Subtle Cueing For Improving Depth Perception in Virtual Reality

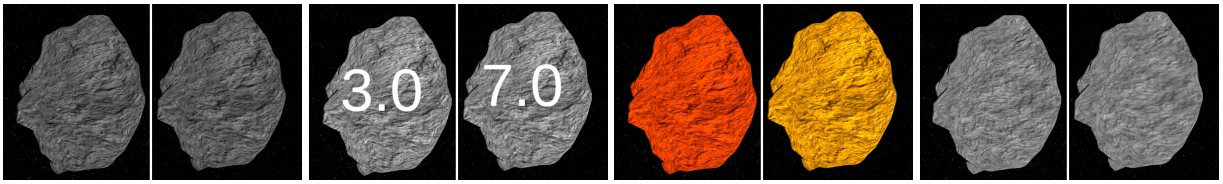
Ammaar Zaman*
Know Center Research GmbH

Ernst Kruijff†
Bonn-Rhein-Sieg University of Applied Sciences

Eduardo Veas‡
Graz University of Technology

Aleksandra Krajnc§
Know Center Research GmbH

Neven ElSayed¶
Know Center Research GmbH
FCAI, Benha University, Egypt



(a) No Cue, as a Baseline. (b) Text, as an explicit cue. (c) Color, as an explicit cue. (d) Blur, as Subtle cue.

Figure 1: VR Depth perception cues at various distances are depicted in the figures, with the left-side image displaying cues at 3 meters, and the right-side image illustrating cues at 7 meters.

Abstract

Understanding an environment relies on human sensory systems, with visual perception as the primary source for defining spatial relationships and estimating distances. The visual system uses natural cues, each offering partial information that can lead to bias and conflicts, especially when ambiguous. Virtual Reality (VR) environments challenge these natural depth cues with discrepancies in perspective and variations in light and shadow depiction, leading to potential confusion in depth perception. However, VR also allows for the isolation and study of these cues. This paper introduces artificial subtle cues (texture blur) to enhance natural depth information in VR. Our results show that augmenting natural depth cues with artificial ones improves depth prediction accuracy and spatial relationship awareness. Subtle blur cues enhance depth estimation without participants' subjective awareness of the augmentation, suggesting that such subtle cueing can effectively enhance depth perception.

Keywords: Virtual Reality, Perception, Depth Perception, Augmented cues, Subtle Cues, Implicit Cues, Explicit Cues.

1 Introduction

Visual perception is essential for defining spatial relationships, enabling precise interaction, and acquiring spatial knowledge [89]. Multiple sources of visual information, known as ***visual depth cues***, including binocular cues (stereopsis), motion cues, and pictorial cues, help estimate depth [52]. Depth perception is fundamental to understand spatial relationships. Visual depth cues aid in estimating distances between objects, enhancing situation awareness and forming a basis for decision-making [64]. These cues combine through a complex process to assess distance to an object, with each being affected differently by the distance between the observer and the object [15]. Our work investigates encoding complementary depth information in VR using minimally invasive cues. Usually, visual cues from our physical reality are used in VR yet are often somewhat distorted, leading to estimation errors. Depth cues interact differently based on egocentric distance: personal space ($< 2m$), action space ($2m - 30m$), and vista space ($> 30m$) [15]. Distance judgements in Virtual Environments (VE) in action space are often underestimated compared with judgments in real world [65]. This discrepancy arises when the interaction of natural depth cues with virtual objects provides conflicting information to the visual system [22]. Furthermore, external factors like atmospheric conditions, physical properties of objects, their spatial locations, and individual visual perception differences influence depth perception [29, 31, 80, 89]. These factors affect monocular cues, such as foreshortening, the reduction in contrast of distant objects, light and shadow, relative size, and proximity-luminance covariance. Decision-making in virtual environments requires understanding spatial organization, often forcing users to adapt their perception from real-world experiences. Dynamic and/or cluttered environments can impair perception in VR [74]. Therefore, additional cues should avoid altering the appearance of objects or the environment's organization.

VR allows manipulation and control of visual properties, enabling supplementary information encoding. We investigate techniques to present depth information, differentiating between natural and augmented depth cues. Natural cues, like binocular vision, motion parallax, and occlusion, are fundamental to human perception. Augmented cues, such as text annotations, color overlays, and blur effects, aim to ***supplement rather than replace natural cues to improve depth perception***. The goal of augmented depth cues is to resolve depth ambiguities without causing cue conflicts. Whenever possible, it would be desirable to avoid altering the appearance of objects. However, this is not always easily possible, especially when using explicit cues. We denote ***explicit cues*** as overt cues that are clearly or apparent to encode the depth information. ***Subtle cues*** encode the information covertly but still in processable form. Note that a subtle cue may be a modified (e.g., exaggerated) natural cue, so long as it is covert. Our study evaluated the effectiveness of subtle blur cues in VR for enhancing depth perception, focusing on smooth integration into VR user experiences. The primary focus was on the ability to catch moving objects, indirectly measuring depth perception, simulating realistic VR interactions, and assessing how subtle cues help users judge the proximity of moving objects. Our study focused on central vision where depth perception is most accurate. We deploy two explicit augmented cues for baseline depth estimation:

*e-mail: azaman@know-center.at
†e-mail: ernst.kruijff@h-brs.de
‡e-mail: eveas@tugraz.at
§e-mail: akrajnc@know-center.at
¶e-mail: nelsayed@know-center.at

2024 IEEE International Symposium on Mixed and Augmented Reality (ISMAR) | 979-8-3315-1647-5/24/$31.00 ©2024 Authors | DOI: 10.1109/ISMAR62088.2024.00088

979-8-3315-1647-5/24/$31.00 ©2024 IEEE
DOI 10.1109/ISMAR62088.2024.00088

text labels with distance information and color gradients that encode depth as a visual feature. Text labels provide explicit distance data but can be challenging to read in dynamic scenes. Color gradients vary in value according to distance, enhancing pre-attentive depth perception. Additionally, we use a subtle cue based on atmospheric factors such as foreshortening, reduction in contrast, and proximity-luminance covariance. This subtle cue increases texture blurriness with distance, resembling natural depth cues. This approach aims to improve user experience and immersion, making interactions more intuitive for users who find overt cues distracting or disorienting. Detailed descriptions of these cues are provided in Section 4.

2 Research Focus and Contributions

Our research focuses on improving depth perception in VR through the use of augmented cues. We employ explicit cues (*text labels and color gradients*) and a subtle cue (*texture blur*) to enhance depth estimation within action distance in both cluttered and non-cluttered environments. We explore how subtle cues, particularly blur, can seamlessly integrate into VR experiences, aiming for natural integration into the user's perceptual framework. This approach covertly enhances depth perception, a relatively understudied aspect in depth cue research. The study addresses two key questions:

RQ1: To what extend do augmented cues improve the precision of distance judgments for moving objects under different perceptual difficulty conditions (e.g., varying levels of environmental clutter and dynamic VR environments?

RQ2: How do subtle cues compare to explicit cues in terms of performance?

This comparative study is conducted in a dynamic VR scene with an object moving towards the observer. Clutter complicates depth perception but provides occlusion cues aiding spatial judgment. The study controls natural depth cues, except for relative size from object motion, using an asteroid in outer space to limit familiar size effects. We compare a baseline without augmented cues, explicit cues (text and color), and a subtle cue (blur). Our findings highlight the effectiveness of augmented cues, with all outperforming natural cues alone. Explicit cues (text and color) provide better quantitative accuracy, while subtle cues (blur) enhance depth perception without being overtly noticeable. By examining the balance between subtlety and perceptual impact, our research explores a less-studied aspect of depth perception and cue integration within dynamic VEs [22, 30]. This provides novel insights into the field of VR, showing that subtle cues can maintain visual integrity while enhancing depth perception. When traditional depth cues (e.g., binocular disparity) become less precise, such as in peripheral vision, blur serves as a robust alternative for depth estimation [32].

3 Background and Related Work

Visual perception interprets spatial information from stimuli [89], shaped by an individual's perceptual processing ability [51]. It involves attention allocation and cognitive resources, influenced by a perceptual set, including personality, experience, and expectations, guiding sensory interpretation through perceptual hypotheses [77]. Studies highlight visual saliency and depth cues in directing attention [71, 75]. Following, we will explain the basics of visual perception and how these cues fit into the process.

Depth Cues – Depth estimation can be ordinal (relative depth) or metric (measured distance) [52]. Depth perception is influenced by light scattering, object interaction, and human sensory response. We will illustrate the role of depth cues in our environment in the user study section (Section 5.1). Mechanisms for depth perception include binocular disparity (slight differences in retinal images), convergence (angle of central view rays), and accommodation (lens shape adjustments). Motion cues like kinetic depth field and motion parallax provide depth information through the relative movement of objects and the observer. Pictorial cues, perceivable with one eye, include linear perspective, occlusion, height in the plane, familiar size, relative size, textural gradients, light and shadow, proximity-luminance covariance, and aerial perspective [91].

