Foveated Walking: Translational Ego-Movement and Foveated Rendering

David Petrescu
david.petrescu@manchester.ac.uk
University of Manchester
Human-Computer Systems
Manchester, UK

Paul A. Warren
paul.warrem@manchester.ac.uk
Virtual Reality Research (VR2)
Facility, Division of Psychology,
Communication and Human
Neuroscience,
University of Manchester
Manchester, UK

Zahra Montazeri
zahra.montazeri@manchester.ac.uk
University of Manchester
Human-Computer Systems
Manchester, UK

Boris Otkhmezuri
boris.pettifer@manchester.ac.uk
Virtual Reality Research (VR2)
Facility, Division of Psychology,
Communication and Human
Neuroscience,
University of Manchester
Manchester, UK

Steve Pettifer
steve.pettifer@manchester.ac.uk
University of Manchester
Human-Computer Systems
Manchester, UK

ABSTRACT

The demands of creating an immersive Virtual Reality (VR) experience often exceed the raw capabilities of graphics hardware. Perceptually-driven techniques can reduce rendering costs by directing effort away from features that do not significantly impact the overall user experience while maintaining a high level of quality where it matters most. One such approach is foveated rendering, which allows for a reduction in the quality of the image in the peripheral region of the field-of-view where lower visual acuity results in users being less able to resolve fine details. 6 Degrees of Freedom tracking allows for the exploration of VR environments through different modalities, such as user-generated head or body movements. The effect of self-induced motion on rendering optimization has generally been overlooked and is not yet well understood. To explore this, we used Variable Rate Shading (VRS) to create a foveated rendering method triggered by the translational velocity of the users and studied different levels of shading Level-of-Detail (LOD). We asked 10 participants in a within-subjects design to report whether they noticed a degradation in the rendering of a rich environment when performing active ego-movement or when being passively transported through the environment. We ran a psychophysical experiment using an accelerated stochastic approximation staircase method and modified the diameter and the LOD of the peripheral region. Our results show that self-induced walking can be used to significantly improve the savings of foveated rendering by allowing for an increased size of the low-quality area in a foveated algorithm compared to the passive condition. After fitting psychometric functions showcasing the percentage of correct responses related to different shading rates in the two types of movements, we also report the threshold severity (75%) point for when participants are able to detect such degradation. We argue such metrics can inform the future design of movement-dependent foveated techniques that could reduce computational load and increase energy savings.

CCS CONCEPTS

• **Computing methodologies** → **Perception**; *Virtual reality*.

KEYWORDS

foveated rendering, psychophysics, variable rate shading, motion

ACM Reference Format:
David Petrescu, Paul A. Warren, Zahra Montazeri, Boris Otkhmezuri, and Steve Pettifer. 2023. Foveated Walking: Translational Ego-Movement and Foveated Rendering. In *ACM Symposium on Applied Perception 2023 (SAP '23), August 05–06, 2023, Los Angeles, CA, USA.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3605495.3605798

SAP '23, August 05–06, 2023, Los Angeles, CA, USA

ACM ISBN 979-8-4007-0252-5/23/08.
https://doi.org/10.1145/3605495.3605798

1 INTRODUCTION

Recent growth in the availability and affordability of Head-Mounted Display (HMD) technology combined with powerful consumer-grade graphics hardware has rekindled interest in Virtual Reality (VR). Generating compelling, fully-immersive VR experiences for leisure or research purposes in real-time is, however, computationally expensive. To maximize the quality of a user's experience, perceptual deficits and characteristics are often exploited in order to optimize the way we render such Virtual Environments. Shading represents one of the most expensive operations required of Graphics Processing Units (GPUs), and new display technologies

with increased refresh rates and resolutions make shading proportionally expensive. One popular approach to perceptually optimize this computational load is called *Foveated Rendering*. This approach rests on the fact that sensitivity to fine detail of the Human Visual System (HVS) decreases with distance from the fovea (the most sensitive part of the retina). In this work, we investigate whether self-induced (active) movement through an environment might permit more severe foveated rendering. More specifically, we consider how user-generated translational movement, a form of interaction specific to VR, might change the sensitivity to spatial changes in visual quality under foveated rendering. For comparison, we consider how sensitivity is affected when the user is stationary but the motion presented on-screen is consistent with active user movement (implied movement).

HMDs generally allow for tracking of user movement with 6 Degrees of Freedom (6DOF) when interacting with an environment. This enables free movement in VR, which users often perform in order to explore their surroundings. Such movement is typically accompanied by a combination of proprioceptive (from the muscles causing movement), vestibular (from the inner ear), motor commands (from the motor areas that drive movement) and purely visual (optic flow) information about self-movement. This complex array of information is important for the correct interpretation of a stable environment and movement of other objects within that environment. Moreover, the availability of these information sources changes depending on the types of movement undertaken [Warren et al. 2022]. In his seminal research, Murphy [Murphy 1978] showed that there is evidence of the HVS's sensitivity being reduced when movement on the retina is present. Other studies show similar results when global motion on the retina from the observer's eye or head movements [Braun et al. 2017] is present. To the best of our knowledge, information about how different types of motion can affect sensitivity to various forms of quality degradation is not well-understood. How this might change sensitivity to the image manipulations performed under foveated rendering and how this might depend on the cause of the movement is therefore of interest. In light of this, the contributions of this research are:

- We test if it is feasible to use the translational movement of the user in 6DOF in order to trigger degradation in a foveated rendering algorithm using Nvidia's VRS (i.e. independent from other HVS properties).
- Using a psychophysical procedure we measure the critical radius of the undegraded (HQ) region such that users cannot identify a drop in quality. We investigate how this radius might change depending on: 1) Whether the movement is active or implied and 2) The severity of the degradation outside this region.

