld et al. [5] use halos and isosceles triangles as clearly visible cues to draw attention to objects outside the field of view in virtual environments. Another technique, proposed by Renner and Pfeiffer [21], uses spherical waves moving towards the POI which, unlike arrows and halos, are not only placed in one specific location, but expand throughout the entire field of view. Additionally, their speed adapts based on the angle between the user's gaze direction to the POI. This technique was also compared with guiding arrows and showed a comparable effectiveness in gaze guidance, while potentially being more subtle. Other overt techniques use radar-like visualizations to indicate and convey the direction, distance, and position of multiple objects and POIs in a scene. Bork et al. [2] evaluates several of these techniques, showing that while they effectively guide users, they also introduce visual clutter, occupying part of the user's field of view.

With a view to attention guidance in multi-user VR environments, there also exist different overt techniques. Peter et al. [18] introduce a specialized tool designed to assist "VR guides", who are users in an asymmetrically immersive environment that guides fully immersed users. In particular, the work focuses on highlighting relevant occluded objects by outlining them. Furthermore, Horst et al. [6] investigate a similar guidance scenario, but focus on view-related visualizations for interactive guidance.

A systematic review of overt gaze guiding techniques is given by Quinn and Gabbard [20]. Particularly, this work examines techniques applicable in augmented reality that work if the POI is outside of the user's current view. It is concluded that, while many of these techniques may offer benefits, like lower search times and reduced cognitive load, the apparent visualizations can lead to perceptual issues such as clutter and occlusion. Another evaluation by Woodworth and Borst [26] reported similar findings, further suggesting that cues that directly connect the user's gaze to a target are more effective for guidance as opposed to an indirectly conveyed direction. To conclude, while overt techniques are quickly visible and easy to understand, they can have a strong negative impact on immersion in virtual realities and thus can be disruptive, affecting the user experience [22].

Subtle Techniques Subtle techniques aim to draw the user's attention without their awareness. The primary goal is to maintain immersion and prevent any disruption in the experience. The *Deadeye* technique [9] uses the principles of preattentive perception. By exploiting the property of stereoscopy of virtual worlds, and only displaying the target object in one eye while remaining invisible to the other, the gaze is directed to said object. This technique has proven to be effective, even in environments with many distracting factors. Furthermore, the initial appearance of objects remains unchanged, as the color and sharpness remain the same and are not overlapped by other objects. A disadvantage is the possibility of one eye seeing areas of the scene behind the target object, which increases the mental load. The perceptual mismatch between both eyes can cause discomfort [10].

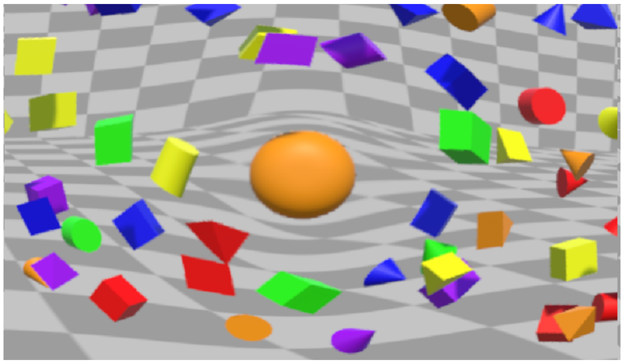
Work by Bailey et al. [1] introduces a technique that combines eye tracking with subtle screen-space modulation to direct a viewer's gaze toward a digital image, referred to as subtle gaze direction. This approach takes advantage of the limited acuity of peripheral vision compared to foveal vision. By presenting brief, subtle modulations in the peripheral field of view, the technique draws the viewer's foveal attention to the modulated area. The changes in contrast involve adjustments in the brightness of a region of the target object, as well as a warm-cool modulation of the corresponding colors. The results demonstrated a strong effect of the technique; however, the subjective perception of image quality was poorer for modulated images. Grogorick et al. [4] combine color modulation with the ability to leverage stereoscopy. The technique utilized takes advantage of the binocular rivalry effect inherent in human stereo vision and has been demonstrated to be effective in static environments. Additionally, they suggest an enhancement of the method to enable reliable guidance toward POIs beyond the initial visual field. However, the disadvantages of this and other subtle techniques in general can be found in the subtlety itself. Although they remain less recognizable compared to obvious techniques, their effectiveness can decrease accordingly [22]. In addition, they often fail to provide proper controls to balance reliability and perceptibility.

Diegetic Techniques There also exist some diegetic techniques for attention guidance in VR environments. These techniques use cues as part of the environment and narrative world, allowing them to be clearly visible but act more unobtrusively, as they may be perceived as less distracting [22]. Diegetic techniques are used in multiple forms, such as swarm movements [7, 11], animals [17], or a firefly [16] to emphasize the POI and/or guide the user to the relevant information. Regardless of the technique used, diegetic methods have two disadvantages. If the cues used are created directly on the POI, they cannot be perceived outside the field of view. Furthermore, no general solution can be created when diegetic techniques are employed, as the narrative can differ between worlds [22].

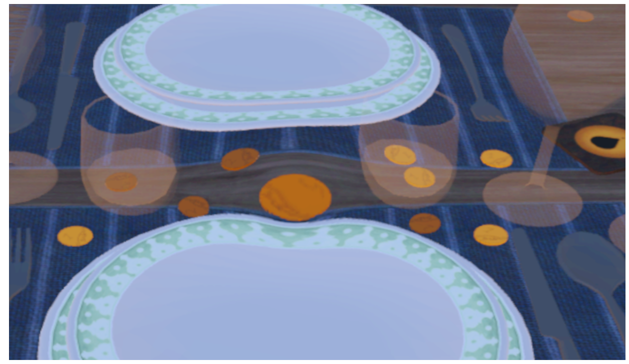
With diegetic techniques being tailored to specific environments, overt techniques being disruptive to the immersion, and subtle techniques being too ineffective, the need prevails to find a reliable and controllable approach to guide the user's gaze and attention in VR—ideally including out-of-view objects or POIs. To address these restrictive properties of existing techniques, we propose *WarpVision* as a new technique to guide attention with the goal of achieving results similar to overt gaze guidance while remaining as subtle as possible. Furthermore, by being non-diegetic and allowing out of view guidance, *WarpVision* is supposed to be applicable in a diverse range of scenarios.

