Towards Foveated Rendering for Gaze-Tracked Virtual Reality

Anjul Patney* Marco Salvi Joohwan Kim Anton Kaplanyan Chris Wyman Nir Benty
David Luebke Aaron Lefohn

NVIDIA

Figure 1: *Our* classroom *scene with eye fixation at the yellow reticle. (Left) Our perceptually-validated target foveated image. (Right) Our proposed foveated rendering system that avoids shading up to 70% of the pixels and closely matches the frequency content of our target by using pre-filtered shading terms, contrast preservation, and applying a new temporal antialising that improves temporal stability by an order of magnitude (providing stability similar to a temporally antialiased non-foveated renderer). The original version of the classroom scene is courtesy of Christophe Seux.*

Abstract

Foveated rendering synthesizes images with progressively less detail outside the eye fixation region, potentially unlocking significant speedups for wide field-of-view displays, such as head mounted displays, where target framerate and resolution is increasing faster than the performance of traditional real-time renderers.

To study and improve potential gains, we designed a foveated rendering user study to evaluate the perceptual abilities of human peripheral vision when viewing today's displays. We determined that filtering peripheral regions reduces contrast, inducing a sense of tunnel vision. When applying a postprocess contrast enhancement, subjects tolerated up to 2× larger blur radius before detecting differences from a non-foveated ground truth. After verifying these insights on both desktop and head mounted displays augmented with high-speed gaze-tracking, we designed a *perceptual target* image to strive for when engineering a production foveated renderer.

Given our perceptual target, we designed a practical foveated rendering system that reduces number of shades by up to 70% and allows coarsened shading up to 30° closer to the fovea than Guenter et al. [2012] without introducing perceivable aliasing or blur. We filter both pre- and post-shading to address aliasing from undersampling in the periphery, introduce a novel multiresolution- and saccade-aware temporal antialising algorithm, and use contrast enhancement to help recover peripheral details that are resolvable by our eye but degraded by filtering.

We validate our system by performing another user study. Frequency analysis shows our system closely matches our perceptual target. Measurements of temporal stability show we obtain quality similar to temporally filtered non-foveated renderings.

*apatney@nvidia.com

SA '16 Technical Papers,, December 05 - 08, 2016, , Macao
ISBN: 978-1-4503-4514-9/16/12
DOI: http://dx.doi.org/10.1145/2980179.2980246

Keywords: foveated rendering, gaze-tracking, perception, virtual reality

Concepts: •**Computing methodologies** → **Graphics systems and interfaces; Perception; Virtual reality;**

1 Introduction

Even with tremendous advances in graphics hardware, computational needs for real-time rendering systems have grown faster. Adoption of realistic lighting and physically based shading [Pharr and Humphreys 2010; Hill et al. 2015] has amplified shading complexity, while rapidly evolving head mounted displays (HMDs) for virtual reality (VR) have increased display resolution and target refresh rates. In addition, the trend toward rendering on low-power devices such as phones, tablets, and portable gaming consoles further motivates the goal of achieving the highest possible image quality using minimal computation.

As a result, algorithms that imperceptibly reduce cost are becoming more important. Interestingly, human visual acuity radially decreases between the retina center (the *fovea*) and the eye's periphery, and for HMDs and large desktop displays a significant percentage of pixels lie in regions viewed with lower visual acuity. *Foveated rendering* algorithms exploit this phenomenon to improve performance, decreasing rendering quality toward the periphery while maintaining high fidelity in the fovea. Coupled with high-quality eye tracking, foveated rendering could drive future wide field-of-view displays targeting higher pixel densities and refresh rates.

Prior foveated renderers like Guenter et al. [2012] and Vaidyanathan et al. [2014] focus on practical near-term techniques to reduce costs without explicitly identifying and minimizing perceptible artifacts introduced by foveation. Our studies show they often exhibit significant head- and gaze-dependent temporal aliasing, distracting users and breaking immersion.

Before proposing our own system, we first ask a fundamental question: what lower-quality imagery maintains perceived visual quality in human peripheral vision?

We start with a full-resolution image and reduce peripheral detail via a postprocess Gaussian blur using progressively increasing filter width based on distance to the fovea. This progressive blur provides a compelling foveated image unless using an aggressive filter radius. Large radius blurs induce a sense of tunnel vision, with peripheral pixels appearing blurry despite having appropriate frequency content based on human visual acuity curves.

We determined this degraded experience arises from missing contrast, as filtering typically reduces image contrast. Enhancing contrast to normalize for the effect of the foveated Gaussian recovers most of the apparent peripheral detail, allowing a $2\times$ larger rate of increasing peripheral blur before inducing a sense of tunnel vision. Our user studies show that a contrast-preserving Gaussian filter with a progressive standard deviation of 1 arcmin per peripheral degree of eccentricity, or even larger, is barely distinguishable from non-foveated rendering. We propose using this contrast-preserving, filtered image as a perceptual target when engineering practical foveated renderers.

We then designed a perceptually-based foveated real-time renderer that approaches our perceptual target by pre-filtering some shading attributes, while undersampling others. To avoid temporal aliasing from undersampling, we apply post-filtering using a novel multiscale version of temporal antialiasing [Karis 2014]. Since both pre- and post-filtering reduce contrast, we normalize with a post-process foveated contrast enhancing filter. We describe algorithmic details and design choices that enable a practical system that approaches our theoretical results.

Our user studies show our system enables coarse rendering up to 30° closer to the fovea than Guenter et al. [2012] without introducing gaze-dependent aliasing of blur. We also provide various numerical metrics to validate our system, including showing the frequencies maintained by our system closely match those of our perceptual target.

Our specific contributions include:

- revalidating, in complex virtual environments, the perceptual studies that show temporal stability and contrast preservation are key requirements for foveated renderers;
- providing a perceptually-validated target image that can serve as an image quality reference when designing practical foveated renderers;
- a real-time gaze-tracked foveated rendering system that closely matches our perceptual target and maintains temporal stability, with demonstrated performance and memory savings;
- novel improvements to temporal antialiasing that enable application to gaze-dependent, multi-resolution shading systems.

2 Overview of Human Peripheral Vision

In order to begin our search for a perceptual target, we started by reviewing literature in the field of the human visual system to help understand fundamental differences between foveal and peripheral vision. We discovered that differences include more than a simple reduction in acuity, e.g. higher than expected ability to detect motion in our peripherial vision.

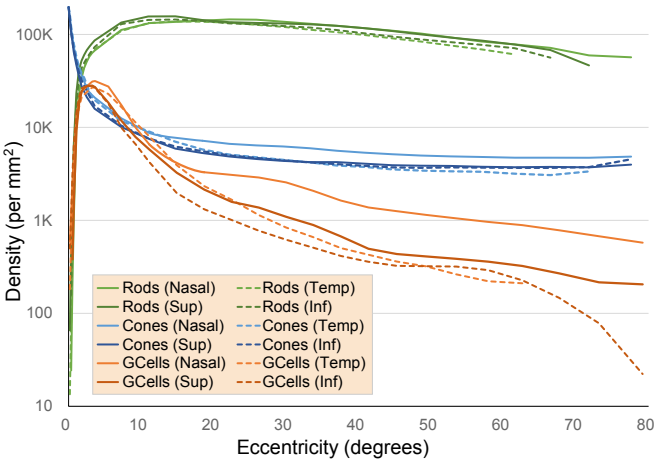

Figure 2: *Spatial distribution of various retinal components, using data from Curcio and Allen [1990] and Curcio et al. [1990]. Ganglion cell (orange) density tends to match photoreceptor density in the fovea (left), but away from the fovea many photoreceptors map to the same Ganglion cell. 'Nasal', 'Temp', 'Sup' and 'Inf' indicate the four directions (nasal, temporal, superior, inferior) away from the fovea. See Section 2.*

Because the vision literature is vast and we focus on work that informed our design decisions, we refer readers to more extensive reviews for additional details on human vision (e.g., [Strasburger et al. 2011; Rosén 2013; Wandell 1995]).

Human visual pathways include the optical components of the eye, retinal structures like photoreceptors, and high-level neural processing (e.g., in the visual cortex) [Strasburger et al. 2011]. The literature suggests differences between foveal and peripheral vision largely stem from variation in these components.

Peripheral image quality is fundamentally lower than foveal quality due to refractive lens effects [Ferree et al. 1931; Thibos 1987]. Beyond 10° of eccentricity, signals begin suffering from defocus, astigmatism, and chromatic aberration. But optical degradation is secondary to neural factors, which degrade still quicker [Banks et al. 1991; Navarro et al. 1993; Levi et al. 1985; Williams et al. 1996].