2 BACKGROUND

2.1 Foveated Rendering

Foveated, or gaze-contingent, rendering aims to exploit the visual sensitivity falloff of the HVS caused by the distribution of photoreceptors on the retina [Curcio and Allen 1990]. Lower photoreceptor density lessens the eye's ability to resolve fine detail and lowers spatio-temporal contrast sensitivity in the peripheral region (i.e. reduced visual acuity). Because less detail is perceived as eccentricity increases, less computational power is required. Foveated techniques can be gaze-contingent or fixed, which build on the human tendency to look towards the center of the screen. For a state-of-the-art report refer to [Wang et al. 2023] and for surveys on perceptual rendering refer to [Spjut et al. 2020; Weier et al. 2017].

Other methods are reducing the spatial image resolution using gaze detection. [Guenter et al. 2012] used a simple linear model of degradation and improved the rendering performance by a factor of five. Other techniques use the detected gaze and switch resolutions based on a multi-level pyramid encoded in the image in real-time [Dorr et al. 2006; Perry and Geisler 2002]. Traditional, resolution-based techniques have been adapted for VR HMDs and show considerable computational savings without a perceptual consequence [Patney et al. 2016; Vaidyanathan et al. 2014]. In order to optimize foveated rendering for content in VR, [Tursun et al. 2019] developed a content-aware method that resulted in no significant perceptual differences from non-foveated rendering. In their method, they analyze the contrast and luminance of lower-resolution images to predict the minimal probability of detecting artifacts in the fully-rendered version of the frame. Weier et al. [Weier et al. 2016] showed foveation can be used to guide ray-tracing algorithms. Other applications of foveated techniques are creating a better user experience by using the user's gaze in order in a tone-mapping method in order to increase the visible dynamic range without inducing after-images [E. Jacobs et al. 2015].

2.2 Variable Rate Shading

Variable Rate Shading (VRS) [NVIDIA 2018] is a recent technology introduced for the Nvidia Turing architecture. This technique is similar to the coarse pixel shading method used for foveated rendering in [Vaidyanathan et al. 2014]. Compared to other multi-resolution techniques, VRS provides granular control over the resolution of shading in the image. Each 16x16 pixel image tile can have a different level of pixel shader calls. Note that the rasterization process will not be altered in order to not introduce additional aliasing artifacts and preserve sharper edges. In order to minimize the calls to the pixel shader, the appearance of pixels in the tile are derived from the same value. The shading rate of the 16x16 tile can have a uniform configuration: {2x2, 4x4} or a non-uniform one: {1x2, 2x1, 2x4, 4x2} which provides more control for vertical and horizontal shading. In our study, we will explore the uniform configurations for simplicity and because of the annular distribution of the degradation area which is symmetrical in both lenses. The consequence of this downsampling results in visual artifacts that can be amplified by the properties of a scene. Ample research has been done in order to maximize the computational benefits of VRS without inducing noticeable artifacts, and will be discussed in Section 2.3.

2.3 Related Work

The interaction between on-screen velocity and eccentricity in the visual field has been explored previously in the context of optimized rendering. For example, Reddy developed a geometrical LOD reduction algorithm that accounted for rotational velocity, but also for eccentricity [Reddy 2001]. Parkhurst and Niebur [Parkhurst and Niebur 2004] implemented a perceptual foveated renderer that

reduced the geometrical Level of Detail (LOD) of the meshes via an edge-collapse algorithm in a search and localization task. In their design, everything outside 1.5°of visual angles of the fixation was rendered with a lower LOD.

Perceptual metrics based on artifact detection caused by movement have since been developed in order to preserve motion quality. [Denes et al. 2020] created a model that explains how the perceived quality of motion varies as a function of refresh rate, velocity, and spatial resolution, and fitted psychometric data describing sensitivity to motion artifacts. Similarly, [Chapiro et al. 2019] created a luminance-aware model that predicts perceived judder in order to maintain visual quality. [Jindal et al. 2021] explored the effect of different on-screen velocities in VR and used adaptive shading using VRS in order to reduce rendering load whilst accounting for motion artifacts. In a series of studies, Suchow and Alvarez showed that movement reduces the ability to detect changes in qualitative properties of a stimulus (luminance, color, size and shape), an effect they called *silencing* [Suchow and Alvarez 2011b]. They showed the same effect to hold true when background motion occurs and foreground stimuli change [Suchow and Alvarez 2011a].

However, less work has considered the impact of self-induced motion on perceived quality. [Petrescu et al. 2023] explored how self-induced head rotations can be used as a metric for LOD reduction. [Ellis and Chalmers 2006] used a dynamic scaling of the foveated region, generating translational movement with a 6-DOF motion pod and scaled the region to the force perceived by the participants, whilst studying multiple sampling rates. In the present paper, we build upon this work, using the fact that 6-DOF user movement information is readily available in VR. We study the extent to which different types of movement (active vs. implied) and different levels of LOD degradation (2x2 vs 4x4) impact on the human ability to detect drops in quality.

3 MATERIALS AND METHODS

In the current study, participants either moved actively (AM condition) or were shown the equivalent visual motion while stationary (IM condition). They were asked whether they detected a difference from a reference scene with no degradation. The foveation algorithm was activated when the camera velocity reached 75% of the velocity of the fixation sphere. A velocity of 1 m/s was chosen, which is within the range of walking speeds usually present in VR environments [Perrin et al. 2019]. We chose to focus on a simple approach to degrading the image. Therefore, we utilized a simple model containing a foveal region (HQ) which was always rendered at the original sampling rate, and a peripheral region (LQ) which was rendered in 2x2 or 4x4 VRS configurations. This meant that the LQ region was rendered at 1/4 and 1/16 the sampling rate of the HQ region. The VRS implementation we used culled everything outside the peripheral region because HMD lenses are not rectangular.