Interrelationships Between Depth Cues – In VR, depth estimation is often underestimated in action space [42]. Dominant depth cues in this space include height in the visual field, motion perspective, and binocular disparity, while occlusion and relative size remain constant, and aerial perspective becomes more significant at greater distances [14]. Natural and augmented depth cues may overlap, conflict, or dominate each other, sometimes leading to the disregard of weaker cues [14]. Previous research has highlighted occlusion for ordinal information, relative size for ratio-scaled information, and stereopsis and motion perspective for absolute distance information [48]. Shadows and motion parallax enhance structural information perception in immersive analytics [34], while texture contrast and area contrast impact depth perception, with depth judgments influenced by texture contrast especially when area contrast is low [36]. In near-field augmented reality, factors like focal distance, brightness, and age affect depth estimation, with collimated targets being overestimated and midpoint targets accurately estimated [76]. Peripheral stimulation experiments show that constant white light in the far periphery improves distance and size estimation accuracy, suggesting VR depth perception may closely resemble the real world [40]. Defocus blurs and focus cues have been proposed to mitigate perceptual conflicts [46]. We examine augmented cues using a motion depth perception task. Due to the dynamic nature of this task, the brain continually integrates various cues in real-time [26]. Natural cues work together but can present conflicting information, leading to depth perception ambiguity, especially with motion parallax in dynamic scenes [15].

Depth Perception Performance in VR – VR environments often lead to compressed distance perceptions compared to real environments, yet maintain coherence in distance perception [54, 63, 92]. Factors influencing VR distance estimation include field of vision and stereoscopic rendering [25]. Egocentric distance misestimation is a key concern, though studies have found that field of view, monocular vs. binocular viewing, and other potential influences do not significantly impact distance perception in VR [13, 24, 45, 63], suggesting that the underestimation of distance in VR may stem from other, less obvious aspects of the VR experience. Various methodologies for estimating distances—verbal estimation, size estimation, relative measures, and directed action methods—each affect perception differently [16, 28, 44, 80]. Research continues to explore the compression or expansion of perceived intervals in VR compared to the physical world [5, 38, 62, 85, 93]. To address depth perception issues in head-mounted displays (HMDs), solutions like focus-tunable lenses and gaze-tracking technology are proposed [68].

4 Augmented Depth Cues

Augmented depth cues are artificial sources of depth information introduced deliberately to enhance the perception of depth. To achieve this, depth information is encoded visually, leveraging VR to arbitrarily modify the scene. The visual processing of the augmented cue is linked to the strategy for visual encoding. The cues are designed to work on an object-by-object basis, affecting a single object in its entirety. We designed and examined three augmented cues: two explicit and one subtle cue.

4.1 Explicit Augmented Cues

Two overt methods were selected to present depth information in an explicit manner. Numeric labels were chosen because they directly present the information clearly. Color encoding was chosen as an apparent overt cue that can encode magnitude. These cues are considered explicit because the modification of the scene is clearly visible and noticeable, despite the fact that they may lead to erroneous decoding of the underlying information.

Text annotations provide depth information through labels on objects, representing the distance between the observer and the object numerically. The effectiveness of text annotations is influenced by intrinsic factors such as prior linguistic knowledge [83], cognitive state and representation [57, 81], situational [97] and mental models [37,98]. While commonly used in VR, they require a higher level of detail, attention, and comprehension, making their continuous perceptual processing complex influenced by strategies employed by users rooted in their prior knowledge, comprehension, and motivation [11, 84]. In our implementation, text annotations display depth information to two decimal places in meters. This precision ensures consistency and allows accurate comparisons across conditions. For an object moving at 1.5m/s the rate of change of the presented text depth information is 0.15 meters, and for an object moving at 2.5m/s corresponds to 0.25 meters per second. The text is centered on the object and scaled to occupy a 15% of it when the original scale size of the object is 30 units. Figure 1-b depicts how this cue appears at 3 meters and 7 meters distance.

The ***color cue***, influenced by luminance contrast, enhances depth perception when combined with other cues, where proximity-luminance covariance applies to color as a pictorial depth cue [31]. Research shows that color augmentation improves depth perception in VR, especially against darker backgrounds [2, 50]. Partial occlusion combined with the color red produces stronger depth perception and faster judgments, making variations in the red spectrum effective depth cues [31]. Our color cue utilizes the RGBA model to control the red color channel within a 10-meter range from the observer. At the objects' farthest point, the overlay is yellow, shifting to red as distance decreases through linear interpolation. For an object moving at 1.5 m/s rate of color change from 1 to 0 in green channel is at 15% and moving at 2.5 m/s corresponds to 25% per second. This means the red spectrum gradient becomes 15% to 25% redder as the distance decreases, which is more influential in what humans typically perceive as a significant change within this range. This gradient is effective as red is perceived as nearer and yellow attracts more attention [2, 79, 95]. Warmer colors evoke a stronger depth response in depth perception [53]. Our choice of the color scale was informed by these findings, ensuring that our approach aligns with established practices in depth encoding. Figure 1-c depicts how this cue appears at 3 meters and 7 meters distance.

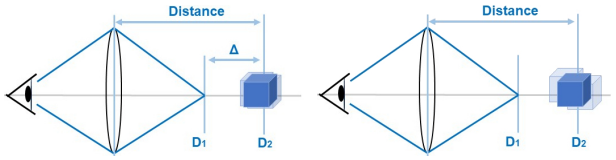
4.2 Subtle Augmented Cue

Our study employs texture blurring on target objects to enhance depth perception in VR. This approach aligns with the natural occurrence where distant objects appear blurrier than those close to the observer, a principle supported by prior research in visual perception [4, 33, 35]. Blur perception is influenced by the eye's inherent optical quality, including aberrations, and extends beyond the retina to involve higher level neurophysiological processes, differing between central and peripheral vision [10,56,60,86–88]. These studies lead to an argument that blur could be intercepted implicitly by our vision without overwhelming the visual system.

Empirical research shows that blur can function as an independent and economical depth cue, providing a pictorial cue for depth [3,9,67, 72]. VR users prefer clearer textures for closer objects and stronger contrast of blur rather than a blur consistent with the background [94]. The manipulation of focus and texture blurring offers depth perception independently of other depth cues; Blur and binocular disparity yield similar depth perception results [58, 59], and blur detection may operate pre-attentively without increasing cognitive load [55].

Our study focused on blur cues in central vision, where depth perception is sharpest. We follow previous research [10], which detailed the differences in processing blur between central and peripheral vision, demonstrating that objects in specific dioptric zones are perceived with varying clarity, affecting depth perception. Our findings confirm that, despite its lesser sensitivity, blur remains integral for depth perception.

The ***blur cue*** uses a texture blur filter that disorients texture pixels in eight directions, thereby creating a multi directional blur effect that simulating atmospheric or distance-induced blurring. Optimal levels of texture blurring were determined through preliminary testing, examining various degrees against user-reported depth perception accuracy and interaction ability. These tests, informed by established models of visual processing [35], enabled us to identify a maximum effective blur of 30% at 10 meters as the optimal threshold for enhancing depth perception without compromising essential object details. Figure 2 illustrates the difference between perception blur adjustment due to human perception and augmented cue blur adjustments. In perception blur, the adjustment is affected by the distance between the object and the focal point disparity planes, combined with the blur adjustment due to the distance between the eye and the object. However, in augmented cue, the blur adjustment is only affected by the distance between the eye and the object. This adjustment exaggerates the effect of a mixture of natural cues that effectively affect the clarity associated with objects at different distances (textural gradients, proximity-luminance covariance, and aerial perspective). For a moving object, the subtle blur cue uses a dynamic blurring scale, decreasing blur as the object approaches. For an object moving at 1.5 m/s, blur decreases by 4.5% per second, and at 2.5 m/s, by 7.5% per second. This range and variation factor per second were refined through iterative feedback during pilot testing. This ensures that, although the blur intensifies with distance, it does not obstruct the participant's ability to interact with the object, thus maintaining a balance between enhancing depth cues and preserving the clarity necessary for task performance. Figure 1-d shows this cue at 3 and 7 meters distance.



(a) Perception Blur (b) Augmented Blur

Figure 2: Blur adjustment in human perception versus augmented blur cues, Where D_1 and D_2 are two depth levels (disparity planes) defined by an object and a focal point at a different depth, while Δ is the distance between eye and the object. Using Augmented Blur, the level of blur is not affected by the difference between the focal disparity plane D1 and the object disparity plane.

