Review Article

Foveated rendering: A state-of-the-art survey

Lili Wang[1,2,3], **Xuehuai Shi**[1] (✉)**, and Yi Liu**[1]

Abstract Recently, virtual reality (VR) technology has been widely used in medical, military, manufacturing, entertainment, and other fields. These applications must simulate different complex material surfaces, various dynamic objects, and complex physical phenomena, increasing the complexity of VR scenes. Current computing devices cannot efficiently render these complex scenes in real time, and delayed rendering makes the content observed by the user inconsistent with the user's interaction, causing discomfort. Foveated rendering is a promising technique that can accelerate rendering. It takes advantage of human eyes' inherent features and renders different regions with different qualities without sacrificing perceived visual quality. Foveated rendering research has a history of 31 years and is mainly focused on solving the following three problems. The first is to apply perceptual models of the human visual system into foveated rendering. The second is to render the image with different qualities according to foveation principles. The third is to integrate foveated rendering into existing rendering paradigms to improve rendering performance. In this survey, we review foveated rendering research from 1990 to 2021. We first revisit the visual perceptual models related to foveated rendering. Subsequently, we propose a new foveated rendering taxonomy and then classify and review the research on this basis. Finally, we discuss potential opportunities and open questions in the foveated rendering field. We anticipate that this survey will provide new researchers with a high-level overview of the state-of-the-art in this field, furnish experts with up-to-date information, and offer ideas alongside a framework to VR display software and hardware designers and engineers.

Keywords foveated rendering; virtual reality (VR); real-time rendering

1 Introduction

In recent years, virtual reality (VR) technology has been widely used in medical [1–3], military [4–6], manufacturing [7–9], entertainment [10–12], and other fields [13–15]. Despite the increasing computational power of devices, rendering overhead continues to increase owing to the diversification of surface materials of virtual objects, the increasing number of dynamic objects, and the higher complexity of physical phenomena to be simulated in VR applications. Moreover, Potter et al. [16] demonstrated that the visual latency tolerance threshold for the human visual system (HVS) is approximately 13 ms [16], making it more difficult for these applications to meet HVS real-time requirements. If rendering results are too delayed, users will observe that the content is inconsistent with the interaction, which creates discomfort. Therefore, improving rendering performance is a critical factor in promoting the practicality of VR technology.

Foveated rendering is an accelerated rendering technology that allocates computing resources based on HVS perceptual models. More computing resources are allocated to the fovea of human eyes, while fewer are allocated to the periphery. The fovea is responsible for clear central vision because approximately half of the optic nerve fibers are distributed in the fovea of the retina, and the remaining half is distributed to the rest of the

periphery [17]. Foveated rendering takes advantage of this inherent feature of human eyes. It performs different rendering qualities in different regions of the image. High-quality rendering is performed in the foveal region (fovea), and low-quality rendering is performed in the peripheral region (periphery). Therefore, foveated rendering can speed up rendering without sacrificing perceived visual quality.

Three challenges must be addressed in foveated rendering: the first is to use the perceptual model of the human visual system to guide foveated rendering, the second is to render different regions with different qualities, and the third is to integrate foveated rendering into existing rendering paradigms to improve rendering performance.

- **Using the perceptual model of the human visual system to guide foveated rendering.** This reduces computational overhead and ensures the user does not experience quality loss from the images generated. The basic idea of foveated rendering is to render the results of different qualities to different regions to accelerate the rendering process; therefore, it is first necessary to evaluate the rendering result quality based on the HVS. A well-designed questionnaire for user studies is a straightforward approach to evaluate the visual quality of rendering results. However, this requires many user experiments to obtain effective results, which is extremely time-consuming. Prior to conducting large-scale user studies, researchers frequently use perceptual models and related metrics to evaluate the visual quality of rendering results and then perform user evaluations based on the results with satisfactory quality, thereby improving evaluation efficiency. Visual quality is related to perceptual sensitivity [18]. The two most representative perceptual models related to the foveated rendering technique are the visual acuity and contrast sensitivity models. The visual acuity models describe the relationship between different regions in the visual field and the spatial resolution of the HVS. When applied to foveated rendering, the visual acuity models can be divided into the fall-off, binocular horopter, and ocular dominance models. Based on these models, foveated rendering allows low-quality rendering in regions with low spatial

resolutions of the HVS and high-quality rendering in regions with high spatial resolutions to improve rendering performance without perceptual loss. The contrast sensitivity models describe the relationship between different contrast levels and the sensitivity of the HVS. The application of foveated rendering mainly includes various contrast sensitivity functions (CSFs), such as the spatial CSF, spatio-temporal CSF, spatio-luminance CSF, spatio-chromatic CSF, and critical flicker fusion. According to the contrast sensitivity model, foveated rendering can allocate less computational resources to the regions with low contrast sensitivity to improve rendering performance without losing visual perception.