3.1 Design

The study was performed in a within-subjects 2x2 fully factorial design. We manipulated the type of movement (active vs implied) as an independent variable. We compared the AM and IM conditions in order to determine how self-induced forward motion (surge) affects the probability of being unaware of quality change. We also manipulated 2 levels of degradation of the peripheral LQ region using Nvidia's VRS technology (2x2, 4x4). Therefore we have 4 experimental conditions. For each of these conditions, data were collected in one experimental block comprising a series of trials in which the radius of the HQ region was varied. The radius varied in line with two interleaved, partially overlapping adaptive staircases (Figure 4). One staircase started with a large diameter (HQ = 90°, which covered the entire visible FOV of our configuration), and the other started with full degradation (i.e. the diameter of the HQ = 0). We wanted to recover the diameter point at which participants would reliably miss the degradation (75% probability of not detecting). Consequently, the two staircases were defined to overlap and converge to points on either side of 75%, where participants had an 87.5% and 62.5% probability of reporting that they did not detect the degradation. For more information about the Kesten staircase, see below (Section 3.2). Each staircase was comprised of 40 trials. To improve participant engagement we also added catch trials (4x4 where HQ diameter = 0°) every 20 trials, resulting in 8 catch trials per session. This meant a total of 336 trials per participant across the 4 blocks of data. In order to minimize fatigue, we ran the experiment in two parts, one in which participants undertook the AM condition by producing self-induced surge motion by forward walking, and one in which they were stationary and were moved through the scene (IM). The order of these two parts of the experiment was randomized across participants to minimize learning and order effects.

3.2 Kesten Adaptive Staircase

Kesten's adaptive staircase, or *Accelerated Stochastic Approximation*, is a fast converging algorithm, widely used in psychophysics in order to adaptively control the levels of a stimulus presented to participants [Treutwein 1995]. Crucially, the Kesten staircase can be designed to converge at a specified value; in this case, a specific probability of not detecting the degradation. Here, we set the convergence value around the 75% point (i.e., the point where participants are three times more likely to detect a difference).

$$x_{n+1} = x_n - \frac{step}{n}(resp_n - \Phi), n \leq 2 \quad (1)$$

Equation 1 represents the Standard Stochastic Approximation staircase. The Kesten method shares the standard form for the first 2 trials in a staircase. In our case, x_n represents the diameter of the HQ area (refer to Figure 2). The initial value x_0 depends on whether the staircase is ascending or descending towards the threshold Φ.

$$diam(HQ) = x_n - \frac{step}{2 + m_{revs}}(resp_n - \Phi), n > 2 \quad (2)$$

Equation 2 represents the accelerated stochastic approximation. The variable m_{revs} refers to the shift in the response category (also called *reversals*). The variable $resp_n$ is a binary value representing the answer the participant gave for the n trial in the staircase. Having two interleaved staircases helps provide more data in the area of perceptual uncertainty. For the ascending staircase, we chose $x_0 = 0, \phi = 0.875$; and for the ascending staircase $x_0 = 0.9, \phi = 0.625$. Note that the values for ϕ were chosen symmetrically around the 75% point in order for the data to converge around the threshold

severity point. The *step* value was set to be 0.3, as piloting data suggested this provided reliable staircase convergence. These values were the same across the four experimental blocks. We present an example visualization of one of the pair of interleaved staircases from one of the participants in Figure 4.

3.3 Stimuli

We adapted Unity's Corridor Lighting Example (Figure 1) and scaled it so it matched the dimensions of the Virtual Reality Research (VR2) Facility (approximately 3x4m). Initially, participants were aligned to the diagonal (5 meters) of the lab before starting the experiment so they are able to walk the 3.5m distance without any risk of injury.

The fixation point was a blue sphere set a meter away from the participant that occupied 2° of visual angle. Because humans are not able to walk at a target velocity instantaneously, we added an acceleration parameter to the fixation point so it reached its peak velocity 0.25 seconds after the motion was activated. Note that the fixation point guides the speed of the observer and its behavior is further discussed in Section 4.2.

During the actual trial, participants were asked to report any artifacts they experienced during the observation (spatiotemporal aliasing or flickering). At the end of the trial, participants were presented with a blank scene and asked to record their answers by pointing at one of two buttons: *yes* or *no*, meaning respectively that they did or did not detect the degradation. In the walking condition, they were asked to rotate and press the trigger on the controller to start the next trial. In the passive condition, they simply pressed the trigger to initiate the next trial.

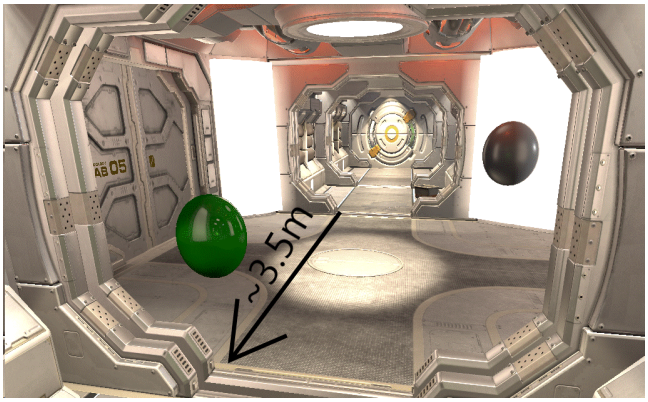

Figure 1: Sample scene used for both experiments (*Unity's Corridor Lighting Example*)

3.4 Participants

Ten participants (all staff and students at The University of Manchester, 9 naive to the study, and one involved in data collection) were recruited. The study was approved by the Ethics Committee of The University of Manchester, Division of Neuroscience and Experimental Psychology. Informed consent was given by all participants. Four participants had expertise in VR and computer graphics. Data were collected in two sessions and participants were given a break every 20 minutes. The effective time of the experiment was about 70 minutes per participant over the two sessions. Eye-tracking data and velocity data were collected in order to confirm that the participants performed the experiment as instructed. All participants had normal or corrected-to-normal vision.

