

# Visual Acuity Consistent Foveated Rendering Towards Retinal Resolution

Zhi Zhang, Meng Gai, and Sheng Li , *Member, IEEE*

**Abstract**—Prior foveated rendering methods often suffer from a limitation where the shading load escalates with increasing display resolution, leading to decreased efficiency, particularly when dealing with retinal-level resolutions. To tackle this challenge, we begin with the essence of the human visual system (HVS) perception and present visual acuity-consistent foveated rendering (VaFR), aiming to achieve exceptional rendering performance at retinal-level resolutions. Specifically, we propose a method with a novel log-polar mapping function derived from the human visual acuity model, which accommodates the natural bandwidth of the visual system. This mapping function and its associated shading rate guarantee a consistent output of rendering information, regardless of variations in the display resolution of the VR HMD. Consequently, our VaFR outperforms alternative methods, improving rendering speed while preserving perceptual visual quality, particularly when operating at retinal resolutions. We validate our approach using both the rasterization and ray-casting rendering pipelines. We also validate our approach using different binocular rendering strategies for HMD devices. In diverse testing scenarios, our approach delivers better perceptual visual quality than prior foveated rendering while achieving an impressive speedup of  $6.5\times - 9.29\times$  for deferred rendering of 3D scenarios and an even more powerful speedup of  $10.4\times - 16.4\times$  for ray-casting at retinal resolution. Additionally, our approach significantly enhances the rendering performance of binocular 8 K path tracing, achieving smooth frame rates.

**Index Terms**—Foveated rendering, retinal resolution, visual acuity, shading rate, log-polar mapping.

## I. INTRODUCTION

**I**N VIRTUAL reality (VR) systems, rendering performance is critical for creating immersive experiences, with modern systems requiring high-quality rendering at frame rates above 90 Hz [22]. However, performance drops can disrupt immersion. Foveated rendering addresses these challenges by optimizing resource allocation based on human visual characteristics,

Received 7 April 2024; revised 27 January 2025; accepted 1 February 2025. Date of publication 6 February 2025; date of current version 5 September 2025. This work was supported by the National Key R&D Program of China under Grant 2022ZD0160805, and in part by the NSFC of China under Grant 62172013. Recommended for acceptance by L. Wang. (*Corresponding author: Sheng Li.*)

Zhi Zhang is with the School of Software and Microelectronics, Peking University, Beijing 100871, China (e-mail: zhangzhi543@stu.pku.edu.cn).

Meng Gai is with the School of Computer Science, Peking University, Beijing 100871, China (e-mail: gaimeng@pku.edu.cn).

Sheng Li is with the School of Computer Science, Peking University, Beijing 100871, China, and also with the National Key Laboratory of Intelligent Parallel Technology, Peking University, Beijing 100871, China (e-mail: lisheng@pku.edu.cn).

This article has supplementary downloadable material available at <https://doi.org/10.1109/TVCG.2025.3539851>, provided by the authors.

Digital Object Identifier 10.1109/TVCG.2025.3539851

thereby improving rendering efficiency. While the human visual field spans approximately  $135^\circ$  vertically and  $160^\circ$  horizontally, high visual acuity is concentrated in the fovea, a small central region covering just  $1.5^\circ$  vertically and  $2^\circ$  horizontally [12]. By focusing computational power on this foveal region, foveated rendering enhances clarity where it is most needed and reduces resources allocated to peripheral areas, lowering overall computational demands. Additionally, real-time gaze tracking allows rendering engines to dynamically adjust detail levels, optimizing fidelity in critical areas.

Log-polar mapping, a transformation that converts Cartesian coordinate into a logarithmic-polar coordinate system, aligns the display space with human visual perception and can be used to approximate cortical excitation [2]. The classic log-polar transformation for 2D image foveation on GPUs leverages the non-uniform distribution of visual acuity, emphasizing detail in the foveal region [1]. As representatives, the log-polar mapping foveated rendering framework (LMFR) [23] linearly reduces the resolution (inverse of the shading rate) outward from the fovea, employing a log-polar (LP) buffer for shading, while the locomotion-aware foveated rendering method (LaFR) [29] dynamically adjusts the average shading rate based on user locomotion patterns. However, these methods generally enhance performance only at lower resolutions (typically below 4 K binocular display). As display resolution increases, the LP buffer size and shading load grow accordingly, leading to performance declines. Moreover, they rely on empirically determined model parameters attached to each HMD device's resolution, often requiring manual tuning for optimal performance.

Human visual acuity near the fovea is extremely sharp, reaching an angular resolution of 1 arcminute (about 60 PPD), which in a  $120^\circ$  field of view could equate to a resolution of  $7200 \times 7200$  pixels [48]. Achieving such high resolutions, even at retinal levels, poses significant challenges for prior methods. To address this, we introduce the visual acuity-consistent foveated rendering (VaFR) technique, designed for efficient rendering at the high resolutions of next-generation HMDs.

Our approach derives a novel log-polar mapping function based on human visual acuity, independent of the device display resolution. Using cortical magnification as a theoretical basis, we scale the log-polar transformation without relying on empirically determined model parameters. Additionally, we introduce variable scaling of the tangent plane in the log-polar coordinate system, ensuring that resolution on the log-polar tangent axis stays within human visual acuity limits.

Our VaFR decouples model parameters from device-specific display resolutions, delivering performance gains without compromising perceptual fidelity. By maintaining a constant-size LP

buffer across resolutions, VaFR removes the need for resolution-specific parameter tuning required by previous methods, enabling consistent and efficient rendering aligned with human visual acuity. Additionally, our approach is adaptable to various VR rendering pipelines, achieving over 131 fps at  $11520 \times 6480$  resolution per eye.

In summary, our contributions include:

- We developed a visual acuity-based log-polar mapping function and derived optimal settings for highly efficient foveated rendering independent of device-specific display resolutions (see Fig. 1), eliminating the need for empirically determined model parameters.
- We proposed a method for computing shading rates in log-polar algorithms, allowing independent adjustment of tangential and radial rates to align with visual acuity and assess perceptual compatibility.
- Our method ensures efficient rendering across various VR displays, especially at near-retinal resolutions, paving the way for ultra-high-definition virtual reality experiences in the future.

## II. RELATED WORK

### A. Foveated Rendering

Foveated rendering methods perform high-resolution sampling for foveal regions and significant user-noticeable areas and low-resolution sampling for peripheral regions.

*Rasterization Paradigm:* In recent years, due to growing 3D model complexity and larger virtual scenes, researchers are increasingly exploring foveated rendering and enhancing geometric mesh rasterization for multi-spatial resolution. Guenter et al. [12] introduced foveated rasterization, finely controlling rendering quality across the foveal, transitional, and peripheral regions. Stengel et al. [30] introduced a sampling method based on the visual acuity fall-off model, spatio-temporal and spatio-luminance CSFs, and subsequently integrated the sampling method into the deferred rendering pipeline. Turner et al. [38] aligned the rendered pixel grid with the content of the virtual scene during rasterization and upsampling to reduce motion artifacts. To enhance speed while maintaining perceptual quality, Patney et al. [26] presented a foveated shading system, drastically reducing shading computation while introducing an anti-aliasing algorithm for peripheral detail recovery. Friston et al. [10] proposed a unified foveated rasterization pipeline for HMDs, employing per-fragment ray casting for VR rendering efficiency. Foveated rendering has been diversified to address various aspects, including accelerating shadow rendering [47] and improving peripheral rendering quality with pixel reuse techniques through reprojection [9], a variable-rate shading pipeline that partitions the image into tiles and dynamically modifies the shading precision to enhance the rasterization performance [13], post-processing methods [41], and foveated rendering via rectangular mapping [46].

*Ray Tracing Paradigm:* Ray tracing allows control over the number of rays per pixel, enhancing rendering quality and naturally supporting spatial multi-resolution rendering. Fujita et al. [11] pioneered foveated ray tracing, using pre-computed sampling maps and K-nearest neighbor algorithms to reconstruct

sparse samples for final image generation. Building on this, Weier et al. [42] incorporated reprojection techniques, recycling ray samples from previous frames. Koskela et al. [16], [17], [18] introduced progressive Monte Carlo path tracing to accelerate rendering in regions of interest. Willberger et al. [44] introduced a hybrid path tracing method to accelerate global illumination in foveated rendering. Further advances include luminance-contrast-aware sampling for optimized ray tracing [39], spatio-temporal reservoir resampling for efficient light sample selection [4], and adaptive selective sampling tailored for HMDs combined with foveated rendering [15].

*Other Extensions:* Recent foveated rendering research explores various innovative approaches. Sun et al. [32] achieved perceptual quality by using fewer rays through 4D light field foveated rendering. Other advances include GAN-based methods to improve the quality of foveated images and videos in peripheral regions [14], as well as NeRF-based foveated rendering [6].

Meng et al. [23] introduced LMFR, which transforms screen-space coordinates into the log-polar space for efficient lighting and emphasizes the importance of the kernel function selection in improving peripheral perception and image quality. X. Shi et al. [29] extended LMFR with LaFR, dynamically adjusting parameters based on user locomotion patterns. Fan et al. [7] stretches the image after the variant log-polar mapping to fill the buffer and then stretch the mapped pixels based on the visual importance map of the scene.

Both LMFR and LaFR employ kernel functions to transform the Cartesian coordinates into a log-polar system, creating a lower-resolution log-polar (LP) buffer for shading. These methods rely on the heuristic selection of kernel functions and a parameter  $\delta$ , where  $\delta^2$  represents the ratio of full-resolution screen pixels to LP buffer pixels, typically  $\delta \approx 2$  [23], [29] based on trial, error, and user studies. Shading occurs at a resolution approximately  $\delta^2 \approx 4$  times lower than the native display resolution, making these approaches efficient for displays up to 4K per eye. However, as display resolution increases, so do the LP buffer size and shading load, particularly at retinal-level resolutions, resulting in significant performance drops and resource inefficiency. While these methods demonstrate empirical effectiveness, they lack a robust theoretical foundation.

We aim to develop more intuitive and mathematically sound kernel functions to enhance understanding of peripheral rendering. Our investigation examines how the mapping function affects shading point distribution, particularly in relation to human visual acuity. This highlights the importance of aligning the shading distribution with visual acuity for accurate and cohesive rendering. However, our research found that using kernel functions to redistribute shading points is impractical and inefficient. Consequently, we have moved away from traditional kernel functions.

### B. Visual Acuity

The human visual system consists of key components, including the eye's optical elements, retinal structures like photoreceptors, and neural processing in the visual cortex [31]. The differences between foveal and peripheral vision are mainly due to these elements. Peripheral vision has a lower image

quality due to refractive lens effects [8], [35], [37]. Beyond  $10^\circ$  eccentricity, visual signals experience defocus, astigmatism, and chromatic aberration. In the fovea, cones and ganglion cells connect one-to-one, preserving spatial detail, while in the periphery, they connect one-to-many, effectively filtering information.

Early studies of peripheral vision indicated that increasing the stimulus size as eccentricity results in a similar perceptual appearance. This phenomenon is linked to the relative neural volume across various regions, known as the cortical magnification factor, represented by  $M$ . In general, this factor scales as an inverse linear function, i.e.  $M^{-1} = M_0^{-1} \cdot (1 + \alpha E)$ , where  $E$  denotes the eccentricity angle,  $M_0$  represents foveal magnification [5] and  $\alpha$  is the slope of this linear function. In contrast, the non-linear magnification model was also investigated [27]. In the early stages, Levoy et al. [20] combined the ray casting method used for volume rendering with the visual acuity fall-off model. Some studies have integrated visual acuity fall-off models with vertex decimation for level-of-detail (LoD) [21], [25]. Guenter et al. [12] simulated acuity reduction using layered rendering around the gaze point. Weier et al. [42] applied the model in ray tracing for HMDs, while Swafford et al. [33] used it for multi-resolution ambient occlusion and tessellation. Tariq et al. [34] enhanced adaptive-resolution foveated rendering by generating image details with spatial frequencies visible but not fully resolvable by the human eye.

Building upon the cortical magnification function, we aim to distribute shading points from the LP buffer in alignment with human visual acuity. Using the idea of log-polar mapping without employing the kernel function and the linear cortical magnification function, we will derive a mapping function that differs in structure from the traditional log-polar mapping approach.

### III. VISUAL ACUITY CONSISTENT FOVEATED RENDERING

First, we analyze the fundamental deficiencies of the prior methods as preliminary (Section III-A). As a solution, we leverage human visual acuity models to derive a novel log-polar mapping function (Section III-B), which maintains a constant-sized LP buffer irrespective of resolution variations, providing a constant stream of visual information consistent with the finite bandwidth of the human brain. Following the segmented nature of the mapping function, a shading rate adaption is proposed to automatically adjust the shading rate for different devices and for further custom tuning (Section III-C). Lastly, we present the pipeline of our algorithm in Section III-D.

