Latency Requirements for Foveated Rendering in Virtual Reality

RACHEL ALBERT, UC Berkeley and NVIDIA Research
ANJUL PATNEY, DAVID LUEBKE, and JOOHWAN KIM, NVIDIA Research

Foveated rendering is a performance optimization based on the well-known degradation of peripheral visual acuity. It reduces computational costs by showing a high-quality image in the user's central (foveal) vision and a lower quality image in the periphery. Foveated rendering is a promising optimization for Virtual Reality (VR) graphics, and generally requires accurate and low-latency eye tracking to ensure correctness even when a user makes large, fast eye movements such as saccades. However, due to the phenomenon of saccadic omission, it is possible that these requirements may be relaxed.

In this article, we explore the effect of latency for foveated rendering in VR applications. We evaluated the detectability of visual artifacts for three techniques capable of generating foveated images and for three different radii of the high-quality foveal region. Our results show that larger foveal regions allow for more aggressive foveation, but this effect is more pronounced for temporally stable foveation techniques. Added eye tracking latency of 80–150ms causes a significant reduction in acceptable amount of foveation, but a similar decrease in acceptable foveation was not found for shorter eye-tracking latencies of 20–40ms, suggesting that a total system latency of 50–70ms could be tolerated.

CCS Concepts: • **Computing methodologies** → **Perception**; **Virtual reality**;

Additional Key Words and Phrases: Foveated rendering, eye-tracking, latency

1 INTRODUCTION

In recent years, the display and rendering requirements for Virtual Reality (VR) have continued to rise toward human perception of the real world, but there is still substantial room for improvement. At the time of this writing, the best VR displays have a resolution of 1080×1200 pixels per eye [16]. In terms of visual angle on the retina this is equivalent to 5 arc-minutes per pixel, whereas the human visual system can resolve gratings as fine as 0.5 arc-minute per pixel [43]. The Field of View (FOV) of current displays is limited to 100° horizontally, which is only half of the binocular horizontal FOV of the human visual system [16, 40]. Many VR applications also require stereo depth information, doubling rendering costs, and perceptual research in flicker sensitivity demonstrated that the minimum frame rate requirement for the far periphery is at least 85Hz [46]. While VR display technology has been rapidly advancing to fill this gap, such hardware improvements represent a
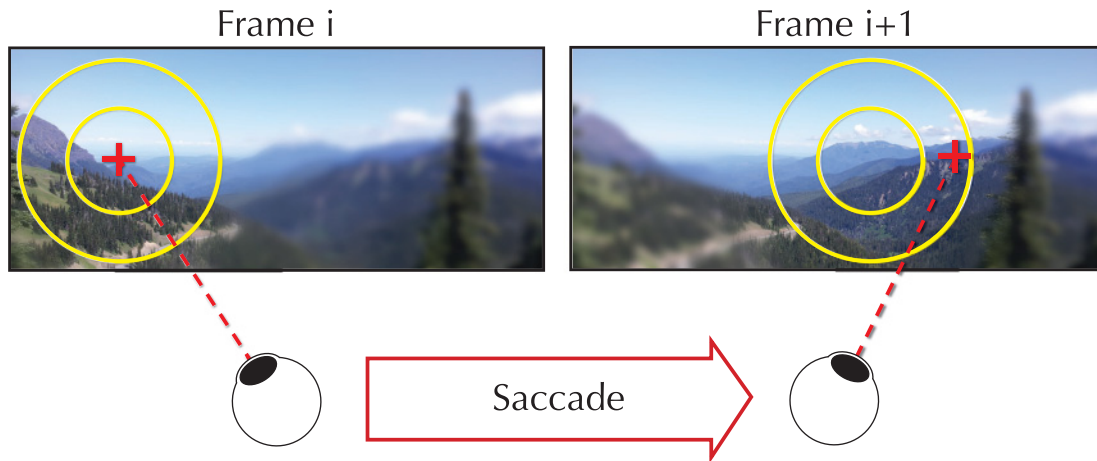
Fig. 1. Eye-tracked foveated rendering is a key performance optimization for future mixed-reality and eye-tracked desktop displays. It is important for such eye trackers to provide low-latency data, as any lag between large eye movements (saccades) and display updates can limit the extent of foveation and corresponding performance gain. In this article, we study the limits of eye-tracking latency on the perceived quality of three foveated rendering techniques.

significant increase in the number of rendered pixels. Furthermore, the rendering costs per pixel have also continued to rise with increasing demand for more realistic computer-generated imagery. For example, more accurate object occlusion boundaries rely on higher polygon counts, dynamic and complex lighting environments necessitate more samples per pixel, and sophisticated surface reflectance behavior is needed for realism in many applications such as subsurface scattering for human skin [49, 51].

However, many of these high-quality pixels are squandered due to the low fidelity of the Human Visual System (HVS) in the periphery. The HVS has a small 1.5° region of high visual acuity called the fovea. Outside the fovea the distribution of retinal components changes rapidly, resulting in decreased visual acuity [43], less sensitivity to color [2, 14], and limited stereoscopic depth discrimination [6, 41], as well as increased sensitivity to flicker [46]. By 20° of eccentricity the HVS can no longer resolve gratings narrower than 7.5 arc-minutes per pixel [50]. However, in wide FOV displays such as VR, 96% of pixels are outside this 20° boundary [31].

In this context, the technique of foveated rendering seems vital to meet the growing computational demands of VR environments. Foveated rendering allows the peripheral image to be degraded in a way that decreases computation and is perceptually unobtrusive. This requires gaze tracking, and therefore foveated systems are very sensitive to latency. There are several parameters that can influence the latency requirements of a foveated system—the size of the full resolution foveal image that follows the gaze, the degree of degradation applied to the image, and the type of degradation method used. Although many previous works have mentioned the importance of latency in foveated rendering, to our knowledge there has been no comprehensive study examining the relationship between these parameters for VR. Figure 1 illustrates how increased latency has the potential to increase the visibility of foveation artifacts. In this article, we measure the amount of foveation that users can tolerate in conjunction with various levels of eye-tracking latency, fovea size, and foveation technique. We show for the first time that there is no significant decrease in acceptable foveation level for total system latencies up to 50–70ms in VR.