3.5 Apparatus

The data for this experiment were collected in VR2 Facility. We used the Oculus Quest Pro headset with eye-tracking enabled. The headset has 1800 x 1920 pixels resolution per eye and a visible horizontal FOV of 106°. We used the Oculus Air Link which transferred data at 90Hz and capped the FPS at 90 on our desktop configuration (Intel i7-7700K CPU and NVIDIA RTX 3080 Ti GPU). Note that HMDs tend to vary in effective resolution over the FOV because of optics. [Beams et al. 2020] showed that effective resolution in VR HMDs decreases rapidly for off-axes angles which could mask visual artifacts. This is a common side-effect of the display technology. Given that we studied a relative effect in terms of the type of movement, we did not calculate the modulation transfer function for our HMD.

The experiment was coded using the *Unity* game engine and verbose data about each trial was collected using the Unity Experiment Framework (UXF), which was designed for the development and control of psychophysical studies [Brookes et al. 2020]. The HTC Implementation of Foveated Rendering was adapted for our configuration and used for a Unity implementation of VRS [ViveSoftware 2020].

4 EXPERIMENTS

4.1 Pre-screening

In order to confirm that participants could reliably detect differences, we started the experiment by asking them if they detected differences between the scene rendered at full quality and fully degraded (i.e. the diameter of the HQ region = 0°, and LQ = 90°) whilst stationary. Participants performed pairwise comparisons for each of these conditions and all of them were able to correctly identify the degraded scene and detect the difference between the two LQ conditions.

4.2 Active ego-movement

4.2.1 Procedure. The position of the fixation point and the orientation of the corridor were informed by the position of the participant. Thus, the ball was always positioned in front of the camera when the participant initiated movement and the corridor was re-aligned to their midsaggital plane in order for the trials to appear consistent, even if there were small deviations from the lab diagonal. A three-second countdown appeared on the screen and the fixation sphere accelerated for 0.25 seconds until it reached the final velocity of 1 m/s. This ensured that the participant had enough time to reach the intended velocity. Participants were warned if they did not achieve the intended velocity by the fixation point turning red whenever they were not within 1.5 meters of it.

At the end of the trial, participants were presented with a blank scene where they recorded the answer and were instructed to turn 180°. To ensure that they had produced a complete rotation, we added alignment cues, more specifically a line on the floor for the alignment of the midsaggital plane and a green sphere that they

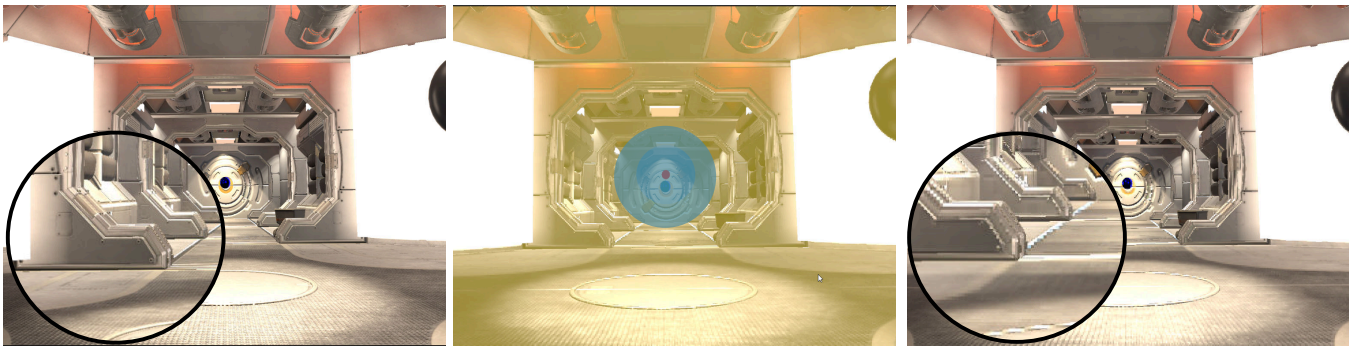

Figure 2: Example of our foveated method applied to our rich environment. Scene rendered at full quality (left), visualization of the foveated regions: blue - HQ; yellow - LQ (middle), 4x4 LQ degradation (right)

had to face in order to be completely re-aligned with the diagonal of our facility. The procedure was repeated 168 times.

4.3 Implied Movement

4.3.1 Procedure. The procedure for the *Passive* condition is very similar to the *Active* one. The participants were sat on a chair and asked to press a trigger to initiate the trial. Similarly to a vehicle in motion, they were moved with the same speed and acceleration as the fixation point. At the end of the movement, they were asked to record their answer by pressing *yes* or *no*.

4.4 Validation

All the participant data was checked post-experiment in order to validate if the participants have followed the instruction accordingly. In the AM condition, we calculated the percentage of frames in which the intended velocity was not reached. We have also used the eye-tracker provided by the Oculus Quest Pro in order to track the participants' gaze. We expect them to spend most of the time fixated on the blue sphere. Note that the eye-tracking hardware does not report data about blinking. The obtained results will be discussed in Section 5.

4.5 Psychometric Function Fitting

In order to analyze the data for each participant and for each of the four conditions, we first plotted the HQ region diameter values against the responses that the participant correctly identified as having a degraded resolution. A psychometric function (PF) represents a model of how physical stimulus parameters (in this case the HQ region diameter) are related to our measure of human perception (in this case the probability of failing to detect the stimulus degradation - the 'unawareness probability'). Subsequently, we recovered the diameter associated with a 75% unawareness probability as a threshold value where participants were unlikely to detect the change. At this point, they were 3 times more likely not to detect the degradation. We then conducted a 2 (type-of-movement) x 2 (degradation level) repeated measures Analysis of Variance (ANOVA) to assess the impact our manipulations had on the unawareness probability.