#### A. Preliminary

LMFR [23] and LaFR [29] aimed to reduce shading load by lowering the LP buffer resolution. They used an empirical mapping function, fine-tuning the  $K$  function and parameters to adjust the pixel distribution and optimize VR visuals. However, as later analysis shows, their approach essentially tries to fit the shading rate to the human visual acuity model, a case of putting the cart before the horse.

Specifically, the initial step of LMFR and LaFR involves converting the screen-space frame buffer pixel coordinate  $(x, y)$  into  $(u, v)$  in the LP buffer in the log-polar coordinate

system as:

$$\begin{cases} u = K\left(\frac{\ln\sqrt{(x-x_0)^2+(y-y_0)^2}}{L}\right) \cdot w \\ v = \left(\frac{1}{2\pi} \tan^{-1} \frac{y-y_0}{x-x_0} + 1[y-y_0 < 0]\right) \cdot h \end{cases} \quad (1)$$

Here,  $(x_0, y_0)$  represents the position of the gaze point in the screen space,  $L$  denotes the longest log distance from the gaze point to the corner of the screen,  $(w, h)$  denotes the size of the LP buffer,  $K(\cdot)$  denotes the kernel function, and  $1[\cdot]$  serves as the indicator function.  $u$  denotes the horizontal ordinate of the LP buffer, while  $v$  represents the vertical ordinate. However, importantly, the kernel function  $K(z) = z^\alpha$  in LMFR [23] and  $K(z) = \frac{z^\alpha}{1-\alpha}$  in LaFR [29] rely on empirical choices, limiting the flexibility of the mapping function and the broader applicability of their methods. Here,  $z$  is short for  $\frac{\ln\sqrt{(x-x_0)^2+(y-y_0)^2}}{L}$  and  $\alpha$  is an empirically determined model parameter. Sections IV-C and IV-D will further elaborate on this issue, highlighting that dynamically aligning shading rates with human visual acuity under changing VR conditions presents even greater challenges.

#### B. Our Formulation

We establish a theoretical foundation for precise shading rate alignment in foveated rendering by deriving a mapping function directly from visual acuity. This ensures a strict correlation between shading rate distribution and visual acuity. Specifically, our mapping functions (5) and (8) rigorously align both radial and tangential shading rates with the acuity model, optimizing rendering workload distribution. While prior methods also seek to optimize shading rate distribution to improve efficiency, they often over- or under-allocate shading rates in certain regions, leading to unnecessary resource consumption. By contrast, our approach effectively eliminates these inefficiencies, ensuring the optimal utilization of rendering resources in terms of shading rate.

When expressed in terms of the minimum angle of resolution (MAR), the reciprocal of visual acuity, a linear model aligns well with anatomical data, such as receptor density and performance outcomes in numerous low-level vision tasks [31], [43]. Acuity models serve as the foundation for the widely recognized theory of cortical magnification or M-scaling, often using a linear model based on the reciprocal of the visual cortex tissue volume allocated to each eccentricity slice.

1) *Mapping Function*: We illustrate the log-polar mapping and its inverse mapping between the screen space and the LP space in Fig. 2, along with the key variables. To derive the mapping function, we begin with the linear model:

$$\omega(e) = me + \omega_0 \quad (2)$$

where  $\omega$  denotes MAR measured in degrees per cycle,  $e$  refers to the eccentricity angle in degree,  $\omega_0$  stands for the smallest resolvable angle that corresponds to the reciprocal of visual acuity at the foveal region ( $e = 0$ ), and  $m$  signifies the slope of the MAR.  $f$  is acuity defined as the reciprocal of  $\omega$  as  $f = \frac{1}{\omega}$ , expressed in cycles per degree (cpd) with each cycle corresponding to two pixels.

To determine a constant ratio  $c_r$  for a known camera and viewport in rendering, divide the camera's film plane height (in

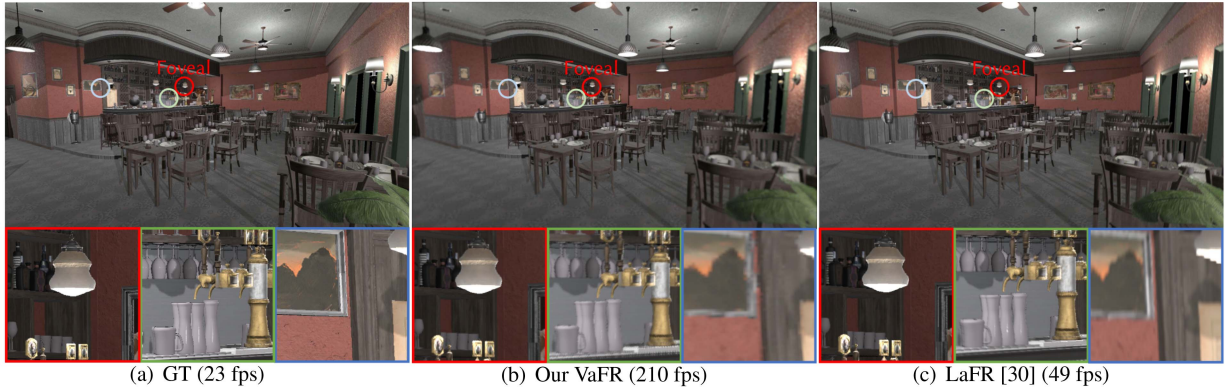


Fig. 1. Results of our visual acuity consistent foveated rendering (VaFR) in the middle, ground truth (GT) on the left, and locomotion-aware foveated rendering (LaFR [29]) on the right, rendered in ray casting mode at 8 K ( $7680 \times 4320$ ) resolution. Our approach, achieving an extraordinarily high frame rate of 210 fps for binocular rendering in VR settings, delivers comparable perceptual quality to that of LaFR (49 fps) and GT (23 fps).

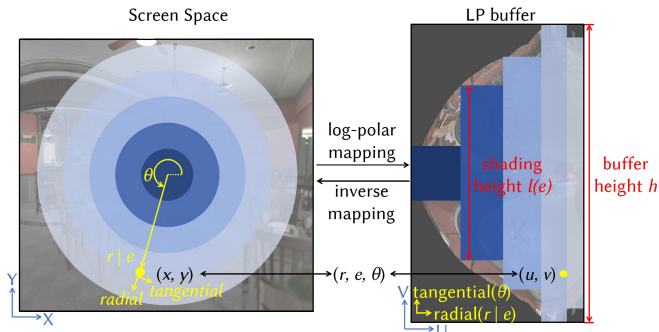


Fig. 2. Schematic diagram of log-polar mapping and its inverse between the screen space and the log-polar space associated with the LP buffer, along with key variables in Section III-B1. Additionally, we highlight the variable shading height  $l(e)$ , which differs from the buffer height  $h$ .

millimeters) by its focal length (also in millimeters). Then, divide this result by the display height in pixels of the native display resolution. The resulting ratio,  $c_r$ , represents the reciprocal of the pixel length corresponding to the focal length and simplifies the conversion equation between  $e$  and  $r$ . It can be used in (3):

$$r = \frac{\tan e}{c_r} = \sqrt{(x - x_0)^2 + (y - y_0)^2} \quad (3)$$

where  $e$  denotes the eccentricity angle in degrees, while  $r$  refers to the pixel distance from any given point  $(x, y)$  to the gaze point  $(x_0, y_0)$ . Therefore, we can infer:  $e = \tan^{-1}(c_r \cdot r)$  from pixel distance  $r$ . Importantly, note that all trigonometric functions in the context are measured in degrees.

We have  $f(e)$  from (2) as:  $f(e) = \frac{1}{\omega(e)} = \frac{1}{me + \omega_0}$ . Subsequently, each pixel in the log-polar space corresponds to half a cycle. Consequently, to ensure proper perception, each unit eccentricity angle should be shaded by  $2f(e)$  pixels. Furthermore,  $\frac{du}{de}$  indicates how many pixels in the LP buffer correspond to one unit of eccentricity angle, expressed as:

$$\frac{du}{de} = 2f(e) = \frac{2}{me + \omega_0} \quad (4)$$

The significance of this equation is that each particular pixel in the LP buffer is mapped to a specific eccentricity angle, regardless of changes in gaze orientation or display resolution. This characteristic is a fundamental distinction between our algorithm

and others. Accordingly, the shading rate can be defined as the actual number of shaded cycles in the log-polar space per unit of eccentricity angle. Specifically, the radial shading rate is the number of shaded cycles per unit of radial eccentricity angle, and the tangential shading rate is the number of shaded cycles per unit of tangential eccentricity angle.

Upon solving this equation, the horizontal coordinate  $u$  in the LP buffer can be computed as follows:

$$u(e) = 2 \frac{\ln(me + \omega_0)}{m} + c_0 \quad (5)$$

To ensure the accuracy of the tangential shading rate, our algorithm initially set the tangential equal to the radial shading rates. Therefore, we can infer:

$$r \frac{d\theta}{dv} = \frac{dr}{du} \quad (6)$$

Two sides of this equation denote the ratio of the tangential pixel length in the screen space to the vertical pixel length in the LP buffer, and the ratio of the radial pixel length in the screen space to the horizontal pixel length in the LP buffer, respectively. For detailed information on the adjustment of the tangential shading rates by timing a ratio factor  $\Delta(e)$ , which varies with the eccentricity angle  $e$  (see Section IX).

By differentiating (3), we can get:

$$\frac{de}{dr} = c_r \cos^2 e \cdot \frac{180}{\pi} \quad (7)$$

Upon solving (6) and (7), the vertical coordinate  $v$  in the LP buffer can be computed as:

$$v(e, \theta) = \theta \cdot \frac{\sin 2e}{me + \omega_0} \quad (8)$$

Next, we introduce the term shading height  $l$  in the LP buffer, which refers to the varying number of shaded pixels along the vertical direction, as shown in Fig. 2. In previous studies, every pixel in each column of the LP buffer was shaded, meaning the shading height was always equal to the buffer's total height. This rigid approach resulted in inaccuracies in the tangential shading rate and unnecessary rendering resource consumption. In contrast, our method allows the shading height to adapt to the eccentricity angle  $e$ , as illustrated in Fig. 2, the total height of shaded pixels in each vertical column adjusts accordingly,

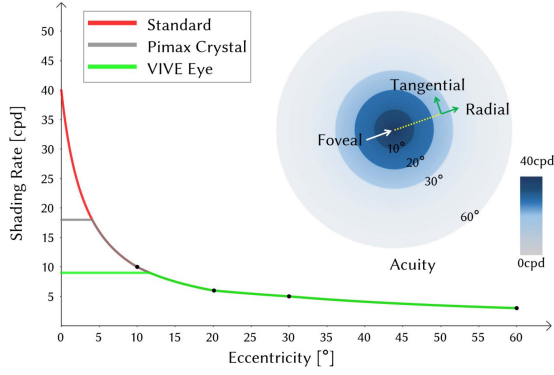


Fig. 3. Our pivotal points for segmented mapping function, Standard shading rate, two curves adapted to VIVE EYE and Pimax Crystal.

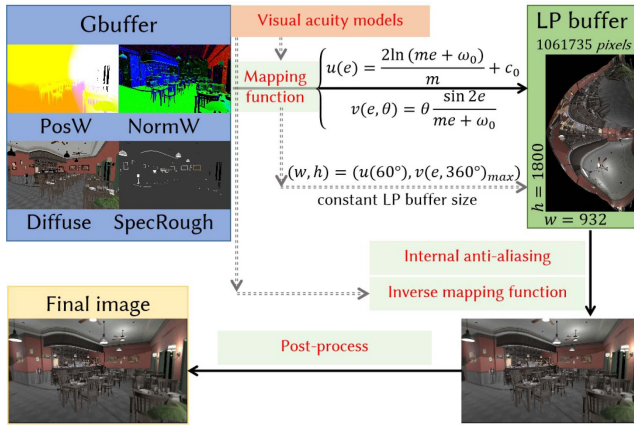


Fig. 4. Working pipeline of our visual acuity consistent foveated rendering in the deferred rendering pipeline. The LP buffer dimensions are  $932 \times 1800$ , yet it includes only 1,061,735 shading points.

reflecting how shading height evolves with changes in the eccentricity angle.

Thus, the shading height at any given eccentricity angle  $e$  is:

$$l(e) = v(e, 360^\circ) = 360 \cdot \frac{\sin 2e}{me + \omega_0} \quad (9)$$

This adaptability enables fine-grained control over the tangential shading rate, setting our algorithm apart from existing approaches. Details of the tangential shading rate in prior methods are presented in Fig. 21, with their shortcomings discussed in Section IX. Unless specifically stated otherwise, all shading rates discussed in the following refer to radial shading rates.

2) *Constant-Sized LP Buffer*: Assuming the maximum perceivable eccentricity as  $e_{max} = 60^\circ$ , a fixed LP buffer size of  $(w, h) = (u_{max}, l_{max}) = (u(60^\circ), v(e, 360^\circ)_{max}) = (932, 1800)$  is determined, where the shading pass is executed. Note that, we calculate the derivative of  $l(e)$  equals zero, obtaining  $e' \approx 45^\circ$  which maximizes  $l(e)$ , and then obtain  $l_{max} \approx 1800$ . From (5) and (9), variables  $u$  and  $v$  cause the shaded image to shape as a semi-ellipse, as illustrated in the LP buffer in Fig. 4.