2 RELATED WORK

The concept of gaze-contingent imagery was first introduced by Stephen Reder in 1973 [33] and was later applied to ray tracing by Levoy and Whitaker in 1990 [19]. Since then there have been many variants of foveated

rendering [11, 13, 26, 29, 53, 54], as well as several reviews of gaze-dependent and perceptually guided rendering in general [12, 21, 34].

More recent work has focused on practical foveated rendering using modern Head-Mounted Displays (HMDs) with retrofitted eye trackers [48]. Such practical implementations often benefit from recent GPU features that enable efficiently varying rendering detail across an image [27, 28]. Some recent work has also explored foveated rendering with an emphasis on perceptually guided foveation [30, 32, 42]. While our exploration of the impact of latency is largely orthogonal to these advances, different techniques may differ in their tolerance to increasing latency, so in this article we independently study three representative methods of foveation, described in Section 3.1 below.

2.1 Latency

There are many contributing factors in determining system latency as well as its perceptual impact, for instance [18, 22]. In this article, we focus on eye-tracking latency in VR, and its influence on the perception of foveated rendering. A number of the works listed above have noted the impact of latency on the effectiveness of foveated rendering, but the reported latencies used have varied significantly. For desktop experiments, Levoy and Whitaker reported a best and worst case system latency of 100–150ms [19], while Guenter et al. reported 23–40ms [13]. Thunström found that a system latency of 42ms was acceptable for 95% of subjects [44]. In head-mounted systems, Stengel et al. reported approximately 53–91ms, and although Patney et al. did not report a complete system latency, the latency of their head-mounted experiment was at least 20–37ms plus render time [32].

There are several techniques previously described for measuring various parts of display and eye-tracking system latency, including using a photodiode to characterize monitor refresh [10, 52], using an oscilloscope [1, 45], using a Light-Emitting Diode (LED) to simulate a saccade in the eye tracker [5, 7], and combining these techniques to measure end-to-end system latency [38]. It has also been a standard practice to use these techniques to estimate and sum the latency of each component of the system [15, 20, 24, 37], which is the method we employ in our study.

2.2 Saccadic Omission

Our expectation of an acceptable latency period is grounded in a large body of research on saccadic omission (also known as saccadic suppression and saccadic blindness). Saccadic eye movements occur when an observer changes their gaze position rapidly from one location to another. These movements follow a stereotypical ballistic motion as described by Bahill [3]. During a short period before, during, and after the saccadic eye movement, saccadic omission prevents some visual input from being processed. In addition to the visual masking effect of the pre- and post-saccadic images, specific signals in the magnocellular neural pathway are also suppressed before and during the saccade, and the neural signal of the post-saccadic image is enhanced [8, 9, 17, 23]. Omission typically begins about 50ms prior to the start of a saccade and sensitivity returns to normal about 100ms after the saccade begins, although the exact time course is dependent on many factors including background luminance and saccade amplitude [25, 36]. A similar mechanism has also been proposed for suppression of visual input during blinks [35]. Since the artifacts introduced in foveated rendering are obvious under direct gaze, we can infer that any unnoticeable image updates after a saccade or blink may be due to perceptual omission. Longer latencies for image update may exceed the omission period, causing foveation artifacts to become visible and therefore lowering the acceptable level of foveation.

3 EXPERIMENTS

We conducted two experiments to determine the relationship between foveation technique, fovea size, eye-tracking latency, and amount of foveation that can be tolerated. The first experiment was conducted using a HMD and the second using a desktop display with head tracking. Our primary goal was to determine practical

latency and foveation guidelines for current state-of-the-art HMD systems. However, we also recognize that HMD hardware is rapidly progressing and therefore included a small parallel experiment on a desktop display with higher resolution and lower system latency. Naturally the desktop stimulus does not provide the same wide FOV as an HMD, but we hope that the second experiment can still provide a window into possible perceptual limitations regarding latency for future HMD hardware.

Both experiments were run on a PC with an Intel i7 3.5GHz CPU, 32GB memory, and an Nvidia TITAN X (Pascal) GPU. Our rendering sandbox was implemented using the Falcor rendering framework [4].

3.1 Foveation Technique

We evaluated the impact of latency for three different methods of visual foveation:

Subsampling. Reducing image resolution with increasing retinal eccentricity, bilinearly upsampled using the texture unit. This method is based on the one proposed by Guenter et al. [13].

Gaussian Blur. Gaussian image-space blur with blur radius increasing with increasing retinal eccentricity.

fCPS. The method described by Patney et al. [32], which employs foveated coarse-pixel shading with prefiltered shading, foveated temporal anti-aliasing, and post-processed contrast enhancement.

All of our experiments used the classroom scene in Figure 2, which shows the same viewpoint without foveation and with foveation applied for each of the above methods (for more information about the spatial frequency characteristics of the scene itself, please see the Appendix). The subsampling technique may produce flicker and jagged edges, and therefore serves as a minimum benchmark for less temporally stable methods. Gaussian blur is a maximally stable upper benchmark, although it does not provide any computational savings and reduces contrast. Finally, the fCPS technique provides an evaluation of recent research in the area of foveated rendering. The visual artifacts induced by the fCPS technique are temporally stable jagged edges, ghosting, and contrast distortion.

Figure 3 shows our three foveation density mappings where the visual field is divided into three regions based on retinal eccentricity:

—**Far Periphery:** A uniformly low-density region beyond a fixed eccentricity of either 5°, 10°, or 20°.
—**Fovea:** A uniformly high-density region from the gaze center to two-thirds of the corresponding far-peripheral eccentricity, i.e., 3.33°, 6.67°, and 13.33°, respectively.
—**Transition Region:** A transition region with linear density falloff between the fovea and far periphery, one-third of the far-peripheral eccentricity in size.

For any given eccentricity, our foveation density mapping returns a blur radius between 0 (foveal) and a maximum (peripheral) chosen differently for each trial. We use the blur radius directly as the standard deviation for the Gaussian Blur method. For the other methods, we compute the subsampling factor as twice the blur radius. Thus, we interpret a blur radius of 1 as a 2×2 resolution downscale for subsampling, and a 2×2 coarse shading rate for fCPS. Note that following the original proposal, we have limited the maximum blur radius for fCPS to 2, corresponding to a 4×4 coarse shading rate.