3 WarpVision

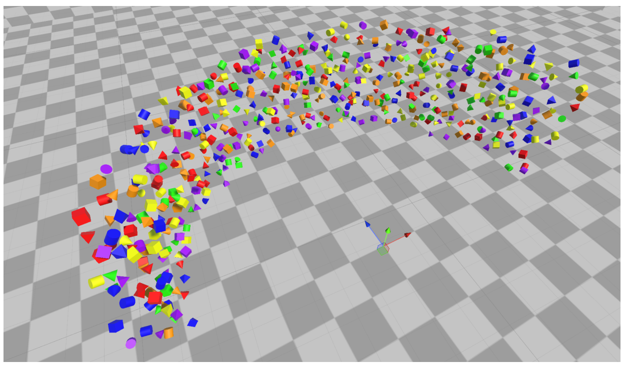
WarpVision utilizes the curvature of space to distort an area around a POI in a similar way to a zoom/fisheye lens, enlarging it and its surroundings as a result. The strength of the effect is dynamically controlled using eye tracking. If the user is not looking at the current POI, the effect becomes stronger. If the user's gaze point approaches the POI, the effect becomes less pronounced. The aim is to stimulate preattentive perception in the peripheral field of vision through the visual variables of movement, position, and shape, thus acting as a direct cue that does not require prior knowledge to interpret. Two examples of this property are shown in Figures 2a and 2b. Located in the center is the object of relevance acting as POI, which is distorted by *WarpVision* together with the surrounding environment. The distortion is strongest in the center and decreases towards the edge, which influences the shape of the room and the position of other objects depending on their distance to the center. In this way, *WarpVision* provides a direct non-diegetic hint that is not itself located in the virtual world, but is always oriented to the coordinates of the relevant object being searched for and is only visible in the user's HMD.



(a) *WarpVision* locally distorts the spatial environment around the point-of-interest (orange sphere). The perceptual impact of this distortion increases with the user's viewing angle relative to the object - zoomed in for illustration.

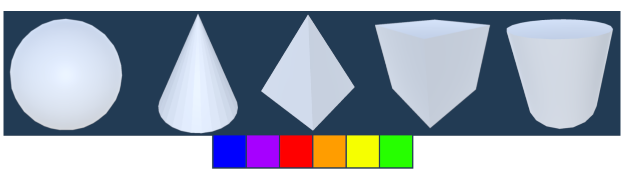


(b) As the user is not looking directly at the point-of-interest (coin) *WarpVision* shows a pronounced effect of the warped space around the point-of-interest - zoomed in for illustration.



(c) Overview of the artificial environment used to evaluate the effectiveness of *WarpVision*. The participant is positioned at the coordinate cross.

(d) Overview of the realistic environment with coins used to evaluate the effectiveness of *WarpVision*. The participant is sitting on the chair in the lower left.



(e) All geometric objects placed within the artificial environment. The point-of-interest (orange sphere) appears exactly once per scene. All objects are assigned a color from the shown color spectrum.

(f) All coins placed within the realistic environment. The point-of-interest (five-cent coin) appears exactly once per scene. The background replicates the table texture.

Fig. 2: *WarpVision* applied in an artificial (a) and a realistic (b) scene. Scene overviews are shown in (c) and (d). Objects used in the user study's search task are shown in (e) and (f).

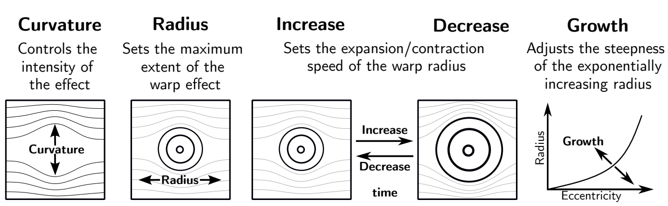


Fig. 3: Overview of the five adjustable *WarpVision* parameters, with descriptions and example images illustrating their effects.

Five customizable parameters allow different settings of *WarpVision*: *Curvature*, *radius*, *increase*, *decrease*, and *growth*. The illustration in Figure 3 provides an overview of the effects of each parameter. The *curvature* determines the intensity of the effect, while a higher *curvature* value reduces the intensity. The *radius* is the basic distance from the center that sets a maximum limit on the radius affected by the warp effect (and the *curvature* parameter). The *radius* of the effect itself can change and is set to its predefined value directly at the start of a search task to trigger immediate gaze guidance. The parameters *increase* and *decrease* describe the speed at which the radius of curvature changes. In order to integrate movement as a visual variable and allow immersion to remain unaffected, the radius of the distortion increases or decreases depending on the eccentricity of the POI in the visual field. Compared with a manipulation of the curvature radius directly proportional to eccentricity, the reliance on the increase and decrease parameters has two concrete advantages: first, the imperfect eye tracking data is implicitly smoothed, and jittering of the distortion effect is thereby prevented. Second, manipulations that directly rely on eccentricity can be subject to saccadic suppression which is avoided with the separation of the reduction as a direct eccentricity-based effect. Figure 4 illustrates this behavior in a simplified example on a two-dimensional grid.

As soon as the viewing direction approaches the center of the effect at the POI, the radius decreases until no distortion can be perceived when focusing directly on it. The radius increases exponentially with the viewing angle between the POI and the user's gaze point. This allows for a more subtle increase in the radius when the gaze is within the immediate vicinity of the POI being focused on. In contrast, the effect becomes stronger the larger the angle, so an intense curvature of the space in the corresponding direction can draw attention to POIs outside the field of view. Finally, the growth parameter provides the option to adjust the steepness of the exponentially increasing radius.

WarpVision is implemented using a custom shader. To initially set

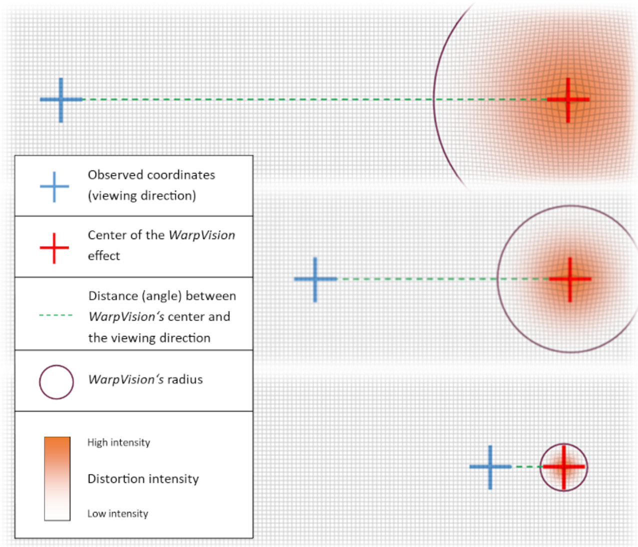


Fig. 4: The change in the radius of *WarpVision* based on the eccentricity (angle between the POI and the viewing direction). For a simplified interpretation, the curvature is shown on a two-dimensional grid. If the angle between the direction of view and the center is smaller (shown here as the distance between the crosses), the radius of the technique is reduced.

the intensity of *WarpVision*, all five parameters are part of the host code and can be adjusted by the user. The current size of *WarpVision's* radius is recalculated per frame using the values of the increase, decrease, and growth parameters. To this end, the angle between the center and the gaze a is used to calculate a normalized angle n as:

$$n = \text{clamp}\left(\frac{\texttt{max}(0, a-t)}{90-t}, 0, 1\right)$$

where t is the minimum threshold before increasing the resulting value. This allows for an undistorted view of the scene at the POI if the gaze is in its direct vicinity and allows for a margin of error in the eye tracking data. For the user study, t is set to a value of 5. This value was chosen empirically after a small initial pre-study and ensures that scene is not distorted when the POI is in central vision (fovea and parafovea). The fraction calculates the delayed increasing value based on t. This value is then clamped between 0 and 1 to ensure that the highest radius will be reached at a viewing angle of 90°.
The final size for the radius r is calculated as:

$$r = r_{\text{base}} \cdot 2^{(n \cdot g)} - 1$$

where r_{base} is the value of the initial radius parameter, r the new target radius, and g the growth parameter. The final subtraction by 1 ensures that r approaches 0 when the POI is being looked at directly (n close to 0). The current radius is calculated by approaching r in the step size given by the increase or decrease parameters for each frame, depending on whether r is greater or smaller than the current radius. The curvature remains unaffected by these changes and is only set initially.