The variables  $u$  and  $l$  in our algorithm are uniquely linked to the variable  $e$ , while all other variables are set before rendering. This ensures a consistently sized LP buffer that remains unaffected by resolution or parameter changes. In contrast, both the LMFR and LaFR algorithms rely on empirical fine-tuning

through user experiments at specific resolutions to determine the LP buffer size. Since their LP buffer size depends on the display resolution and parameter  $\delta$ , these algorithms can experience erratic shading rate fluctuations under changing conditions. Detailed issues will be discussed in Section IV.

A key feature of our algorithm is that its shading rate is solely determined by human visual acuity, unlike other methods that approximate this rate through coarse, resolution-specific tuning to fit the visual acuity model. This unique characteristic, absent in prior algorithms, ensures a stable shading load across various display resolutions. Consequently, even at extremely high resolutions, such as  $11520 \times 6480$  pixels per eye, our method can maintain robust performance while preserving retinal details in the foveal region. The constant LP buffer size aligns with the intuitive understanding that human visual processing is limited by the brain's finite bandwidth.

3) *Inverse Mapping Function*: To invert the image from the LP buffer to the screen space, we derive the inverse function using the mapping function mentioned above, known as the inverse mapping function. Equations (10) and (11) depict the inverse transformation process of our approach, with the former depicting the calculation of the eccentric angle  $e$  from the LP buffer's horizontal coordinate  $u$ :

$$e = \frac{\exp[(u - c_0) \cdot m/2] - \omega_0}{m} \quad (10)$$

and the derivation of the screen-space position  $x, y$  from  $r$  (see (3)) and  $v$ :

$$\begin{cases} x = r \cos \frac{v}{l(e)} 360^\circ + x_0 \\ y = r \sin \frac{v}{l(e)} 360^\circ + y_0 \end{cases} \quad (11)$$

### C. Shading Rate Adaption

Examining (5) and (8), we find that  $\omega(e)$  is a common element in both. However, recent studies [34], [36] suggest that visual acuity does not follow a single linear path, but is better approximated by segmented functions. Consequently, we identify specific pivot points and assign variable acuity values  $f$  to precisely control the shading rate, enabling adjustments to meet the preferences of user programmers. The detailed calculation of the shading rate and its correlation with the visual acuity will be elucidated in Section IV.

The shading rate adaption offers a range of functions. Firstly, it enables customized shading rate control, allowing designers or users to define a tailored set of parameters for specific scenes or individual eye conditions in advance. Secondly, it supports automatic parameter tuning by detecting the VR HMD in use. Thus, (5) and (8) are reformulated as:

$$\begin{cases} u(e) = 2 \frac{\ln(m_1 e + \omega_1)}{m_1} + c_1, & v(e) = \frac{\theta \cdot \sin 2e}{m_1 e + \omega_1}, & e \in [0, 10] \\ u(e) = 2 \frac{\ln(m_2 e + \omega_2)}{m_2} + c_2, & v(e) = \frac{\theta \cdot \sin 2e}{m_2 e + \omega_2}, & e \in [10, 20] \\ u(e) = 2 \frac{\ln(m_3 e + \omega_3)}{m_3} + c_3, & v(e) = \frac{\theta \cdot \sin 2e}{m_3 e + \omega_3}, & e \in [20, 30] \\ u(e) = 2 \frac{\ln(m_4 e + \omega_4)}{m_4} + c_4, & v(e) = \frac{\theta \cdot \sin 2e}{m_4 e + \omega_4}, & e \in [30, 60] \end{cases} \quad (12)$$

Regarding the inverse transformation, (10) is also divided into four segments.

In our approach, we have defined  $u_{max} = u(60^\circ)$  and  $l_{max} = v(e, 360^\circ)_{max}$  as constants. For our testing purposes, we have

set fixed values for  $f_{0^\circ} = 40 \text{ cpd}$ ,  $f_{10^\circ} = 10 \text{ cpd}$ ,  $f_{20^\circ} = 6 \text{ cpd}$ ,  $f_{30^\circ} = 5 \text{ cpd}$ ,  $f_{60^\circ} = 4 \text{ cpd}$ . These values are determined through a combination of data from [34] and [36] with a series of experimental trials. It's important to note that, at a given eccentricity, there exists a range of frequencies that are perceptible to the observer but not fully resolvable. To further enhance the perception of details, we will discuss Anti-aliasing and compatible Post-process below. Moreover, the value of these pivotal points can be increased to enhance the image quality.

The standard shading rates for our tests, based on the values above, are shown in Fig. 3. These parameters are adjustable, allowing us to design two specific curves for the HTC VIVE EYE and Pimax Crystal headsets, each maintaining a constant foveal shading rate aligned with their maximum display densities. As shown in Fig. 3, the HTC VIVE EYE has a lower display resolution of 9 cpd, while the Pimax Crystal offers a higher resolution of 18 cpd. Consequently, the HTC VIVE EYE curve achieves a 20% reduction in shading load, while the Pimax Crystal curve yields a 8% reduction. Overall, this adaptation provides users and developers with an intuitive and convenient adjustment option.

#### D. Pipeline and Technical Detail

In our pipeline (see Fig. 4), the G-buffer pass runs at native resolution, with extracted data mapped to a fixed-size LP buffer, independent of display resolution. The inverse mapping function then reconstructs the screen-space image at native resolution, with internal anti-aliasing to ensure high quality. Finally, optional post-processing steps, such as temporal anti-aliasing (TAA), tone mapping, and others, can be applied to further enhance image quality.

*Anti-aliasing:* Given the low-resolution nature of the LP buffer, potential artifacts may occur in peripheral regions after the inverse transformation. We employ the Fast Approximate Anti-Aliasing (FXAA) within the LP buffer. In peripheral areas, where a few pixels in the LP buffer may cover a larger number of pixels in the screen space, precise smoothing is essential to minimize aliasing artifacts.

LMFR [23] uses Gaussian filters to reduce peripheral artifacts in the LP buffer, followed by foveal-aware Gaussian filters for anti-aliasing the final output. However, we found that this approach tends to blur the image without significantly enhancing visual quality. Therefore, we opt for FXAA instead.

### IV. SEMANTIC ANALYSIS AND COMPARISONS

In this section, we conduct a semantic analysis and comparison of shading rate curves to identify deficiencies in prior work, along with our advantages. We first establish the methodology for calculating shading rates in log-polar-based methods in Section IV-A. Using parameters from two representative studies, LMFR [23] and LaFR [29], we derive and visualize shading rates from their mapping functions alongside our algorithm's shading rate. In Section IV-B, we evaluate the alignment of each method with visual acuity. Section IV-C examines the adaptability of these methods to various resolutions, emphasizing the generality and extensibility of our approach. Finally, Section IV-D investigates the availability of each method during gaze

shifts, such as saccades or smooth pursuits, in immersive VR environments.

#### A. Shading Rate Analysis

From the linear model illustrated in (2), it is imperative to express the shading rate in cpd as human visual acuity model:

$$SR(e) = f(e) = \frac{1}{2} \times \frac{du}{de} \quad (13)$$

rather than assuming one percent of the unit area in LaFR [29]. Here, a cycle corresponds to two pixels, where a pixel is defined as a valid shading point within the LP buffer. Consequently, the shading rate  $SR(e)$  is determined by the number of cycles per unit angle at any angle  $e$  within the field of view, calculated as  $\frac{du}{2de}$  where  $u$  represents the abscissa of the LP buffer and  $e$  denotes the eccentricity angle.

The mapping equation proposed in LMFR [23] is as:

$$u_{\text{LMFR}}(z) = (z(e))^4 w \quad (14)$$

and the mapping equation proposed in LaFR [29] is

$$u_{\text{LaFR}}(z) = \begin{cases} (z(e)(1 - Fa))^{\frac{1}{F_a}}, & e \in [0, 4.89] \\ \left( \frac{2 \arccos(Za) - \pi}{\pi} \right)^{\frac{\beta}{0.7}} (1 - U_{ef}) + U_{ef}, & e \in (4.89, 55] \end{cases} \quad (15)$$

where  $z$  ranges from 0 to 1, representing the log ratio in the screen-space image as:  $z(e) = \frac{\ln r}{\ln L} = \frac{\ln \frac{\tan e}{c_r}}{\ln L}$ , and  $Fa$  is short for  $f'(a, \alpha)$  as:  $Fa = f'(a, \alpha) = 1 - \log_{Z_{ef}} \frac{\alpha Z_{ef}^{1-a}}{1-a}$ , and a parameter of LaFR for its kernel function is as:  $K_f(\alpha, z(e)) = \frac{z(e)^{f'(a, \alpha)}}{1 - f'(a, \alpha)}$ . Differently,  $a$  of LaFR is  $\alpha$  of LMFR, and  $\alpha$  of LaFR is a new parameter ranging from 0 to 1. Please refer to the original paper [29] for detailed information. The  $U_{ef}$  is the  $U$  value at foveal boundary  $e_{\text{foveal}}$  of LaFR:  $U_{ef} = (Z_{ef}(1 - Fa))^{\frac{1}{F_a}}$ . The  $Za$  is just a shorthand for a bunch of calculation formulas, where  $Z_{ef}$  is the  $Z$  value at the foveal boundary  $e_{\text{foveal}}$  of LaFR:  $Za = \frac{\frac{1}{z(e)} - \frac{1}{Z_{ef}}}{\frac{1}{Z_{ef}} - 1}$ , and  $LW(e)$  is also a shorthand for a bunch of calculation formulas as  $LW(e) = \frac{1}{\ln L} \frac{w\pi}{180 \sin e \cos e}$ , where  $L$  represents the farthest distance from the gaze point to the corners, and  $w$  stands for their reduced-sized LP buffer width.

By taking the derivative of  $e$  on both sides of the functions in both (14) and (15), then multiplying by 0.5, we obtain the shading rate for LMFR:

$$SR_{\text{LMFR}}(e) = \frac{1}{2} \times \frac{du_{\text{LMFR}}(z(e))}{de} = 2 \left( \frac{\ln \frac{\tan e}{c_r}}{\ln L} \right)^3 LW(e) \quad (16)$$

and shading rate for LaFR:

$$SR_{\text{LaFR}}(e) = \begin{cases} \frac{1 - Fa}{2F_a} (z(e) - z(e)Fa)^{\frac{1 - Fa}{F_a}} LW(e), & e \in [0, 4.89] \\ \frac{2(1 - U_{ef})}{-\pi \sqrt{1 - Za^2}} \frac{1}{Z_{ef} - 1} \frac{-1}{z(e)^2} LW(e), & e \in (4.89, 55] \end{cases} \quad (17)$$

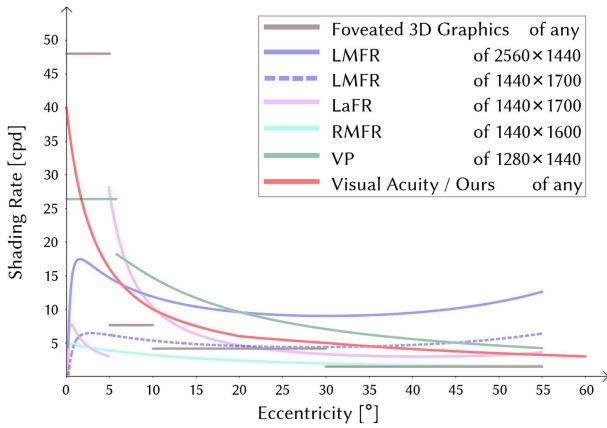


Fig. 5. Foveated 3D graphics [12] uses a piecewise step function, causing shading rate discontinuities. LMFR [23] introduces a continuous shading rate but still fails to align with human visual acuity. LaFR identifies LMFR’s suboptimal quality and refines it by segmenting the kernel function and fine-tuning parameters to better adapt shading distribution. RMFR [46] plots all points along the x-axis ( $f_x = 0.2W$ ,  $f_y = 0.2H$ ,  $\delta = 2.6$ , gaze-centered, as per the original paper). Similarly, Visual-Polar [19] exhibits a shading rate far below visual acuity in the foveal region.

*Our Shading Rate:* Taking the derivative of  $e$  on both sides of (5), the mapping function of our approach, and subsequently multiplying by 0.5, yields the shading rate of our VaFR:

$$SR_{Ours}(e) = \frac{1}{me + \omega_0} = f(e) \quad (18)$$

where  $f(e)$  is exactly the human visual acuity model.

As our mapping function is derived from the acuity model, our approach precisely aligns the shading rate with the human visual acuity. Overall, we address several issues found in previous algorithms. Our approach ensures a shading rate that fully adheres to the human visual acuity model, providing users with both a comfortable experience and powerful performance.