To fit the psychometric function we first plotted the HQ region diameter against the binary response data, with 1 coded as being unaware of the degradation and 0 coded as being aware of the degradation (Figure 3). We then assumed the psychometric function had a cumulative Gaussian form such that as the diameter increased the probability of being unaware (a 1 response) increased. The location and slope parameters of the PF were fitted using the *quickpsy* package [Linares and López-Moliner 2016] in R (version 4.2.3, [R Core Team 2023]), which is designed to undertake maximum likelihood psychometric function fitting to data of this form. Example data and fits for three observers are illustrated in Figure 3. From the fitted parameters we can recover the diameter that corresponds to a 75% unawareness probability. This threshold is taken to reflect the diameter at which participants tolerate the degradation (they are three times less likely to be aware of it than be aware of it). Lower thresholds would indicate that participants are willing to tolerate a greater area of degradation (i.e a smaller HQ and larger LQ).

5 RESULTS

Catch trials were removed from the data set prior to the analysis. Kolmogorov-Smirnov tests of normality were conducted and revealed all distributions to be normal. Figure 5 illustrates the mean thresholds recovered across our four conditions. Note that the 75% threshold diameter (y-axis) denotes the diameter of the HQ region for which participants are unlikely to notice the degradation. Smaller values suggest that degradation over a larger portion of the field is possible without people being aware. The first interesting feature of this data is that across all our conditions significant degradation could occur over a large portion of the field (outside a central range with a diameter between around 25 and 60 degrees, depending on the condition). The second feature is that participants are largely unaware of the degradation over a larger portion of the field when that degradation is less severe, i.e. in the 2x2 relative to 4x4 degradation conditions. This is not necessarily surprising but provides a nice manipulation check. The third feature is that participants are largely unaware of degradation over a larger portion of the field when they are actively moving relative to when the movement is implied. These conclusions were supported by a 2x2 repeated measures ANOVA which revealed a main effect of Degradation ($F(1,9) = 70.2$, $p < 0.001$) and a main effect of Type of Movement ($F(1,9) = 16.2$, $p = 0.003$). We also found evidence for an interaction between these factors ($F(1,9) = 5.4$, $p < 0.05$). Posthoc analyses of the

interaction were conducted using a Bonferroni-Holm correction for multiple comparisons (six in our case) we found statistical evidence for significant differences between all conditions except for the IM2x2 vs AM2x2 and IM2x2 vs AM4x4 pairs. This suggests that the interaction is driven by a considerably stronger effect of movement type when scene degradation is more severe and indicates that severe degradations are less detectable during active movement.

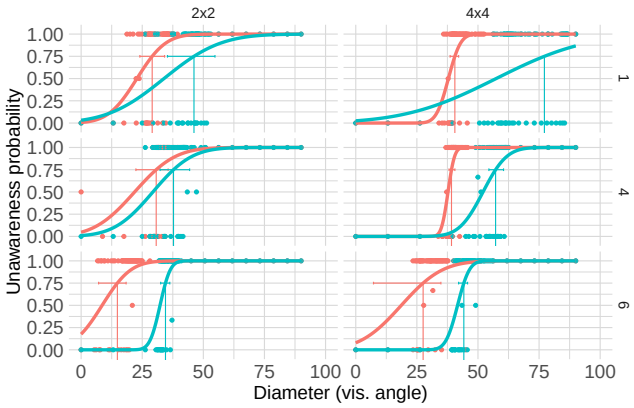

Figure 3: Example psychometric functions for three of the participants. In blue - the seated condition, in red - the active condition. The bars represent standard deviations.

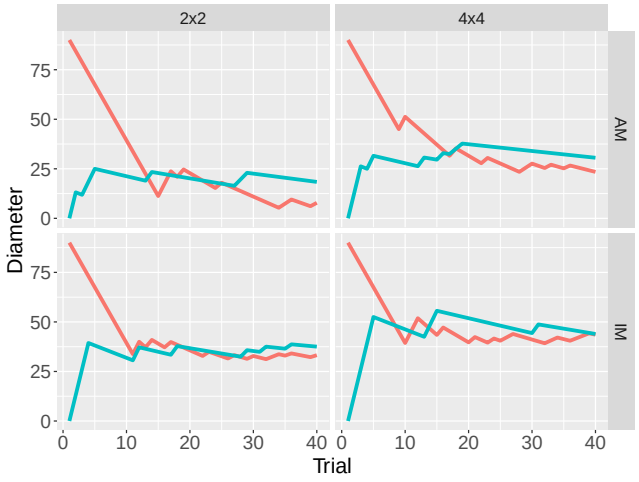

Figure 4: Example of the manipulation of the HQ diameter using the 2 interleaved staircases described in Section 3.2.

The data in Figure 5 suggest that in the 2x2 and 4x4 conditions respectively, participants are willing to accept an area of degradation approximately 23.5% and 27% bigger when actively moving. This finding is supported by the interaction reported above. The difference between the 2x2 vs 4x4 configuration is approximately 30% in the IM condition in the AM condition 36.4%.

Table 1: Summary statistics for our ANOVA analysis

Conditions	F stat	p-value
Degradation	70.156	$p < 0.001$
Type of Movement	16.171	$p = 0.003$
Deg * TypeMov	5.428	$p < 0.05$

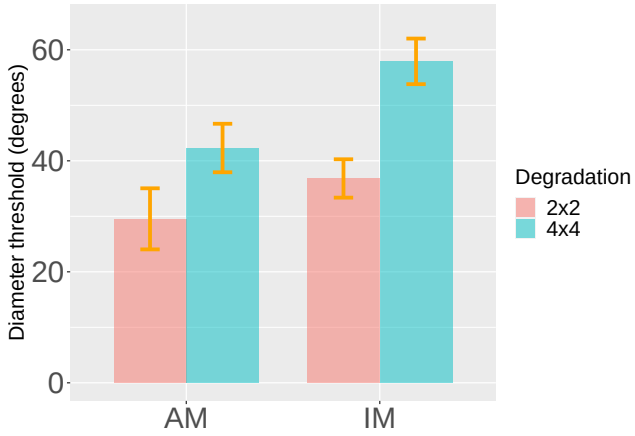

Figure 5: Mean bar plots representing the conditional means. The error bars represent standard error.