The shader used to create the spatial warping of the rendered output image is based on a modified version of the whirl and pinch algorithm taken from the Gimp source code [14]. Based on the radius r and curvature c of the effect, the shader calculates a distortion factor f per pixel in a distortion render pass.

$$f = ((r-d)/r)^c$$

with d being the distance from the current uv position to the position of the POI p. Finally, the factor f is used to warp the uv position as follows

$$uv_{warp} = p + (uv - p) \cdot (1 - f);$$

Fig. 5: Study setup. The participant sits on a chair that cannot be turned on adjusted in height to minimize unwanted movement.

To ensure *WarpVision* does not distort empty space (black pixels outside the rendered area) towards the field of view, the effect is limited to POIs in the forward facing hemisphere of the camera. A sample Unity project, which includes the shader code and other relevant files, is openly available [1].

4 User Study

In order to evaluate *WarpVision*, we conducted a VR study with a within-subjects design using the VIVE Pro Eye VR-Headset. A total of 20 participants took part (demographics in Table 2). All of them had normal or corrected-to-normal vision. The study aims to address the following two research questions:

RQ1 Does *WarpVision* significantly reduce search times without disrupting immersion compared to an existing guidance technique and the unassisted search?

RQ2 Does *WarpVision* have an effect on cybersickness or discomfort?

To benchmark *WarpVision* against an existing technique, we implemented *Deadeye* [10], a preattentive gaze-guidance method that renders the target object monocularly. *Deadeye* was chosen for its conceptual similarity and suitability for direct comparison within our experimental design. Three hypotheses **H** are formulated to answer **RQ1** and **RQ2**:

H1 *WarpVision* significantly reduces search time in comparison to *Deadeye* and the unassisted search.

H2 *WarpVision* settings with significantly reduced search time compared to *Deadeye* cause significantly less disruption of immersion.

H3 *WarpVision* has a significantly lower impact on cybersickess and discomfort compared to *Deadeye*.

4.1 Design

The study was conducted in two virtual environments: an artificial scene and a realistic bistro environment. The artificial environment (Fig. 2c), inspired by Grogorick et al. [3], features 500 geometric objects distributed within a 180° field-of-view, at a radius of 7–8 virtual meters and a height of 5 meters. Objects vary in shape, color, and orientation, and are randomly placed in each trial. Participants are positioned at the center of the scene and are instructed to find the one target sphere (the POI), presented once per run in varying positions and colors (Fig. 2e).

The realistic environment (Fig. 2d) is a bistro scene [13] and depicts a table setting with randomly placed one- and two-cent coins. The search task requires locating a single five-cent coin among them, characterized by its distinctive texture ("5") and intermediate color and brightness (Fig. 2f). Coin positions remain fixed across trials, providing multiple visual cues for target identification.

In order to find parameters that allow *WarpVision* to remain unrecognized and influence immersion as little as possible, four different

[1] https://github.com/XRG-HSRM/WarpVision

Table 1: Settings used to test *WarpVision* in the user study. *WarpVision* is abbreviated to "WV", increase to "inc.", decrease to "dec".

Setting	Curvature	Radius	Inc.	Dec.	Growth
WV-Min	100	0.2	0.01	0.015	5
WV-Low	100	0.5	0.01	0.010	5
WV-Med	100	2.0	0.05	0.250	5
WV-High	100	3.0	0.05	0.250	5

Table 2: Demographic data. Overview of the participants of the user study to evaluate *WarpVision*

Category	Details
Age Groups	Between 11 and 55 ($\bar{x} = 29.5$, SD=10.3)
Gender Distribution	13 male, 7 female
VR Experience	9 low, 4 medium, 7 high
Eye Impairments	12 none, 7 shortsighted, 1 farsighted

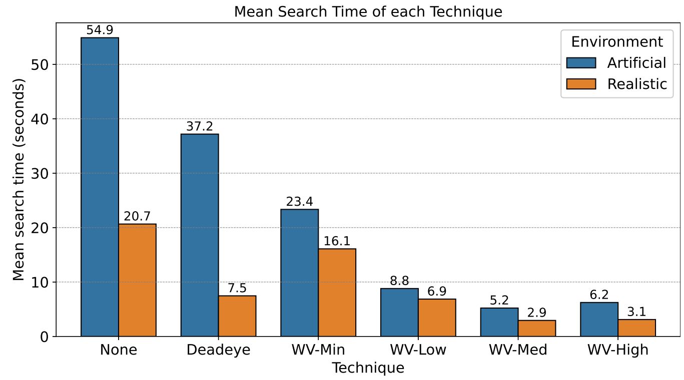


Fig. 6: Bar plot of the mean search times of the techniques in both environments. Due to the search duration being limited to a maximum of 3 minutes, the mean is slightly skewed.

settings are examined: "WV-Min", "WV-Low", "WV-Med", and "WV-High". The parameters for each setting are shown in Table 1. "WV-Min" is the variant with the smallest radius and therefore the least distortion and should remain as unrecognizable as possible to allow only for a subtle recognition. The other three settings offer more recognizable versions of the distortion effect, with "WV-Low" as a setting that is easier to perceive and "WV-Med" and "WV-High" as clearly recognizable settings. In addition to the radius, the four settings differ in the rate at which the radius increases and decreases. For example, the highest setting "WV-High" decreases faster than "WV-Low" when there is direct eye contact with the object being searched for.

Gaze-guidance techniques must balance reliability and perceptibility. Depending on the application, different positions along this trade-off may be desirable. We selected *WarpVision*'s parameters to sample this continuum uniformly, informed by an empirical pre-study, to evaluate its gaze-directing performance and potential impact on immersion across diverse application scenarios.

4.2 Procedure

Prior to the experiment, participants were informed about the procedure and provided their written consent. Based on previous experiments, the risks were assessed as minimal. For this reason, and in accordance with the legal framework of the institution, no ethical review board was involved in the review of the study. Participation was voluntary, data was anonymized, and potential risks (cybersickness) were discussed individually and in the declaration of consent.