On the contrary, LMFR and LaFR will face major issues: their shading rates do not align with human visual acuity, lack smooth adaptability to different resolutions, and cannot adjust to 3D gaze orientation changes. In the following subsections, we analyze these limitations using visual plots with different resolutions. In contrast, our algorithm will maintain a consistent shading rate that matches human visual acuity, and adapt seamlessly to resolution changes and gaze shifts common in VR. In addition, it provides user-friendly control over tangential shading rates, detailed in Section IX.

### B. Evaluation of Consistency With Acuity

Taking the display resolutions specified in two respective papers as examples, we conduct visual analysis of shading rates. Fig. 5 showcases the shading rate of foveated 3D graphics by Guenter et al. [12] and LMFR, utilizing the resolution (2560 × 1440) defined in LMFR. Switching to another resolution (1440 × 1700) defined in LaFR, Fig. 5 depicts the shading rate of several comparable algorithms, exposing issues associated with both LMFR and LaFR.

*Guenter et al. [12]:* Typical multi-resolution rendering involves partitioning the field of view and rendering at different resolutions, followed by synthesizing the final image. Coarse

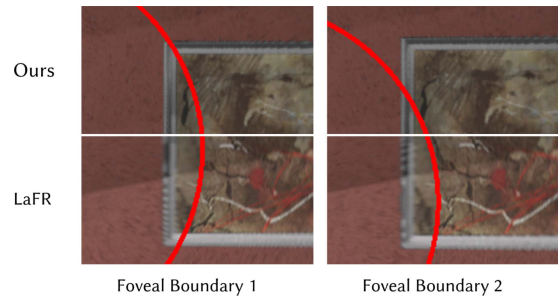


Fig. 6. LaFR’s shading rate discontinuity and artifacts are illustrated through ROI changes in two consecutive frames. The foveal boundary ( $4.89^\circ$  in LaFR), marked in red, varies between left and right images. Inside this boundary, LaFR exhibits blurriness and aliasing (notably in the picture frame on the wall), while areas beyond it appear smoother. In contrast, our method eliminates these issues.

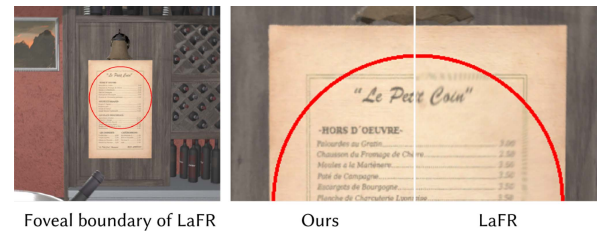


Fig. 7. Side-by-side comparison of LaFR and VaFR, with the foveal boundary (defined as  $4.89^\circ$  in LaFR) delineated in red. LaFR exhibits obvious blurriness and distortion within the foveal region, while our approach shows superior visual quality.

Pixel Shading (CPS) [40] introduces coarse-grained pixel rendering to share shading values among pixels within a defined range. Methods based on Variable Rate Shading (VRS), akin to CPS, transfer coarse-grained partitioning from software to hardware. The shading rates generated by these methods are discrete, roughly illustrated in Fig. 5. Clear boundaries between sharp and blurred blocks significantly impact the user’s sensory experience in VR.

*LMFR [23]:* LMFR approximates human visual acuity in specific resolution settings, with its fitting being more accurate within the  $3^\circ$ - $15^\circ$  range. However, beyond this range, the shading rate in the peripheral field exceeds human visual acuity, resulting in significant computational resource waste. Furthermore, the manually configured kernel function used to redistribute shading points limits the shape of the shading rate curve, narrowing its fitting range.

*LaFR [29]:* At LMFR’s specified resolution, its shading rate closely approximates the visual acuity model. However, this alignment breaks down when switching to the resolution used in LaFR. LaFR segments the kernel function and introduces new parameters to adjust the shading points, closely approximating the human visual acuity within  $8^\circ - 55^\circ$  (see Fig. 5). However, truncating the shading rate results in discontinuities that degrade image quality in the foveal region, and impact user experience, as evidenced by our experiment (see Figs. 6 and 7).

LMFR’s shading rate parameters are not adaptable to different resolutions, leading to misalignment with the acuity curve. For both LMFR and LaFR, this reveals the limited universality of the parameters derived from user studies. Additionally, LaFR requires fine-tuning, such as the kernel function segmentation and new parameters, to recalibrate shading distribution for new

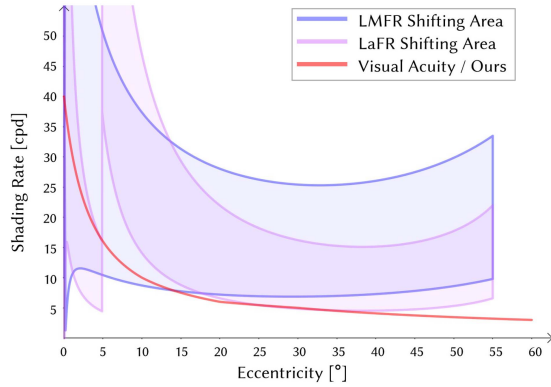


Fig. 8. As the resolution increases from  $1920 \times 1080$  to  $7680 \times 4320$ , the shading rates of LMFR and LaFR deviate significantly from human acuity, indicating resolution dependence. In contrast, our method consistently aligns with the acuity curve across resolutions, showcasing its adaptability and robust capabilities.

HMD resolutions. This need for case-by-case adjustments highlights the complexity of achieving optimal shading rates across varying resolutions.

**RMFR [46]:** Rectangular mapping foveated rendering (RMFR) also involves transformations between two Cartesian coordinate systems; however, our shading rate calculation in (13) based on the log-polar approach, no longer applies for capturing shading rates in all directions for RMFR. Therefore, we visualize only the radial and tangential shading rates along the x-axis, using ratios relative to window height and width as defined in the original paper. At RMFR's resolution of  $1440 \times 1600$  with a  $\delta$  of 2.6, the shading rates fall significantly below the human visual acuity curve, especially in the tangential direction (see Fig. 21). Since rectangular mapping follows a different technical approach and has distinct features from ours, we will not include comparisons with RMFR in the following sections.

**VP [19]:** Visual-Polar method introduces a log-polar approach with linear fall-off in the foveal region, performing partial rendering of the LP buffer to create a cropped triangular area, ensuring uniform shading rates within the foveal region. Outside of this area, the method employs a different visual acuity model and a gaze-related parameter  $d$ , which can lead to inconsistencies during saccades. Interestingly, the LP buffer size used here is similar to ours, resulting in comparable shading rate curves in the peripheral region. However, it shows a shading rate that falls far below visual acuity in the foveal region. In addition, this approach still does not decouple the radial and tangential shading rates, causing the tangential shading rate to depend on the radial rate and to be lower.

### C. Adaptability Across Different Resolution

The shading rates of LMFR and LaFR are sensitive to changes in resolution due to the term  $w = \frac{W}{\delta}$  in their shading rate functions, where  $W$  represents the width of the display resolution and  $\delta$  is the reduction ratio (see details in Section II-A). As shown in Fig. 8, transitioning from  $1920 \times 1080$  to  $7680 \times 4320$  results in significant changes in the shading rate, represented by the colored shifting area. This indicates that LMFR and LaFR are optimized for specific resolutions, making them unsuitable

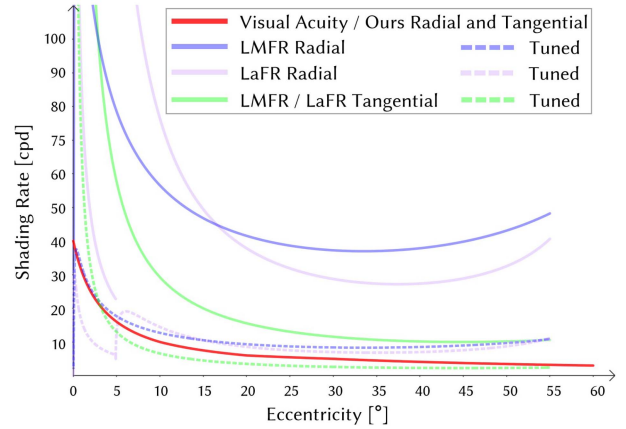


Fig. 9. At a retinal resolution of  $11520 \times 6480$ , prior methods inefficiently use computing resources, as their shading rate unnecessarily exceeds the actual requirements of the human eye. We tuned the parameters for LMFR and LaFR, setting  $\delta = 8$ ,  $\alpha = 0.78$ , and  $\beta = 0.9$ , in an attempt to optimize their shading rate closer to human visual acuity, though they can hardly achieve this alignment.

for use on different devices where their parameters may lose effectiveness. The requirement of using a kernel function  $K(\cdot)$  to redistribute shading points to fit human visual acuity presents inherent challenges. Mapping all display resolution pixels into a proportionally reduced shading buffer is closely tied to the display resolution and reduction ratio, despite efforts to fine-tune parameters through user experiments.

Assuming a visual field of  $100^\circ$  vertically and  $120^\circ$  horizontally, with a foveal acuity of 40 cpd (where each cycle equals two pixels), the theoretical retinal resolution amounts to an astonishing  $100 \times 80 \times 120 \times 80 = 76,800,000$  pixels per eye. For practical purposes, we approximate this as  $11520 \times 6480 = 74,649,600$  pixels per eye. As illustrated in Fig. 9, at a retinal-level resolution of  $11520 \times 6480$ , LMFR and LaFR exhibit markedly elevated shading rates, surpassing practical requirements and resulting in inefficient resource utilization.

In contrast, our method consistently conforms to the visual acuity and remains unaffected by changes in resolution. We avoid the need for kernel functions and instead derive our approach directly from the acuity model. This involves calculating the information density perceivable by the human eye in a single frame, allowing us to establish a fixed LP buffer size that is directly correlated with visual acuity. This fixed buffer size remains unchanged regardless of device or resolution variations, as the spatial information density perceptible to the human eye remains constant.

### D. Shading Rate During Gaze Shift

Typically, an individual's gaze point dynamically shifts, naturally transitioning to different regions of interest. This gaze transfer behavior is pivotal in comprehending how our visual perception adjusts to diverse stimuli and environments. Moreover, variations in gaze orientation significantly affect tasks involving interaction with digital interfaces.

As shown in Fig. 10, illustrating the rendering at a resolution of  $2560 \times 1440$  in VR, a noticeable shift occurs in the shading rate when the gaze point shifts from the center to the corner of the screen. The colored shifting area delineates the variation

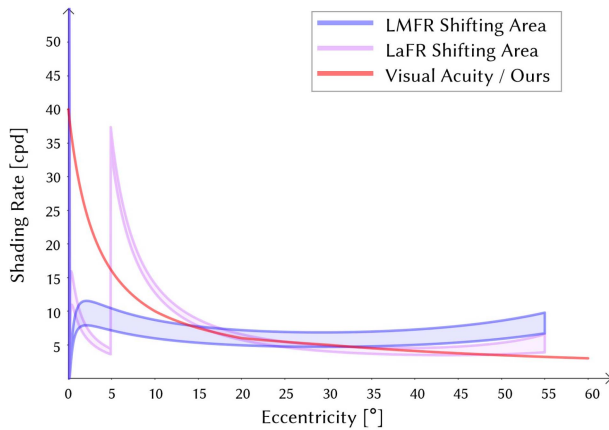


Fig. 10. Shading rates of LMFR and LaFR vary during gaze transfer from looking straight ahead to another direction (gaze deviation), resulting in perceptual inconsistencies. In contrast, our approach maintains a consistent alignment with visual acuity, ensuring stability regardless of gaze shift.

in shading rate, illustrating the potential for sensory discontinuity. This phenomenon, where image quality varies with eye movement, arises from the use of  $\ln(L)$  in the denominators of their log-polar transformations, as indicated in (14) and (15), inherently correlating with gaze orientation.

In contrast, our method deviates from the traditional log-polar transformation used by LMFR and LaFR. By exclusively harnessing the human visual acuity model, our algorithm maintains independence from the gaze orientation. This ensures uniformity in the VR experience with a stable shading rate across different gaze orientations.

### E. Parameters Tuning

In addition to the parameter values  $\delta = 1.8$ ,  $\alpha = 0.85$ , and  $\beta = 0.7$  used in the original paper, we further tuned the parameters for LMFR and LaFR, setting  $\delta = 8$ ,  $\alpha = 0.78$ , and  $\beta = 0.9$  to allow for more rigorous comparisons. These values aim to optimize the shading rates of LMFR and LaFR to closer align with human visual acuity at retinal resolution.

As shown in Fig. 9, although parameter tuning allows the LMFR and LaFR methods to reduce the number of shading pixels for practical use at extremely high resolutions, their lack of generalizability remains a drawback. Furthermore, parameter tuning cannot alter the curve shape dictated by the mapping equations, meaning that these curves can only approximate, rather than perfectly fit, the human visual acuity model. Lastly, the tangential shading rate curves produced by these methods remain suboptimal.