5 User Study

The goal of the study is to evaluate the effect of augmented depth cues (explicit and subtle) in estimating depth in VR. It was designed in the form of a VR game, where participants catch asteroids using a catching net. The study was conducted under four conditions: *no cue* (baseline) (see Figure 1-a), *text cue* (see Figure 1-b), *color cue* (see Figure 1-c), and *blur cue* (see Figure 1-d). Under the ***no cue*** condition, objects retain their natural state without alteration to appearance or texture. The ***text*** condition shows the numerical distance on the object surface to indicate depth. For the ***color*** condition, a dynamic color gradient from yellow to red is applied to the object's surface, signifying depth with yellow representing the furthest and red the nearest point. Lastly, the ***blur*** condition applies varying levels of blur to the object's texture; minimal blur is used for objects that are close, while a more pronounced blur signifies objects that are further away. The study comprises three distinct sessions: a ***training session***–to familiarize the user with the game mechanics, a ***non-clutter depth estimation session***–to assess the user's ability

to judge moving objects in empty VR environment, and a ***clutter depth estimation session***–to evaluate depth perception in clutter VR environment. In each of these sessions, participants were introduced to all four cue conditions. To ensure consistency in the results, users are instructed to remain seated in a static position and to limit their head movements. This controlled setting is essential for isolating the variables being tested and ensuring the accuracy of the study's findings.

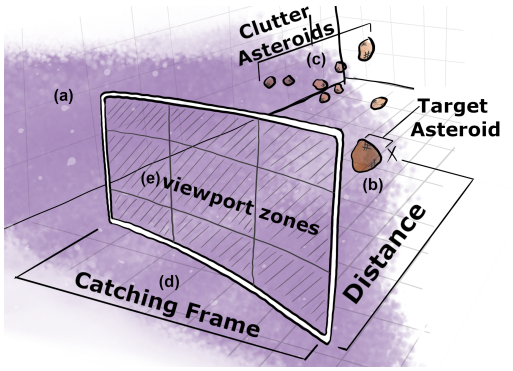


Figure 3: Visual depiction of (a) Scenery, providing the backdrop for VE; (b) Target Asteroid, the focal point of participants' attention; (c) Clutter Asteroids, populating the scene to create occlusion; (d) Catching Frame, a planar reference for the depth estimation; and (e) View-port zones, ensuring coverage of all possible object entry points within the observer's field of vision.

5.1 Virtual Reality Testing Environment

Our VR testing environment is designed to control factors like lighting, atmospheric effects, spatial misalignment, and dynamic changes to prevent biases in depth perception, ensuring more accurate spatial judgment.

We chose a space-themed VR game environment to minimize extraneous visual variables, allowing us to isolate and evaluate the impact of both subtle and explicit depth cues on perception. This setting enables systematic observation of the effect of augmented cues on depth perception, testing their efficacy without interference from typical real-world depth cues. This is a necessary preliminary step before studying their effect in real-world settings. We used limited virtual objects to enhance the generalizability of our findings, ensuring reliable results by reducing potential confounding factors, crucial for initial validation of the subtle cues. We eliminated potential confounding variables by disabling directional lights, texture compression, and shadows. Figure 3 illustrates the key components of the developed virtual testing environment. The game's appearance ensures a high level of contrast between the scenery and the target objects. The game scenery (Figure 3(a)) features a dark background, with skybox colors selected from the blue channel to avoid conflicts with the implemented color cues. The asteroids (Figure 3(b)) serve as the primary objects of interest. Their inherent ambiguity in shape and size eliminates any biases from participants' prior knowledge, focusing the study solely on depth perception.

We intentionally selected an irregular asteroid **shape** for control purposes, maintaining consistency to control for shape variability. The asteroid's texture increases perceived visual complexity and overall perception of the 3D shape [49]. While the size remained constant, we dynamically varied the asteroid's movement speed, as size discrimination is more precise when the object appears closer to the observers [1]. The asteroids move towards the participant from a fixed point 10 meters away, incorporating the kinetic depth effect—a critical cue requiring an object in motion while the observer remains static [19]. The movement speeds range from 1.5 m/s to 2.5 m/s, important for assessing depth perception in dynamic environments. Note that the speed of target asteroids was a control parameter to isolate the effects of depth cues and catching frame distance, not a dependent variable. The **space** theme has no horizon, with only asteroids visible and no sharp parallel lines. This design controls for cues such as linear perspective, familiar size, relative size, light and shadow, and proximity-luminance covariance. Head movements were not controlled, allowing for motion parallax, but the game's speed made significant head movement impractical. Our analysis did not observe any apparent head movements that would affect perception due to motion parallax.

To examine the performance of the augmented depth cues, we implemented both non-clutter and clutter environment modes. The **clutter** factor, introduced via a group of asteroids using particle systems (Figure 3(c)), measures cue performance in complex and occluded environments. This setup allows us to test depth perception under varying visual complexities. Additionally, we integrated a catching frame into the environment, providing ordinal depth information supported by the same type of depth cue applied to the target asteroid. The frame is invisibly subdivided into nine zones, forming *view-port zones* (Figure 3(e)), to ensure coverage of all possible object entry points within the observer's field of vision. Distances for the catching frame (Figure 3(d)) were informed by research on VE distance perception [45], suggesting the robustness of perceived distances at 2m, 6m, and 12m, which guided our settings for middle (5m) and peripheral (3m and 7m) distances. Iterative feedback from pilot testing validated that augmented cues' visibility and variability beyond 10m did not effectively alter perception, as changes in cue strength at these distances were indiscernible. This controlled VR testing environment offers a relevant setting for examining the effectiveness of augmented depth cues.

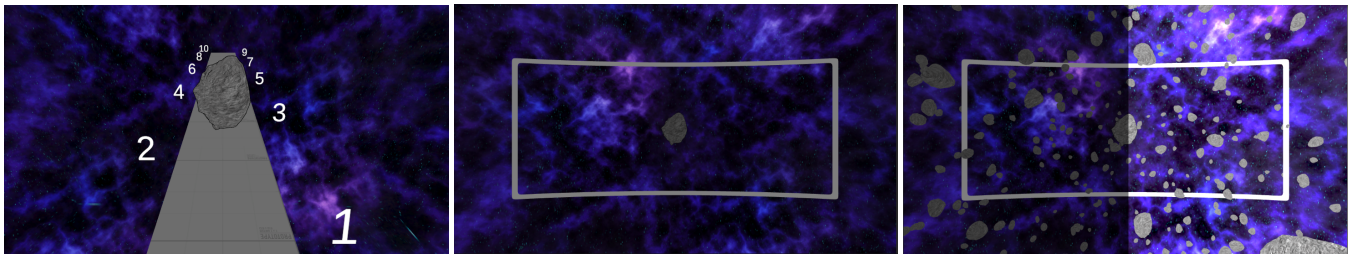
Our study was deployed on a Meta Quest Pro VR headset device, the experiment was developed using Unity3D 2021.3.15f1 and conducted via Oculus Link. The render resolution was set at 4128 x 2096 at 72 Hz, utilizing a cable connection to ensure real-time visualization.

5.2 Study Procedure

Training Session: In the training session, users were introduced to all four conditions and familiarized with depth cues across a range of 0 to 10 meters (see Figure 4-a). The environment featured a floor grid ruler marking distances at 1-meter intervals. The primary study object, an asteroid, remained at eye level throughout the training. Depth cues were augmented on the asteroid, with intensity varying over the 0 to 10 meter range. Users adjusted the asteroid's position within this range using the VR hand controller's toggle button, moving it linearly by pushing the thumb toggle forward or backward. The experimenter instructed users to move the asteroid to specific, randomly selected distance markers, allowing them to observe changes in augmentation at different distances. Each training session lasted about 1 minute.

Dynamic Depth Estimation Sessions: The Dynamic Depth Estimation Session in the VR game introduced users to the primary objective, i.e., tagging a flying asteroid at the precise depth indicated by the frame position. Participants engaged in two distinct sessions: ***depth estimation in non-clutter environment*** with a single asteroid approaching the user (see Figure 4-b), and ***depth estimation in cluttered environment*** where numerous asteroids move dynamically whereas only one primary asteroid is the target moving towards the user's view (see Figure 4-c).