3.2 HMD Experimental Design

We used an HTC Vive HMD equipped with a factory-calibrated SMI eye tracker. The system latency for our setup is illustrated in Figure 4. The SMI eye tracker had a minimum latency of approximately 6ms and a sampling rate of 250Hz, asynchronous to the rest of the system. To remove any effect of variable render time we chose to synchronize eye-tracker sampling time across all techniques. Our HMD has a framerate of 90Hz, so we designed our stimulus such that render time was less than approximately 11.11ms for any technique, and could therefore be synchronized to VSYNC. We used OpenVR "running start" to synchronize our frame submission to approximately
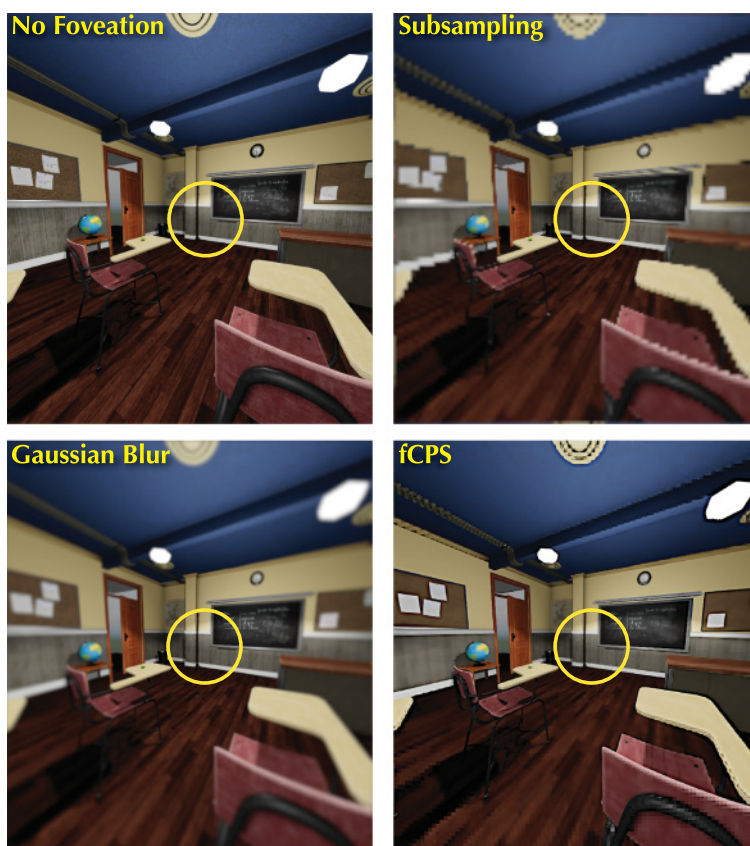
Fig. 2. The scene used in our experiments, original and the three foveation techniques used. We have exaggerated the magnitude of foveation for illustration. Center of gaze is shown in yellow.
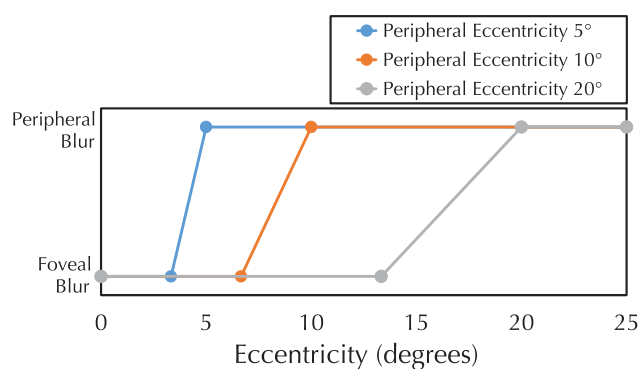

Fig. 3. We use the above foveation mappings for our user studies. For each peripheral blur (5°, 10°, and 20°), we linearly increase the foveation blur radius from 0 to the maximum in a transition region whose width is 33% of the far-peripheral eccentricity. We vary the maximum/peripheral blur in our experiment to compute perceptual thresholds.

HMD Experiment

Total Latency: 31-35ms

| Eye-tracker latency (6-10ms) | Rendering time (up to 11.11ms) | Submit to photons latency (3 + 11.11ms) |

VSYNC Events (90 Hz)

Desktop Experiment

Total Latency: 22-25 ms

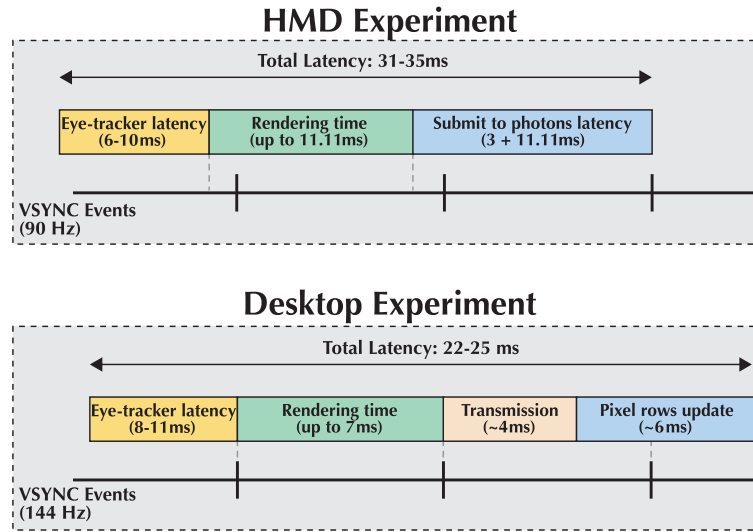| Eye-tracker latency (8-11ms) | Rendering time (up to 7ms) | Transmission (~4ms) | Pixel rows update (~6ms) |

VSYNC Events (144 Hz)

Fig. 4. A diagram estimating the system latency for the HMD experiment (top) and desktop experiment (bottom). Total system latency is calculated from time of most recent eye position sample until the corresponding image is displayed on screen. For HMD rendering, we use running start which synchronizes submissions to approximately 3ms before VSYNC. Desktop latency is calculated relative to pixel update at the midpoint of the screen.