In contrast, our approach, without any parameter tuning, maintains a consistent shading rate aligned with human visual acuity, enabling retinal-level detail in the foveal region while providing satisfactory quality in peripheral areas. Furthermore, VaFR's radial and tangential shading rates are identical because of our setting in (6). Further details on the tangential shading rate are discussed in Section IX. Even at retinal resolutions where conventional methods falter, our approach remains robust. Tailored for forthcoming high-end VR headsets boasting retinal-level displays, our methodology aspires not only to sustain a

seamless frame rate but also to elevate image fidelity to the retinal level, thus advancing the frontiers of immersive experiences in virtual reality.

## V. VALIDATION ON DEFERRED RENDERING PIPELINE

To evaluate both the rendering quality and performance of our method, we have designed a comprehensive testing framework using C++ with the NVIDIA Falcor framework [3]. Our setup includes a PC workstation equipped with a 3.2 GHz Intel(R) Core(TM) i7-8700 CPU, 8 GB of memory, and an NVIDIA GeForce GTX 2080 SUPER graphics card.

### A. Integration With Reprojection

Extensive experiments have shown that rasterization is the primary bottleneck in high-resolution deferred rendering, especially in stereo rendering. Reprojection for binocular VR can reduce the rasterization workload, increasing performance when integrated into our deferred rendering pipeline.

To test the upper limit and compatibility of our VaFR, we integrate it with Wissmann et al.'s reprojection strategy [45], creating VaFR<sub>Rep</sub>. This approach reprojects left-eye images to the right perspective using forward grid warping. Unlike Wissmann's method, we perform reprojection directly from the left log-polar buffer, reducing computational load and significantly enhancing performance. In contrast, the non-reprojection counterpart, VaFR<sub>Dup</sub>, renders the left- and right-eye images independently using the VaFR method, performing separate renderings twice. Detailed results are presented in Section V-B and Section V-C, with additional benefits discussed in Section X.

### B. Performance Evaluation

1) *Setup and Procedure*: Experiments are conducted using our VaFR series methods as the experimental group and several existing algorithms as the control group, including LaFR(2.2), LaFR(1.8), LMFR(1.8), and GT. In this context, GT refers to the native rendering algorithm used as the baseline, identical to the other comparable algorithms but without foveated rendering acceleration. According to the results of X. Shi [29], LaFR achieves different rendering ratios by adjusting the parameter  $\delta$  within the range of 1.8 to 2.2, which means that the shading resolution equals the display resolution divided by  $\delta^2$ . LaFR(1.8) achieves better rendering quality, while LaFR(2.2) achieves higher rendering performance. All parameters of LMFR and LaFR are used exactly as stated in the original papers. Anti-aliasing, as specified in LMFR [23], is implemented accordingly. Since LaFR does not mention anti-aliasing, we assume that it uses an approach analogous to that of LMFR.

NVIDIA's Bistro and Sun Temple are used for the testing scenes. In BistroExterior, there is one directional light for the daylight scene, 12 light sources for the nighttime scene, 21 lights for the BistroInterior scene, and 31 lights for the Temple scene (see Fig. 13). Note that the complexity of the scene and the number of lights can heavily affect the rendering workload. During the rendering process, performance metrics are recorded for all approaches, encompassing VaFR<sub>Rep</sub>, VaFR<sub>Dup</sub>, LaFR(2.2), LaFR(1.8), LMFR(1.8), GT. The performance is evaluated by

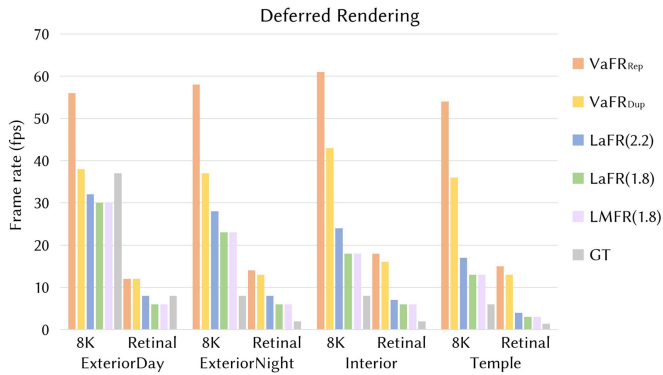


Fig. 11. Performance comparisons of different approaches in various scenes, 8 K and retinal resolutions are evaluated, respectively. VaFR<sub>Rep</sub> achieves over 54 fps at 8 K, whereas other modes fall below 30 fps.

the averaged frame rate across various resolutions, from 2K per eye to retinal resolution.

We begin our analysis with the BistroExteriorNight scene, which features 12 lights, followed by the BistroExteriorDay scene, illuminated by a single directional light source, representing scenarios with minimal lighting complexity. To further investigate the performance, we use the BistroInterior scene with 21 lights, which has one-third of the triangles and twice lights compared to BistroExteriorNight. This setup allows us to explore the impact of reducing triangle count and increasing light sources on rasterization and lighting loads. Furthermore, given that VaFR<sub>Dup</sub>, LMFR, and LaFR aim to enhance performance by reducing shading computations, we extend our tests to the Temple scene, which has fewer triangles than BistroInterior but features 31 lights.

2) *Result:* Comparing BistroExteriorDay, which has one directional light, it is evident that LMFR, LaFR, and VaFR<sub>Dup</sub> all lead to a decrease in performance. This is because the lighting pass has minimal influence on overall performance, due to only one light in the scene, with the main bottleneck arising from rasterization. However, VaFR<sub>Rep</sub> achieves speedup owing to the elimination of rasterization processing of the right-eye image. Detailed performance is shown in Fig. 11.

In contrast, BistroInterior presents more lights and fewer triangles. As resolution increases, LMFR and LaFR maintain a relatively constant speedup of 2 $\times$ , while VaFR<sub>Dup</sub> demonstrates a much higher speedup. At the highest resolution, VaFR<sub>Dup</sub> achieves a 8 $\times$  speedup and VaFR<sub>Rep</sub> attains a 9 $\times$  speedup.

In the most challenging scenario, involving the Temple scene with 31 lights and 606,376 triangles, LaFR(2.2) achieves a speedup of 2.9 $\times$  to 3 $\times$ , while LaFR(1.8) and LMFR(1.8) achieve a speedup of 2.1 $\times$  to 2.3 $\times$ . VaFR<sub>Dup</sub> exhibits a remarkable 9.3 $\times$  speedup and VaFR<sub>Rep</sub> demonstrates an outstanding 10.7 $\times$  speedup.

However, due to rasterization demands in the deferred rendering, which require generating G-buffers at the display resolution and creating several large textures for the lighting pass, this process represents a significant portion of the rendering time. As a result, methods based on the deferred rendering often show relatively low efficiency, particularly at extremely high resolutions, such as 8 K or retinal levels. Notably, as the resolution exceeds

TABLE I  
NUMBER OF PIXELS/RAYS OF EACH RENDERING MODE

Quantity	VaFR <sub>Rep</sub>	VaFR <sub>Dup</sub>	LaFR(2.2)	LaFR(1.8)	GT
2K	637K	1,061K	761K	1,137K	3,686K
4K	743K	1,061K	1,713K	2,560K	8,294K
6K	849K	1,061K	3,855K	5,760K	18,662K
8K	955K	1,061K	6,854K	10,240K	33,177K
Retinal	1,035K	1,061K	15,423K	23,040K	74,649K

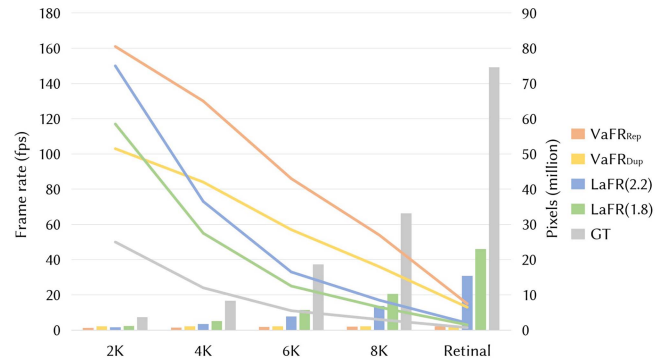


Fig. 12. Performance of deferred rendering declines when the resolution is increased from 2 K to retinal level, using the Temple scene for testing. The histogram illustrates the number of pixels rendered per eye.

4 K, LaFR loses its performance advantage and undergoes a marked decline.

Since LaFR(1.8) and LMFR(1.8) generally exhibit similar frame rates, we only show LaFR(1.8)'s performance as a representative for simplicity. Table I lists the number of pixels rendered per eye of each rendering mode across different resolutions. Fig. 12 illustrates this drop in frame rate when the resolution increases from 2 K to the retinal level in the Temple scene. In contrast, the ray casting pipeline, not requiring the rasterization of very large resolution G-buffers, proves to be more suitable and performant at these high resolutions, as we will discuss in Section VI.

### C. User Study

1) *Pilot User Study:* We conducted a user study, testing it in an immersive virtual reality environment equipped with various types of head-mounted displays (HMDs) across different scenarios. This is done to compare real users' preferences for the experience provided by our method against alternative algorithms.

*Setup:* An HTC VIVE PRO EYE is used, with a built-in gaze tracking system, to precisely monitor user's gaze orientation. This HMD features a display resolution of 1440  $\times$  1600 pixels per eye, while the actual rendered texture resolution (also called rendering resolution) is 2468  $\times$  2740 pixels per eye. Additionally, a Pimax Crystal is used, also equipped with a gaze tracking system, boasting a much higher display resolution of 2880  $\times$  2880 pixels per eye, achieving up to 35pixels per degree, with the actual rendered texture resolution of 3108  $\times$  2624 pixels per eye.

*Participant:* 20 participants (17 males and 3 females, aged 22 to 28) were recruited, all of whom were graduate students at our university with normal or corrected-to-normal vision.



Fig. 13. Four benchmark scenes are employed for testing, including BistroExteriorDay, BistroExteriorNight, BistroInterior, and SunTemple.

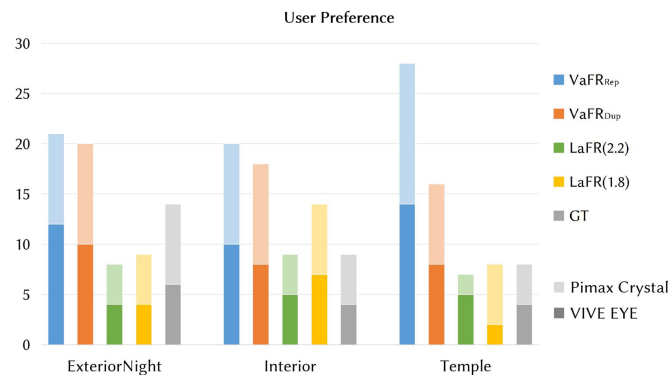


Fig. 14. User preference using VIVE EYE HMD with a rendering resolution of  $2468 \times 2740$  per eye and Pimax Crystal headset with a rendering resolution of  $3108 \times 2624$  per eye, VaFR<sub>Rep</sub> and VaFR<sub>Dup</sub> receive the most votes.

**Procedure:** Each participant's session lasted approximately 25 to 30 minutes, involving free exploration of three scenes in five modes from different approaches. To minimize bias, we avoided providing information about our research or the rendering algorithm. Once participants comfortably wear the headset, they will be exposed to various rendering approaches in random order and select their top two preferences. They could choose only a single preference if they could not decide on two. Additionally, after evaluating each approach, participants could revisit specific rendering methods before finalizing their decision. Subsequently, they moved on to assess the next scene for testing. Two HMDs listed above were used for our user study.

2) **Results and Analysis:** Fig. 14 illustrates the user preferences observed in the VIVE EYE and Pimax Crystal HMDs. Evidently, VaFR<sub>Dup</sub> and VaFR<sub>Rep</sub> consistently received top user preferences across scenes. However, users favored GT in the BistroExteriorNight scene for its smooth frame rate and superior image quality. Conversely, in the Temple scene with complex lighting, performance of LaFR and GT was notably affected, whereas VaFR<sub>Rep</sub> and VaFR<sub>Dup</sub> perform better in handling such conditions, further highlighting our algorithm's capability in challenging scenarios.

In addition, we established null hypotheses for the user study: There is no significant difference in user preference data for specific rendering modes between any two groups. As indicated in Table II, the Kruskal-Wallis test was applied for pairwise

TABLE II  
KRUSKAL-WALLIS TEST FOR PAIRWISE COMPARISONS WITH  
HOLM-BONFERRONI CORRECTION

P values	VaFR <sub>Rep</sub>	VaFR <sub>Dup</sub>	LaFR(2.2)	LaFR(1.8)	GT
VaFR <sub>Rep</sub>	/	0.177	0.028	0.023	0.025
VaFR <sub>Dup</sub>	/	/	0.031	0.031	0.032
LaFR(2.2)	/	/	/	0.562	0.525
LaFR(1.8)	/	/	/	/	0.806

At a significance level of 0.05, both VaFR<sub>Rep</sub> and VaFR<sub>Dup</sub> show p values smaller than 0.05, signifying that our approaches are significantly superior to the alternatives.

comparisons at a significance level of 0.05, with adjustments using the Holm-Bonferroni correction. The corrected P values for VaFR<sub>Rep</sub> and VaFR<sub>Dup</sub> are less than 0.05, signaling a significant difference compared to other methods. This suggests a violation of the null hypothesis, indicating that our approach outperforms alternative approaches. In contrast, the P values for LaFR(2.2), LaFR(1.8), and GT following the Kruskal-Wallis test exceed 0.05, implying no significant differences among them.