Participants remained stationary, tagging the asteroid with the controller trigger when they perceive it to be within the frame, which was randomly positioned at distances of 3, 5, or 7 m; further referred to as ***near***, ***medium***, and ***far***, respectively. The asteroid's starting point was fixed for all trials, maintaining a linear path towards the participants' view, following nine different linear trajectories to offer



(a) Training session (b) Non-clutter session (c) Clutter session

Figure 4: The user study consisted of sessions for conditions: baseline, text, color, and blur. In the clutter session figure, the lighting exposure on the right side of the image was adjusted to improve visibility for the purpose of paper presentation. However, the clutter session was clear in VR.

varying view-port angles. The trajectories themselves were also randomized. The eye-tracker of the device was used to measure participants' gaze-on-target time for precise focal attention assessment. Prior to the study, participants self-reported their vision status and used corrective lenses or glasses if needed. After the study, participants completed a post-task assessment form rating how demanding they found the task of judging depth for each cue condition on a scale from 0 ("Not demanding at all") to 10 ("Extremely demanding"). In total, there were 81 trials for each augmented depth cue within each environment, culminating in 648 trials for the session, encompassing both clutter and non-clutter conditions with all four depth cues.

6 Results

A total of 25 participants took part in the study (19M, 6F), aged 20 - 44 years ($M = 29, SD = 5.58$). 21 participants reported having prior experience with VR. All participants had normal or corrected-to-normal vision. Overall, 16200 trials were analysed.

Depth estimation performance during the (non-)clutter sessions was measured as allocation error rate calculated as the number of misallocated asteroids divided by the total number of asteroids. To achieve better comparability of the data, the raw allocation error rate values were z-score normalised within the participants prior to performing any further analyses. To measure depth estimation performance over time, a slope of the regression line across all trials was calculated for each participant, using the z-score normalised error rates. Exploratory, we performed multiple one sample t-tests against zero to test whether the slopes and thus depth estimation performance during different cue conditions (i.e., no cue, text, color, and blur) changed - increased or decreased - over time. None of the tests showed statistical significance for any of the slopes in the clutter or the non-clutter session, suggesting a stable performance over time, irrespective of the cue or the clutter manipulation. Additionally, four polynomial contrasts were conducted using the z-score normalised error rates to compare depth estimation performance in different cue conditions between the non-clutter and the clutter sessions. Planned comparisons revealed significantly higher error rates for the text in the non-clutter ($M = -.43, SD = .23$) compared to the clutter ($M = -.32, SD = .22$) session, with a mean difference of -.11 ($SE = .03$; $F[1,24] = 13.23, p < .05$, partial $\eta^2 = .355$). No statistically significant differences in error rates between the non-clutter and the clutter session were found for the no cue ($M = .60, SD = .30$ and $M = .51, SD = .22$, respectively; $MD = .09, SE = .05$; $F[1,24] = 3.36$, ns., partial $\eta^2 = .123$), the color ($M = -.20, SD = .21$ and $M = -.22, SD = .17$; $MD = .02, SE = .03$, respectively; $F[1,24] = 0.28$, ns., partial $\eta^2 = .011$), and the blur ($M = .03, SD = .25$ and $M = .03, SD = .22$, respectively; $MD = .00, SE = .05$; $F[1,24] = 0.00$, ns., partial $\eta^2 = .000$). ***Focal attention*** during the (non-)clutter sessions was measured as the total percentage of time the participant was looking at the (target) asteroid during each trial, i.e., gaze-on-target time. ***Subjective mental demand*** during different cue conditions was measured as the percentage of the self-reported mental demand perception score.

To further test differences in depth estimation performance and focal attention during (non-)clutter sessions, we carried out 4×3 repeated-measures ANOVAs with the within-subject factors *cue* (no cue vs text vs color vs blur), and *distance* (near vs medium vs far; see Table 1, and Sections 6.1 and 6.2). As for subjective mental demand, we performed a repeated-measures ANOVA with the within-subject factor *cue* (no cue vs text vs color vs blur; see Section 6.3). The Kolmogorov-Smirnov test was used to test the assumption of normal distribution. If not stated otherwise, residuals were normally distributed. Mauchly's test was used to test the assumption of sphericity. In case of sphericity violation, the Greenhouse-Geisser correction was used. The outliers were identified using boxplots. The same parametric procedure was performed twice, i.e. once without and once with outliers. When reporting statistics, the values are by default reported for the sample without the outliers. In case of discrepancies between the statistical results on the samples without vs with outliers, the affected main factors and/or interactions include the reports of the statistics from the latter as well. In such cases, only results from the sample without the outliers are further interpreted. As post-hoc tests, Bonferoni-adjusted t-tests for dependent sample were performed. When reporting the results of the post-hoc tests on interactions between the main effects for individual ANOVAs, only statistically significant comparisons are being explicitly reported.

6.1 Non-Clutter Depth Estimation Sessions

Depth Estimation Performance – The results of the 4×3 repeated-measures ANOVA are presented in Table 1. Four participants were identified as outliers. The analysis showed a statistically significant main effect of the ***cue***. The post-hoc tests showed that error rate was significantly higher during no cue ($M = .64, SD = .06$) when compared to text ($M = -.41, SD = .05$), color ($M = -.26, SD = .03$) and blur ($M = .01, SD = .05$; all $p < .05$). A significantly higher error rate was also observed during blur when compared to text ($p < .05$), whereas the difference between text and color was not significant (ns.). In addition, there was a statistically significant main effect of the ***distance***. The post-hoc tests showed that error rate was significantly lower at the near ($M = -.26, SD = .05$) compared to the medium ($M = -.08, SD = .03$) and the far ($M = .32, SD = .05$; all $p < .05$) distance. Moreover, the medium distance revealed significantly lower error rate in comparison to the far distance ($p < .05$). The analysis further revealed a statistically significant interaction between ***cue × distance*** (see Figure 5). The post-hoc tests showed that during no cue and text, error rate was significantly higher at the far distance compared to the medium and the near distance (both $p < .05$). During color and blur, comparisons revealed significant differences between all three distances (all $p < .05$), showing an increase in error rate with an increasing distance and thus replicat-

Table 1: Results of the 4×3 ANOVAs on depth estimation performance and focal attention during the (non-)clutter depth estimation sessions

	df	df_{error}	F	p	$\eta^2_{partial}$
Non-Clutter - Depth Estimation Performance					
Cues	3	60	71.68	.000	.782
Distance	1.38	27.50	30.30	.000	.602
Cues x Dist.	2.01	41.75	4.23	.001	.174
Non-Clutter - Focal Attention					
Cues*	1.05	23.20	233.60	.000	.914
Distance*	1.48	32.57	5.34	.016	.195
Distance*	[1.42]	[34.08]	[3.27]	[.066]	[.120]
Cues x Dist.*	2.19	48.18	55.94	.000	.718
Clutter - Depth Estimation Performance					
Cues	3	51	75.47	.000	.816
Distance*	1.27	21.63	12.56	.000	.425
Cues x Dist.*	1.95	33.21	1.44	.251	.078
Cues x Dist.	[4]	[96]	[4.35]	[.009]	[.154]
Clutter - Focal Attention					
Cues*	1.07	23.37	638.75	.000	.971
Distance*	1.37	26.09	3.83	.049	.168
Cues x Dist.*	1.60	30.40	51.22	.000	.729

The values in the brackets [] represent the outcomes of the analysis performed with the outliers, only for those parts of the analysis were the outcome differs from the analyses without the outliers. Asterisks * are indicating Greenhouse-Geisser correction applied to the *df*s to adjust for the sphericity violation.

ing the trend previously reported for the main effect *distance*. The post-hoc tests further revealed that at the near distance, error rate was significantly higher during no cue compared to text, color and blur (all $p < .05$) as well as during blur when compared to text ($p < .05$). At the medium distance, significantly higher error rates were observed during no cue compared to text, color and blur (all $p < .05$), during blur compared to text and color (all $p < .05$), and during color compared to text ($p < .05$). As for the far distance, post-hoc tests replicated the trend reported for the main effect *cue*.