5.1 Validation Results

Participants reliably detected the catch trials 100% of the time. Data in the AM shows an almost perfect tracking speed (approximately 0% failure to reach 1 m/s across all ten participants). These data are irrelevant in the IM condition.

Eye-tracking data across all participants showed that on average, participants tracked the fixation sphere 81.5% of the time across all trials in the AM condition and 82% in the IM condition. We believe this is an acceptable range, given that the eye-tracker does not account for blinking. Note that the eye-tracking data for one of the participants only shows 11% success rate. These omissions could mean participants would be more likely to detect the change as they were looking directly at the degraded area. If this was the case, it would only amplify the effect of our findings.

6 DISCUSSION

Taken together, the results presented in Section 5 show that:

- The type of motion experienced (Active vs. Implied) has a significant effect on the perception of visual artifacts in a foveated rendering algorithm.
- Participants are significantly better at detecting differences in the 4x4 configuration than in the 2x2 configuration
- The type of motion has a significant impact on the extent to which we can degrade the scene. When actively moving, participants are less likely to detect more severe degradation.

Our method represents a potential approach to enhance foveated rendering algorithms. We show that during self-generated (surge)

user movements, we can significantly degrade large portions of the visual field without participants reliably detecting the degradation. Although our metrics can inform foveated rendering designs, it is important to mention that this work does not aim to produce a stand-alone foveated rendering system. However, with the rise in popularity of VR arcades and dynamic VR experiences, locomotion is becoming increasingly common and sought after. Critically, these experiences are usually computationally expensive due to the increasing demand for photorealistic graphics. Therefore, studying how different types of movement affect visual perception is important for future perceptual optimization methods. We suggest that in future work it would be useful to consider the extent to which different types of user movement (e.g. surge vs. heave vs. sway) permit similar optimizations.

In the VRS 4x4 configuration, spatiotemporal aliasing is present due to the increased scene motion during walking. [Albert et al. 2017] showed that in foveated algorithms, peripheral vision is more sensitive to flickering artifacts. There is evidence for significant interactions between the vestibular, visual, and other systems in the processing of motion [Holten and MacNeilage 2018; Warren et al. 2022]. In particular, recent work examining the interplay between different systems which signal scene relative movement in active vs passive movement conditions, hints at the complexity of this processing which can attenuate artifact detection. Moreover, we know that retinal motion is minimized in order to preserve depth information during surge [Liao et al. 2010]. This could create additional masking for spatiotemporal aliasing. We suggest that in the active movement condition, the additional processing required to integrate different sensory signals might make VRS degradations harder to detect.

[Ellis and Chalmers 2006] explored the cross-modal interaction of the visual-vestibular system. They produced an ego-motion map based on vestibular activation. Their method, however, did not use psychometric testing for determining what region of degradation the participants were sensitive to and they used a lower bound for their foveal HQ region. Using their dynamic ego-motion map method, the HQ pixels occupied 65% of the FOV. The values we found are lower which means that we are rendering fewer HQ pixels which shows our method provides a greater scope for optimization (for this HMD, 38.8% of the rendered FOV for 4x4 AM and 27% for 2x2 AM).

The effects that self-motion has on various forms of quality degradation in VR HMDs are not yet well-explored. Jindal et al. [Jindal et al. 2021] explored how information about horizontal velocity can be used to generate VRS State maps of a rich environment. They did not find a significant effect of speed in VR, but the scope of the study was to investigate Smooth Pursuit Eye Movement. It is important to mention that in their model, the participants looked at 2 LCD screens rendered in VR and the motion was generated by camera rotations rather than induced by the participants themselves. We show here that the type of movement matters in the perception of visual artifacts during VRS degradation.

Limitations: The method used here does have some limitations. First, we are using a fixed-foveated approach (thereby controlling eye movement). Moving to a fully-foveated approach is a goal for future work, but in this first experiment, we deliberately focused on the fixation case in order to minimize potential confounding effects arising from individual differences in eye movements. Note that, as mentioned above, the relatively few failures in fixation observed in our data could only have made detecting the degradation more likely (i.e. it would reduce the range measured in which degradation was undetected). If, instead, we had used a fully foveated approach, this would not be the case (since the degraded region would move with the eye) and so we anticipate that there would have been even less chance of detecting the degradation (i.e. we could have degraded an even larger portion of the field). Second, we used only two layers of degradation whereas most foveated rendering models have at least three layers of foveation. While we are interested in how the effects extend to other multiple layers, the amount of additional data required from each participant would have been prohibitive in this first study. Third, our degradation was activated when the user reaches a certain velocity. Studying how surge motion can affect popping artifacts caused by the sudden activation of a foveated rendering method could be useful to predict an optimal relative velocity for activating the degradation. Moreover, the popping effect caused by the sudden degradation could be used as a cue for when the scene is being degraded. When asked after the trial, participants reported that they generally did not notice the popping which could further build on the theory that the additional processing the brain is undertaking could mask VRS artifacts.

Future Work. Our results show promising results and create an interesting avenue for future research. Due to the standalone VR HMDs becoming state-of-art, environments that users can walk through, will become increasingly more common. This calls for the investigation of similar effects at different velocities in a wider variety of environments. Using a fully-foveated method together with our user velocity-based approach could also yield significant performance benefits. The effect type of movement has on such an implementation is not well understood. Furthermore, dynamically adjusting the area of degradation similar to the method in [Ellis and Chalmers 2006] could also benefit our implementation. Metrics recovered about how ego-movement can affect spatiotemporal sensitivity such as ours can be used in the future in order to enhance already existing models such as [Denes et al. 2020; Mantiuk et al. 2022; Tursun et al. 2019].