The setup of the main study is shown in Figure 5. The participant sits on a chair that cannot be turned or adjusted in height in order to minimize movement away from the relevant search area in the virtual world, while movements of the body and head were permitted. Following the calibration of the eye tracking system within the VR headset, participants were instructed to remain seated throughout the study. The study takes between 20 and 35 minutes.

The study began with an empty virtual scene, during which demographic and personal information was collected. Subsequently, a demonstration scene for the search task within the first environment is shown. Participants had the opportunity to ask potential questions, and a test run was performed. The demonstration scene is simplified and does not use a supporting technique. Once the demonstration had been successfully completed, all runs of the first environment were conducted. Before transitioning to the second environment, an additional demonstration scene was presented to introduce the new target object.

In each environment, we tested the four *WarpVision* settings, *Deadeye*, and an unassisted search as a control condition. A full factorial combination of the techniques and environments was used, including one repetition each. To increase reliability and avoid sequence effects, all combinations are randomized, whereas environments are always presented separately.

At the beginning of a trial, the scene displays a central fixation cross in front of a black background. The search task starts after fixating the cross for 1.5 seconds. In order to successfully complete the search task, participants must press a predefined trigger button while fixating the object. An indicator confirms whether the selection was correct, and if successful, the participants are presented with the following subjective statements about the previous trial:

(1) I felt supported in the search task. [*0 - 5*]
(2) The supporting technique affected my immersion. [*0 - 5*]

The first question assesses the perceived helpfulness and noticeability of the technique, while the second evaluates its impact on immersion, which should ideally remain unaffected by subtle guidance methods. A forced-choice scale without a neutral option was used to encourage clear tendencies. To ensure accurate results, all support is deactivated immediately after the object has been found. An incorrect selection does not end the task. If the target object is not found within three minutes, the subjective statements are skipped, and the next trial starts with the fixation cross. At the end of the study, participants were shown all techniques and completed a Virtual Reality Sickness Questionnaire (VRSQ) to assess comfort and potential adverse effects [8].

5 Results

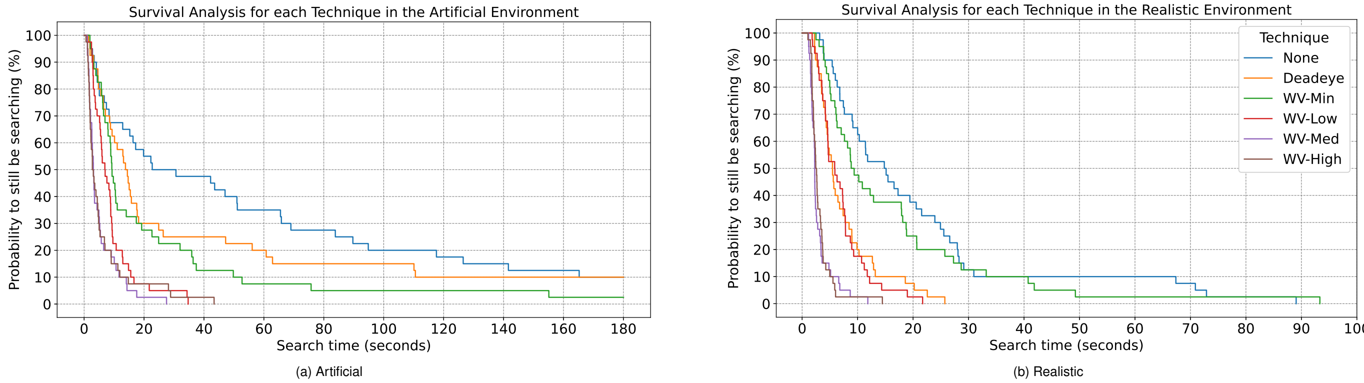
This section provides insights into the objective and subjective data of our user study. Each VR environment is examined individually.

5.1 Objective Data

The focus of this analysis is the search duration, i.e. the time before the searched object was found. Figure 6 shows the mean search time required to find the searched object for each technique and both environments. This time is highest when participants were not assisted in their search ("None"). The three *WarpVision* settings "WV-Low", "WV-Med", and "WV-High" show the highest reduction in mean search duration in both environments compared to unassisted search. *Deadeye* and "WV-Min" lie between these settings and the unassisted search and show differences between the two environments. While *Deadeye* has a higher mean search duration then "WV-Min" in the artificial environment, it is vice versa in the realistic environment.

However, analyzing only mean search times is insufficient, particularly due to the imposed three-minute time limit, which introduces right-censoring and distorts the mean. To address this, we employ survival analysis to examine the search duration. This approach allows for evaluation of the time until an event occurs, in this case, successfully finding the target object. The analysis uses the Kaplan-Meier estimator, which enables a non-parametric estimation of a survival function and accepts right-censored data. The survival function represents the probability with which an event has not yet occurred at a certain point in time, i.e. the object being searched for has not yet been found.

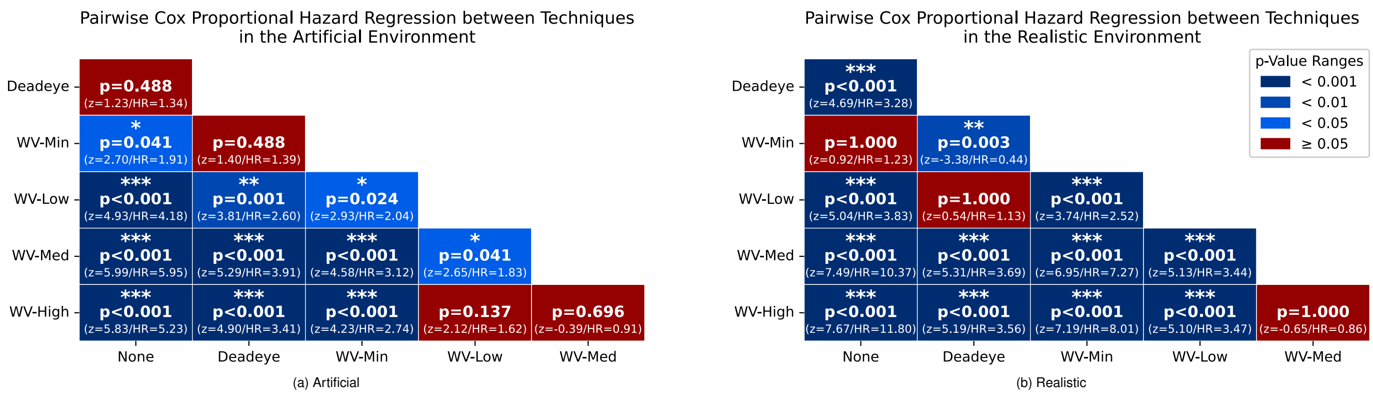
Figure 7a shows the respective functions of the techniques in the artificial environment. The effectiveness of the techniques is indicated by the respective strength of the curve slopes. The three strongest settings of *WarpVision* ("WV-Low", "WV-Med", and "WV-High")



(a) Artificial

(b) Realistic

Fig. 7: Survival functions of the techniques within the artificial environment (a) and realistic environment (b). Confidence intervals are not shown to simplify interpretation.