After optical transport, the rods and cones in the retina capture photons. Their outputs are combined in retinal ganglion cells which output signals from the retina. The relative densities of rods, cones, and ganglion cells varies non-uniformly with eccentricity (see Figure 2). In the fovea cone and ganglion connections are one-to-one so little loss of spatial information occurs. In the periphery, connections are many-to-one so ganglion cells provide a form of filtering. Additionally, this pooling of photoreceptor contributions reduces visual acuity beyond that expected by measuring their distribution.

2.1 Cortical Magnification Theory

Early studies on peripheral vision concluded that increasing stimuli size with increased eccentricity gives similar perceptual appearance. This magnification factor relates closely to the relative neural volume for varios regions of our visual field, and is thus known as the cortical magnification factor M. Largely, this factor scales as an inverse linear function, i.e., $M^{-1} = M_0^{-1} \cdot (1 + aE)$ for M_0 the foveal magnification [Cowey and Rolls 1974], while some have proposed non-linear magnification models [Rovamo and Virsu 1979].

Recent foveated renderers build on cortical magnification theory, reducing resolution [Guenter et al. 2012; Swafford et al. 2016], shading rate [Vaidyanathan et al. 2014], and screen-space ambient

occlusion, tessellation, and ray-casting steps [Swafford et al. 2016] with increasing eccentricity. However, cortical magnification alone fails to describe all aspects of peripheral vision. Different visual tasks can follow different scaling factors, and some even exhibit no scalibility (see Strasburger et al. [1994], Makela et al. [2001] and Strasburger et al. [2011]).

2.2 Spatial Aliasing in Peripheral Vision

In peripheral regions, the many-to-one relationship between photoreceptors and ganglions causes a mismatch between optical, retinal, and final neural sampling resolutions. Numerous researchers found this mismatch leads to aliasing in our peripheral vision [Thibos et al. 1996; Wang et al. 1996]. Using spatial gratings of varying frequencies, we can measure when we visually perceive a grating (the *detection* point) and when we can resolve its orientation (the *resolution* point). While both points are similar in the fovea, resolution acuity is lower than detection in the periphery (see Figure 3), where the grating becomes recognizable well before its orientation [Thibos et al. 1987a; Thibos et al. 1987b]. Resolution acuity follows angular density of ganglions. The gap between the detection and resolution thresholds is often called the *aliasing zone*—the range of frequencies where one can detect but not resolve the stimulus. Presence of the aliasing zone suggests we can perceive higher frequencies in the periphery even if we cannot discern the content [Wang et al. 1997].

Compared to resolution acuity, detection acuity decays slowly. After correcting for refraction, the detection acuity at 30° eccentricity is 30 cycles per degree, compared to 5 cycles per degree for resolution acuity [Thibos et al. 1996]. Thus, *detection* rather than *resolution* should serve as a conservative acuity estimate for foveated rendering. Foveated rendering targeting resolution acuity will eliminate frequencies in the aliasing zone, which can cause apparent loss of contrast in peripheral vision. We exploit our ability to detect but not resolve by enhancing peripheral contrast to maintain apparent detail.

2.3 Peripheral Motion and Flicker Perception

Performance of some tasks remains uniform across the visual field. For example, our ability to differentiate two objects' velocities is quite uniform across the visual field [McKee and Nakayama 1984]. And after scaling stimuli according to cortical magnification theory, our flicker sensitivity remains roughly uniform across the visual field [Kelly 1984; Koenderink et al. 1978b; Koenderink et al. 1978a; Koenderink et al. 1978c].

Consistent motion and temporal sensitivity requires foveated renderers avoid artifacts like temporal aliasing, as these are easily perceptible even in the periphery. Foveated renderers that fail to account for temporal stability either break immersion or must be run conservatively in the periphery.

2.4 Other Factors

We have anisotropic visual acuity, with detection and resolution varying between frequencies in radial and tangential directions. We are more sensitive to gratings oriented radially [Rovamo et al. 1982], by roughly a factor of 2× [Thibos et al. 1996] at 30° eccentricity. Our experiments applying anisotropic blurring did not provide conclusive results, so we were unable to identify ways to leverage anisotropy into further performance savings, however this merits further investigation.

Our color perception also degrades with increasing eccentricity. But even at high eccentricities we maintain some color differentiation [Hansen et al. 2009; Noorlander et al. 1983; Solomon et al. 2005]. Foveated rendering could exploit this behavior by reducing the precision of chroma computations in the periphery, but successfully leveraging this observation likely requires hardware with varying precision math.

3 Emulated Foveated Rendering

Next we aim to build on our knowledge from Section 2 to create a perceptual sandbox to help evaluate peripheral image quality in isolation from rendering design choices. This sandbox can be considered an "emulated foveated renderer," as we perform foveation as a postprocess to allow easy prototyping. The goal of this exercise is to obtain a *perceptual visual target* that can help inform design decision for our final, cost-reducing foveated renderer.

We evaluate the resulting perceptual target via a user study using both a gaze-tracked VR HMD and a gaze-tracked high-resolution desktop display. We assess the perceptual performance, perform a wavelet analysis to establish wavelet domain characteristics (see Section 4.4.2), and analyze contrast preservation of the resulting images.

3.1 Design of Perceptual Sandbox

Our sandbox is a VR application which renders all images at full-resolution and level-of-detail, and uses supersampled antialiasing (SSAA) to achieve high-quality and temporally-stable images. Due to the performance requirements of VR rendering, we were limited to using 2× SSAA. Then, it applies any foveation as a post-process, allowing us to study perceived image quality. Foveation is only applied outside of a fixed central foveal region, increasing linearly with retinal eccentricity. Our sandbox is implemented using the Falcor rendering framework [Benty 2016].

Our sandbox naturally underperforms non-foveated rendering, but provides a framework to rapidly explore and evaluate image quality of foveated rendering approaches. Within this framework, we are able explore the following strategies for generating peripheral images.

3.1.1 Emulating Aliased Foveation

Motivation Due to the enhanced perception of motion and flicker in peripheral vision (Section 2.3), we want to explore the impact of temporal aliasing on foveated rendering.

Implementation To compare our non-foveated peripheral image with a temporally aliased one, we artificially inject temporal instability by subsampling the full resolution image. Further, to recreate the image quality from a multi-resolution foveated renderer, we bilinearly filter subsamples and bias the texture level-of-detail to match the subsampled rate. This results in a spatially smooth but temporally unstable peripheral image, corresponding to an upsampled lower resolution image. The size of the subsampling filter increases linearly with retinal eccentricity.

The impact of the resulting temporal aliasing on user experience was surprisingly prominent. Previous work in foveated rendering also identifies this problem [Guenter et al. 2012].

3.1.2 Emulating Temporally Stable Foveation

Motivation To compare against the above, we also explored temporally stable foveation.

Implementation We apply radially-increasing Gaussian or bilateral blurs to our full resolution rendering. Unlike Section 3.1.1, this image is temporally stable. While using a bilateral filter preserves details in the periphery and is perceptually an improved experience, we discovered that tuning the bilateral filter to avoid inducing "tunnel vision" required setting the edge threshold values so high that the bilateral filter removed only a small amount of information in the periphery.

Thus we primarily use a radially-progressive Gaussian blur, because it represents the type of pre-filtering common in rendering systems. As evaluated in Section 3.2, this simple system works for gaze-tracked foveated rendering, but has low detectability limits (users again report a "tunnel vision" experience.)

3.1.3 Preserving Peripheral Image Contrast

In our final sandbox strategy, we augment our emulation of temporally stable foveation with contrast preservation.

Motivation For a non-emulated foveated renderer, achieving temporally stable peripheral image like the one obtained using a Gaussian blur in Section 3.1.2, requires careful filtering during rendering. This unavoidably bandlimits the various rendering terms, effectively eliminating frequencies above a threshold.

To maximize performance without impacting perceived image quality, we should select the bandlimiting frequency based on the visual detection threshold. This bandlimiting is shown by the orange plot in Figure 3. In other words, the visual detection threshold limits the rate of peripheral image degradation in a foveated renderer, since frequencies beyond the detection threshold are imperceptible and can be safely removed via filtering.

Frequencies lower than the detection threshold but higher than the resolution threshold are special. They lie in the aliasing zone (Section 2.2), where the visual system is capable of detection but lacks resolution. The resolution threshold degrades much faster with increasing eccentricity, so it is a much more attractive target for foveated renderers. In other words, we expect higher performance with a foveated renderer which can target the resolution threshold (green plot in Figure 3).

To achieve this target, we wanted to explore foveation techniques which generate images with details in the aliasing zone, i.e. those details can be detected, but not resolved. Since it is well known that in comparison to filtered images (e.g. beyond the detection threshold), subsampled / aliased images (e.g. between the detection and resolution thresholds) tend to preserve image contrast [Öztireli and Gross 2015], we propose that foveated rendering using contrast-preserving filtering can target the resolution threshold for peripheral image degradation.