Focal Attention – The results of the 4×3 repeated-measures ANOVA are presented in Table 1. Two participants were identified as outliers. The analysis revealed a statistically significant main effect of the ***cue***. The post-hoc tests showed that participants spent significantly more time looking at target during text ($M = 48.14\%, SE = 2.89$) compared to color ($M = 1.12\%, SE = 0.20$), blur ($M = 1.63\%, SE = 0.29$), and no cue ($M = 3.00\%, SE = 0.45$; all $p < .05$). Moreover, they spent significantly less time looking at target during color when compared to no cue ($p < .05$). No significant differences in gaze-on-target time were found in blur compared to color and no cue (both ns.). Furthermore, the analysis showed a statistically significant main effect of the ***distance***. The post-hoc tests showed that participants spent significantly more time looking at target at the near ($M = 14.09\%, SE = 0.65$) compared to the far distance ($M = 12.81\%, SE = 0.73, p < .05$). No significant differences in gaze-on-target time were found at the medium ($M = 13.51\%, SE = 0.75$) distance compared to the near and the far distance (both ns.). The analysis further revealed a statistically significant interaction between ***cue*** × ***distance*** (see Figure 6 (a)). The post-hoc tests showed that during no cue, text and blur, comparisons between all three distances were statistically significant (all $p < .05$), showing an increase in gaze-on-target time with an increasing distance. As for color, participants spent more time looking at target during the near and the medium distance compared to the far distance (both $p < .05$). Further post-hoc tests revealed that at the

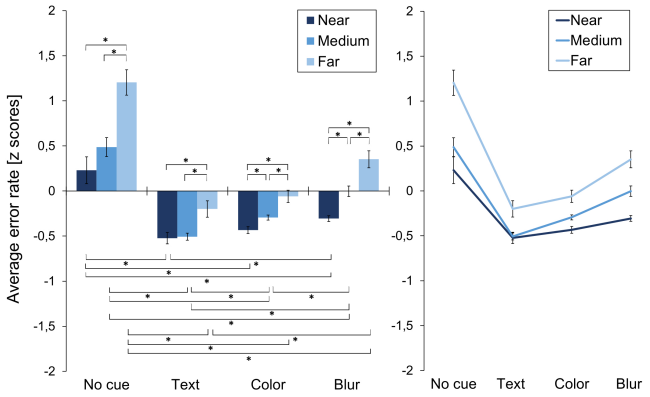


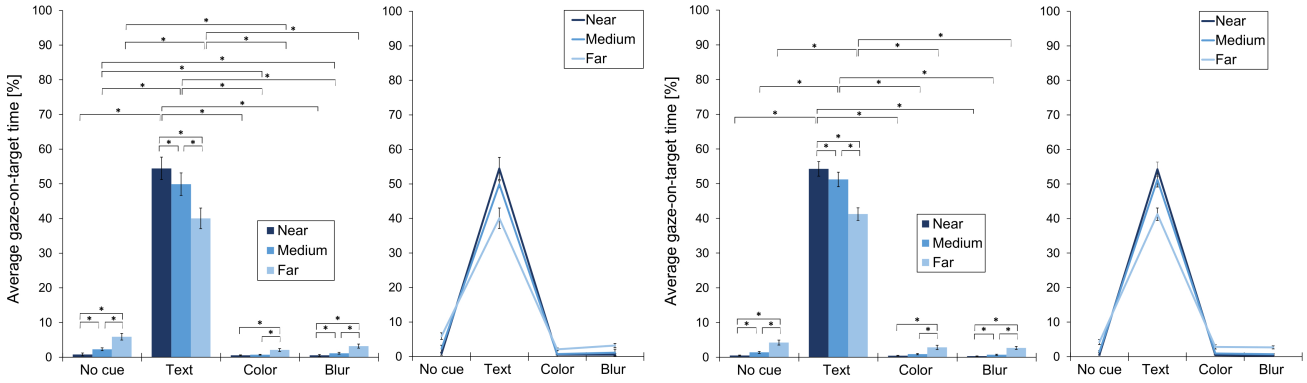
Figure 5: Average error rate for depth cues (no cue, text, color, and blur) and catching distances (near, medium, and far) in the non-clutter session. Error bars represent standard errors, and asterisks indicate statistically significant differences ($^*p < .05$). Both the bar chart and the line chart present the same values. The bar chart shows statistically significant differences and the line chart visualises interaction trends.

near distance, participants spent significantly more time looking at target during text compared to color, blur, and no cue (all $p < .05$). At the medium distance, gaze-on-target time was again significantly higher during text compared to color, blur, and no cue (all $p < .05$), as well as during no cue compared to color and blur (both $p < .05$). As for the far distance, participants spent significantly more time looking at target during text compared to color, blur, and no cue (all $p < .05$), and during color compared to no cue ($p < .05$), which is in line with the results reported for the main effect cue *cue*.

6.2 Clutter Depth Estimation Sessions

Depth Estimation Performance – The results of the 4×3 repeated-measures ANOVA are presented in Table 1. Eight participants were identified as outliers. The analysis showed a statistically significant main effect of the ***cue***. Similarly to the non-clutter sessions, the post-hoc tests revealed significantly higher error rate during no cue ($M = 564, SD = .046$) compared to text ($M = -.33, SD = .033$), color ($M = -.23, SD = .03$) and blur ($M = -.10, SD = .05$; all $p < .05$), as well as during blur compared to text ($p < .05$), while showing no significant difference between text and color (ns.). The main effect of the ***distance*** was shown to be statistically significant as well. The post-hoc tests revealed a significantly lower error rate at the near distance ($M = -.162, SD = .056$) compared to the far ($M = .20, SD = .05$; $p < .05$) distance. Furthermore, the medium ($M = -.11, SD = .03$) distance showed significantly lower error rate compared the far distance ($p < .05$), whereas no significant differences were found between the near and the medium distance (ns.). No statistically significant interactions were found between ***cue*** × ***distance***.

Focal Attention – The results of the 4×3 repeated-measures ANOVA are presented in Table 1. Five participants were identified as outliers. The analysis revealed a statistically significant main effect of the ***cue***. The post-hoc tests revealed that participants spent significantly more time looking at target during text ($M = 48.90\%, SE = 1.82$) compared to color ($M = 1.39\%, SE = 0.25$), blur ($M = 1.23\%, SE = 0.20$), and no cue ($M = 2.04\%, SE = 0.32$; all $p < .05$). No significant differences in gaze-on-target time were found between other cue conditions (all ns.). In addition, the analysis showed a statistically significant main effect of the ***distance***, however, the post-hoc tests revealed no statistically significant comparisons (all ns.) Furthermore, the analysis revealed a statistically



(a) Non-clutter session (b) Clutter session

Figure 6: Average gaze-on-target time for different depth cues (no cue, text, color, and blur) and catching distances (near, medium, and far) during the study sessions. The error bars represent standard errors, and the asterisks indicate statistically significant differences ($^*p < .05$). Note that the bar chart and the line chart are both representing the same values, whereas the bar chart shows statistically significant differences and the line chart visualises interaction trends.

significant interaction between ***cue* × *distance*** (see Figure 6 (b)). The same as during the non-clutter session, the post-hoc tests showed that during no cue, text and blur , comparisons between all three distances were statistically significant (all $p < .05$), showing an increase in gaze-on-target time with an increasing distance. During color, participants spent more time looking at target at the near and the medium distance compared to the far distance (both $p < .05$). Furthermore, the post-hoc tests revealed that during all three distances (i.e., near, medium and far), participants spent significantly more time looking at target during text compared to color, blur, and no cue (all $p < .05$), which is in line with the trend previously reported for the main effect *cue*.

6.3 Subjective Mental Demand Assessment

One participant was identified as an outlier. A repeated-measures ANOVA revealed a statistically significant effect of the ***cue*** ($F[3,69] = 12.46, p < .05$, partial $\eta^2 = 0.351$). The post-hoc tests showed that participants' reported significantly higher mental demand for blur ($M = 63.20, SE = 5.12$) compared to text ($M = 31.20, SE = 5.55$), color ($M = 32.00, SE = 5.00$), and no cue ($M = 40.40, SE = 5.05$; all $p < .05$).

7 Discussion

We examined the impact of augmented depth cues (subtle and explicit) in VR, both in non-cluttered and cluttered environments, focusing on enhancing depth perception. Before discussing our results, it is important to acknowledge certain limitations that contextualize our findings and guide future research. Our study primarily focused on texture blur as a subtle cue, limiting the exploration of other subtle cues like opacity, luminance, or distortion. The controlled VR environment excluded many natural factors to isolate the effects of the subtle cue. Future work should assess these cues in more realistic settings by adding more natural cues. Additionally, the controlled nature of our study meant disabling physics for asteroid trajectories in cluttered scenes to avoid collisions. Our results highlight the importance of subtle cues, like blur, in enhancing depth perception while preserving the visual integrity of virtual objects. These unobtrusive cues are suitable for VR applications where aesthetics are crucial. While we focused on blur, combining it with other cues like opacity or luminance could provide a more comprehensive depth cue system. Testing these combinations in diverse VR scenarios will help understand their effectiveness. Sensitivity to blur decreases in peripheral vision due to reduced visual acuity. Future experiments will vary the visual field to assess the effectiveness of blur cues across different gaze areas, including objects not directly in the user's line of sight. By initially focusing on central vision, we controlled for visual acuity, ensuring the subtle cues' effectiveness was accurately assessed. Focusing on single moving objects simplified the initial investigation. Future work will test both subtle and explicit cues with multiple moving objects to better inform the design of VR systems supporting complex user interactions. This staged approach allows systematic evaluation of multiple variables, ensuring thorough understanding of each aspect of depth perception. To address non-uniform and non-familiar objects in the study, we used an asteroid with a non-uniform shape and texture, preventing participants from relying on familiar understanding and adding complexity to the task. This controlled setting established a foundational understanding of the cues' effectiveness before introducing more variables. Future studies will evaluate our subtle cue in indoor, outdoor, and more complex environments to demonstrate their robustness and versatility in enhancing depth perception across different VR experiences. Additionally, future research will test our method across diverse scenarios and object types, including various sizes, textures, colors, shapes, and lighting, to broaden the applicability of our findings. To gauge the stability and effectiveness of subtle cues over time, more trials are necessary, which we plan to include in future studies. Future research will also include experiments to measure depth perception accuracy using techniques like verbal estimation or distance matching, providing direct quantification of perceived depth. Comprehensive vision tests will ensure participants do not have undiagnosed depth perception problems, recognizing the prevalence of 3D and depth perception issues. While depth of field (DoF) effects are underutilized in VR HMDs without eye-tracking capabilities, incorporating DoF through eye-tracking could enhance future perception studies. Within these limitations, the following are the most significant findings from our study.