- **Rendering different regions with different qualities.** Foveation principles should be considered to address this challenge. Level of detail (LoD) techniques in computer graphics provide a solution to render 3D scenes composed of geometric meshes with different qualities. This increases rendering efficiency by decreasing geometric mesh complexity and maintaining unnoticed visual quality reduction. LoD techniques select different levels of details according to the viewpoint position and orientation. When using LoD technology to render geometric meshes by foveated rendering [19, 20], the user's fovea is detected first, then the meshes that must be tessellated according to the fovea are finely controlled, and finally refined meshes are used to generate high-quality rendering results in the foveal region. LoD technology is not only suitable for geometric meshes but also for other data representations, such as point cloud data [21]. In addition to the degree of mesh tessellation, the rendering sampling rate in rendering is also an essential factor that directly affects the quality of the resulting image. User behavior and performance have been evaluated in user studies [18, 22, 23]. The results showed that users could not distinguish images with a reduced sampling rate below the perceptual thresholds in the peripheral regions from full resolution images. Multi-spatial resolutions based foveated rendering methods perform high-resolution sampling for foveal regions and some important regions that users may notice, and low-resolution sampling

for peripheral regions. Alongside the concept of multi-spatial resolution, multi-temporal, multi-luminance, and multi-color resolution can also be used to accelerate foveated rendering. In this survey, these foveation principles are essential factors in our taxonomy of foveated rendering technologies.

- **Integrating foveated rendering into existing rendering paradigms to improve rendering performance.** Rasterization is the most widely studied rendering paradigm in foveated rendering [24, 25]. To rasterize the image with different resolutions in screen space, early research first rasterized the full resolution image and then reduced the image resolution in the desired region with time-consuming filters, which opposed the goal of foveated rendering. Since 2012, foveated rendering using rasterization has only performed high-resolution rendering in foveal regions and some important regions that users may notice, and low-resolution rendering in peripheral regions [26–30]. Because implementing rasterization into foveated rendering may create multiple rendering passes, a general rendering pipeline to rasterize pixels with foveated rendering in a single render pass was introduced, thereby further improving rendering efficiency [23]. The ray tracing rendering paradigm can control the number of rays emitted by each pixel. This directly supports the multi-spatial resolution. Therefore, many researchers implemented this approach with foveated rendering [31–34]. Besides rasterization and ray tracing, some studies focus on implementing other rendering paradigms into foveated rendering, such as ray casting, instant radiosity, and neural rendering [35–38]. Hence, the rendering paradigm is also an essential factor in our taxonomy.

This survey aims to review the state-of-the-art in the field of foveated rendering, and to discuss foveated rendering methods with different input data types, foveation principles, and rendering paradigms in design and implementation, especially 3D foveated rendering methods that emerged in the past 10 years.

This section briefly introduces foveated rendering concepts and challenges. Section 2 discusses the application of HVS perceptual models to foveated rendering. Section 3 proposes foveated rendering method taxonomies and classifies previous methods. Section 4 revisits early related foveated rendering research from 1990 to 2011 based on the taxonomy. Section 5 reviews methods that emerged over the past decade based on the taxonomy. Research conducted in the first 20 years and the last 10 years are separated because the focus of foveated rendering research has changed. Finally, Section 6 discusses foveated rendering open questions and opportunities.

2 Applying visual perceptual models in foveated rendering

First, the HVS visual features involved in foveated rendering are briefly summarized. Then, perceptual models are introduced after which we discuss the application of these models in foveated rendering. We recommend Weier et al.'s survey [39] to those who wish to establish a more comprehensive understanding of perception-based rendering techniques.

2.1 HVS features involved in foveated rendering

Currently, HVS primary visual features involved in foveated rendering include visual acuity and contrast sensitivity. Both are described as follows.