3ms before VSYNC [49]. Synchronized transmission and pixel flash on the display was approximately 11.11ms. Thus, except in rare cases, each submitted frame was displayed approximately 14.11ms after submission. As a result, the end-to-end system latency, calculated from the time of eye image capture to displaying the next frame, was usually between 31 and 35ms, except in the case of a dropped frame where it was approximately 11.11ms longer. Only 11 out of 8,851 trials had more than 1% of dropped frames and were therefore excluded. For our experiments an additional latency was added only to the eye-tracking sample by using an older sample of the eye position at render time.

Nine subjects participated in the HMD experiment for a total of 3 hours per subject. Each subject performed the standard SMI eye-tracking calibration at the start of the experiment and the calibration error was within the typical range for all subjects. Participants viewed a classroom scene from a standard viewpoint shown in Figure 2. A total of 55 randomly interleaved foveation conditions were presented, including 15 no-foveation control trials and every combination of foveation technique (subsampling, Gaussian blur, fCPS), fovea size eccentricity (5°,10°,20°), and added tracker latency (0,10,20,40,80,150ms). For each experimental condition the blur radius of the far periphery was varied using a one-up/one-down adaptive staircase procedure to determine the 50% detection threshold. The stimulus duration for each trial was 4s, and participants were instructed to use head and eye motion as needed to observe any unusual quality artifacts, including flicker, blur, tunnel vision, or motion artifacts. At the end of each stimulus presentation subjects responded "yes" if they observed any artifacts or "no" if the scene appeared normal. To disambiguate from visual artifacts caused by the display itself, subjects were provided with a 2 minute period to observe the scene with no foveation at the start of the experiment, as well as 10 minutes of practice with various techniques and levels of foveation applied.

3.3 Desktop Experimental Design

Our desktop display was an Acer XB270HU with a refresh rate of 144Hz. We used a Tobii TX300 eye and head tracker with a minimum latency of 8ms and a sampling rate of 300Hz. Head tracking allowed us to test more complex subject motion in a way that is more comparable to an HMD. As in the case of the first experiment,

Table 1. Actual End-to-end System Latency Values for Each Added
Latency Condition in the HMD Experiment

Actual HMD System Latency Per Added Latency Condition		
Added latency	Proportion excluded	Mean (SD) latency (ms)
0ms	0.0380	34.0 (1.5)
10ms	0.0387	38.0 (1.5)
20ms	0.0403	47.9 (1.5)
40ms	0.0389	68.0 (1.5)
80ms	0.0408	108.0 (1.5)
150ms	0.0450	178.0 (1.5)

Column 2 shows the proportion of dropped frames that were excluded from the
calculation of mean and standard deviation shown in column 3.

we synchronized eye-tracker sampling time across all techniques to remove any effect of variable render time (see Figure 4). Render time was again 7ms or less and synchronized with the monitor VSYNC. We measured the time from monitor VSYNC to change in pixel luminance on the display using a photometer connected to the display computer. Pixel display time was consistently 4ms for the top of the monitor and 10ms for the bottom of the monitor. Since we do not know the exact gaze position at the time of pixel update, we report an end-to-end system latency for the midpoint of the display. The minimum system latency was usually between 22 and 25ms except in the case of a dropped frame where it was 7ms later.

The experimental design was identical to the HMD experiment; however, several parameters were modified to accommodate the desktop hardware. To provide better synchronization with the longer minimum eye-tracker latency, added latency values were 8, 15, 22, 40, 80, and 150ms. Fovea size eccentricities were also reduced to 3°, 5°, and 10° to fit within the size of the monitor. For the fCPS technique, temporal anti-aliasing level of detail was set at 1 to reduce motion blur artifacts. Stimulus duration was also decreased to 2.5 seconds. Two subjects with extensive psychophysical experience participated in the desktop experiment for a total of 2 hours per subject. Each subject followed the standard Tobii calibration procedure at the start of the experiment and the calibration error was within the typical range.

4 RESULTS

4.1 Actual Latency

We calculated the actual latency of stimulus presentation relative to time of eye-image capture for each frame and grouped frames by intended added latency. Latency statistics for the HMD experiment are shown in Table 1. The distribution of latencies was normal with a heavy tail of very large latencies in the case of dropped frames. We therefore excluded dropped frames from calculations of actual latency, but report the percentage of excluded frames for each condition. Dropped frames were defined to be latencies greater than three times the scaled median absolute deviation relative to the median. Results for both experiments in Figure 5 are plotted relative to the actual mean system latency for each latency condition.