In addition to the above observation, we also draw inspiration from recent thumbnail generation algorithms [Öztireli and Gross 2015], where similar problems arise from filtering to minify the image. Further, recovering contrast to account for details lost via filtering has previously been shown to be effective [Grundland et al. 2006; Kim et al. 2011]. These works suggest that by carefully preserving image contrast, we can improve image perception even for filtered images.

Implementation To achieve contrast preservation, we employ a post-process kernel that enhances the contrast of a foveated image obtained after applying a radially-progressive Gaussian blur (Section 3.1.2). We observed that this kernel indeed improves the image contrast in the medium frequencies of the power spectrum, which are likely to lie in the aliasing zone (see Figure 4). Please see supplementary material for details of our contrast-preserving post process.

3.1.4 Other Strategies

In addition to temporal stability and contrast preservation, we also experimented with foveation strategies based on blur anisotropy and peripheral color reduction. Our experience with the former was inconclusive as the superiority of radial or tangential blur seemed to depend on the dominant edge in the image. Our experience with the latter was encouraging as subjects were often unable to tell

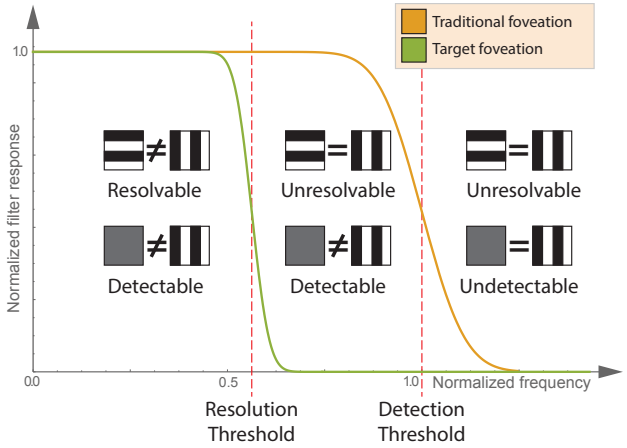

Figure 3: *An illustration motivating the use of post-process contrast enhancement in the periphery. We plot angular frequency on the x-axis, normalized to the detection threshold. We divide the x-axis into three regions around the resolution and detection thresholds, which are distinct in peripheral vision. On the y-axis, we plot the bandlimiting targeted by a conservative filter (orange) which follows detection acuity, and an aggressive filter (green) which follows resolution acuity. The latter loses frequencies in the region between the two thresholds, where we can detect but not resolve details, so we perceive loss of contrast due to filtering. To recover this loss of contrast, we propose the use of contrast-enhancement filter.*

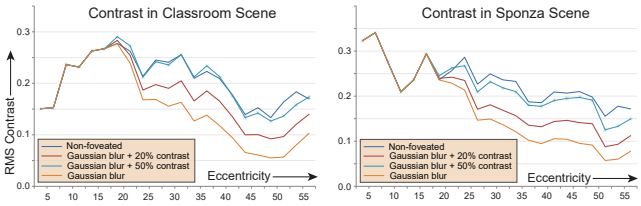

Figure 4: *Root mean square contrast as a function of retinal eccentricity in two different scenes. We show contrast of a non-foveated image and a radially varying Gaussian blur. We also enhance the contrast of our Gaussian blur with two different parameters, showing we can normalize to account for contrast degradation due to filtering.*

significant magnitudes of chromaticity reduction. However, we were unable to translate that insight into sensible savings in a rendering system.

3.2 User Study

We validated our observations of contrast-preserving filtering by performing a user study to compare the relative detectability of the three foveation strategies: aliased foveation (Section 3.1.1), temporally stable foveation (Section 3.1.2), and temporally stable and contrast preserving foveation (Section 3.1.3). We measured detectability using the threshold rate of change of peripheral blur with increasing retinal eccentricity, which is an indicator of the maximum acceptable extent of foveation. For each strategy, we estimate the threshold rate of blur change, above which the participant would easily perceive the presence of foveation.

Hardware Setup

We used two display configurations for our study (also see Figure 7):

Figure 5: *Scenes used in our evaluation. Top row shows desktop scenes, bottom row shows HMD scenes. Left column is* Crytek-Sponza *(courtesy of Frank Meinl and Efgeni Bischoff), middle column shows* Classroom *(original model courtesy of Christophe Seux), and right column shows* San Miguel *(courtesy of Guillermo M. Leal Llaguno).*

HMD Setup We use an Oculus Development Kit 2 Virtual Reality HMD, retrofitted by SensoMotoric Instruments (SMI) with a high-speed gaze tracker. This display has a resolution of 1920×1080 pixels (to which we render at 1184×1464 pixels per eye), a field-of-view of 110°, and the gaze tracker runs at 250 Hz with a response latency of 6.5 ms. For this setup we use a central foveal radius of 15°, which is larger than desktop estimates due to the low angular resolution of the HMD.

Desktop Setup Due to low angular resolution in the first setup, we also test our results on a desktop display. We used a 27 inch LCD monitor with resolution of 2560×1440 (Acer XB270HU), with the subject seated at a distance of 81 cm (chosen to approximate 1 pixel/arcminute at the center of the display.) The subject wore a head-mounted gaze tracker (Eyelink2 by SR Research), which estimates gaze at 250 Hz with an accuracy of < 1°, and a photon-to-machine latency of 5 ms. We set the central foveal radius for this setup to 7.5°.

Procedure

The basic procedure of our study was to show our users the Sponza scene (Figure 5) rendered with and without foveation, and to ask them to pick which looked better. Participants responses helped determine the threshold foveation rate for each strategy.

In each trial of our two-alternative forced choice (2AFC) test, we sequentially presented a foveated version and a non-foveated version of the scene in randomized order, with a blank frame of 0.75 second between the two presentations. Then we asked each participant to choose the better-looking one between the two stimuli. We conducted several such trials for each participant.

In the desktop setup, we maintained the viewpoint—camera location and orientation—within each trial, but randomly cycled between three predetermined viewpoints in the scene.

We instructed each participant to maintain their gaze near the center of the display all the time, even though our emulated foveated renderer used gaze tracking to correctly update the foveated image. For the HMD setup, we also asked them to try to maintain a nearly constant viewing direction. Consequently, the participants experienced a similar visual field across all the trials.

We determined the rate of change of peripheral downsampling in each trial according to a 1-up 2-down staircase method [Levitt 1971]. To minimize "training effect" in our study, we interleaved staircases for the three strategies within each experiment. Participants performed up to 200 trials to estimate the threshold of one foveation method. We tuned the trial counts to include at least 40 reversals for each staircase to ensure narrow confidence intervals. It took about one hour to complete the experiment including orientation and calibration of the tracker.

Participants

We conducted our experiment on four participants, aged 26 to 46, with 20/20 corrected vision, and no history of visual deficiency. One participant was an author and the other three were not aware of the experimental hypothesis. We chose only four participants due to the rigorous nature of our psychophysical study which includes 40 reversals per participant. Studies with such few participants are common in low-level psychophysics [Banks et al. 2004; Kelly and Savoie 1973]. It is also an accepted practice for authors to participate in such studies.

Proper consent and debriefing were provided according to the Declarations of Helsinki. Before the experiment, we briefly described to each participant the motivation behind foveated rendering but not the details of our strategies. We allowed all participants some training time to get familiar with the experimental setup.

Analysis

To analyze participant responses, we fitted a psychometric function to the user performance for each strategy, measured as a function of rate of change of peripheral blur using Bayesian inference methods [Schütt et al. 2016; Wichmann and Hill 2001a; Wichmann and Hill 2001b]. Using the psychometric fits, we computed the threshold blur rate, defined as the rate at which a participant's performance, measured as the correctness rate, meets 75%, i.e. midway between a random response rate (50% performance) and a perfect response rate (100% performance). We have included the measured psychometric functions in our supplementary material.

We compared the three thresholds to evaluate which strategy can provide the most savings. Because the trend was very consistent among all the individual participants, we aggregated all the responses together and performed the analysis.

Results

Figure 6 shows the resulting thresholds for each strategy and setup. Since the stimulus for our desktop setup was static, we did not include the "Aliased Foveation" strategy in that setup. Higher thresholds are better, since they indicate our ability to foveate more without perceptual degradation. Our results demonstrate that while the threshold blur for our first and second strategies was similar, the threshold for contrast-preserving foveation strategy was significantly higher. For the HMD setup, threshold for contrast-preserving foveation was approx. 2× better than temporally stable and approx. 3× better than aliased foveation. For the desktop setup, threshold for contrast-preserving foveation was approx. 2.3× better than foveation without contrast preservation. Contrast-preserving temporally stable foveation is harder to perceive than its alternatives. Consequently, we confirmed our hypothesis that temporally stable and contrast-preserving peripheral images are perceptually superior to temporally unstable as well as non-contrast-preserving alternatives.