7 CONCLUSION

In this paper, we examined if the type of movement undertaken in VR can have an impact on the probability of the detection of the increased diameter of an under-sampled peripheral region in a foveated rendering method. We created a foveated algorithm using Nvidia's VRS technology in the 2x2 and 4x4 configurations. We show that there is a clear difference in detection probability between implied movement through an environment and active movement and that there is more scope for degradation in the walking condition. Active movement yields greater benefits and this is particularly so when more severe degradation is present. Due to 6-DOF tracking in VR HMDs becoming readily available, we hope these findings could provide a promising future avenue for research.

ACKNOWLEDGMENTS

This research has been funded and supported by UKRI.

REFERENCES

Rachel Albert, Anjul Patney, David Luebke, and Joohwan Kim. 2017. Latency Requirements for Foveated Rendering in Virtual Reality. *ACM Transactions on Applied Perception* 14, 4 (Sept. 2017), 25:1–25:13. https://doi.org/10.1145/3127589

Ryan Beams, Brendan Collins, Andrea S. Kim, and Aldo Badano. 2020. Angular Dependence of the Spatial Resolution in Virtual Reality Displays. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 836–841. https://doi.org/10.1109/VR46266.2020.00108 ISSN: 2642-5254.

Doris I. Braun, Alexander C. Schütz, and Karl R. Gegenfurtner. 2017. Visual sensitivity for luminance and chromatic stimuli during the execution of smooth pursuit and saccadic eye movements. *Vision Research* 136 (July 2017), 57–69. https://doi.org/10.1016/j.visres.2017.05.008

Jack Brookes, Matthew Warburton, Mshari Alghadier, Mark Mon-Williams, and Faisal Mushtaq. 2020. Studying human behavior with virtual reality: The Unity Experiment Framework. *Behavior Research Methods* 52, 2 (April 2020), 455–463. https://doi.org/10.3758/s13428-019-01242-0

Alexandre Chapiro, Robin Atkins, and Scott Daly. 2019. A Luminance-aware Model of Judder Perception. *ACM Transactions on Graphics* 38, 5 (July 2019), 142:1–142:10. https://doi.org/10.1145/3338696

Christine A. Curcio and Kimberly A. Allen. 1990. Topography of ganglion cells in human retina. *Journal of Comparative Neurology* 300, 1 (1990), 5–25. https://doi.org/10.1002/cne.903000103 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cne.903000103.

Gyorgy Denes, Akshay Jindal, Aliaksei Mikhailiuk, and Rafał K. Mantiuk. 2020. A perceptual model of motion quality for rendering with adaptive refresh-rate and resolution. *ACM Transactions on Graphics* 39, 4 (July 2020). https://doi.org/10.1145/3386569.3392411

Michael Dorr, Martin Böhme, Thomas Martinetz, and Erhardt Barth. 2006. Gaze-Contingent Spatio-temporal Filtering in a Head-Mounted Display. In *Perception and Interactive Technologies (Lecture Notes in Computer Science)*, Elisabeth André, Laila Dybkjær, Wolfgang Minker, Heiko Neumann, and Michael Weber (Eds.). Springer, Berlin, Heidelberg, 205–207. https://doi.org/10.1007/11768029_24

David E. Jacobs, Orazio Gallo, Emily A. Cooper, Kari Pulli, and Marc Levoy. 2015. Simulating the Visual Experience of Very Bright and Very Dark Scenes. *ACM Transactions on Graphics* 34, 3 (May 2015), 25:1–25:15. https://doi.org/10.1145/2714573

Gavin Ellis and Alan Chalmers. 2006. The effect of translational ego-motion on the perception of high fidelity animations. In *Proceedings of the 22nd Spring Conference on Computer Graphics (SCCG '06)*. Association for Computing Machinery, New York, NY, USA, 75–82 PAGE@7. https://doi.org/10.1145/2602161.2602170

Brian Guenter, Mark Finch, Steven Drucker, Desney Tan, and John Snyder. 2012. Foveated 3D graphics. *ACM Transactions on Graphics* 31, 6 (Nov. 2012), 164:1–164:10. https://doi.org/10.1145/2366145.2366183

Vivian Holten and Paul R. MacNeilage. 2018. Optic flow detection is not influenced by visual-vestibular congruency. *PLoS ONE* 13, 1 (Jan. 2018), e0191693. https://doi.org/10.1371/journal.pone.0191693

Akshay Jindal, Krzysztof Wolski, Karol Myszkowski, and Rafał K. Mantiuk. 2021. Perceptual model for adaptive local shading and refresh rate. *ACM Transactions on Graphics* 40, 6 (Dec. 2021), 281:1–281:18. https://doi.org/10.1145/3478513.3480514

Ke Liao, Mark F. Walker, Anand C. Joshi, Millard Reschke, Michael Strupp, Judith Wagner, and R. John Leigh. 2010. The linear vestibulo-ocular reflex, locomotion and falls in neurological disorders. *Restorative Neurology and Neuroscience* 28, 1 (Jan. 2010), 91–103. https://doi.org/10.3233/RNN-2010-0507 Publisher: IOS Press.