4.2 HMD Experiment

For each subject in each condition (combination of technique, eccentricity, and added latency), we obtained the blur radius corresponding to a 50% artifact detection threshold. Thresholds were calculated by fitting a Gaussian psychometric function using the *psignifit* toolbox version 3 for MATLAB which implements the maximum-likelihood method described by Schütt et al. [39]. Figure 5(a) shows the mean across subjects for each condition grouped by technique, and Figure 5(b) shows the same data grouped by eccentricity. Error bars represent 95% confidence intervals of the mean.
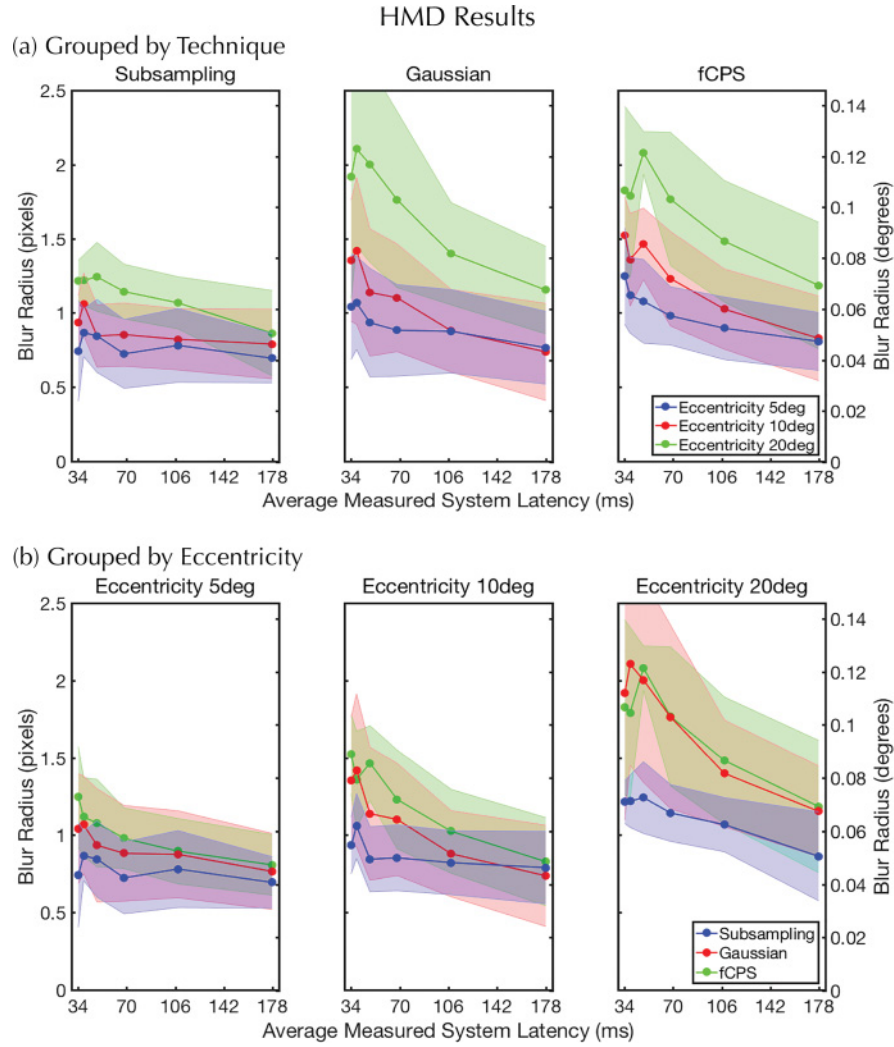
Fig. 5. HMD experiment results showing average peripheral blur radius threshold corresponding to 50% artifact detection across all subjects (a) grouped by foveation technique, and (b) grouped by eccentricity. Error bars show 95% confidence interval of the mean.

Using the fitted thresholds, a three-way ANOVA was conducted to compare the effects of technique, eccentricity, and added latency on blur radius threshold. All three main effects were significant at the 0.05 significance level: for technique $F_{(2,432)} = 34.06$, $p \approx 0$ for eccentricity $F_{(2,432)} = 77.33$, $p \approx 0$ and for latency $F_{(5,432)} = 13.67$, $p \approx 0$. The interaction between technique and eccentricity was also significant with $F_{(4,432)} = 4.05$, $p = 0.0031$. No significant interaction was found between technique and latency or between latency and eccentricity.

Follow-up analysis using unpaired t-tests with Bonferroni correction for multiple comparisons was conducted for both the technique-eccentricity interaction and the main effect of added latency. Within each technique separately, no interaction effects were apparent for eccentricity and all three techniques showed a significant
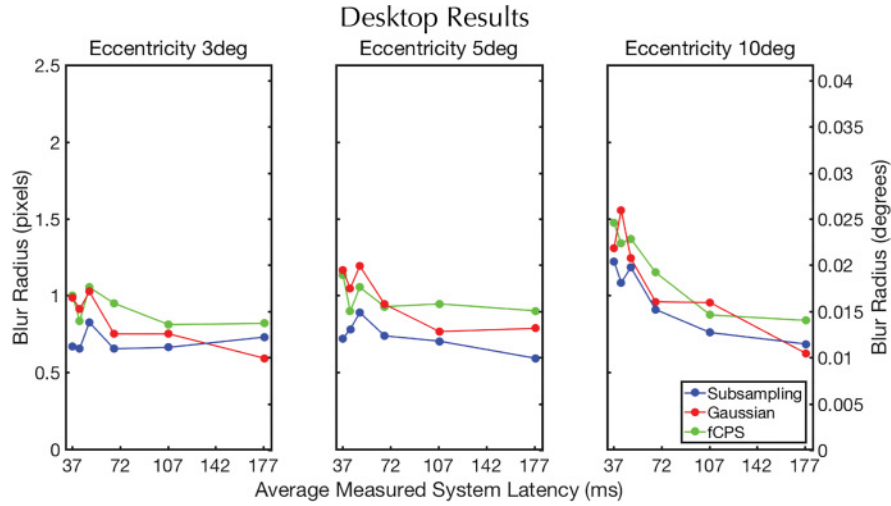
Fig. 6. Desktop experiment results showing average peripheral blur radius threshold corresponding to 50% artifact detection across both subjects grouped by eccentricity. Error bars not shown due to small sample size.

difference between 5° and 20° of eccentricity (Subsampling: p = 0.0025, Gaussian Blur: p ≈ 0, fCPS: $p \approx 0$). However, comparisons across techniques differed with eccentricity. At 5° of eccentricity, no significant difference was found between any pair of techniques. At 10° of eccentricity, a significant difference was found between subsampling and fCPS only (p = 0.0021). At 20° of eccentricity, a significant difference was found between Subsampling and both Gaussian Blur ($p \approx 0$) and fCPS ($p \approx 0$). No significant difference was found between Gaussian Blur and fCPS for any eccentricity.

For the main effect of added latency pooled across all other conditions, no significant difference was found between the 0, 10, 20, and 40ms added latency conditions. For the 80ms condition, a significant difference was found between each of the following conditions: 0ms (p = 0.0012), 10ms (p = 0.0003), and 20ms (p = 0.0029). Similarly, for the 150ms condition, a significant difference was found between each of the following conditions: 0ms ($p \approx 0$), 10ms ($p \approx 0$), 20ms ($p \approx 0$), and 40ms (p = 0.0007).

4.3 Desktop Experiment

Average per condition thresholds were calculated in the same manner as in the HMD experiment. An identical three-way ANOVA was conducted, but no significant main effects or interactions were found. Figure 6 shows the average results across both subjects grouped by eccentricity. The data show similar trends overall to the HMD experiment results.

4.4 Head and Eye Motion

An analysis of head and eye motion was also performed to rule out the possibility that increased blur radius thresholds could be attributed to decreased exploration of visual artifacts. Four types of subject motion were considered: head translation, head rotation, frequency of saccades, and saccade amplitudes. Subject motion data was quite noisy, so we discarded physically implausible values beyond a speed of 1m/s for head translation, 200°/s for head rotation, and saccades with amplitude greater than 500°/s. Distributions for all four types of subject motion were positively skewed, so a Mann-Whitney-U test was used to check for significantly different distributions. No significant difference was found between any pair amongst all 54 conditions (i.e., all unique

combinations of technique, eccentricity, and latency). We are therefore confident that differences in exploration of artifacts could not have influenced foveation thresholds.