4 Designing a Foveated Renderer

Guided by the perceptual target image described in Section 3, we now strive to design a foveated renderer for gaze-tracked VR that achieves variable-rate sampling without temporal aliasing, preserves contrast, and improves performance from the reduced sampling rates in the periphery. We previously presented a demo of our renderer [Patney et al. 2016]. Here we provide its design details.

We evaluate the perceptual performance of our renderer with a user study, assess performance, measure temporal stability, and compare the wavelet analysis to non-foveated and perceptual target images.

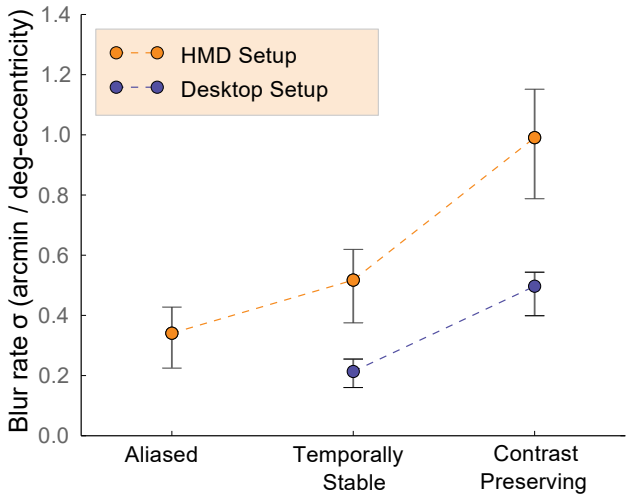

Figure 6: *Results of our user study to compare different foveation strategies (Section 3). The y-axis shows the rate of peripheral blur change with increasing retinal eccentricity, measured by the standard deviation of Gaussian blur (in arcminutes) per unit degree of eccentricity beyond the central foveal region. For both our gaze-tracked setups (a VR HMD and a desktop display), we see a significantly higher threshold of detectability for temporally-stable and contrast-enhanced peripheral blur, as compared to a strategy that is temporally stable but not contrast-preserving. For the HMD we also compare against a temporally aliased peripheral image, and observe significant improvement.*

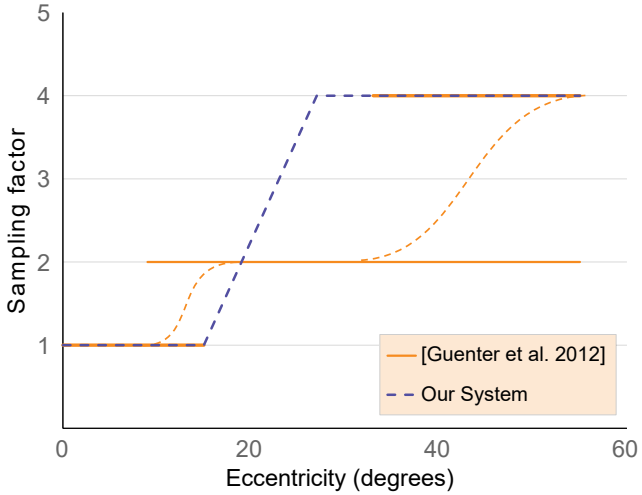

Figure 8: *Comparison of foveation parameters used in our evaluation of the two systems. We use the transition size as determined by our user study in Section 4.4.1, using our best user as an aggressive target. The rendering system proposed by Guenther et al. [2012] renders in three foveal layers (solid yellow), and blends between them using a smoothstep (dashed yellow) function. Our system (blue) is based on Vaidyanathan et al. [2014] and uses a piecewise linear variation of shading rate within a single layer.*

4.1 Design Decisions

Foveated renderers seek to reduce sampling rates to save work and thereby produce a blurry result in the undersampled regions. Reducing sampling rate without introducing aliasing is only possible for rendering terms that can be pre-filtered.

Visibility

There are no known methods to pre-filter primary visibility, which in rasterization occurs by z-buffering, comparing a fragment's depth with prior fragments. Although the visibility rate can be lowered to account for image distortion introduced by HMD lenses [Toth et al. 2016], our system samples visibility at full resolution throughout the entire image. This is a different decision than in Guenter et al. [2012], and we evaluate the impact of this decision for VR in our user studies.

Note that foveated alpha-tested geometry is problematic; while it uses a filtered mipmap hierarchy, the binary alpha threshold is not filterable and creates objectionable gaze-dependent aliasing. This may be solvable using alpha compositing with approximate order-independent transparency (e.g., Salvi et al. [2014]) or by filtering signed distance fields [Green 2007], but we did not address this challenge in our renderer and hence currently do not support alpha tested materials.

Figure 7: *Our user study Setups. Left shows our desktop setup with a high-speed gaze tracker, and right is our HMD setup with a gaze-tracker integrated within the HMD.*

Pixel Shading Rate

Guenter et al. [2012] generate foveation by sampling visibility and shading at the same rate but in a multi-resolution fashion. We seek to hold the visibility rate constant while varying the pixel shading rate, applying existing algorithms that pre-filter material attributes and lighting terms.

Literature contains numerous systems that vary pixel shading rate independent from the visibility sampling rate. Very few of these are implemented in current graphics hardware, though most can be implemented at various levels of emulation. One class of decoupled shading systems uses object-space or texture-space sampling, such as Clarberg et al. [2014], Baker [2016], and Hillesland et al. [2016]. These mechanisms are attractive for foveated rendering because they are temporally stable, but require complex memory management and substantial changes to art assets. Alternatively, coarse pixel shading (CPS) [Vaidyanathan et al. 2014; He et al. 2014] decouples shading and visibility rates in screen space. This approach works with existing triangle-based art assets, does not require complex memory management, though the possible shading rates are limited by triangle size and they require careful pre-filtering of shading terms to avoid introducing aliasing. We opt to start with Vaidyanathan et al. [2014] because it is the simplest system that meets our requirement of fixed-visibility and variable-rate pixel shading rate.

However, Vaidyanathan's coarse shading system was designed for high-resolution displays in small form-factor devices such as phones and tablets, not VR HMDs. Using this system in an HMD presents two problems: the effective pixel size in current HMDs is very large (not small pixels like in Vaidyanathan's design point) and head- and gaze-dependent motion exaggerates the impact of any artifacts. We address both of these challenges with a new post-processing anti-aliasing algorithm (see Section 4.2).

Material and Light Shading

We carefully choose material and lighting models that can be pre-filtered to make it possible for our system to vary the shading rate based on eccentricity without introducing aliasing. In practice, the pre-filtered material and light shading problem means a renderer can evaluate shading terms across variable-size pixel footprints in constant time. Texture mipmapping provides texture pre-filtering [Williams 1983], Kaplanyan [2016]'s specular lighting formulation integrates specular over variable-size pixel footprints, we use LEAN mapping [Olano and Baker 2010] to pre-filter normal maps, and we pre-filter shadows using exponential-variance shadow maps [Lauritzen et al. 2011] (EVSM).

4.2 Post-process Antialiasing

Although we sample visibility at a constant rate, our sampling rate is not high enough to eliminate aliasing, so we add post-process anti-aliasing to further improve image quality. In addition, coarse pixel shading (CPS) produces a constant color for all pixels within a coarsely shaded pixel. This is an important simplification that makes CPS more practical for hardware implementation than other approaches; however, this produces gaze-dependent artifacts in a gaze-tracked foveated HMD renderer. Rather than filter the shading samples in the pipeline, as He et al. [2014] propose, we introduce a new post-process image reconstruction algorithm.

Our new algorithm, variance sampling, is derived from the temporal anti-aliasing (TAA) algorithm described by Karis [2014], but introduces variable-size sampling and saccade-aware reconstruction. First, we briefly summarize how TAA works. Then we present variance sampling and show how it can be used to reduce or even eliminate temporal artifacts over arbitrarily large regions of the image.

Background Reuse of samples from prior frames requires temporal reprojection to maintain image stability. Guenter et al. [2012] relies on such projections. State-of-the-art temporal anti-aliasing (TAA) variants [Karis 2014; Jimenez et al. 2012] have minimal computational and memory overheads and provide a major improvement to quality. TAA can be tuned to reduce all kinds of noise and aliasing, including specular aliasing, geometric aliasing, shader aliasing, and stochastic noise. In exchange for reduced aliasing, it tends to remove high frequency details or introduce ghosting as details from prior frames persist beyond when they can be correctly reprojected.