(a) Artificial

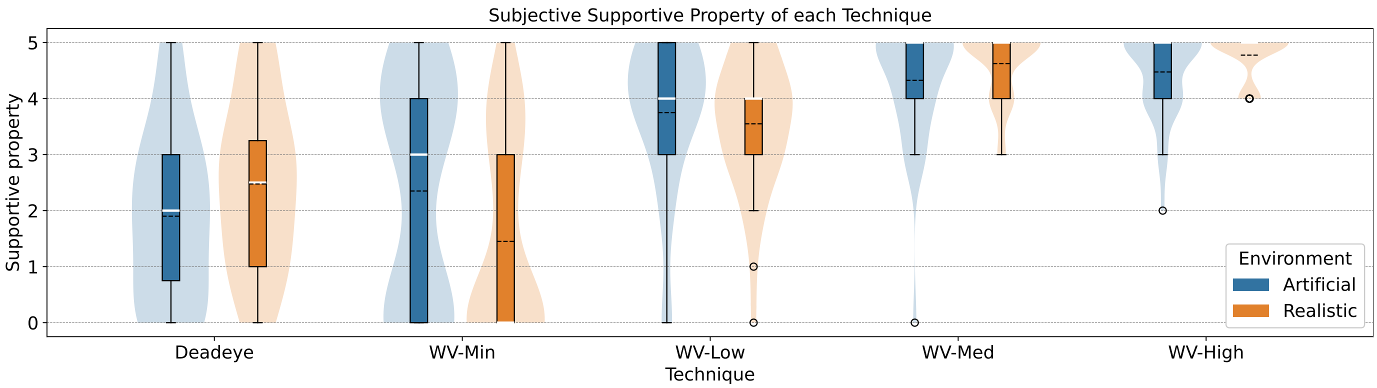
(b) Realistic

Fig. 8: Pairwise comparisons of techniques within the artificial (a) and realistic environment (b) in relation to search duration using a Cox proportional hazard regression. Each cell displays the hazard ratio (HR), z-value, and Holm-Bonferroni adjusted p-value for the comparison between the row and column techniques. An HR >1 indicates a higher hazard (i.e. shorter search time) for the row technique relative to the column technique.

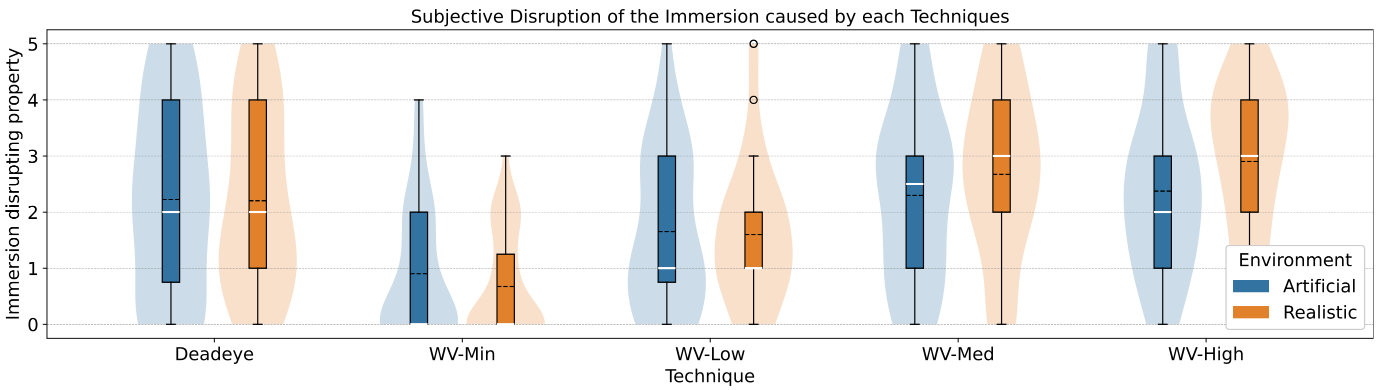
show the highest objective supporting properties. In particular, "WV-Med" and "WV-High" allow for quick success in finding the correct object in the artificial environment. After ten seconds of searching, the probability of not having found the object is approximately 15%. This probability is almost as low with "WV-Low" at 25%, although a search supported by this technique takes, on average, around four seconds longer. In none of the runs in which one of these three *WarpVision* settings were used was right-censored data available, meaning that the object searched for could always be found within the maximum time given (even well less than a minute in some cases). The curve with the next strongest downward trend is "WV-Min". When using this technique, with a probability of approx. 50% the target object is found after 10 seconds, and with a probability of over 90% after one minute. *Deadeye* shows a similar efficiency, but the detection time is delayed by approximately five seconds compared to "WV-Min". A premature flattening of the survival function after just under 20 seconds indicates longer search times, of which 20% last longer than one minute.

Figure 7b shows the survival functions of the techniques in the realistic environment. As in the artificial environment, "WV-Med" and "WV-High" are the two techniques under which the searched object was found the fastest. Both curves are almost identical and with a probability greater than 95% the search task was successfully completed after the first ten seconds. The next best techniques are "WV-Low" and *Deadeye*, with one in five cases still searching for the object after ten seconds. Finally, "WV-Min" closely resembles the unassisted search, with a slightly lower curve that suggests a modest effect below the performance of the other techniques. Significant differences between survival curves are assessed using the Cox proportional hazards model with Holm-Bonferroni correction for p-value adjustment. Figure 8a shows the heatmap of the obtained p-values and hazard ratios for all pairs of techniques in the artificial environment using the Cox proportional hazard regression. It is noticeable that only *Deadeye* is insignificantly different to trials without support ($p > 0.05$). Thus, this technique did not contribute to a significant reduction in search time. While "WV-Min" and *Deadeye* also do not significaly differ ($p > 0.05$, HR = 1.39), "WV-Min" significantly improves search compared to unsupported trials ($p = 0.041$, HR = 1.91). Since the Cox proportional hazard model compares groups (or curves) in pairs, there is no transitive property of significance implied by the p-value (discrepancies in combined hazard ratios can occur due to statistical variability). All other techniques show significant differences to the comparisons, *Deadeye* and unsupported search. The results of the Cox proportional hazard model within the realistic environment are shown in Figure 8b. In this environment, only "WV-Min" provides no significant support (HR = 1.23). All other techniques are highly significant compared to unassisted search ($p \leq 0.001$), with"WV-Med" and "WV-High" even exhibiting hazard ratios greater than 10, indicating that the target was found at least ten times faster using these techniques. "WV-Low" and *Deadeye* form a group of almost identical survival functions with $p > 0.05$ and a hazard rate of 1.13, providing similarly effective assistance. Conversely, compared to *Deadeye*, "WV-Min" has a hazard ratio of 0.44 ($p = 0.003$), indicating a 56% lower rate of finding the target object at any given time. Analogously to the artificial environment, "WV-High" and "WV-Med" are statistically identical. As the two environments represent fundamentally different scenarios and search tasks, the survival functions are not examined between the environments.