5 DISCUSSION AND FUTURE WORK

Our results confirm previous findings that more foveation is possible in the periphery, especially for certain techniques. The fCPS technique supported significantly more foveation than the Subsampling method even at 10° eccentricity. It is interesting to note that Gaussian Blur did not show the same effect at 10°, suggesting that the contrast-preserving feature of the fCPS technique provides more opportunity for foveation at closer eccentricities. Both fCPS and Gaussian Blur allowed for much greater foveation at 20° eccentricity, illustrating the severe limitations of temporally unstable techniques in the far periphery.

All techniques were negatively affected by large amounts of added latency (80 and 150ms). No significant difference was found between the 0, 10, 20, and 40ms added latency conditions. This suggests that there is some amount of acceptable latency below which artifact detection is only limited by peripheral perception. Given our measured system latency values, this corresponds to an eye-to-image latency of between 50 and 70ms. This finding is even more remarkable in light of the fact that subjects were specifically tasked with finding visual artifacts, and therefore tended to move their head and eyes more and pay more attention to the periphery. It is likely that less demanding tasks would have even less stringent latency requirements.

We note that there were individual differences in the detection and classification of artifacts. As shown in Figure 2, each technique degraded the periphery in a unique way, resulting in different image artifacts like flicker, specular aliasing, blur, loss of contrast, and ghosting. With increasing latency, subjects likely observed and responded to these artifacts in different ways. Yet, results across subjects were consistent and thus we believe will apply to other techniques not studied in this work.

In this article, we have explored the impact of increasing eye-tracking latency on three foveated rendering systems. From our results we infer that foveated rendering is tolerant to eye-to-image latencies of about 50—70ms, beyond which it is easily perceived and hence loses effectiveness. We also infer that perceptual improvements to foveated rendering, e.g., temporal stability, also improve its latency tolerance.

We believe that ours is one of the first studies to evaluate the perceptual impact of adding latency to foveated rendering in VR, and in the future, we would like to further explore the reasons behind latency tolerance of foveated rendering. For example, we would like to understand the specific role of saccadic blindness in our ability to detect artifacts. Further, in addition to latency, we would also like to explore the tolerance of foveated rendering to eye-tracking accuracy and robustness toward blinking and miscalibration.

APPENDIX

To verify that images rendered in our study are representative of realistic VR content, we compared the power spectrum of our rendered classroom image against the power spectra of four photographs of real classrooms. Figure 7 shows the images we used in our comparison, and the 2D power spectra of these images, and Figure 8 shows the 1D power spectral density of each of these images. Our computation of power spectra is based on the description by van der Schaaf and van Hateren [47]. As shown by the figures, all resulting spectra, including reference and rendered images, are very similar. Further, on a log-log scale, all spectra are nearly linear, which indicates that the reference as well as our rendered images follow the well-known power-law relationship between spectral power and spatial frequency in natural images. It should be noted that scenes with significantly different power spectra may either accentuate or mask certain foveated rendering artifacts; high-contrast, high-frequency edges tend to make artifacts more conspicuous.
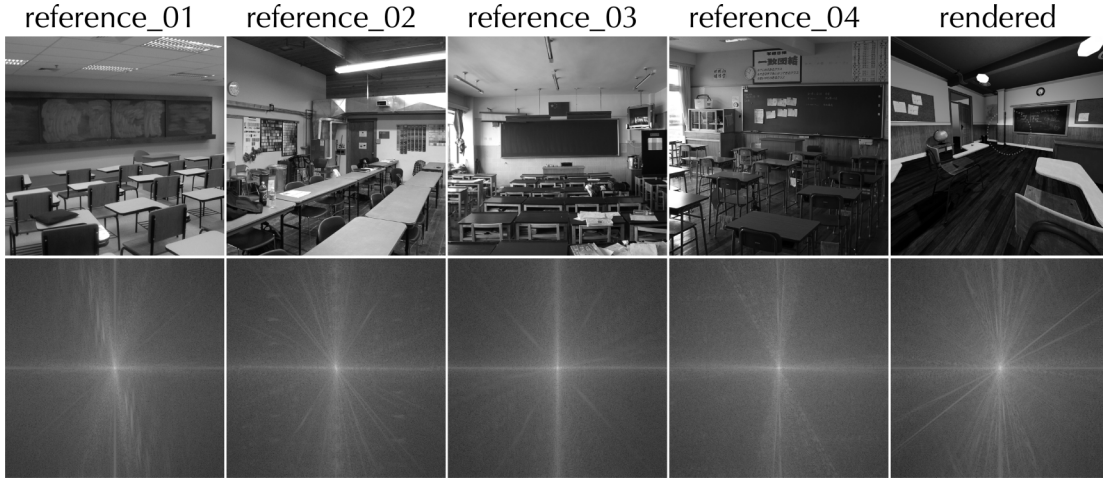
Fig. 7. Images used for verifying that the spectral content of our rendered images matches that of real-life content. We used the above four reference photographs of classrooms (obtained courtesy of Wikipedia), and compared their 2D power spectra (below) against the 2D power spectrum of our rendered image. Please see Figure 8 for a comparison of 1D power spectral density.
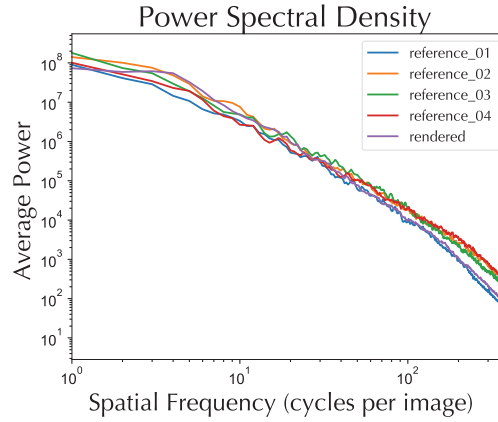

Fig. 8. 1D power spectral densities of images shown in Figure 7. The density plots are all nearly identical and linear on the log-log scale. We can conclude from this plot that the spectral content of our rendered images is similar to that of real-life classroom images and obeys the power law found in natural image spectra.