Continuously integrating color samples from previous frames is a common way to amortize the cost of supersampling [Yang et al. 2009], but it can easily generate ghosting artifacts when stale samples are used. Karis' [2014] TAA technique reduces ghosting by conditioning color samples from previous frames that are inconsistent with samples in the current frame.

First, this method computes an axis-aligned bounding box in color space containing all color samples within a local 3×3 pixel window from the current frame. The 3×3 window is centered around the current pixel, with color C_t. Second, back-projected and re-sampled color C_{t-1} from the previous frame is clamped against this bounding box. The data reused from the past frame is not affected if it is enclosed in the bounding box. Conversely, after the clamping operation the reused data will lie on the edge on the bounding box, making it more consistent with the present color data. Finally, temporal supersampling is achieved by integrating the newly shaded data and the post-clamping color sample C'_{t-1} by using an exponential moving average:

$$C_t = \alpha \cdot C_t + (1 - \alpha) \cdot C'_{t-1}, \tag{1}$$

where α is typically set to 0.1. Note that when color clamping has no effect 90% of the current pixel color is effectively contributed by the previous frame.

Variance Sampling The method described above is widely popular among game developers but still suffers from significant ghosting artifacts. This is due to outlying color samples that significantly expand the bounding box, making it a poor representation of local color distribution. Moreover, previous foveated rendering techniques do not work well with this approach, as color samples are replicated over many pixels impoverishing our knowledge of local color distribution. Computing a color space bounding box over a larger pixel window significantly improves our knowledge of the local color distribution but it becomes expensive.

We introduce variance sampling, a method that explicitly computes the first two raw moments of the local color distribution for each color component i:

$$m_{1i} = \frac{1}{n} \sum_j c_{ij}$$

$$m_{2i} = \frac{1}{n} \sum_j c_{ij}^2$$

Enabling us to derive the distribution mean $\mu_i = m_{1i}$ and standard deviation $\sigma_i = \sqrt{(m_{2i} - m_{1i}^2)}$ to generate a bounding box centered around μ_i and with extents determined by σ_i, up to a scaling factor.

This approach provides two major benefits. First, our color bounding box tends to naturally exclude outliers and it better encloses the region of color space where most of the current frame color samples are, effectively reducing ghosting artifacts over previous methods, without increasing computational and memory bandwidth costs. Second, the raw moments of a distribution are linear quantities that can be pre-filtered over large image regions (e.g. with mipmapping). This property enables multiscale TAA implementations that can efficiently determine the statistical properties of large and undersampled image regions, for instance rendered with coarse pixel shading, and increase their temporal stability (see Section 4.3 for more details).

Saccade Detection and Recovery After saccadic eye movement, previously peripheral regions may suddenly be located in the fovea. Due to temporal integration, a previously blurry peripheral region might require several frames to converge to its sharp foveal counterpart, causing an effect similar to camera focusing on an object. To eliminate this *focus lag* we locally accelerate the rate of convergence by increasing α:

$$\alpha' = \omega \cdot \alpha_{max} + (1 - \omega) \cdot \alpha. \tag{2}$$

α_{max} is the maximum value α' can take and it is controlled by the normalized difference between a pixel previous and current normalized shading rates ω:

$$\omega = \max\left(0, \min\left(1, \frac{(S_{t-1} - S_t)}{\beta}\right)\right) \tag{3}$$

Storing the normalized shading rate S_t in the frame buffer alpha channel, alongside color information, makes it possible to automatically detect and recover from saccadic eye movements in postprocessing. By definition $S_{\text{fovea}} = 0$ and $S_{\text{periphery}} = 1$ and for optimal results we use $\alpha_{max} = 1/3$ and $\beta = 1/4$, that guarantee a rapid convergence to the *correct* image, without incurring significant temporal artifacts. Note that in the apparently symmetrical case of a pixel transitioning from fovea to periphery there is no need modify α and we have $\omega = 0$ and $\alpha' = \alpha$, leaving the temporal integration rate unaffected and thus reducing the likelihood of introducing gaze-dependent temporal artifacts in the peripheral regions of the image.

	Pixel Writes (%)	Shaded Quads (%)		Average Temporal Difference (%)
	Guenter et al. 2012	Guenter et al. 2012	Our method	Our method
Sponza	44.8	48.4	26.6	9.4
Classroom	46.1	49.9	31.8	13.2
San Miguel	45.4	50.7	38.1	12.3

Table 1: *Comparison of foveated rendering algorithms performed by running three 300 frames long gaze-tracked traces captured on an Oculus DK2 HMD customized with an eye tracker. All reported values are percentages relative to data captured by rendering the traces with standard non-foveated rendering techniques. We do not count the shaded quads beyond the edge of the lens, since they are not visible in the HMD. Our method significantly lowers the number of shaded quads while reducing by an order of magnitude our qualitative indicator of temporal instability.*

4.3 Renderer Implementation

We tested three real-time forward rendering algorithms:

1. Non-foveated rendering (NONFOV),
2. Guenter et al. [2012] augmented with our TAA algorithm (MULTIRES),
3. Coarse pixel shading [Vaidyanathan et al. 2014] with shading pre-filtering, contrast preservation, and multiscale TAA (OUR).

Since no hardware implementations of coarse pixel shading (CPS) are currently available, we emulate CPS in a GLSL program. Our emulation code, built using Falcor [Benty 2016], supports three rates of shading, 1 (one shade per pixel), $1/2$ (one shade per 2×2 pixels) or $1/4$ (one shade per 4×4 pixels), grouped correspondingly into 2×2, 4×4 and 8×8 coarse pixel quads. Our emulation can be used on an high-end GPU to validate image quality at 75 Hz on a Oculus DK2 HMD for reasonably complex scenes. Moreover we use this emulator to precisely count the number of shaded coarse quads.

All rendering methods generate 8 bytes of per-pixel data for:

- color C_t and normalized shading rate S_t (4 bytes),
- backprojection motion vector (4 bytes).

To drive our multiscale TAA algorithm a fullscreen pass generates the first two raw color moments for each pixel within a 3×3 window and pre-filters the result of this operation by creating a full mipmap chain. This data is stored in half precision textures and requires 12 bytes per pixel.

The coarser shading rate supported by OUR can shade one sample and replicate its color over 4×4 pixels. Therefore the 3×3 pixel window required to compute color moments can be *virtually stretched* to up to 12×12 pixels in the eye periphery. We account for this stretch by re-sampling the pre-filtered color moments via tri-linear filtering using:

$$\text{LOD} = 4 \cdot S_t \tag{4}$$

as texture level-of-detail. When $S_t = 0$ we have LOD $= 0$ (fovea) causing our TAA implementation to sample the highest resolution color moments. When $S_t = 1$ we have LOD $= 4$ (periphery) and we sample pre-filtered moments computed over 16×16 pixels in order to account for the coarse pixels. Since the normalized shading rate changes smoothly from the fovea to the periphery we can continuously reconstruct the local color distribution and use it to clamp backprojected color samples from the previous frame. Note that NONFOV and MULTIRES always render 1 to the alpha channel since they do not need to re-sample the color moments over a larger image region. Finally, our TAA algorithm is fast and runs in roughly 1.5ms on an NVIDIA GTX Titan X GPU.

For contrast preservation we implemented an image-space postprocess that executes after all other post effects (TAA, tone mapping, etc.) and recovers the contrast of the filtered image. We use a filter inspired from recent contrast enhancement filter from Grundland et al. [2006] designed to preserve contrast after manipulating the image. The filter is based on the variance interpretation of the contrast and applies a variant of unsharp masking (Eq. 9 in [Grundland et al. 2006]) to the image.

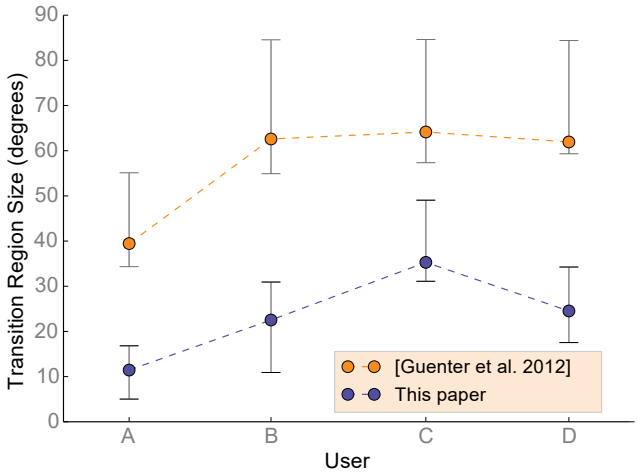

Figure 9: *Results of our user study comparing our renderer with the system proposed by Guenter et al. [2012]. The y-axis shows the size of the intermediate region, i.e. the angular distance between fovea and periphery. Hence, lower thresholds are better, since they enable more peripheral pixels. In our study we set the maximum threshold in our user study to 60°, which can be seen to be hit for many users. Please see Section 4.4.1 for details of this study.*

4.4 Renderer Evaluation

4.4.1 User Study

We performed a second user study to verify that our rendering system indeed achieves the superior image quality predicted by our perceptual target from Section 3. For this study, we compare OUR rendering system against the MULTIRES foveated rendering system [Guenter et al. 2012]. Our setup is similar to our study from Section 3.2, with the following changes:

- We only evaluate the HMD setup, since the primary goal of our desktop study in Section 3.2 was to confirm our hypothesis for a higher density display.
- We also run this experiment using four subjects, but all of these subjects were different from the ones for our previous experiment in Section 3.2.