2.1.1 Visual acuity

Visual acuity refers to the ability to discern shapes and details of objects [40]. As the main HVS feature widely used in foveated rendering, it has the following properties:

- *Foveal\peripheral vision.* Human visual acuity is not uniform over the whole visual field. When a person looks at an object, the foveal vision scene details can be recognized; however, the peripheral vision scene cannot be clearly recognized [43]. That is, the HVS has higher visual acuity in the fovea of the human visual region and is the basis of foveated rendering. Figure 1(a) shows a schematic illustration of the foveal\peripheral vision.
- *Fusional vision.* The movement of both eyes enables the fusion of monocular images producing binocular vision. In fusional vision, the area where objects are perceived as single unified objects when viewed with both eyes is called Panum's fusional area [44]. The scene out of the Panum's fusional area is recognized as a
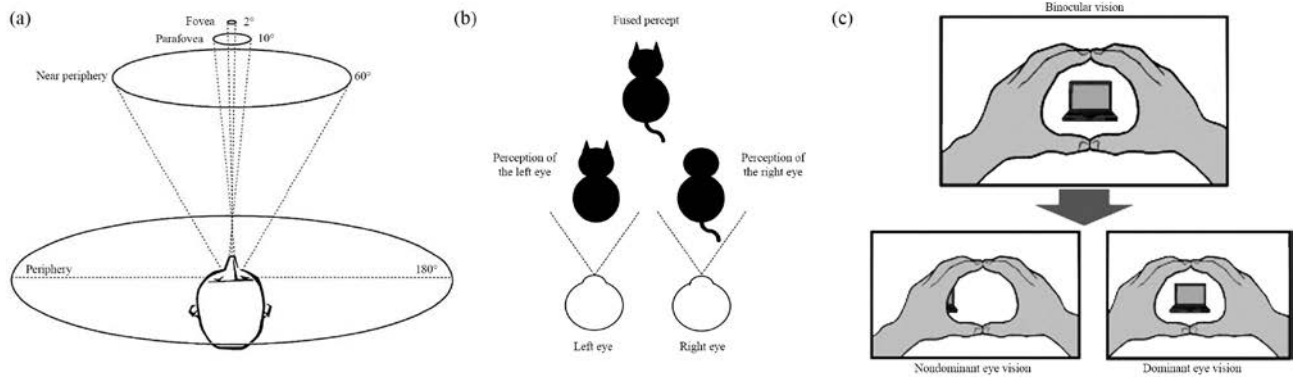
Fig. 1 (a) Schematic illustration of foveal\peripheral vision from Ivančić Valenko et al. [41]. The eccentric angle of the foveal region is very small, while the eccentric angle of the parafoveal region is up to 10°. The eccentric angle of the near peripheral region is about 60° and that of the peripheral region area is 180°. (b) Schematic illustration of fusional vision from Schaadt [42]. Two laterally placed eyes provide us two horizontally shifted and disparate images of the visual scene, which are continuously integrated into a single percept. (c) Schematic illustration of dominant eye. Compared with the nondominant eye, the dominant eye contributes more to the binocular vision. Reproduced with permission from Ref. [41], © Portal hrvatskih znanstvenih i stručnih čsopisa–Hrčak 2019; Ref. [42], © Universität des Saarlandes 2015.

"double image" with lower image quality and less visual realism [45]. It can be used to simplify the scene out of Panum's fusional area for efficient foveated rendering. Figure 1(b) shows a schematic illustration of fusional vision.

- *Dominant eye.* Both eyes have different sensitivity to visual stimuli in the HVS, i.e., one is more sensitive than the other, and the eye with higher sensitivity is called the dominant eye [46]. Less computational resources can be allocated for the non-dominant eye to speed up rendering when performing foveated rendering for binoculars. Figure 1(c) shows a schematic illustration of the dominant eye.

2.1.2 Contrast sensitivity

Contrast sensitivity refers to the ability to distinguish between foreground objects and background [47]. This varies from individual to individual, reaching a maximum at approximately the age of 20, and subsequently decreases with age. Other factors (such as cataracts and diabetic retinopathy) can also cause a decrease in contrast sensitivity. Contrast sensitivity can be considered from the following distinct aspects:

- *Spatial contrast sensitivity.* This refers to the HVS sensitivity in recognizing patterns at different frequencies [48]. For certain scene regions to be rendered, where the HVS frequency is less sensitive, it is possible to perform lower quality rendering in these regions to improve efficiency.
- *Spatio-temporal contrast sensitivity.* This refers to the HVS spatial contrast sensitivity at different retinal velocities [49]. Rendering quality can be dynamically adjusted to the current retinal

velocity to improve foveated rendering quality and performance.