ACKNOWLEDGMENTS

REFERENCES

[1] Robert Scott Allison, Jens Schumacher, Shabnam Sadr, and Rainer Herpers. 2010. Apparent motion during saccadic suppression periods. *Experimental Brain Research* 202, 1 (2010), 155–169.

[2] Stephen J. Anderson, Kathy T. Mullen, and Robert F. Hess. 1991. Human peripheral spatial resolution for achromatic and chromatic stimuli: Limits imposed by optical and retinal factors. *The Journal of Physiology* 442, 1 (1991), 47–64.

[3] A. Terry Bahill, Michael R. Clark, and Lawrence Stark. 1975. The main sequence, a tool for studying human eye movements. *Mathematical Biosciences* 24, 3–4 (1975), 191–204.

[4] Nir Benty. 2016. The Falcor Rendering Framework. Retrieved from https://github.com/NVIDIA/Falcor.

[5] Jean-Baptiste Bernard, Scherlen Anne-Catherine, and Castet Eric. 2007. Page mode reading with simulated scotomas: A modest effect of interline spacing on reading speed. *Vision Research* 47, 28 (2007), 3447–3459.

[6] Colin Blakemore. 1970. The range and scope of binocular depth discrimination in man. *The Journal of Physiology* 211, 3 (1970), 599.

[7] Christopher J. Bockisch and Joel M. Miller. 1999. Different motor systems use similar damped extraretinal eye position information. *Vision Research* 39, 5 (1999), 1025–1038.

[8] David C. Burr, M. Concetta Morrone, John Ross, and others. 1994. Selective suppression of the magnocellular visual pathway during saccadic eye movements. *Nature* 371, 6497 (1994), 511–513.

[9] Mark R. Diamond, John Ross, and Maria C. Morrone. 2000. Extraretinal control of saccadic suppression. *The Journal of Neuroscience* 20, 9 (2000), 3449–3455. http://sci-hub.cc http://www.jneurosci.org/content/20/9/3449.short.

[10] Michael Dorr and Peter J. Bex. 2011. A gaze-contingent display to study contrast sensitivity under natural viewing conditions. In *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 78650Y–78650Y. http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=730733

[11] Andrew T. Duchowski and Arzu Çöltekin. 2007. Foveated gaze-contingent displays for peripheral LOD management, 3D visualization, and stereo imaging. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)* 3, 4 (2007), 6.

[12] Andrew T. Duchowski, Nathan Cournia, and Hunter Murphy. 2004. Gaze-contingent displays: A review. *CyberPsychology & Behavior* 7, 6 (2004), 621–634.

[13] 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. DOI : http://dx.doi.org/10.1145/2366145.2366183

[14] Thorsten Hansen, Lars Pracejus, and Karl R. Gegenfurtner. 2009. Color perception in the intermediate periphery of the visual field. *Journal of Vision* 9, 4 (2009), 26–26.

[15] John M. Henderson, Karen K. McClure, Steven Pierce, and Gary Schrock. 1997. Object identification without foveal vision: Evidence from an artificial scotoma paradigm. *Attention, Perception, & Psychophysics* 59, 3 (1997), 323–346.

[16] Cale Hunt. 2016. Field of view face-off: Rift vs Vive vs Gear VR vs PSVR. Retrieved from https://www.vrheads.com/field-view-faceoff-rift-vs-vive-vs-gear-vr-vs-psvr.

[17] Michael R. Ibbotson and Shaun L. Cloherty. 2009. Visual perception: Saccadic omission—Suppression or temporal masking? *Current Biology* 19, 12 (June 2009), R493–R496. DOI : http://dx.doi.org/10.1016/j.cub.2009.05.010

[18] Jason Jerald and Mary Whitton. 2009. Relating scene-motion thresholds to latency thresholds for head-mounted displays. In *Proceedings of the Virtual Reality Conference (VR '09)*. IEEE, 211–218.

[19] Marc Levoy and Ross Whitaker. 1990. Gaze-directed volume rendering. *ACM SIGGRAPH Computer Graphics* 24, 2 (1990), 217–223.

[20] Lester C. Loschky and George W. McConkie. 2000. User performance with gaze contingent multiresolutional displays. In *Proceedings of the 2000 Symposium on Eye Tracking Research & Applications*. ACM, 97–103. http://dl.acm.org/citation.cfm?id=355032

[21] David Luebke and Benjamin Hallen. 2001. Perceptually driven simplification for interactive rendering. In *Rendering Techniques 2001*. Springer, 223–234.

[22] Katerina Mania, Bernard D. Adelstein, Stephen R. Ellis, and Michael I. Hill. 2004. Perceptual sensitivity to head tracking latency in virtual environments with varying degrees of scene complexity. In *Proceedings of the 1st Symposium on Applied Perception in Graphics and Visualization*. ACM, 39–47.

[23] Ethel Matin. 1974. Saccadic suppression: A review and an analysis. *Psychological Bulletin* 81, 12 (1974), 899. http://sci-hub.cchttp://psycnet.apa.org/psycinfo/1975-06562-001

[24] George W. McConkie. 1981. Evaluating and reporting data quality in eye movement research. *Behavior Research Methods & Instrumentation* 13, 2 (1981), 97–106. DOI : http://dx.doi.org/10.3758/BF03207916

[25] George W. McConkie and Lester C. Loschky. 2002. Perception onset time during fixations in free viewing. *Behavior Research Methods, Instruments, & Computers* 34, 4 (2002), 481–490.

[26] Hunter A. Murphy, Andrew T. Duchowski, and Richard A. Tyrrell. 2009. Hybrid image/model-based gaze-contingent rendering. *ACM Transactions on Applied Perception (TAP)* 5, 4 (2009), 22.

[27] NVIDIA. 2016. VRWorks—Lens Matched Shading. Retrieved from https://developer.nvidia.com/vrworks/graphics/lensmatchedshading.

[28] NVIDIA. 2016. VRWorks—Multi-Res Shading. Retrieved from https://developer.nvidia.com/vrworks/graphics/multiresshading.