Subtle cues significantly improved depth perception, but explicit cues performed slightly better. Our results demonstrated that all augmented cues outperformed the absence of cues. Subtle cues enhanced performance but were less effective than explicit cues (text and color), even though participants did not consciously notice them, aligning with previous findings [30, 66]. These findings address **RQ1**, indicating that augmented cues improve distance judgment precision under various perceptual difficulty conditions. Text cues

performed best by aiding spatial memory and reducing ambiguity, thereby enhancing depth judgments [69]. However, subtle cues have high potential as they retain contextual information without being visually invasive, blending seamlessly with the surroundings and maintaining immersion. Explicit cues enhance depth perception by directing attention to spatial attributes [39, 92], though they can disrupt immersion if not designed carefully [5, 7]. When examining cue integration, the blur cue works in conjunction with natural depth cues such as disparity, convergence and accommodation, motion, and aerial perspective, as well as relative density, to enhance overall depth perception. Textual cues usually combine with other natural depth cues like binocular disparities and motion parallax, improving spatial understanding. Similarly, color cues integrate with spatial cues like texture, shape, and atmospheric effects, creating cohesive visual perception. Further research is needed to address the interplay between augmented and natural depth cues precisely.

Text draws more attention to the target object compared to color and blur, yet mental demand is highest for blur. Our study examined the focused attention level for each augmented cue. Text cues received the most attention, explicitly conveying information but potentially causing tunnel vision, reducing situational awareness and engagement in VR [6, 8, 30]. This addresses **RQ2**, showing how explicit and subtle cues differ in impact on user performance. Excessive focus on text can detract from immersion, limiting exploration and interaction in VR. Balancing attentional demands with broader situational awareness is crucial for creating engaging VR experiences. Dynamic attention guidance strategies may help mitigate the drawbacks of text cues dominating attention [66]. In contrast, blur cues subtly influence attention without dominating the scene [43], and color cues had a comparably low level of attention, aiding spatial navigation in virtual environments [47, 82]. Attention is usually interrelated with task demand [41]. Our results showed that blur had the highest mental demand, which was unexpected. This suggests that the effectiveness of blur as a subtle cue can be impacted by cognitive demands [17]. Textual cues also vary in their influence on perception depending on cognitive factors, aligning with insights into visual working memory [71]. Similarly, the effectiveness of color in cognitive tasks is closely related to cognitive demands and motivational states [61]. Participants did not consciously recognize the blur cue but still found it demanding, indicating a significant cognitive load and complex perceptual processing in VR. This suggests that cues can impact cognitive workload and task performance even without conscious awareness, possibly due to active search and processing efforts [55]. In contrast, textual cues explicitly direct focus to specific objects, influencing both conscious and subconscious states of perception [70]. Further research is needed to assess cues in depth judgment tasks without informing users about additional cues, impacting subtle cues like blur.

Both subtle cues and explicit cues exhibited similar levels of effectiveness across cluttered and non-cluttered environments. We examined results in both cluttered and non-cluttered environments, addressing **RQ1** and showing that augmented cues enhance depth perception regardless of environmental clutter. This finding somewhat contradicts previous work indicating reduced performance in cluttered environments [18]. Future research should explore occlusion as a depth cue in cluttered settings, as our study focused on clutter difficulty. Explicit cues outperformed subtle cues overall, but subtle cues led to better performance than no cues in cluttered environments, demonstrating their potential. There was no significant difference in focused attention and distance impact between conditions. Previous research suggests that subtle cues can provide necessary depth information in cluttered environments without adding to cognitive load, unlike explicit cues, which may increase visual clutter [27, 55]. Both cue types maintained effectiveness in both settings, emphasizing the need for thoughtful design of depth cues in VR experiences [7, 74]. Subtle cues were unaffected by the clutter factor, while text cues were impacted. Subtle cues remained visible due to inherent aspects of visual processing [12], whereas text readability was affected by clutter. This shows subtle cues' scalability and persistence even in large, dynamic VR environments.

Subtle cues performed better in near and medium distance than far distances, similarly to color cues. When analyzing results across different distances, text cues performed best at near distances. Color and blur cues both outperformed the absence of any cue. At medium distances, all augmented cues significantly outperformed no cues. However, at far distances, while text and color cues remained effective, subtle cues showed a significant decrease in performance. These findings suggest the need for adaptive cue utilization based on distance, addressing **RQ2**. Subtle cues like blur diminish in performance with increasing distance, while explicit cues maintain effectiveness. Text cues, being explicit and direct, provide clear comprehension beneficial for near distances where detail and precision are crucial [68]. Color and blur cues, considered subtler than text, did not differ significantly from each other but contributed to depth perception at closer ranges [59, 90]. The effectiveness of subtle cues decreases with increasing distance, likely due to their limited capacity to convey depth information over longer distances as visual acuity diminishes [65, 89]. Text and color cues maintain their effectiveness at far distances, as they can draw attention and convey depth information without relying heavily on visual acuity [20, 21].

The performance was stable over time in all cues. After analyzing depth perception performance across all cues, we found consistent stability in their effectiveness. Interestingly, during post-study discussions, participants admitted they did not recognize the blur cue but found it demanding, likely due to prior indication prompting them to search for its effects consciously. This may have influenced mental demand and attention, suggesting the need for further research. Despite this, subtle cues demonstrated reliable performance over time. The stability of all cues is significant, indicating robustness across different temporal contexts within a VR experience, at least within our study's duration. Longitudinal effects warrant further investigation. This stability is relevant for VR design, implying that once users adapt to specific cues, their effectiveness remains consistent [73].

8 Conclusion

This work developed and evaluated augmented depth cues in virtual environments, demonstrating their effectiveness in enhancing depth estimation. Subtle cues, particularly blur, improve depth estimation and immersion without overwhelming users, effective in both cluttered and non-cluttered settings. Their effectiveness is distance-dependent, and combining them with color cues could enhance performance at greater distances. Our findings show subtle cues like blur can enhance depth perception with minimal contextual impact. Future studies should explore additional subtle cues, such as opacity, illumination, and distortion, and assess long-term effects. Testing these cues in varied scenarios is crucial for validating and generalizing our findings. Future research will extend to diverse environments and dynamic settings, exploring the interaction of these cues with complex natural environments in VR. This approach provides foundational insights and marks initial steps in exploring depth cues in VR.

9 Acknowledgements

The Know-Center is funded within the Austrian COMET Program, supported by the BMVIT, BMWFJ, and the State of Styria, and managed by the FFG.

References

[1] T. Ahsan, K. Bolton, L. M. Wilcox, and E. Freud. Perceived depth modulates perceptual resolution. In *Psychonomic Bulletin & Review*, pp. 1–12. Springer, 2021. 4

[2] R. Bailey, C. Grimm, C. Davoli, and R. Abrams. The effect of object color on depth ordering. 2007. 3

[3] G. Berkeley. *An essay towards a new theory of vision*. IndyPublish.com, 1709. 3

[4] M. S. Bittermann, I. S. Sariyildiz, and Ö. Ciftcioglu. Blur in human vision and increased visual realism in virtual environments. In 2007, Proc., Part I 3, pp. 137–148. Springer, 2007. 3

[5] B. Bodenheimer, J. Meng, H. Wu, G. Narasimham, B. Rump, T. P. McNamara, T. H. Carr, and J. J. Rieser. Distance estimation in virtual and real environments using bisection. In *Proc. of the ACM SAP*, pp. 35–40, 2007. 2, 8

[6] G. F. Briggs, G. J. Hole, and J. A. Turner. The impact of attentional set and situation awareness on dual tasking driving performance. *Transportation research part F: traffic psychology and behaviour*, 57:36–47, 2018. 8

[7] T. Chandler, M. Cordeil, T. Czauderna, et al. Immersive analytics. In *2015 IEEE BDVA*, pp. 1–8. IEEE, 2015. 8