Altogether, these results are consistent with the original hypothesis H1, strongly supporting the assumption that *WarpVision* significantly improves search performance compared to *Deadeye* if the employed setting is properly calibrated.



(a) Subjective supportive characteristic



(b) Subjective immersion disrupting characteristic

Fig. 9: Subjective supportive (a) and immersion disrupting (b) characteristic of the techniques, divided into both environments. The median is highlighted by a white line, the mean by a dashed black line. Values on the y-axis are presented on a semantic differential scale, where 0 indicates strong disagreement 5 indicates strong agreement.

5.2 Subjective Data

Figure 9a shows a violin plot that provides an overview of the subjective supportive property of each technique for both environments. The perceived support for trials without an assisting technique is omitted from the violin plot, as all values–except for two outliers–are 0 and show no meaningful variation. For *Deadeye*, most ratings in the artificial environment cluster around a median of 2, reflected by a symmetrical interquartile range. A similar distribution is observed in the realistic environment, though with slightly higher values (median around 2.5), suggesting a marginally higher perceived support by the technique. The lowest setting of *WarpVision* ("WV-Min") does not show a consistent trend between environments. In the artificial environment, ratings are more positive (median of 3), though the lower quartile reveals many lower ratings. In contrast, in the realistic environment, over half of the participants did not perceive the guidance (median 0), though the upper quartile extends to 3, indicating a wide spread. "WV-Low" shows small differences between environments, with both medians at 4, but a wider distribution in the artificial environment. The two strongest settings, "WV-Med" and "WV-High," consistently yield the highest perceived support (median 5) in both environments, with only minor differences in data spread. Overall, the perceived support strongly correlates with the objectively measured effectiveness of the techniques across both environments.

For further evaluation, we use the Wilcoxon signed-rank test with Holm-Bonferroni correction for p-value adjustment to assess the supportive properties for significant differences. This test was chosen because the data is not normally distributed (see Figure 9a). All techniques show a significant impact on the subjective supportive property compared to runs without any technique. Among the *WarpVision* settings, only "WV-Min" exhibits a subjective supportive property comparable to or lower than *Deadeye* (artificial: $p = 0.582$, $z = 1.05$, realistic: $p = 0.015$, $z = 2.63$). All other techniques were perceived to be significantly more helpful.

Figure 9b shows the subjective immersion disruption ratings for each technique. Trials without assistance are omitted from the plot, as no immersion disruption was reported aside from the previous outliers. *Deadeye* caused a comparable level of immersion disruption in both environments (median = 2, "rather non-disrupting"), with only minor differences in the interquartile range and a slightly higher first quartile in the realistic environment. "WV-Min" has a lower negative influence on immersion in the realistic environment, with the median in both environments being 0, which indicates the absence of a recognizable influence on immersion. In the artificial environment, the values are more widely distributed, with no data points reaching the maximum value of 5 in either environment. As the next highest intensity of *WarpVision*, "WV-Low" has comparable deviations between the environments. Although the median in both cases has a value of 1, the third quartile in both environments has a similar distribution, showing that the responses are roughly one step above "WV-Min" on the semantic differential scale. Consequently, values of up to 5 ("very disrupting") can be observed in the artificial environment. The first quartile is located close to the median with only slight deviation, resulting in a limited number of values at 0. The immersion disrupting properties of "WV-Med" and "WV-High" are mostly comparable, with a median of 2.5 and 2 in the artificial environment, respectively, and 3 in the realistic environment. However, data points in both environments can lie far outside the interquartile range and assume values over the entire spectrum from 0 to 5.

For the evaluation of H2, the Wilcoxon signed-rank test with Holm-Bonferroni correction for p-value adjustment was employed to assess the immersion disrupting properties for significant differences. All techniques demonstrated a significant increase in immersion disruption when compared to unassisted trials. Compared with *Deadeye*, only "WV-Min" caused significantly less disruption to immersion across both environments (artificial: p = 0.008, z = 3.23, realistic: p <0.001, z = 3.95). No significant differences were observed between *Deadeye* and the remaining *WarpVision* settings.

These results challenge hypothesis H2, not strictly supporting the assumption that *WarpVision's* disruption of immersion is significantly lower compared to *Deadeye* while providing significantly higher support. Only particular settings (e.g., "WV-Low") show the tendency to support this hypothesis.

5.3 Cybersickness and Discomfort

In order to provide insights into **RQ2**, cybersickness and discomfort were assessed using the VRSQ [8]. Figure 10 shows a box plot of the ratings divided into the components "oculomotor", "disorientation", and the total VRSQ score for all techniques. As the focus of this subjective survey was on the individually presented techniques, the environments are not considered separately. We conducted a Wilcoxon signed-rank test with Holm-Bonferroni correction for p-value adjustment to evaluate significant differences. The results demonstrate that all *WarpVision* settings have significantly lower total VRSQ scores than *Deadeye* except for "WV-High" ($p = 0.112$, $z = -2.11$), indicating that comfort is better preserved with our technique. The median oculomotor score of *Deadeye* with 25 is even more than 10 points higher than all other techniques. In the "WV-Min" setting, the subjective rating of discomfort is statistically insignificant to the baseline ("None"), indicating that no measurable effect could be detected. Also the VRSQ scores for the other *WarpVision* settings only vary slightly from the baseline average. These findings confirm H3 and support the assumption that *WarpVision* causes significantly less cybersickness and discomfort than *Deadeye*.

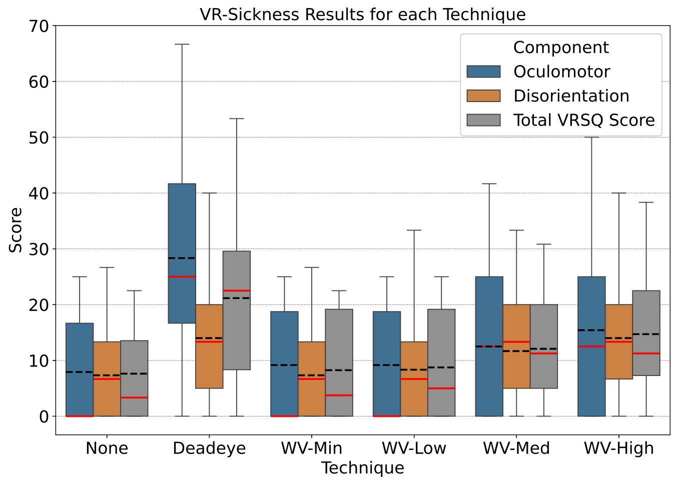


Fig. 10: Results of the VRSQ score by technique, regardless of the environment used. The median is highlighted by a red line, the mean by a dashed black line.