Following Vaidyanathan et al. [2014], OUR rendering system only supports three discrete shading rates. The MULTIRES system [Guenter et al. 2012] also uses three discrete rendering reso-

lutions. Thus, instead of estimating a threshold rate of foveation, we estimate the threshold size of the intermediate region between the inner (foveal) and the outer (peripheral) regions. Note that the latter is also a measure of the rate of foveation, since it represents the angular distance between two fixed shading rates. However, it is an inverted measure, so a lower threshold is better than a higher one.

Also note that the two rendering systems have a slightly different definition of transition regions, based on the corresponding original proposals (see Figure 8).

Other than the above changes, trials of this study were identical to those in our first user study. During each trial, we showed the participants a sequence of foveated and non-foveated rendering in randomized order, and asked them to identify the one that looked better.

Figure 9 shows the results of this user study, using the same analysis procedure as in Section 3.2. Due to higher individual variation in the thresholds, we have listed them separately. The estimated threshold size of the transition region for OUR rendering system is 1.8–3.5× lower than that for a MULTIRES foveated renderer. This lets us conclude that our renderer is capable of generating images with demonstrably higher quality, and verifies our hypothesis.

We would also like to note here that the thresholds obtained from both our user studies are generally conservative, since our study estimates ability to distinguish between a non-foveated and a foveated image. In practice, we find that we can often reduce the transition region to around 5° before artifacts become prominent. While we can tell differences between non-foveated and foveated images in such a configuration, the foveated images are in acceptable in isolation.

4.4.2 Wavelet Analysis

To help understand the behavior of our postprocess foveation in the perceptual sandbox as well as our new rendering system, we performed a wavelet analysis to isolate which frequencies remained in different areas of our images. We used a 2D Haar wavelet transform with five wavelet levels (representing $2^1 \times 2^1$ to $2^5 \times 2^5$ pixel regions). We combined wavelet coefficients for regions of the image with varying eccentricity, allowing us to analyze frequency changes for our techniques as a function of angle from the gaze direction.

Within various eccentricity bands, we computed both maximum and average wavelet coefficients. We found the average wavelet coefficients highlight eccentricities where various foveation methods *discard* information whereas the maximum wavelet coefficients highlight where the most important features are *kept*.

Figure 10 shows average wavelet coefficients at three different frequencies for a representative view in the Sponza scene. As expected, all algorithms tested leave low frequency details essentially untouched though contrast enhancement augment these frequencies somewhat in the periphery to account for lost energy in higher frequencies.

At high frequencies, our perceptually validated foveated blur maintains most high frequencies out to 30°, significantly further than prior research renderers (including our implementation of Guenter et al. [2012] with high quality temporal antialiasing). Our new temporally-stable system does a much better job maintaining frequencies in a similar range, though without contrast enhancement it also underrepresents frequencies at eccentricities between 25° and 40°. At mid range, our new system maintains a similar frequencies profile to our perceptual target and with contrast enhancement, frequency content closely resembles our non-foveated renderer.

4.4.3 Performance and Temporal Stability Evaluation

For our evaluation we captured on an Oculus DK2 300 frames long gaze-tracked traces at the resolution of 1184×1464 pixels per eye from three different scenes: Sponza, Classroom and San Miguel. We use the total number of shaded quads and pixel writes as a proxy for potential performance improvements attainable with foveated rendering methods. While our method does not reduce the visibility rate and therefore the number of pixel writes, we report in Table 1 a reduction of over 50% in pixel writes for Guenter et al. [2012] relative to non-foveated rendering. On the other hand our method is up to 50% more effective at lowering the shading cost than Guenter et al. [2012], as shown in the Sponza trace, where $3/4$ of all pixel quads shaded by non-foveated rendering are eliminated. Figure 11 reiterates this finding by illustrating how the total number of shaded quads can change over time for all tested algorithms.

As mentioned in Sections 2 and 3 temporal stability is a very important property for a foveated renderer, although it is hard to quantify. To guide our study we developed a qualitative evaluation of temporal stability based on tracking temporal differences under motion. We do so by computing the difference in post tone-mapping luminance of two consecutive frames after backprojecting the previous frame onto the present frame. The resulting image is reduced to a single average value yielding an *average temporal difference* indicator (ATD) for the whole frame. In Figure 12 we report ATD values for a non-foveated rendering method and our method, with and without temporal supersampling. Simply reducing the shading rate in the periphery of the image leads to more temporal instability (dark blue line). Our multiscale temporal supersampling approach is effective and it matches or even exceed non-foveated rendering temporal stability, lowering ATD by an order of magnitude (see light blue line in Figure 12 and rightmost column in Table 1).

5 Conclusions and Future Work

We began our research after discovering objectionable gaze- and head-motion-dependent artifacts in our implementations of Guenter et al. [2012] and Vaidyanathan et al. [2014] and realizing we lacked a principled approach to reduce these artifacts. To that end, we took a step back and built a perceptual sandbox that allowed experimentation on various foveation techniques without considering performance impact. This led to development of a *perceptual target image* that we validated as perceptually closer to non-foveated results inside a gaze-tracked HMD, using user studies as well as frequency and contrast analysis. The key ideas behind our perceptual target are use of a blur with filter width varying with retinal eccentricity to avoid temporal and spatial aliasing and a post-process contrast enhancement to normalize filtered image contrast.

New Foveated Renderer. With a new perceptual target, we designed a practical foveated rendering system to match our target that reduces the rendering workload. Frequency analysis shows our system closely matches our perceptual target and user studies verify it has measurable quality improvements, allowing coarsened shading up to 30° closer to the fovea than Guenter et al. [2012] without objectionable perceptual impact.

Key design decisions in our foveated renderer include keeping visibility sampling rate constant, **use of pre-filtered shading** terms wherever possible, including using texture mipmapping, LEAN mapping to filter normals, normal distribution function (NDF) filtering to reduce specular aliasing [Kaplanyan et al. 2016], and exponential variance shadow maps [Lauritzen et al. 2011]. A **postprocess contrast enhancement** normalizes for lost contrast due to filtering shading attributes. We also focus on temporal stability, using a **new temporal antialiasing algorithm** that supports multi-resolution renderings and avoids gaze-dependent blurring artifacts caused by eye saccades. This provides a $10\times$ reduction in temporal instability, with similar temporal qualities to a non-foveated, temporally filtered renderer.

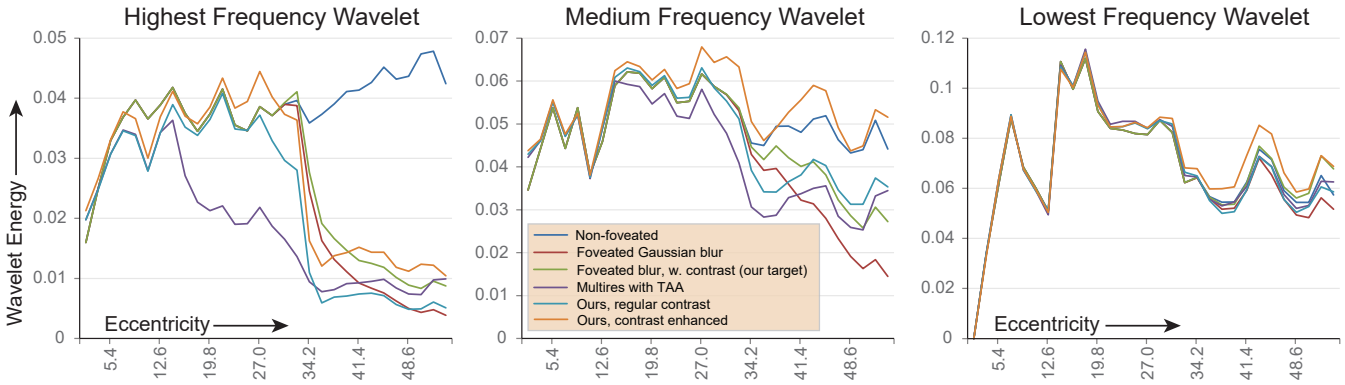

Figure 10: *Wavelet analysis of foveated rendering in Sponza, showing wavelet energy as a function of retinal eccentricity (in degrees). We show three different wavelet coefficients, representing details in (left) 2×2 , (center) 8×8, and (right) 32×32 pixel windows. All foveated techniques lose high frequency details in the far periphery and maintain low frequencies throughout the view. With contrast enhancement, our system better matches the high frequency curve of our perceptual target than Guenter et al. [2012], and nearly matches mid-range frequencies of a non-foveated renderer. Without contrast enhancement, our system has a perceptible loss in energy, especially at high frequencies.*