3) **2AFC User Study:** To further validate the findings of the pilot user study, we conducted an additional two-alternative forced choice (2AFC) test.

**Participant:** 25 participants (17 males and 8 females, aged 22 to 32, with a mean age of 25.56) were recruited, primarily graduate students from our university, all of whom were graduate students from our university with normal or corrected-to-normal vision. Some participants in the pilot study were included again, and the setup remained the same as in the pilot study.

**Procedure:** In each trial of 2AFC test, they are exposed to two rendering approaches and select only one as their preference. Participants could revisit any rendering methods before finalizing their decision. Each pair of rendering modes is tested, resulting in a total of 10 combinations. Each combination presents 3 times, with the order within each combination being randomized, as well as the order between combinations. Thus, each scene has 10 (combinations)  $\times$  3 (times) = 30 trials. After completing each set, the participants proceed to the next scene for testing. Two HMDs listed above were employed in the user study. In total, 3 (scenes)  $\times$  30 (trials per scene)  $\times$  25 (participants)  $\times$  2 (HMDs) = 4500 trials were conducted.

Regarding preferences, participants were informed in advance to base their preferences solely on their subjective visual experience while navigating the 3D scenes, without any specific guidelines. After completing the experiment, each participant was asked to explain their choices. Most preferred a balanced experience, valuing both frame rate and image quality. A minority, potentially more susceptible to 3D motion sickness, prioritized frame rate, while only two participants highlighted the importance of image quality.

4) **Results and Analysis:** We establish null hypotheses for the user study: There is no significant difference in user preference data for specific rendering modes between any two groups. We performed the Kruskal-Wallis test for pairwise comparisons at a significance level of 0.05, applying the Holm-Bonferroni correction. As indicated in Table III and Fig. 15, the corrected p-values for VaFR<sub>Rep</sub> and VaFR<sub>Dup</sub> are below 0.05, rejecting the null hypothesis and indicating that our approach significantly outperforms alternative approaches. In contrast, the p values for

TABLE III  
KRUSKAL-WALLIS TEST FOR PAIRWISE COMPARISONS WITHIN 2FAC  
EXPERIMENT

P values	VaFR <sub>Rep</sub>	VaFR <sub>Dup</sub>	LaFR(2.2)	LaFR(1.8)	GT
VaFR <sub>Rep</sub>	/	0.522	0.004	0.016	0.006
VaFR <sub>Dup</sub>	/	/	0.006	0.025	0.006
LaFR(2.2)	/	/	/	0.873	0.631
LaFR(1.8)	/	/	/	/	0.749

Both VaFR<sub>Rep</sub> and VaFR<sub>Dup</sub> show p values smaller than 0.05 when compared to the alternatives, signifying that our approaches are significantly superior to these alternatives.

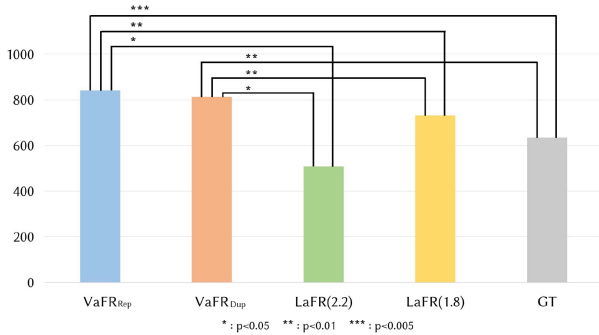


Fig. 15. User preference of 2AFC experiment. VaFR<sub>Rep</sub> and VaFR<sub>Dup</sub> receive the most votes and consistently outperformed other solutions with statistical significance.

LaFR(2.2), LaFR(1.8), and GT exceed 0.05, showing no significant differences among them. These findings are consistent with the results of the pilot study, confirming similar conclusions. In all, VaFR<sub>Dup</sub> and VaFR<sub>Rep</sub> all offer superior perceptual outcomes compared to the other methods.

## VI. VALIDATION ON RAY CASTING PIPELINE

To further demonstrate the compatibility and superiority of our algorithm, we evaluate the performance of our algorithm and the alternatives in the ray casting pipeline (without rasterization).

### A. Ray Casting Pipeline

In our new binocular rendering system for ray casting, a ray generation shader serves as the initial entry point to initiate ray casting. Texture and dispatch sizes are adjusted to match the constant LP buffer size. For each pixel in the LP buffer, its corresponding position in the screen space is calculated, and a ray is generated from this position. Shading occurs at the hit point using a closest hit shader, while a missing shader is applied when no geometry is hit. Upon successful triangle hit detection, the same shading code used in the fragment shader of the deferred rendering pipeline is executed. Following internal anti-aliasing, inverse mapping, and post anti-aliasing passes, the left image is generated. Subsequently, the right image is produced using the same set of passes as applied to the left image.

### B. Performance Evaluation

Four scenes are tested using five rendering approaches, including VaFR<sub>Dup</sub>, LaFR(2.2), LaFR(1.8), LMFR(1.8), and GT,

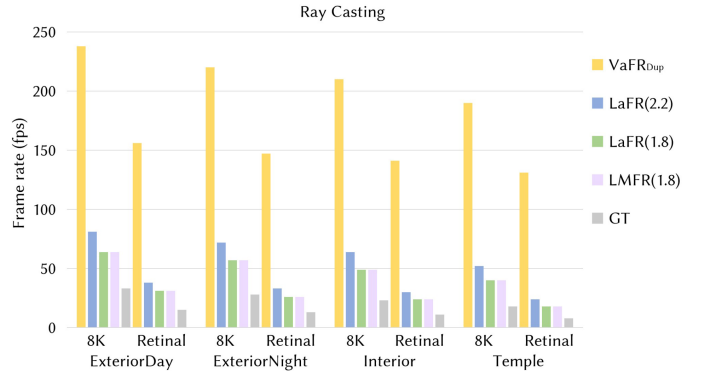


Fig. 16. Performance comparison across different approaches are evaluated in various testing scenes, at both 8 K and retinal resolutions. It reveals that VaFR<sub>Dup</sub> achieves superior performance, with over 131 fps at retinal resolution, while LaFR lags behind, remaining below 40 fps.

all integrated into the ray tracing pipeline. The setup and procedure are the same as described in Section V-B1.

Fig. 16 illustrates average frame rates in four scenes for various rendering approaches at 8 K per eye resolution and retinal resolution. In the absence of rasterization bottlenecks in the ray casting pipeline, LMFR and LaFR achieve approximately a speedup of  $2\times$  in all scenes. Meanwhile, our method exhibits remarkable speedup, ranging from  $7.2\times$  to  $10.6\times$  across all scenes at 8 K resolution, exceeding the range  $1.9\times$  to  $2.9\times$  observed in other approaches. Similarly, at retinal resolution, our method excels with speedup ranging from  $10.4\times$  to  $16.4\times$ , maintaining frame rates between 131 and 156 fps, outperforming the range  $2.0\times$  to  $3.0\times$  seen in other methods.

Compared to the deferred rendering pipeline that includes a rasterization stage, our method in the ray casting pipeline demonstrates an even more powerful ability to accelerate rendering. Our method's significant speedup advantage in the ray casting pipeline comes from a key process: emitting and shading rays from all pixels in the LP buffer, and then writing them back into it. This efficiency is further enhanced because our algorithm's LP buffer is sized based on the human visual acuity model, making it much smaller than typical display resolutions such as 8 K. As a result, our approach maintains a smooth frame rate, demonstrating a powerful advantage in speedup.

Fig. 17 shows the average time cost of the three stages for each approach at different resolutions within a single frame of the ray casting pipeline. Fig. 18 depicts the frame rate fall-off as resolution increases from 2 K to retinal level in the Temple scene. Despite the performance drop, VaFR<sub>Dup</sub> consistently maintains a frame rate of above 131 fps at retinal resolution, while others cannot. A heavy rasterization workload can hinder rendering efficiency, whereas a ray-tracing pipeline fully exploits the advantages of our algorithm. In the ray casting pipeline, VaFR<sub>Dup</sub> achieves significantly higher frame rates at retinal resolution than other methods, effectively enabling retinal rendering in VR while ensuring a smooth visual experience.

## VII. ABLATION STUDY ON ANTI-ALIASING

Analysis of shading rate and performance plots reveals minimal differences among the comparative algorithms at a display resolution of around 4 K. Therefore, we consider the significant

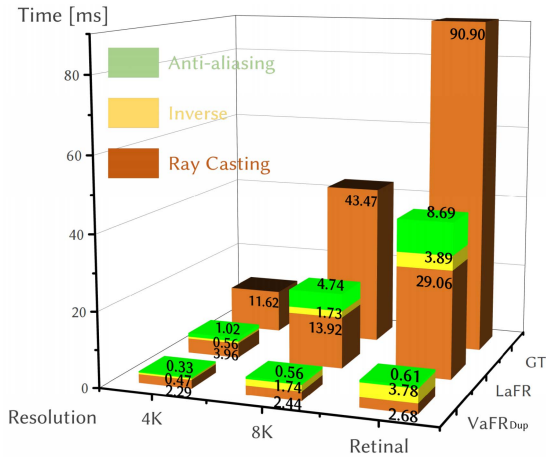


Fig. 17. The average time cost of three stages for each approach across different resolutions.

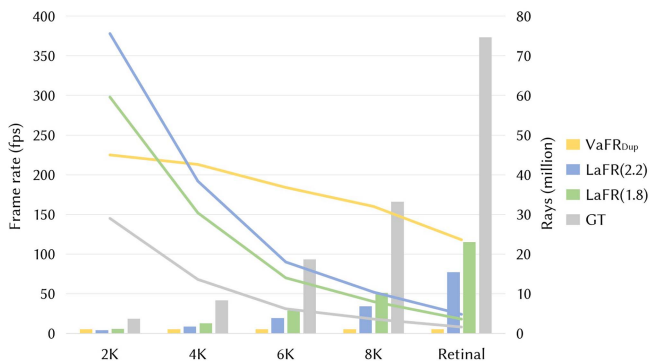


Fig. 18. As resolution increases from 2 K to retinal resolution, all methods experience a decline in performance. However, our approach maintains a more gradual decline, consistently achieving over 131 fps at retinal resolution, outperforming others. The histogram shows the number of rays emitted per eye.

impact of anti-aliasing algorithms on final image quality and frame rate through an ablation study on anti-aliasing. We compare image quality across various rendering modes within the ray casting pipeline at both 4 K and 8 K per eye resolution, as shown in Fig. 19.

LMFR employs two anti-aliasing methods: internal anti-aliasing with a  $3 \times 3$  Gaussian filter on the LP buffer, and post anti-aliasing with foveal-aware Gaussian filters on the final image. To ensure a fair comparison, we evaluate LMFR both with and without anti-aliasing, applying the same configurations to LaFR. Additionally, we compare the versions of VaFR<sub>Dup</sub> with and without FXAA.

Zoomed-in results are cropped at eccentric angles of  $0^\circ$  and  $40^\circ$  for a clearer quality comparison. The anti-aliasing methods used in LMFR offer only marginal improvement, as the Gaussian filter tends to blur the image. In contrast, applying FXAA to the LP buffer in our approach significantly enhances image quality with minimal impact on performance. This also demonstrates that our method consistently achieves superior performance, with or without anti-aliasing, across different rendering resolutions. The results also show that LaFR exhibits slightly better quality outside the foveal region compared to VaFR<sub>Dup</sub>; however,

after applying internal anti-aliasing and post anti-aliasing, it falls slightly behind VaFR<sub>Dup</sub>. Notably, at 8 K per eye resolution, LaFR demonstrates significantly superior image quality outside the foveal region compared to VaFR<sub>Dup</sub>, but its frame rate is only about one-quarter of VaFR<sub>Dup</sub>'s. This indicates that our method, VaFR, allocates computing resources more efficiently than the alternative, achieving higher performance without compromising perceptual visual quality.

## VIII. VALIDATION ON PATH TRACING PIPELINE

To validate the availability of our method to highly realistic rendering pipelines, we also integrated our approach into NVIDIA's PathTracing SDK [24] on a workstation featuring a robust 3.5 GHz Intel(R) Core(TM) i5-13600KF CPU, 32 GB memory, and an NVIDIA GeForce GTX 4070Ti GPU.