6 Discussion

The results of the search times analysis demonstrate that even the most subtle *WarpVision* settings (e.g., "WV-Min") achieve performances comparable to *Deadeye*. At the same time, this setting affects immersion significantly less. If a slight increase in immersion disruption is acceptable, "WV-Low" is recommended. In medium and high settings, *WarpVision* significantly outperforms *Deadeye* in search efficiency. As expected, all *WarpVision* settings reduced search times compared to unassisted trials (H1). Based on participant feedback, the lower scores in subjective support of *Deadeye* can likely be attributed to feelings of confusion due to the visual mismatch the technique creates in the stereoscopic image. Please note that a lower subjective perception of support should not necessarily be interpreted as a drawback in subtle guidance techniques, as effective gaze steering may occur without conscious recognition. Given *WarpVision's* low disruption of immersion (H2) and mostly unaffected VR sickness (H3), while providing significantly more support than *Deadeye* (H1), it emerges as a promising alternative, particularly when subtle guidance is desired.

The different search times of the techniques in the two environments can stem from many factors, including differences in color representations, the distance of the POI, and the manifold of surrounding information. These factors can introduce varying complexities to the search task that could affect the gaze-guiding property of the technique. Higher visual contrast and texture variety may increase the perceptibility of *WarpVision's* distortion. This property may also explain the difference in mean search durations for *Deadeye* and "WV-Min" between the two environments. The high-luminance colors of the objects and the parallel lines of the background pattern in the artificial environment may have enabled better recognition of the warping effect compared to the less saturated texture of the table in the realistic environment. Additionally, the shorter distance to the POI and its background in the second environment might have reduced the effectiveness of *WarpVision* further, since the reduced depth of the scene leads to less distortion in physical space. However, since the artificial environment represented the more complex search task, the results obtained support the interpretation that *WarpVision* does not significantly lose reliability with increasingly complex search and enables faster attention guidance in these cases compared to other techniques. It follows that the parameters of *WarpVision* should be adapted to the visual properties of a scene—or, conversely, scenes may be designed to better reveal screen-space warping effects.

Since the gaze-dependent adaptation of *WarpVision* continuously alters the image, it engages multiple visual channels that can lead to preattentive perception of direction and target location in the search task [19]. Specifically, visual attributes like continuous gaze-contingent changes in the orientation and curvature of parallel lines, as well as alterations in the size and shape of scene features, are known to trigger these perceptual mechanisms [25, pp. 155f]. To further investigate the role of preattentive mechanisms in *WarpVision*, we conducted a focused analysis on a subset of our study data. When comparing the time from the start of the search to the moment the target object enters the area of the fovea centralis (within 5° of visual angle), we observe significant differences between the use of no technique and the various *WarpVision* conditions: "WV-Med" (p <0.001), and "WV-High" (p <0.001). Looking at the median time for Deadeye (3349ms) compared to "WV-Med" (1343ms) and "WV-High" (1297ms), it becomes apparent how quickly *WarpVision* facilitated successful target acquisition. For context, preattentive processing typically occurs within 50–500 ms [23], and the average duration of an eye fixation on a target is around 300–350 ms with a distribution reported to have a long tail [15]. It is important to note that our study was not explicitly designed to isolate preattentive effects, but rather to measure the overall effectiveness of *WarpVision* in improving search performance. Still, the results strongly suggest that *WarpVision* supports early-stage perceptual guidance. However, a more detailed analysis of initial eye movements within the first 1000 ms would be needed to validate this hypothesis—especially given that trials started blind with the targets not in the FOV.

7 Limitations

Several limitations stem from the current implementation. Attributable to the restriction of guidance to POIs in the forward-facing hemisphere of the camera, the current implementation of *WarpVision* cannot guide toward POIs directly behind the user. This limitation could be avoided by integrating guidance through intermediate "waypoints" that sequentially redirect the gaze of the user towards the backside. Another limitation stems from the disparity of the stereoscopic visualization in the HMD. Due to the scene being presented from a slightly different angle to both eyes, the distortions could affect the screen-space unequally. In our study, this imperfection was only recognizable in scenarios containing POIs very close to the camera or in cases where POIs are cut off in the center of the distortion. An additional constraint can be found in scenes with large untextured areas or very dim lighting conditions. If the user, for example, faces an untextured white wall, the distortion will not be visually noticeable. This also holds true for very dim spaces, where such a distortion would remain unseen. While this is a theoretical limitation, such scenarios do not hold much appeal for gaze retargeting in practice, as targets are usually chosen on vivid, high-contrast areas. A limitation of the study is the extent of comparison techniques. While *Deadeye* and the unassisted search show an indication of *WarpVision's* effectiveness, further investigations are required to confirm these findings against other gaze guidance techniques. Since only four settings of *WarpVision* were examined in the study, it can be assumed that the effectiveness of this technique can be further optimized when adjusted to a specific scenario. Since participants in this study do not move (beyond simple body/head movements on a chair), the influence of *WarpVision*'s curvature on spatial understanding and navigation remains unclear.

8 Conclusion and Future Work

This paper introduced *WarpVision*, a novel gaze-guidance technique that subtly warps screen-space around a POI to direct user attention in VR. We conducted a study with a search task in two different virtual scenarios to compare *WarpVision* with related work and unassisted searching. Our results show *WarpVision* to be an effective alternative with higher user comfort and better preservation of immersion. Future investigations may explore space-distorting effects beyond a zoom-like curvature, such as vortex effects on or around the POI and for more diverse environments. In addition, scenarios that reflect not only static but also dynamic environments are interesting to investigate. These include scenarios where the POIs are clearly visible (not hidden) or moving, like in panoramic 360° videos. Since *WarpVision* works entirely in screen-space and does not require a diegetic understanding of the scene to develop matching guidance metaphors, our implementation is well-suited for these application scenarios. By applying the technique to the images of mixed reality video-see-through HMD's, such scenarios could further include real-life environments. Given these opportunities and the experimental results, we are confident that *WarpVision* is a significant step in the process of effective VR guidance methods that are versatile, robust, and subtle.