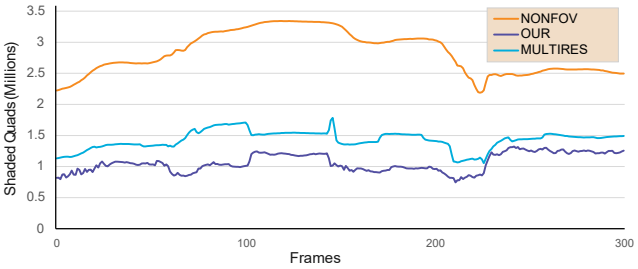

Figure 11: *Number of shaded pixel quads per frame in the San Miguel trace. Our method exhibits the lowest shading rate without impacting the visibility rate.*

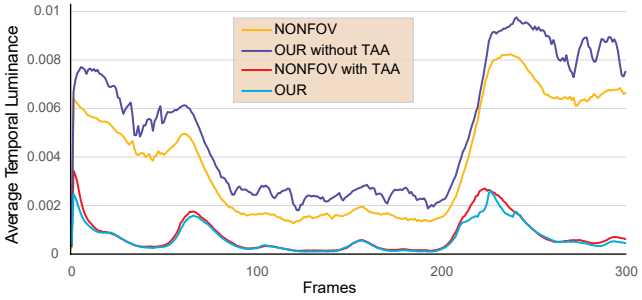

Figure 12: *Average temporal stability for the San Miguel trace. Our multiscale TAA algorithm enables to almost completely eliminate temporal artifacts introduced by severe undersampling in the periphery of the image.*

We also demonstrate our foveated rendering system reduces rendering costs, shading up to 70% fewer pixel quads without significant perceptual degradation.

Limitations. We designed our system to match a perceptually-validated target rather than optimizing for highest performance on current hardware. This means our quality is demonstrably better, but hardware changes are necessary to realize the performance gains described in our system.

Poor blink detection in the gaze tracker can introduce artifacts when blinking due to brief loss of gaze tracking. We expect this to improve with improved gaze tracking hardware and software.

From a rendering perspective, temporal filtering can introduce ghosting and flickering, especially when trying to compensate for severely undersampled signals. Handling high frequency visibility due to alpha tests is not filterable, so we disable alpha testing in our experiments; a high quality pre-filterable alpha test may be interesting future work. Additionally, choosing Vaidyanathan et al. [2014] as a base system limits how far we can coarsen shading, due to triangle sizes and block artifacts.

Future Work. While we develop *a* perceptual target, we do not claim it as an ideal target. We believe future improvements to the perceptual target may provide additional insights enabling further reductions in shading work.

Due to limitations in our prototype gaze-tracker and HMD, our conclusions have only been validated up to 110° field of views (i.e. maximum eccentricity of 55°). While we also validated on a desktop-based eye-tracker, experimentation on wide field of view HMDs may provide further insights to reduce work in the far periphery.

Exploring decoupled shading methods including those in texture-space is also an attractive direction for future work. Finally, an open problem for variable-rate rendering is the need to accurately compute visibility. Exploring methods to pre-filter visibility would provide benefits in many rendering contexts, not just foveated rendering.

Acknowledgments

The authors would like to thank Jaakko Lehtinen, Kayvon Fatahalian, and Tim Foley for their helpful discussions and insights during our experiments as well as in writing this manuscript. Thanks also to Prof. Marty Banks for background discussion and feedback on our literature survey.

We are grateful to SMI for providing HMDs with integrated high-speed eye-trackers. Thanks to Frank Meinl and Efgeni Bischoff for remodeling *Crytek Sponza*, Guillermo M. Leal Llaguno for *San Miguel* model Christophe Seux for the original *Classroom* scene. We downloaded the *Crytek Sponza* from McGuire's repository [2011]. Thanks also to Bill Wagner and Gavriil Pasqualino for their help in developing our assets.

Finally, thanks to the participants of our user studies, and to the reviewers for their valuable feedback in improving this manuscript.

References

BAKER, D., 2016. Object space lighting – following film rendering 2 decades later in real time, 03. Game Developers Conference Talk.

BANKS, M. S., SEKULER, A. B., AND ANDERSON, S. J. 1991. Peripheral spatial vision: limits imposed by optics, photoreceptors, and receptor pooling. *Journal of the Optical Society of America A 8*, 11, 1775–1787.

BANKS, M. S., GEPSHTEIN, S., AND LANDY, M. S. 2004. Why is spatial stereoresolution so low? *The Journal of Neuroscience 24*, 9, 2077–2089.

BENTY, N., 2016. The Falcor rendering framework, 08. https://github.com/NVIDIA/Falcor.

CLARBERG, P., TOTH, R., HASSELGREN, J., NILSSON, J., AND AKENINE-MÖLLER, T. 2014. Amfs: adaptive multi-frequency shading for future graphics processors. *ACM Transactions on Graphics 33*, 4, 141:1–141:12.

COWEY, A., AND ROLLS, E. T. 1974. Human cortical magnification factor and its relation to visual acuity. *Experimental Brain Research 21*, 5, 447–454.

CURCIO, C. A., AND ALLEN, K. A. 1990. Topography of ganglion cells in human retina. *Journal of Comparative Neurology 300*, 1, 5–25.

CURCIO, C. A., SLOAN, K. R., KALINA, R. E., AND HENDRICKSON, A. E. 1990. Human photoreceptor topography. *Journal of Comparative Neurology 292*, 4, 497–523.

FERREE, C. E., RAND, G. G., AND HARDY, C. C. 1931. Refraction for the peripheral field of vision. *Archives of Ophthalmology 5*, 5, 717–731.

GREEN, C. 2007. Improved alpha-tested magnification for vector textures and special effects. In *ACM SIGGRAPH Courses*, SIGGRAPH, 9–18.

GRUNDLAND, M., VOHRA, R., WILLIAMS, G. P., AND DODGSON, N. A. 2006. Cross Dissolve Without Cross Fade: Preserving Contrast, Color and Salience in Image Compositing. *Computer Graphics Forum 25*, 3, 577–586.

GUENTER, B., FINCH, M., DRUCKER, S., TAN, D., AND SNYDER, J. 2012. Foveated 3D graphics. *ACM Transactions on Graphics 31*, 6, 164:1–164:10.

HANSEN, T., PRACEJUS, L., AND GEGENFURTNER, K. R. 2009. Color perception in the intermediate periphery of the visual field. *Journal of Vision 9*, 4, 26:1–26:12.

HE, Y., GU, Y., AND FATAHALIAN, K. 2014. Extending the graphics pipeline with adaptive, multi-rate shading. *ACM Transactions on Graphics 33*, 4, 142:1–142:12.

HILL, S., MCAULEY, S., BURLEY, B., CHAN, D., FASCIONE, L., IWANICKI, M., HOFFMAN, N., JAKOB, W., NEUBELT, D., PESCE, A., AND PETTINEO, M. 2015. Physically based shading in theory and practice. In *ACM SIGGRAPH Courses*, SIGGRAPH, 22:1–22:8.

HILLESLAND, K. E., AND YANG, J. C. 2016. Texel Shading. In *EG 2016 - Short Papers*, The Eurographics Association, T. Bashford-Rogers and L. P. Santos, Eds.

JIMENEZ, J., ECHEVARRIA, J. I., SOUSA, T., AND GUTIERREZ, D. 2012. SMAA: Enhanced morphological antialiasing. *Computer Graphics Forum (Proc. EUROGRAPHICS 2012) 31*, 2.

KAPLANYAN, A., HILL, S., PATNEY, A., AND LEFOHN, A. 2016. Filtering distributions of normals for shading antialiasing. In *Proceedings of the Symposium on High-Performance Graphics*.

KARIS, B. 2014. High-quality temporal supersampling. In *Advances in Real-Time Rendering in Games, SIGGRAPH Courses.*

KELLY, D. H., AND SAVOIE, R. E. 1973. A study of sine-wave contrast sensitivity by two psychophysical methods. *Perception & Psychophysics 14*, 2, 313–318.