### A. Path Tracing Pipeline

In the ray tracing stage, mirroring our approach in the ray casting pipeline, we align the dispatch size and texture size with the dimensions of the LP buffer. Rays are then emitted from the screen-space position of each pixel in the LP buffer, producing a semi-elliptical-shaped image, as depicted in Fig. 4. Finally, in the inverse stage, the screen-space image is reconstructed.

Two typical scenes are tested, *Programmer-art* and *Convergence-test* from Path Tracing SDK [24], configuring all settings to their default values, including 1 sample per pixel (spp), real-time noise, denoiser, 15 maximum bounces, 3 maximum diffuse bounces, and the use of Next Event Estimate. For simplicity, Russian Roulette, ReSTIR DI, and ReSTIR GI are disabled. We integrate our VaFR and LaFR(2.2) algorithms into the path tracing pipeline and make the necessary code modifications to output images at 8 K resolution.

### B. Performance Evaluation

As shown in Fig. 20, the gaze point is positioned in the center of the screen. It can be observed that the images near the gaze point are noticeably clear, while the shading rate falls in the peripheral field of view. From this we can conclude that our algorithm remains applicable to path tracing, a rendering method known for its high realism, showcasing strong universality.

At a monocular 8 K resolution, VaFR<sub>Dup</sub> achieves a frame rate of approximately 144 fps, while LaFR(2.2) reaches only 14 fps. When rendering for VR with binocular setting, we reasonably estimate that VaFR<sub>Dup</sub> has a smooth frame rate of around 72 fps, whereas LaFR(2.2) lags at 7 fps. In comparison, ground truth (GT) real-time rendering at 2K can only achieve a binocular frame rate of 15 fps, and rendering at 8 K will lead to system crashes.

Achieving 72 fps in 8 K binocular rendering already approximates the initial perceptual requirements of VR users, though it falls slightly short of the commonly accepted comfort frame rate of 90 fps. However, this gap is not unbridgeable; a more advanced GPU can further improve the frame rate. This performance indicates that our VaFR significantly enhances path tracing in a high-resolution VR headset, demonstrating considerable potential for highly realistic rendering.

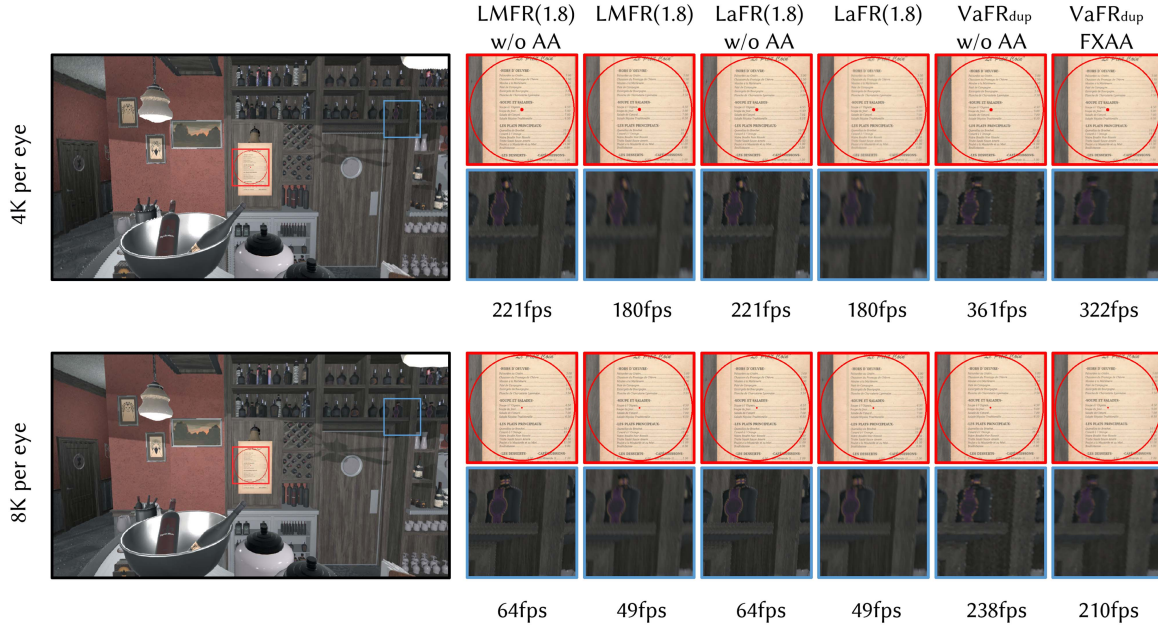


Fig. 19. We compare image quality across three rendering algorithms, with and without anti-aliasing, at 4 K and 8 K resolutions, respectively. LMFR(1.8) and LaFR(1.8) use internal-AA and post-AA as described in Section VII. Zoomed-in results are cropped at eccentric angles of  $0^\circ$  and  $40^\circ$ , respectively. Our method consistently delivers superior performance, with or without anti-aliasing, across various rendering resolutions, achieving higher performance without compromising perceptual visual quality.

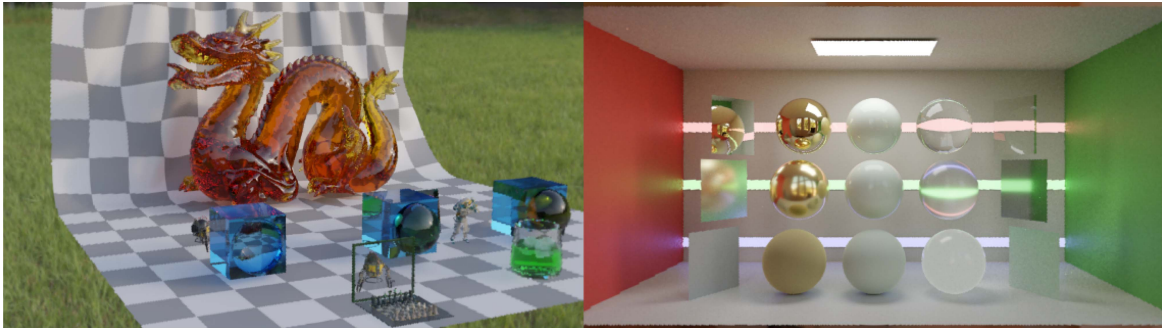


Fig. 20. Path tracing examples (“Programmer-art” and “Convergence-test”) employing our approach, which renders at 8 K per-eye resolution with a frame rate of 72 fps. The ROI is at the center of the image.

### IX. FURTHER DISCUSSION OF ANISOTROPIC VISUAL ACUITY

Humans have anisotropic visual acuity, with detection and resolution varying between frequencies in radial and tangential directions. We are more sensitive to radially oriented gratings [28], approximately twice as much at  $30^\circ$  eccentricity [36]. Although our primary focus does not revolve around this particular issue, it is worth noting that our approach has the inherent advantage of independently adjusting tangential shading rates.

In Section III-B, we keep the tangential shading rate equal to the radial shading rate as 6 for simplicity. If we set the ratio factor of tangential to radial as  $\Delta(e)$  that varies with the eccentricity angle  $e$ , (8) only needs to multiply by it as:

$$v(e, \theta) = \Delta(e) \cdot \theta \cdot \frac{\sin 2e}{me + \omega_0} \quad (19)$$

This conveniently enables the incorporation of anisotropy for additional performance gains, which is an aspect not addressed by previous foveated rendering algorithms.

As illustrated in Fig. 21, both LMFR(1.8) and LaFR(1.8) use the same parameter  $\delta = 1.8$ , resulting in tangential shading rate curves that perfectly coincide. Their tangential shading rates are calculated as:

$$SR_{tangential}(e) = \frac{h/1.8}{2\pi r \frac{de}{dr}} = \frac{h/1.8}{720 \cos(e \frac{\pi}{180}) \sin(e \frac{\pi}{180})} \quad (20)$$

where  $h$  represents display resolution height. As depicted in green, these tangential shading rates deviate from both the visual acuity and their radial shading rates. This discrepancy is unreasonable.

Referring to (6), we infer that our radial and tangential shading rates are equivalent, resulting in identical curves. Our algorithm exhibits a significantly reasonable tangential shading rate and can be independently tuned.

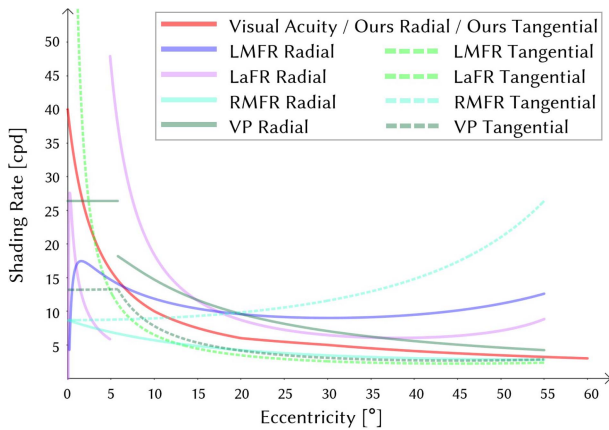


Fig. 21. Our algorithm precisely aligns radial and tangential shading rates with human visual acuity, ensuring consistent performance under all conditions. In contrast, other methods often exhibit deviations in similar scenarios.

## X. CONCLUSIONS, LIMITATIONS, AND FUTURE WORK

We have presented a visual acuity consistent mapping approach for highly efficient foveated rendering, which is applicable to various binocular rendering strategies in VR settings. Our approach exhibits promising performance at high resolutions while maintaining a better perceptual quality than competing methods. With a deferred rendering pipeline, our approach achieves up to  $7.6\times$  speedup over GT method, reaching 61 fps at 8 K per-eye resolution. Using a ray casting pipeline, our approach also delivers an impressive  $10.4\times$  to  $16.4\times$  speedup compared to the GT method at retinal resolution, achieving frame rates between 131 and 156 fps. Additionally, our method enhances path tracing performance, boosting the frame rate of binocular 8 K rendering to approximately 72 fps, enabling highly realistic VR rendering.

Our approach still has some limitations. Temporal flickering, as a known issue associated with the log-polar mapping method, still requires further investigation. As stated in Meng et al. [23], “However, in fly-through of the scene with glossy objects, we notice that view-dependent specular reflection changes before and after applying KFR. Foveation amplifies the specular reflection regions, and makes the specular highlight flicker more.” Foveation exacerbates these specular reflection variations, causing specular highlights to flicker more. Without reprojection as the second stage, the log-polar mapping method essentially re-renders the right-eye image, and due to disparities in camera position and gaze orientation, log-polar mapping may introduce slight mismatches in specular highlights between the left and right views. Using reprojection as the second stage helps mitigate this issue by addressing the mismatch between the left and right views.

Additionally, Patney et al. [26] developed a saccade-aware multi-resolution temporal anti-aliasing method, enhancing peripheral details through contrast enhancement. Tariq et al. [34] introduced a perceptual technique, substituting indiscernible frequencies with procedural noise, customized to image and perception specifics. These post-processing methods, compatible with our approach that provides screen-space images, can be seamlessly integrated into our algorithm to heighten efficiency.

Another challenge associated with foveated rendering is the Moire patterns, particularly in approximately  $30^\circ$  peripheral regions with a lower spatial sampling frequency. The spatial sampling frequency that aligns with the spatial stimuli frequency of the scene will lead to the appearance of Moire patterns. Notably, these patterns evolve with changes in gaze point or head movement. While the foveal and far peripheral regions are less susceptible to Moire patterns, addressing this issue is still essential.

Furthermore, implementing physical brightness-based edge detection and smoothing represents an important challenge for our future research. Anti-aliasing in the LP buffer significantly enhances the final image quality. To achieve this, we use FXAA to smooth the edges. However, tone-mapping within the log-polar space before FXAA poses challenges, as average luminance calculations are not straightforward in this space. Consequently, alternative physical-based edge detection and smoothing methods are necessary.

## ACKNOWLEDGMENT

The authors would like to thank all anonymous reviewers for their constructive feedback and valuable suggestions.