Daniel Linares and Joan López-Moliner. 2016. quickpsy: An R Package to Fit Psychometric Functions for Multiple Groups. *The R Journal* 8, 1 (2016), 122. https://doi.org/10.32614/RJ-2016-008

Rafał K. Mantiuk, Maliha Ashraf, and Alexandre Chapiro. 2022. stelaCSF: a unified model of contrast sensitivity as the function of spatio-temporal frequency, eccentricity, luminance and area. *ACM Transactions on Graphics* 41, 4 (July 2022), 1–16. https://doi.org/10.1145/3528223.3530115

Brian J. Murphy. 1978. Pattern thresholds for moving and stationary gratings during smooth eye movement. *Vision Research* 18, 5 (Jan. 1978), 521–530. https://doi.org/10.1016/0042-6989(78)90196-7

NVIDIA. 2018. VRWorks - Variable Rate Shading (VRS). https://developer.nvidia.com/vrworks/graphics/variablerateshading

Derrick Parkhurst and Ernst Niebur. 2004. A feasibility test for perceptually adaptive level of detail rendering on desktop systems. In *Proceedings of the 1st Symposium on Applied perception in graphics and visualization (APGV '04)*. Association for Computing Machinery, New York, NY, USA, 49–56. https://doi.org/10.1145/1012551.1012561

Anjul Patney, Marco Salvi, Joohwan Kim, Anton Kaplanyan, Chris Wyman, Nir Benty, David Luebke, and Aaron Lefohn. 2016. Towards foveated rendering for gaze-tracked virtual reality. *ACM Transactions on Graphics* 35, 6 (Dec. 2016), 179:1–179:12. https://doi.org/10.1145/2980179.2980246

Théo Perrin, Hugo A. Kerhervé, Charles Faure, Anthony Sorel, Benoit Bideau, and Richard Kulpa. 2019. Enactive Approach to Assess Perceived Speed Error during Walking and Running in Virtual Reality. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 622–629. https://doi.org/10.1109/VR.2019.8798209 ISSN: 2642-5254.

Jeffrey S. Perry and Wilson S. Geisler. 2002. Gaze-contingent real-time simulation of arbitrary visual fields. In *Human Vision and Electronic Imaging VII*, Vol. 4662. SPIE, 57–69. https://doi.org/10.1117/12.469554

David Petrescu, Paul A. Warren, Zahra Montazeri, and Steve Pettifer. 2023. Velocity-Based LOD Reduction in Virtual Reality: A Psychophysical Approach. In *Eurographics 2023 - Short Papers*, Vahid Babaei and Melina Skouras (Eds.). The Eurographics Association. https://doi.org/10.2312/egs.20231010

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Martin Reddy. 2001. Perceptually optimized 3D graphics. *IEEE Computer Graphics and Applications* 21, 5 (July 2001), 68–75. https://doi.org/10.1109/38.946633 Conference Name: IEEE Computer Graphics and Applications.

Josef Spjut, Ben Boudaoud, Jonghyun Kim, Trey Greer, Rachel Albert, Michael Stengel, Kaan Akşit, and David Luebke. 2020. Toward Standardized Classification of Foveated Displays. *IEEE Transactions on Visualization and Computer Graphics* 26, 5 (May 2020), 2126–2134. https://doi.org/10.1109/TVCG.2020.2973053 Conference Name: IEEE Transactions on Visualization and Computer Graphics.

Jordan W. Suchow and George A. Alvarez. 2011a. Background motion silences awareness of foreground change. In *ACM SIGGRAPH 2011 Posters (SIGGRAPH '11)*. Association for Computing Machinery, New York, NY, USA, 1. https://doi.org/10.1145/2037715.2037750

Jordan W. Suchow and George A. Alvarez. 2011b. Motion Silences Awareness of Visual Change. *Current Biology* 21, 2 (Jan. 2011), 140–143. https://doi.org/10.1016/j.cub.2010.12.019

Bernhard Treutwein. 1995. Adaptive psychophysical procedures. *Vision Research* 35, 17 (Sept. 1995), 2503–2522. https://doi.org/10.1016/0042-6989(95)00016-X

Okan Tarhan Tursun, Elena Arabadzhiyska-Koleva, Marek Wernikowski, Radosław Mantiuk, Hans-Peter Seidel, Karol Myszkowski, and Piotr Didyk. 2019. Luminance-contrast-aware foveated rendering. *ACM Transactions on Graphics* 38, 4 (July 2019), 98:1–98:14. https://doi.org/10.1145/3306346.3322985

Karthik Vaidyanathan, Marco Salvi, Rpbert Toth, Tim Foley, Tomas Akenine-Möller, Jim Nilsson, Hacob. Munkberg, Jon Hasselgren, Masamichi Sugihara, Petrick Clarberg, Tomasz Janczak, and Aaron Lefohn. 2014. Coarse pixel shading. *High-Performance Graphics 2014, HPG 2014 - Proceedings* (Jan. 2014), 9–18.

HTC ViveSoftware. 2020. Vive Foveated Rendering - Developer Resources. https://developer.vive.com/resources/vive-sense/tools/vive-foveated-rendering/

Lili Wang, Xuehuai Shi, and Yi Liu. 2023. Foveated rendering: A state-of-the-art survey. *Computational Visual Media* 9, 2 (June 2023), 195–228. https://doi.org/10.1007/s41095-022-0306-4

Paul A. Warren, Graham Bell, and Yu Li. 2022. Investigating distortions in perceptual stability during different self-movements using virtual reality. *Perception* (Aug. 2022), 03010066221116480. https://doi.org/10.1177/03010066221116480 Publisher: SAGE Publications Ltd STM.

Martin Weier, Thorsten Roth, Ernst Kruijff, André Hinkenjann, Arsène Pérard-Gayot, Philipp Slusallek, and Yongmin Li. 2016. Foveated Real-Time Ray Tracing for Head-Mounted Displays. *Computer Graphics Forum* 35 (Oct. 2016), 289–298. https://doi.org/10.1111/cgf.13026

Martin Weier, M Stengel, Thorsten Roth, Piotr Didyk, Elmar Eisemann, M Eisemann, Steve Grogorick, André Hinkenjann, Ernst Kruijff, M Magnor, Karol Myszkowski, and Philipp Slusallek. 2017. Perception-driven Accelerated Rendering. *Computer Graphics Forum (Proc. of Eurographics)* 36 (April 2017). https://doi.org/10.1111/cgf.13150