[8] A. Chaparro and R. Young. Reading with rods: the superiority of central vision for rapid reading. *Investigative ophthalmology & visual science*, 34(7):2341–2347, 1993. 8

[9] C.-Y. Chen, C.-H. Chuang, T.-L. Tsai, H.-W. Chen, and P.-J. Wu. Reducing cybersickness by implementing texture blur in the virtual reality content. *Virtual Reality*, pp. 1–12, 2022. 3

[10] K. J. Ciuffreda, B. Wang, and B. Vasudevan. Conceptual model of human blur perception. *Vision Research*, 47(9):1245–1252, 2007. 3

[11] F. A. Conners. Attentional control and the simple view of reading. *Reading and Writing*, 22:591–613, 2009. 3

[12] L. Copeland and T. Gedeon. Visual distractions effects on reading in digital environments: a comparison of first and second english language readers. In *Proc. of the Annual Meeting of the Australian SIGCHI*, pp. 506–516, 2015. 8

[13] S. H. Creem-Regehr, P. Willemsen, A. A. Gooch, and W. B. Thompson. The influence of restricted viewing conditions on egocentric distance perception: Implications for real and virtual indoor environments. *Perception*, 34(2):191–204, 2005. 2

[14] J. Cutting. *Reconceiving perceptual space*, pp. 215–238. 01 2003. 2

[15] J. E. Cutting and P. M. Vishton. Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In *Perception of space and motion*, pp. 69–117. Elsevier, 1995. 1, 2

[16] J. A. Da Silva. Scales for perceived egocentric distance in a large open field: Comparison of three psychophysical methods. *The American Journal of Psychology*, pp. 119–144, 1985. 2

[17] B. B. De Koning, H. K. Tabbers, R. M. Rikers, and F. Paas. Attention guidance in learning from a complex animation: Seeing is understanding? *Learning and instruction*, 20(2):111–122, 2010. 8

[18] M. Delmas, L. Caroux, and C. Lemercier. Searching in clutter: Visual behavior and performance of expert action video game players. *Applied Ergonomics*, 99:103628, 2022. 8

[19] D. Drascic and P. Milgram. Perceptual issues in augmented reality. In *Stereoscopic displays and virtual reality systems III*, vol. 2653, pp. 123–134. Spie, 1996. 4

[20] B. Dresp, S. Durand, and S. Grossberg. Depth perception from pairs of overlapping cues in pictorial displays. 2002. 8

[21] H. Egusa. Effects of brightness, hue, and saturation on perceived depth between adjacent regions in the visual field. *Perception*, 12(2):167–175, 1983. 8

[22] F. El Jamiy and R. Marsh. Survey on depth perception in head mounted displays: distance estimation in virtual reality, augmented reality, and mixed reality. *IET Image Processing*, 13(5):707–712, 2019. 1, 2

[23] A. Field. *Discovering statistics using IBM SPSS statistics*. sage, 2013.

[24] J. M. Foley, N. P. Ribeiro-Filho, and J. A. Da Silva. Visual perception of extent and the geometry of visual space. *Vision Research*, 44(2):147–156, 2004. 2

[25] P. Fuchs. Interfaces visuelles. In *Encyclopédie des Techniques de l'ingénieur*, volume Télécommunication, p. vol. TE 5906. éditeur de l'Encyclopédie des Techniques de l'ingénieur, May 2003. 2

[26] J. J. Gibson. The perception of the visual world. 1950. 2

[27] C. Gkoumas and A. Shimi. "the more you explore, the less you remember": unraveling the effects of scene clutter on learning and memory for targets. In *2022 ACM ETRA*, pp. 1–7, 2022. 8

[28] W. C. Gogel and J. D. Tietz. Absolute motion parallax and the specific distance tendency. *Perception & Psychophysics*, 13(2):284–292, 1973. 2

[29] R. L. Gregory. *Eye and brain: The psychology of seeing*, vol. 80. Princeton university press, 2015. 1

[30] S. Grogorick and M. Magnor. *Subtle Visual Attention Guidance in VR*, pp. 272–284. Springer Int. Publishing, Cham, 2020. 2, 7, 8

[31] C. R. Guibal and B. Dresp. Interaction of color and geometric cues in depth perception: When does "red" mean "near"? *Psychological Research*, 69(1):30–40, 2004. 1, 3

[32] R. T. Held, E. A. Cooper, and M. S. Banks. Blur and disparity are complementary cues to depth. In *Current Biology*, volume 22, number 5, pages 426–431. Elsevier, 2012. 2

[33] S. Hillaire, A. Lécuyer, R. Cozot, and G. Casiez. Depth-of-field blur effects for first-person navigation in virtual environments. In *Proc. of the 2007 ACM VRST*, pp. 203–206, 2007. 3

[34] P. Hu, S. Boorboor, S. Jadhav, J. Marino, S. Mirhosseini, and A. E. Kaufman. Spatial perception in immersive visualization: A study and findings. In *2022 IEEE ISMAR-Adjunct*, pp. 369–372. IEEE, 2022. 2

[35] R. Hussain, M. Chessa, and F. Solari. Modelling foveated depth-of-field blur for improving depth perception in virtual reality. In *2020 IEEE 4th IPAS*, pp. 71–76. IEEE, 2020. 3

[36] S. Ichihara, N. Kitagawa, and H. Akutsu. Contrast and depth perception: Effects of texture contrast and area contrast. *Perception*, 36(5):686–695, 2007. 2

[37] P. N. Johnson-Laird. Mental models in cognitive science. *Cognitive science*, 4(1):71–115, 1980. 3

[38] J. A. Jones, J. E. Swan, G. Singh, and S. R. Ellis. Peripheral visual information and its effect on distance judgments in virtual and augmented environments. In *Proc. of the ACM SAP*, pp. 29–36, 2011. 2

[39] J. A. Jones, J. E. Swan, G. Singh, E. Kolstad, and S. R. Ellis. The effects of virtual reality, augmented reality, and motion parallax on egocentric depth perception. In *Proc. of the 5th SAP*, pp. 9–14, 2008. 8

[40] J. A. Jones, J. E. Swan II, and M. Bolas. Peripheral stimulation and its effect on perceived spatial scale in virtual environments. *IEEE TVCG*, 19(4):701–710, 2013. 2

[41] D. Kahneman. *Attention and Effort*. Prentice-Hall, 1973. 8

[42] J. W. Kelly. Distance perception in virtual reality: A meta-analysis of the effect of head-mounted display characteristics. In *IEEE Transactions on Visualization and Computer Graphics*, IEEE, 2022. 2

[43] R. A. Khan, E. Dinet, and H. Konik. Visual attention: Effects of blur. In *2011 18th IEEE CIP*, pp. 3289–3292. IEEE, 2011. 8

[44] E. Klein, J. E. Swan, G. S. Schmidt, M. A. Livingston, and O. G. Staadt. Measurement protocols for medium-field distance perception in large-screen immersive displays. In *2009 IEEE VR*, pp. 107–113. 2

[45] J. M. Knapp and J. M. Loomis. Limited field of view of head-mounted displays is not the cause of distance underestimation in virtual environments. *Presence: Teleoperators & Virtual Environments*, 13(5):572–577, 2004. 2, 4

[46] R. Konrad and T. Kong. Depth cues in vr head mounted displays with focus tunable lenses. *Standford University, published Aug*, 2015. 2

[47] Ö. Kumoğlu and N. Olguntürk. The aid of colour on spatial navigation: A study in a virtual hospital environment. In *Proc. of the European Conf. on Cognitive Ergonomics*, pp. 1–3, 2016. 8

[48] M. S. Landy, L. T. Maloney, and M. J. Young. Psychophysical estimation of the human depth combination rule. In *Sensor Fusion III: 3D Perception and Recognition*, vol. 1383, pp. 247–254. SPIE, 1991. 2

[49] L. Li, T. Yamada, and W. Choi. The effect of depth information on visual complexity perception in three-dimensional textures. In *Applied Sciences*, 2020. doi:10.3390/app10155347. 4

[50] Z. Li, Y. Cui, T. Zhou, Y. Jiang, Y. Wang, Y. Yan, M. Nebeling, and Y. Shi. Color-to-depth mappings as depth cues in virtual reality. In *Proc. of the 35th ACM UIST*, pp. 1–14, 2022. 3

[51] P. Liu, J. Forte, D. Sewell, and O. Carter. Cognitive load effects on early visual perceptual processing. *Attention, Perception, & Psychophysics*, 80:929–950, 2018. 2

[52] M. A. Livingston, A. Dey, C. Sandor, and B. H. Thomas. *Pursuit of "X-Ray Vision" for Augmented Reality*, pp. 67–107. Springer New York,