## REFERENCES

- [1] M. Antonelli, F. D. Igual, F. Ramos, and V. J. Traver, “Speeding up the log-polar transform with inexpensive parallel hardware: Graphics units and multi-core architectures,” *J. Real-Time Image Process.*, vol. 10, pp. 533–550, 2015.
- [2] H. Araujo and J. M. Dias, “An introduction to the log-polar mapping [image sampling],” in *Proc. II Workshop Cybernetic Vis.*, 1996, pp. 139–144.
- [3] N. Benty et al., “The falcor rendering framework,” 2020. [Online]. Available: <https://github.com/NVIDIAGameWorks/Falcor>,
- [4] B. Bitterli, C. Wyman, M. Pharr, P. Shirley, A. Lefohn, and W. Jarosz, “Spatiotemporal reservoir resampling for real-time ray tracing with dynamic direct lighting,” *ACM Trans. Graph.*, vol. 39, no. 4, pp. 148:1–148:17, 2020.
- [5] A. Cowey and E. Rolls, “Human magnification factor and its relation to visual acuity,” *Exp. Brain Res. Experimentelle Hirnforschung. Experimentation cérébrale*, vol. 21, pp. 447–454, 1974, doi: [10.1007/BF00237163](https://doi.org/10.1007/BF00237163).
- [6] N. Deng et al., “FoV-NeRF: Foveated neural radiance fields for virtual reality,” *IEEE Trans. Vis. Comput. Graph.*, vol. 28, no. 11, pp. 3854–3864, Nov. 2022.
- [7] R. Fan, X. Shi, K. Wang, Q. Ma, and L. Wang, “Scene-aware foveated rendering,” *IEEE Trans. Vis. Comput. Graph.*, vol. 30, no. 11, pp. 7097–7106, Nov. 2024, doi: [10.1109/TVCG.2024.3456157](https://doi.org/10.1109/TVCG.2024.3456157).
- [8] C. E. Ferree, G. Rand, and C. Hardy, “Refraction for the peripheral field of vision,” *Arch. Ophthalmol.*, vol. 5, pp. 717–731, 1931.
- [9] L. Franke, L. Fink, J. Martschinke, K. Selgrad, and M. Stamminger, “Time-warped foveated rendering for virtual reality headsets,” *Comput. Graph. Forum*, vol. 40, pp. 110–123, 2020, doi: [10.1111/cgf.14176](https://doi.org/10.1111/cgf.14176).
- [10] S. Friston, T. Ritschel, and A. Steed, “Perceptual rasterization for head-mounted display image synthesis,” *ACM Trans. Graph.*, vol. 38, pp. 1–14, 2019, doi: [10.1145/3306346.3323033](https://doi.org/10.1145/3306346.3323033).
- [11] M. Fujita and T. Harada, “Foveated real-time ray tracing for virtual reality headset,” *Light Transport Entertainment Res.*, vol. 3, 2014.
- [12] B. Guenter, M. Finch, S. Drucker, D. Tan, and J. Snyder, “Foveated 3D graphics,” *ACM Trans. Graph.*, 31, vol. 11, 2012, Art. no. 164, doi: [10.1145/2366145.2366183](https://doi.org/10.1145/2366145.2366183).
- [13] A. Jindal, K. Wolski, K. Myszkowski, and R. K. Mantiuk, “Perceptual model for adaptive local shading and refresh rate,” *ACM Trans. Graph.*, vol. 40, no. 6, pp. 1–18, 2021.
- [14] A. S. Kaplanyan, A. Sochenov, T. Leimkühler, M. Okunev, T. Goodall, and G. Rufo, “DeepFovea: Neural reconstruction for foveated rendering and video compression using learned statistics of natural videos,” *ACM Trans. Graph.*, vol. 38, no. 6, pp. 1–13, 2019.

- [15] Y. Kim, Y. Ko, and I. Ihm, "Selective foveated ray tracing for head-mounted displays," in *Proc. 2021 IEEE Int. Symp. Mixed Augmented Reality*, 2021, pp. 413–421, doi: [10.1109/ISMAR52148.2021.00058](https://doi.org/10.1109/ISMAR52148.2021.00058).
- [16] M. Koskela, "Foveated path tracing with fast reconstruction and efficient sample distribution," Tampere Univ. Dissertations - Tampereen yliopiston väitöskirjat, Tampere Univ., Tampere, Finland, 2020.
- [17] M. Koskela, K. Immonen, T. Viitanen, P. Jääskeläinen, J. Multanen, and J. Takala, "Foveated instant preview for progressive rendering," in *SIGGRAPH Asia 2017 Technical Briefs*, New York, NY, USA: Association for Computing Machinery, 2017, doi: [10.1145/3145749.3149423](https://doi.org/10.1145/3145749.3149423).
- [18] M. Koskela, K. Immonen, T. Viitanen, P. Jääskeläinen, J. Multanen, and J. Takala, "Instantaneous foveated preview for progressive monte carlo rendering," *Comput. Vis. Media*, vol. 4, pp. 1–10, 2018, doi: [10.1007/s4109501801130](https://doi.org/10.1007/s4109501801130).
- [19] M. Koskela, A. Lotvonen, M. Mäkitalo, P. Kivi, T. Viitanen, and P. Jääskeläinen, "Foveated real-time path tracing in visual-polar space," in *Proc. Eurographics Symp. Rendering, Eurographics Assoc.*, 2019. [Online]. Available: <https://doi.org/10.2312/sr.20191219>
- [20] M. Levoy and R. Whitaker, "Gaze-directed volume rendering," in *Proc. 1990 Symp. Interactive 3D Graph.*, pp. 217–223, 1990.
- [21] D. Luebke and B. Hallen, "Perceptually driven simplification for interactive rendering," in *Proc. Eurographics Workshop London Rendering Techn.*, U.K., 2001, pp. 223–234.
- [22] A. Maimone, A. Georgiou, and J. Kollin, "Holographic near-eye displays for virtual and augmented reality," *ACM Trans. Graph.*, vol. 36, pp. 1–16, 2017, doi: [10.1145/3072959.3073624](https://doi.org/10.1145/3072959.3073624).
- [23] X. Meng, R. Du, M. Zwicker, and A. Varshney, "Kernel foveated rendering," *Proc. ACM Comput. Graph. Interactive Techn.*, vol. 1, pp. 1–20, 2018, doi: [10.1145/3203199](https://doi.org/10.1145/3203199).
- [24] NVIDIA, "NVIDIA path tracing SDK," 2023. [Online]. Available: <https://github.com/NVIDIAGameWorks/Path-Tracing-SDK>
- [25] T. Ohshima, H. Yamamoto, and H. Tamura, "Gaze-directed adaptive rendering for interacting with virtual space," in *Proc. IEEE 1996 Virtual Reality Annu. Int. Symp.*, 1996, pp. 103–110.
- [26] A. Patney et al., "Towards foveated rendering for gaze-tracked virtual reality," *ACM Trans. Graph.*, vol. vol. 35, no. 6, Dec. 2016, Art. no. 179, doi: [10.1145/2980179.2980246](https://doi.org/10.1145/2980179.2980246).
- [27] J. Rovamo and V. Virsu, "Estimation and application of the human cortical magnification factor," *Exp. Brain Res. Experimentelle Hirnforschung Experimentation Cérébrale*, vol. 37, pp. 495–510, 1979, doi: [10.1007/BF00236819](https://doi.org/10.1007/BF00236819).
- [28] J. M. Rovamo, V. Virsu, P. I. Laurinen, and L. Hyvärinen, "Resolution of gratings oriented along and across meridians in peripheral vision," *Invest. Ophthalmol. Vis. Sci.*, vol. 23, no. 5, pp. 666–70, 1982.
- [29] X. Shi, L. Wang, J. Wu, W. Ke, and C. T. Lam, "Locomotion-aware foveated rendering," in *Proc. 2023 IEEE Conf. Virtual Reality 3D User Interfaces*, 2023, pp. 471–481, doi: [10.1109/VR55154.2023.00062](https://doi.org/10.1109/VR55154.2023.00062).
- [30] M. Stengel, S. Grogorick, M. Eisemann, and M. Magnor, "Adaptive image-space sampling for gaze-contingent real-time rendering," *Comput. Graph. Forum*, vol. 35, no. 4, pp. 129–139, Jul. 2016.
- [31] H. Strasburger, I. Rentschler, and M. Jüttner, "Peripheral vision and pattern recognition: A review," *J. Vis.*, vol. 11, 2011, Art. no. 13, doi: [10.1167/11.5.13](https://doi.org/10.1167/11.5.13).
- [32] Q. Sun, F.-C. Huang, J. Kim, L.-Y. Wei, D. Luebke, and A. Kaufman, "Perceptually-guided foveation for light field displays," *ACM Trans. Graph.*, vol. 36, no. 6, pp. 1–13, 2017.
- [33] N. T. Swafford, J. A. Iglesias-Guitian, C. Koniaris, B. Moon, D. Cosker, and K. Mitchell, "User, metric, and computational evaluation of foveated rendering methods," in *Proc. ACM Symp. Appl. Percep.*, 2016, pp. 7–14.
- [34] T. Tariq, C. Tursun, and P. Didyk, "Noise-based enhancement for foveated rendering," *ACM Trans. Graph.*, vol. 41, pp. 1–14, 2022, doi: [10.1145/3528223.3530101](https://doi.org/10.1145/3528223.3530101).
- [35] L. Thibos, F. Cheney, and D. Walsh, "Retinal limits to the detection and resolution of gratings," *J. Opt. Soc. America. A. Opt. Image Sci.*, vol. 4, pp. 1524–1529, 1987, doi: [10.1364/JOSAA.4.001524](https://doi.org/10.1364/JOSAA.4.001524).
- [36] L. Thibos, D. Still, and A. Bradley, "Characterization of spatial aliasing and contrast sensitivity in peripheral vision," *Vis. Res.*, vol. 36, pp. 249–58, 1996, doi: [10.1016/0042-6989\(95\)00109-D](https://doi.org/10.1016/0042-6989(95)00109-D).
- [37] L. Thibos, D. Walsh, and F. Cheney, "Vision beyond the resolution limit: Aliasing in the periphery," *Vis. Res.*, vol. 27, pp. 2193–2197, 1987, doi: [10.1016/0042698987901349](https://doi.org/10.1016/0042698987901349).
- [38] E. Turner, H. Jiang, D. Saint-Macary, and B. Bastani, "Phase-aligned foveated rendering for virtual reality headsets," in *Proc. 2018 IEEE Conf. Virtual Reality 3D User Interfaces*, 2018, pp. 1–2.
- [39] O. T. Tursun et al., "Luminance-contrast-aware foveated rendering," *ACM Trans. Graph.*, vol. 38, no. 4, pp. 1–14, 2019.
- [40] K. Vaidyanathan et al., "Coarse pixel shading," in *Proc. High-Performance Graph.*, 2014, pp. 9–18.
- [41] D. Walton et al., "Beyond blur: Real-time ventral metamers for foveated rendering," *ACM Trans. Graph.*, vol. 40, pp. 1–14, 2021, doi: [10.1145/3450626.3459943](https://doi.org/10.1145/3450626.3459943).
- [42] M. Weier et al., "Foveated real-time ray tracing for head-mounted displays," *Comput. Graph. Forum*, vol. 35, pp. 289–298, 2016, doi: [10.1111/cgf.13026](https://doi.org/10.1111/cgf.13026).
- [43] F. Weymouth, "Visual sensory units and the minimal angle of resolution," *Amer. J. Ophthalmol.*, vol. 46, pp. 102–13, 1958, doi: [10.1016/0002-9394\(58\)90042-4](https://doi.org/10.1016/0002-9394(58)90042-4).
- [44] T. Willberger, C. Musterle, and S. Bergmann, "Deferred hybrid path tracing," in *Ray Tracing Gems: High-Quality and Real-Time Rendering With DXR and Other APIs*. Berlin, Germany: Springer, 2019, pp. 475–492.
- [45] N. Wissmann, M. Misiak, A. Fuhrmann, and M. Latoschik, "Accelerated stereo rendering with hybrid reprojection-based rasterization and adaptive ray-tracing," in *Proc. 2020 IEEE Conf. Virtual Reality 3D User Interfaces*, 2020, pp. 828–835, doi: [10.1109/VR46266.2020.1580932797283](https://doi.org/10.1109/VR46266.2020.1580932797283).
- [46] J. Ye, A. Xie, S. Jabbireddy, Y. Li, X. Yang, and X. Meng, "Rectangular mapping-based foveated rendering," in *Proc. 2022 IEEE Conf. Virtual Reality 3D User Interfaces*, 2022, pp. 756–764, doi: [10.1109/VR51125.2022.00097](https://doi.org/10.1109/VR51125.2022.00097).
- [47] A. Young, C. Ho, and J. R. Stafford, "Optimized shadows in a foveated rendering system," US Patent 10,650,544, May 12, 2020.
- [48] Y. Zhao et al., "Retinal-Resolution Varifocal VR," in *ACM SIGGRAPH 2023 Emerging Technologies*. New York, NY, USA: Association for Computing Machinery, 2023, doi: [10.1145/3588037.3595389](https://doi.org/10.1145/3588037.3595389).



**Zhi Zhang** received the BSc degree from Nanjing University. He is currently working toward the master's degree with the School of Software and Microelectronics, Peking University. His research focuses on foveated rendering, indirect illumination prediction, and the software development of rendering engines.



**Meng Gai** received the PhD degree from the School of Electronics Engineering and Computer Science, Peking University. He is currently an associate professor with the School of Computer Science, Peking University. His research interests include computer graphics, virtual reality, and real-time rendering.



**Sheng Li** (Member, IEEE) is currently a professor with the School of Computer Sciences, Peking University. He serves as vice director of Graphics & Interactive Technology Lab. and has published more than 50 refereed papers in prestigious journals and conferences, including *ACM Transactions on Graphics*, *Siggraph/Asia*, *IEEE Transactions on Visualization and Computer Graphics*, *CGF*, etc. His research interests include computer graphics and virtual reality. He is a member of the ACM.