[29] Toshikazu Ohshima, Hiroyuki Yamamoto, and Hideyuki Tamura. 1996. Gaze-directed adaptive rendering for interacting with virtual space. In *Proceedings of the IEEE Virtual Reality Annual International Symposium 1996*. IEEE, 103–110.

[30] Anjul Patney. 2017. Perceptual insights into foveated virtual reality. In *Proceedings of the NVIDIA GPU Technology Conference 2017 Talks*. https://gputechconf2017.smarteventscloud.com/connect/sessionDetail.ww?SESSION_ID=110195&tclass=popup.

[31] 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. Retrieved from http://research.nvidia.com/sites/default/files/publications/foveated-siga-16-v1-for-web.pdf.

[32] 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 (Nov. 2016), 179:1–179:12. DOI : http://dx.doi.org/10.1145/2980179.2980246

[33] Stephen M. Reder. 1973. On-line monitoring of eye-position signals in contingent and noncontingent paradigms. *Behavior Research Methods* 5, 2 (1973), 218–228. http://www.springerlink.com/index/4413450650Q75507.pdf.

[34] Eyal M. Reingold, Lester C. Loschky, George W. McConkie, and David M. Stampe. 2003. Gaze-contingent multiresolutional displays: An integrative review. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 45, 2 (2003), 307–328. http://hfs.sagepub.com/content/45/2/307.short.

[35] William H. Ridder III and Alan Tomlinson. 1997. A comparison of saccadic and blink suppression in normal observers. *Vision Research* 37, 22 (Nov. 1997), 3171–3179. DOI : http://dx.doi.org/10.1016/S0042-6989(97)00110-7

[36] John Ross, M. Concetta Morrone, Michael E. Goldberg, and David C. Burr. 2001. Changes in visual perception at the time of saccades. *Trends in Neurosciences* 24, 2 (Feb. 2001), 113–121. DOI : http://dx.doi.org/10.1016/S0166-2236(00)01685-4

[37] Fabrizio Santini, Gabriel Redner, Ramon Iovin, and Michele Rucci. 2005. A general purpose system for eye movement contingent display control. *Journal of Vision* 5, 8 (2005), 594–594.

[38] Daniel R. Saunders and Russell L. Woods. 2014. Direct measurement of the system latency of gaze-contingent displays. *Behavior Research Methods* 46, 2 (June 2014), 439–447. DOI : http://dx.doi.org/10.3758/s13428-013-0375-5

[39] Heiko H. Schütt, Stefan Harmeling, Jakob H. Macke, and Felix A. Wichmann. 2016. Painfree and accurate Bayesian estimation of psychometric functions for (potentially) overdispersed data. *Vision Research* 122 (2016), 105–123.

[40] Robert Sekuler and Randolph Blake. 1985. Perception. Alfred A. Knopf, New York, NY.

[41] John Siderov and Ronald S. Harwerth. 1995. Stereopsis, spatial frequency and retinal eccentricity. *Vision Research* 35, 16 (1995), 2329–2337.

[42] Michael Stengel, Steve Grogorick, Martin Eisemann, and Marcus Magnor. 2016. Adaptive image-space sampling for gaze-c real-time rendering. In *Computer Graphics Forum*, Vol. 35. Wiley Online Library, 129–139.

[43] L. N. Thibos, F. E. Cheney, and D. J. Walsh. 1987. Retinal limits to the detection and resolution of gratings. *Journal of the Optical Society of America A* 4, 8 (1987), 1524–1529.

[44] Robin Thunström. 2014. *Passive Gaze-Contingent Techniques Relation to System Latency*. Retrieved from http://www.diva-portal.org/smash/record.jsf?pid=diva2:829880

[45] Jochen Triesch, Brian T. Sullivan, Mary M. Hayhoe, and Dana H. Ballard. 2002. Saccade contingent updating in virtual reality. In *Proceedings of the 2002 Symposium on Eye Tracking Research & Applications*. ACM, 95–102.

[46] Christopher W. Tyler. 1987. Analysis of visual modulation sensitivity. III. Meridional variations in peripheral flicker sensitivity. *Journal of the Optical Society of America A* 4, 8 (1987), 1612–1619.

[47] A. van der Schaaf and J. H. van Hateren. 1996. Modelling the power spectra of natural images: Statistics and information. *Vision Research* 36, 17 (1996), 2759–2770. DOI : http://dx.doi.org/10.1016/0042-6989(96)00002-8

[48] Peter Vincent and Ritchie Brannan. 2017. S7797 Tobii Eye Tracked Foveated Rendering for VR and Desktop. Retrieved from https://gputechconf2017.smarteventscloud.com/connect/sessionDetail.ww?SESSION_ID=115360&tclass=popup.

[49] Alex Vlachos. 2015. Advanced VR Rendering. Retrieved from http://alex.vlachos.com/graphics/Alex_Vlachos_Advanced_VR_Rendering_GDC2015.pdf.

[50] Andrew B. Watson. 2014. A formula for human retinal ganglion cell receptive field density as a function of visual field location. *Journal of Vision* 14, 7 (2014), 15.

[51] Nick Whiting and Nick Donaldson. 2016. Lessons from Integrating the Oculus Rift into Unreal Engine 4. Retrieved from http://static.oculus.com/connect/slides/OculusConnect_Epic_UE4_Integration_and_Demos.pdf.

[52] Stefan Wiens, Peter Fransson, Thomas Dietrich, Peter Lohmann, Martin Ingvar, and Öhman Arne. 2004. Keeping it short: A comparison of methods for brief picture presentation. *Psychological Science* 15, 4 (2004), 282–285.

[53] Hongbin Zha, Yoshinobu Makimoto, and Tsutomu Hasegawa. 1999. Dynamic gaze-controlled levels of detail of polygonal objects in 3-D environment modeling. In *Proceedings of the 2nd International Conference on 3-D Digital Imaging and Modeling, 1999.* IEEE, 321–330.

[54] Xin Zhang, Wei Chen, Zhonglei Yang, Chuan Zhu, and Qunsheng Peng. 2011. A new foveation ray casting approach for real-time rendering of 3D scenes. In *Proceedings of the 2011 12th International Conference on Computer-Aided Design and Computer Graphics (CAD/Graphics)*. IEEE, 99–102.