KELLY, D. H. 1984. Retinal inhomogeneity. i. spatiotemporal contrast sensitivity. *Journal of the Optical Society of America A 1*, 1, 107–113.

KIM, M. H., RITSCHEL, T., AND KAUTZ, J. 2011. Edge-aware color appearance. *ACM Transactions on Graphics 30*, 2, 13:1–13:9.

KOENDERINK, J. J., BOUMAN, M. A., BUENO DE MESQUITA, A. E., AND SLAPPENDEL, S. 1978. Perimetry of contrast detection thresholds of moving spatial sine patterns. II. The far peripheral visual field (eccentricity 0 degrees-50 degrees). *Journal of the Optical Society of America A 68*, 6, 850–854.

KOENDERINK, J. J., BOUMAN, M. A., BUENO DE MESQUITA, A. E., AND SLAPPENDEL, S. 1978. Perimetry of contrast detection thresholds of moving spatial sine wave patterns. I. The near peripheral visual field (eccentricity 0 degrees-8 degrees). *Journal of the Optical Society of America A 68*, 6, 845–849.

KOENDERINK, J. J., BOUMAN, M. A., BUENO DE MESQUITA, A. E., AND SLAPPENDEL, S. 1978. Perimetry of contrast detection thresholds of moving spatial sine wave patterns. III. The target extent as a sensitivity controlling parameter. *Journal of the Optical Society of America A 68*, 6, 854–860.

LAURITZEN, A., SALVI, M., AND LEFOHN, A. 2011. Sample distribution shadow maps. In *Symposium on Interactive 3D Graphics and Games*, 97–102.

LEVI, D. M., KLEIN, S. A., AND AITSEBAOMO, P. 1985. Vernier acuity, crowding and cortical magnification. *Vision Research 25*, 7, 963–977.

LEVITT, H. 1971. Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical society of America 49*, 2B, 467–477.

MCGUIRE, M., 2011. Computer graphics archive, August. http://graphics.cs.williams.edu/data.

MCKEE, S. P., AND NAKAYAMA, K. 1984. The detection of motion in the peripheral visual field. *Vision Research 24*, 1, 25–32.

MÄKELÄ, P., NÄSÄNEN, R., ROVAMO, J., AND MELMOTH, D. 2001. Identification of facial images in peripheral vision. *Vision Research 41*, 5, 599–610.

NAVARRO, R., ARTAL, P., AND WILLIAMS, D. R. 1993. Modulation transfer of the human eye as a function of retinal eccentricity. *Journal of the Optical Society of America A 10*, 2, 201–212.

NOORLANDER, C., KOENDERINK, J. J., OLDEN, R. J. D., AND EDENS, B. W. 1983. Sensitivity to spatiotemporal colour contrast in the peripheral visual field. *Vision Research 23*, 1, 1–11.

OLANO, M., AND BAKER, D. 2010. Lean mapping. In *Symposium on Interactive 3D Graphics and Games*, 181–188.

ÖZTIRELI, A. C., AND GROSS, M. 2015. Perceptually based downscaling of images. *ACM Transactions on Graphics 34*, 4, 77:1–77:10.

PATNEY, A., KIM, J., SALVI, M., KAPLANYAN, A., WYMAN, C., BENTY, N., LEFOHN, A., AND LUEBKE, D. 2016. Perceptually-based foveated virtual reality. In *ACM SIGGRAPH 2016 Emerging Technologies*, ACM, New York, NY, USA, SIGGRAPH '16, 17:1–17:2.

PHARR, M., AND HUMPHREYS, G. 2010. *Physically Based Rendering, Second Edition: From Theory to Implementation*, 2nd ed. Morgan Kaufmann Publishers, Inc.

ROSÉN, R. 2013. *Peripheral Vision: Adaptive Optics and Psychophysics*. PhD thesis, Royal Institute of Technology, Stockholm, Sweden.

ROVAMO, J., AND VIRSU, V. 1979. An estimation and application of the human cortical magnification factor. *Experimental Brain Research 37*, 3, 495–510.

ROVAMO, J., VIRSU, V., LAURINEN, P., AND HYVÄRINEN, L. 1982. Resolution of gratings oriented along and across meridians in peripheral vision. *Investigative Ophthalmology & Visual Science 23*, 5, 666–670.

SALVI, M., AND VAIDYANATHAN, K. 2014. Multi-layer alpha blending. In *Proceedings of the 18th meeting of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games*, 151–158.

SCHÜTT, H. H., HARMELING, S., MACKE, J. H., AND WICHMANN, F. A. 2016. Painfree and accurate bayesian estimation of psychometric functions for (potentially) overdispersed data. *Vision Research 122*, 105 – 123.

SOLOMON, S. G., LEE, B. B., WHITE, A. J., RUTTIGER, L., AND MARTIN, P. R. 2005. Chromatic organization of ganglion cell receptive fields in the peripheral retina. *Journal of Neuroscience 25*, 18, 4527–4539.

STRASBURGER, H., RENTSCHLER, I., AND HARVEY, L. O. 1994. Cortical magnification theory fails to predict visual recognition. *European Journal of Neuroscience 6*, 10, 1583–1588.

STRASBURGER, H., RENTSCHLER, I., AND JÜTTNER, M. 2011. Peripheral vision and pattern recognition: A review. *Journal of Vision 11*, 5, 13:1–13:82.

SWAFFORD, N. T., IGLESIAS-GUITIAN, J. A., KONIARIS, C., MOON, B., COSKER, D., AND MITCHELL, K. 2016. User, metric, and computational evaluation of foveated rendering methods. In *Proceedings of the ACM Symposium on Applied Perception*, ACM, New York, NY, USA, SAP '16, 7–14.

THIBOS, L. N., CHENEY, F. E., AND WALSH, D. J. 1987. Retinal limits to the detection and resolution of gratings. *Journal of the Optical Society of America A 4*, 8, 1524–1529.

THIBOS, L., WALSH, D., AND CHENEY, F. 1987. Vision beyond the resolution limit: Aliasing in the periphery. *Vision Research 27*, 12, 2193–2197.

THIBOS, L. N., STILL, D. L., AND BRADLEY, A. 1996. Characterization of spatial aliasing and contrast sensitivity in peripheral vision. *Vision Research 36*, 2, 249–258.

THIBOS, L. N. 1987. Calculation of the influence of lateral chromatic aberration on image quality across the visual field. *Journal of the Optical Society of America A 4*, 8, 1673–1680.

TOTH, R., NILSSON, J., AND AKENINE-MOLLER, T. 2016. Comparison of projection methods for rendering virtual reality. In *Proceedings of the Symposium on High-Performance Graphics*.

VAIDYANATHAN, K., SALVI, M., TOTH, R., FOLEY, T., AKENINE-MOLLER, T., NILSSON, J., MUNKBERG, J., HASSELGREN, J., SUGIHARA, M., CLARBERG, P., JANCZAK, T., AND LEFOHN, A. 2014. Coarse pixel shading. In *Proceedings of the Symposium on High-Performance Graphics*.

WANDELL, B. A. 1995. *Foundations of Vision*. Sinauer Associates, Inc.

WANG, Y.-Z., THIBOS, L. N., AND BRADLEY, A. 1996. Undersampling produces non-veridical motion perception, but not necessarily motion reversal, in peripheral vision. *Vision Research 36*, 12, 1737–1744.

WANG, Y.-Z., BRADLEY, A., AND THIBOS, L. N. 1997. Aliased frequencies enable the discrimination of compound gratings in peripheral vision. *Vision Research 37*, 3, 283–290.

WICHMANN, F. A., AND HILL, N. J. 2001. The psychometric function: I. fitting, sampling, and goodness of fit. *Perception & Psychophysics 63*, 8, 1293–1313.

WICHMANN, F. A., AND HILL, N. J. 2001. The psychometric function: Ii. bootstrap-based confidence intervals and sampling. *Perception & Psychophysics 63*, 8, 1314–1329.

WILLIAMS, D. R., ARTAL, P., NAVARRO, R., MCMAHON, M. J., AND BRAINARD, D. H. 1996. Off-axis optical quality and retinal sampling in the human eye. *Vision Research 36*, 8, 1103–1114.

WILLIAMS, L. 1983. Pyramidal parametrics. *SIGGRAPH Comput. Graph. 17*, 3, 1–11.

YANG, L., NEHAB, D., SANDER, P. V., SITTHI-AMORN, P., LAWRENCE, J., AND HOPPE, H. 2009. Amortized supersampling. In *ACM SIGGRAPH Asia 2009 Papers*, ACM, New York, NY, USA, SIGGRAPH Asia '09, 135:1–135:12.