NY, 2013. 1, 2
[53] M. Livingstone and D. H. Hubel. Vision and art: The biology of seeing. 2002. 3
[54] J. M. Loomis, J. M. Knapp, et al. Visual perception of egocentric distance in real and virtual environments. *Virtual and adaptive environments*, 11:21–46, 2003. 2
[55] L. C. Loschky, R. V. Ringer, A. P. Johnson, A. M. Larson, M. Neider, and A. F. Kramer. Blur detection is unaffected by cognitive load. *Visual cognition*, 22(3-4):522–547, 2014. 3, 8
[56] G. Maiello, L. Walker, P. J. Bex, and F. A. Vera-Diaz. Blur perception throughout the visual field in myopia and emmetropia. *Journal of vision*, 17(5):3–3, 2017. 3
[57] A. J. Marcel. Conscious and unconscious perception: An approach to the relations between phenomenal experience and perceptual processes. *Cognitive psychology*, 15(2):238–300, 1983. 3
[58] G. Mather. Image blur as a pictorial depth cue. *Proc. of the Royal Society of London. Series B: Biological Sciences*, 263(1367):169–172, 1996. 3
[59] G. Mather. The use of image blur as a depth cue. *Perception*, 26(9):1147–1158, 1997. 3, 8
[60] G. Mather and D. R. Smith. Blur discrimination and its relation to blur-mediated depth perception. *Perception*, 31(10):1211–1219, 2002. 3
[61] R. Mehta and R. Zhu. Blue or red? exploring the effect of color on cognitive task performances. *Science*, 323(5918):1226–1229, 2009. 8
[62] B. J. Mohler, S. H. Creem-Regehr, and W. B. Thompson. The influence of feedback on egocentric distance judgments in real and virtual environments. In *Proc. of the 3rd SAP*, pp. 9–14, 2006. 2
[63] A. Murgia and P. M. Sharkey. Estimation of distances in virtual environments using size constancy. *Int. Journal of Virtual Reality*, 8(1):67–74, 2009. 2
[64] A. Naceri, R. Chellali, S. Toma, and F. Dionnet. What visual cues do we use to perceive depth in virtual environments? In *ICIRA 2009. Proc. 2*, pp. 812–823. Springer, 2009. 1
[65] P. E. Napieralski, B. M. Altenhoff, et al. Near-field distance perception in real and virtual environments using both verbal and action responses. *ACM Trans. on Applied Perception (TAP)*, 8(3):1–19, 2011. 1, 8
[66] E. Nonino, J. Gisler, V. Holzwarth, C. Hirt, and A. Kunz. Subtle attention guidance for real walking in virtual environments. In *2021 IEEE ISMAR-Adjunct*, pp. 310–315, 2021. 7, 8
[67] R. P. O'Shea, D. G. Govan, and R. Sekuler. Blur and contrast as pictorial depth cues. *Perception*, 26(5):599–612, 1997. 3
[68] N. Padmanaban, R. Konrad, T. Stramer, E. A. Cooper, and G. Wetzstein. Optimizing virtual reality for all users through gaze-contingent and adaptive focus displays. *Proc. of the National Academy of Sciences*, 114(9):2183–2188, 2017. 2, 8
[69] S. D. Peterson, M. Axholt, M. Cooper, and S. R. Ellis. Evaluation of alternative label placement techniques in dynamic virtual environments. In *10th Int. Symp., SG 2009, Proc. 10*, pp. 43–55. Springer, 2009. 8
[70] S. D. Peterson, M. Axholt, and S. R. Ellis. Comparing disparity based label segregation in augmented and virtual reality. In *Proc. of the 2008 ACM VRST*, pp. 285–286, 2008. 8
[71] J. Qian, J. Li, K. Wang, S. Liu, and Q. Lei. Evidence for the effect of depth on visual working memory. *Sci. reports*, 7(1):6408, 2017. 2, 8
[72] J. C. Read. Visual perception: Understanding visual cues to depth. *Current Biology*, 22(5):R163–R165, 2012. 3
[73] R. S. Renner, B. M. Velichkovsky, and J. R. Helmert. The perception of egocentric distances in virtual environments-a review. *ACM CSUR*, 46(2):1–40, 2013. 8
[74] R. Rosenholtz, Y. Li, and L. Nakano. Measuring visual clutter. *Journal of vision*, 7(2):17–17, 2007. 1, 8
[75] V. Santangelo. Forced to remember: when memory is biased by salient information. *Behavioural brain research*, 283:1–10, 2015. 2
[76] G. Singh, S. R. Ellis, and J. E. Swan. The effect of focal distance, age, and brightness on near-field augmented reality depth matching. *IEEE TVCG*, 26(2):1385–1398, 2018. 2
[77] R. M. Spielman, K. Dumper, W. Jenkins, A. Lacombe, M. Lovett, and M. Perlmutter. *5.6 Gestalt Principles of Perception - Psychology | OpenStax*. Texas, Dec 2014. 2
[78] J. E. Swan, A. Jones, E. Kolstad, M. A. Livingston, and H. S. Smallman. Egocentric depth judgments in optical, see-through augmented reality. *IEEE TVCG*, 13(3):429–442, 2007.
[79] W. Tantanatewin and V. Inkarojrit. Effects of color and lighting on retail impression and identity. *Journal of Environmental Psychology*, 46:197–205, 2016. 3
[80] R. Teghtsoonian and M. Teghtsoonian. Range and regression effects in magnitude scaling. *Perception & Psychophysics*, 24(4):305–314, 1978. 1, 2
[81] A. Tourimpampa, A. Drigas, A. Economou, and P. Roussos. Perception and text comprehension. it'sa matter of perception! *Int. Journal of Emerging Technologies in Learning (Online)*, 13(7):228, 2018. 3
[82] K.-I. Tsutsui, H. Sakata, T. Naganuma, and M. Taira. Neural correlates for perception of 3d surface orientation from texture gradient. *Science*, 298(5592):409–412, 2002. 8
[83] T. A. Van Dijk, W. Kintsch, et al. Strategies of discourse comprehension. 1983. 3
[84] L. Verhoeven, W. Schnotz, and F. Paas. Cognitive load in interactive knowledge construction, 2009. 3
[85] D. Waller and A. R. Richardson. Correcting distance estimates by interacting with immersive virtual environments: effects of task and available sensory information. *Journal of Experimental Psychology: Applied*, 14(1):61, 2008. 2
[86] B. Wang and K. J. Ciuffreda. Depth-of-focus of the human eye in the near retinal periphery. *Vision Research*, 44(11):1115–1125, 2004. 3
[87] B. Wang and K. J. Ciuffreda. Blur discrimination of the human eye in the near retinal periphery. *Optometry and Vision Science*, 82(1):52–58, 2005. 3
[88] B. Wang and K. J. Ciuffreda. Foveal blur discrimination of the human eye. *Ophthalmic and Physiological Optics*, 25(1):45–51, 2005. 3
[89] L. R. Wanger, J. A. Ferwerda, D. P. Greenberg, et al. Perceiving spatial relationships in computer-generated images. *IEEE CG and A*, 12(3):44–58, 1992. 1, 2, 8
[90] S. J. Watt, K. Akeley, M. O. Ernst, and M. S. Banks. Focus cues affect perceived depth. *Journal of vision*, 5(10):7–7, 2005. 8
[91] C. Wickens, W. Helton, J. Hollands, and S. Banbury. *Engineering Psychology and Human Performance (5th ed.)*. Routledge, 2022. 2
[92] P. Willemsen, A. A. Gooch, W. B. Thompson, and S. H. Creem-Regehr. Effects of stereo viewing conditions on distance perception in virtual environments. *Presence: Teleoperators and Virtual Environments*, 17(1):91–101, 2008. 2, 8
[93] B. Williams, D. Johnson, L. Shores, and G. Narasimham. Distance perception in virtual environments. In *Proc. of the 5th ACM SAP*, pp. 193–193, 2008. 2
[94] L. Xueting and T. Ogawa. Blur with depth: A depth cue method based on blur effect in augmented reality. In *2013 IEEE ISMAR*, pp. 1–6. IEEE, 2013. 3
[95] K. Yildirim, A. Akalin-Baskaya, and M. Hidayetoglu. Effects of indoor color on mood and cognitive performance. *Building and Environment*, 42(9):3233–3240, 2007. 3
[96] Y. Zhang, H. Liu, S.-C. Kang, and M. Al-Hussein. Virtual reality applications for the built environment: Research trends and opportunities. *Automation in Construction*, 118:103311, 2020.
[97] R. A. Zwaan. Situation models, mental simulations, and abstract concepts in discourse comprehension. *Psychonomic bulletin & review*, 23:1028–1034, 2016. 3
[98] R. A. Zwaan, M. C. Langston, and A. C. Graesser. The construction of situation models in narrative comprehension: An event-indexing model. *Psychological science*, 6(5):292–297, 1995. 3