References

[1] R. Bailey, A. McNamara, N. Sudarsanam, and C. Grimm. Subtle gaze direction. *ACM Trans. Graph.*, 28(4), article no. 100, 14 pages, Sept. 2009. doi: 10.1145/1559755.1559757 2

[2] F. Bork, C. Schnelzer, U. Eck, and N. Navab. Towards Efficient Visual Guidance in Limited Field-of-View Head-Mounted Displays. *IEEE Transactions on Visualization and Computer Graphics*, 24(11):2983–2992, Nov. 2018. doi: 10.1109/TVCG.2018.2868584 2

[3] S. Grogorick, M. Stengel, E. Eisemann, and M. Magnor. Subtle gaze guidance for immersive environments. In *Proceedings of the ACM Symposium on Applied Perception*, SAP '17, pp. 1–7. Association for Computing Machinery, 2017. doi: 10.1145/3119881.3119890 4

[4] S. Grogorick, J.-P. Tauscher, N. Heesen, S. Castillo, and M. Magnor. Stereo inverse brightness modulation for guidance in dynamic panorama videos in virtual reality. *Computer Graphics Forum*, 39, 08 2020. doi: 10.1111/cgf.14091 2

[5] U. Gruenefeld, A. E. Ali, S. Boll, and W. Heuten. 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*, pp. 1–11. ACM, 2018. doi: 10.1145/3229434.3229438 2

[6] R. Horst, F. Klonowski, L. Rau, and R. Dörner. The shared view paradigm in asymmetric virtual reality setups. *i-com*, 19(2):87–101, 2020. 2

[7] S. Kergaßner, N. Doerr, M. Wieland, M. Fuchs, and M. Sedlmair. Hivefive360: Extending the vr gaze guidance technique hivefive to highlight out-of-fov targets. In *Proceedings of Mensch Und Computer 2024*, MuC '24, 10 pages, p. 11–20. Association for Computing Machinery, New York, NY, USA, 2024. doi: 10.1145/3670653.3670662 2

[8] H. K. Kim, J. Park, Y. Choi, and M. Choe. Virtual reality sickness questionnaire (VRSQ): Motion sickness measurement index in a virtual reality environment. 69:66–73, 2018. doi: 10.1016/j.apergo.2017.12.016 5, 8

[9] A. Krekhov, S. Cmentowski, A. Waschk, and J. Kruger. Deadeye Visualization Revisited: Investigation of Preattentiveness and Applicability in Virtual Environments . *IEEE Transactions on Visualization & Computer Graphics*, 26(01):547–557, Jan. 2020. doi: 10.1109/TVCG.2019.2934370 2

[10] A. Krekhov and J. Kruger. Deadeye: A Novel Preattentive Visualization Technique Based on Dichoptic Presentation . *IEEE Transactions on Visualization & Computer Graphics*, 25(01):936–945, Jan. 2019. doi: 10.1109/TVCG.2018.2864498 2, 4

[11] D. Lange, T. C. Stratmann, U. Gruenefeld, and S. Boll. HiveFive: Immersion preserving attention guidance in virtual reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, pp. 1–13. Association for Computing Machinery, 2020. doi: 10.1145/3313831.3376803 2

[12] Y.-C. Lin, Y.-J. Chang, H.-N. Hu, H.-T. Cheng, C.-W. Huang, and M. Sun. 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, pp. 2535–2545. Association for Computing Machinery, 2017. doi: 10.1145/3025453.3025757 2

[13] A. Lumberyard. Amazon lumberyard bistro, open research content archive (orca), July 2017. `http://developer.nvidia.com/orca/amazon-lumberyard-bistro`. 4

[14] M. Natterer and M. Muré. Gimp-python, 2008. `https://github.com/MichaelMure/gimp-plugins/blob/master/pygimp/plug-ins/whirlpinch.py`. 4

[15] S. Negi and R. Mitra. Fixation duration and the learning process: an eye tracking study with subtitled videos. *Journal of eye movement research*, 2020. doi: doi:10.16910/jemr.13.6.1 8

[16] L. T. Nielsen, M. B. Møller, S. D. Hartmeyer, T. C. M. Ljung, N. C. Nilsson, R. Nordahl, and S. Serafin. 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, 4 pages, p. 229–232. Association for Computing Machinery, New York, NY, USA, 2016. doi: 10.1145/2993369.2993405 2

[17] N. Norouzi, G. Bruder, A. Erickson, K. Kim, J. Bailenson, P. Wisniewski, C. Hughes, and G. Welch. Virtual animals as diegetic attention guidance mechanisms in 360-degree experiences. *IEEE Transactions on Visualization and Computer Graphics*, 27(11):4321–4331, 2021. doi: 10.1109/TVCG.2021.3106490 2

[18] M. Peter, R. Horst, and R. Dörner. Vr-guide: A specific user role for asymmetric virtual reality setups in distributed virtual reality applications. In *Proceedings of the 10th Workshop Virtual and Augmented Reality of the GI Group VR/AR*, pp. 83–94, 2018. 2

[19] M. Pomplun. *Analysis and Models of Eye Movements in Comparative Visual Search*. Dissertation, University Bielefeld, Faculty of Technology, May 1998. 8

[20] K. Quinn and J. L. Gabbard. Augmented Reality Visualization Techniques for Attention Guidance to Out-of-View Objects: A Systematic Review. In *2024 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 826–835, Oct. 2024. ISSN: 2473-0726. doi: 10.1109/ISMAR62088.2024.00098 2

[21] P. Renner and T. Pfeiffer. Attention guiding techniques using peripheral vision and eye tracking for feedback in augmented-reality-based assistance systems. In *2017 IEEE Symposium on 3D User Interfaces (3DUI)*, pp. 186–194, Mar. 2017. doi: 10.1109/3DUI.2017.7893338 2

[22] S. Rothe, D. Buschek, and H. Hußmann. Guidance in cinematic virtual reality-taxonomy, research status and challenges. 3(1):19, 2019. Number: 1 Publisher: Multidisciplinary Digital Publishing Institute. doi: 10.3390/mti3010019 1, 2

[23] C. Schmitt, J. C. Schwenk, A. Schütz, J. Churan, A. Kaminiarz, and F. Bremmer. Preattentive processing of visually guided self-motion in humans and monkeys. *Progress in Neurobiology*, 205:102117, 2021. doi: 10.1016/j.pneurobio.2021.102117 8

[24] A. Schmitz, A. MacQuarrie, S. Julier, N. Binetti, and A. Steed. Directing versus attracting attention: Exploring the effectiveness of central and peripheral cues in panoramic videos. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 63–72, 2020. doi: 10.1109/VR46266.2020.00024 2

[25] C. Ware. *Information visualization: perception for design*. Interactive technologies. Elsevier, Cambridge, 4th ed ed., 2021. 8

[26] J. W. Woodworth and C. W. Borst. Visual cues in VR for guiding attention vs. restoring attention after a short distraction. *Computers & Graphics*, 118:194–209, Feb. 2024. doi: 10.1016/j.cag.2023.12.008 2

[27] A. Yoshimura, A. Khokhar, and C. W. Borst. Visual cues to restore student attention based on eye gaze drift, and application to an offshore training system. In *Symposium on Spatial User Interaction*, SUI '19, article no. 30, 2 pages. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3357251.3360007 2