- *Spatio-luminance contrast sensitivity.* This refers to the HVS spatial contrast sensitivity at different luminances. Adjusting rendering quality according this aspect in environments with different luminances can also improve foveated rendering quality and performance.
- *Spatio-chromatic contrast sensitivity.* This refers to the HVS contrast sensitivity through grating stimulation with sinusoidally changing colors [50]. In particular environments, such as bars, fog, and other scenes with prominent theme colors, foveated rendering can also use this aspect to improve perceptual quality.

2.2 Perceptual models and foveated rendering applications

Based on the HVS features, perceptual models were proposed and used in foveated rendering to approximate the HVS functions and features through mathematical descriptions. These models could guide foveated rendering design and determine the perceptual quality of the rendering result. This section reviews the perceptual models, and their application in foveated rendering based on the HVS features discussed.

2.2.1 Visual acuity models

Visual acuity models describe the function of visual acuity with neural and optical factors. Various visual acuity models have been developed based on *foveal/peripheral vision*, *fusional vision*, and *dominant eye*.

Visual acuity fall-off model. This is the psychophysical model that shows the degradation behavior of visual acuity with eccentricity [53]. Weymouth [54] demonstrated that acuity could be measured in terms of MAR (minimum angular resolution). A linear model matches both anatomical data and performance results on many vision tasks. Daniel and Whitteridge [55] proposed the cortical magnification factor (CMF), which provides the mapping from the visual angle to a cortical diameter in millimeters. The magnification factor is the largest in 0°–20° and decreases with eccentricity for the periphery. Levi et al. [56] stated that MAR increases linearly with eccentricity in the first 20°–30°. The higher the eccentricity, the faster the angular dimension rises. From the center of the visual vision to the peripheral vision, the spatial sensitivity is reduced by 35× [57]. Figure 2 shows an example of the visual acuity fall-off model from Geisler and Perry [51].

In early foveated rendering research, Levoy and Whitaker [35] combined the ray casting method used for volume rendering with the visual acuity fall-off model. For each pixel on the image plane, they first calculated the eccentricity of the pixel, then obtained the visual acuity of this pixel based on the eccentricity and the visual acuity fall-off model, and finally modulated the number of rays casting on this pixel and the number of samples per unit length of each ray based on acuity to generate the rendering result. Some studies combined vertex decimation
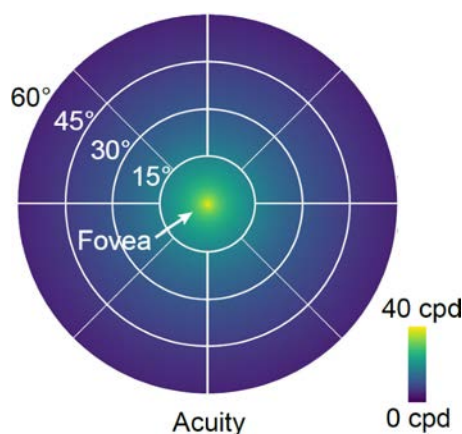

Fig. 2 Visual acuity fall-off model proposed by Geisler and Perry [51]. The HVS visual acuity is the largest (40 cpd) when the eccentricity angle is 0°. With the increase of eccentricity, visual acuity decreases linearly. When the eccentricity angle exceeds 45°, visual acuity decreases to 0 cpd. "cpd" means cycles per degree. Reproduced with permission from Ref. [52], © ACM 2021.

algorithms with the visual acuity fall-off model to dynamically adjust the number of vertices extracted based on the visual sensitivity corresponding to each pixel to achieve LoD, i.e., higher accuracy for face slices in the gaze point region and lower accuracy for face slices in the surrounding region [19, 58–60]. The behavioral performance cost of a series of perceptual experimental surface gaze level-of-detail techniques can be offset by the behavioral performance gain from increased rendering speed. In more recent research on the topic, Guenter et al. [26] simulated the acuity drop by rendering three nested layers of increasing angular diameters and decreasing resolution around the gaze direction. These layers were fused into the final result image. This work employed the CMF to decrease resolution, achieved significant shading reductions, and introduced overhead by repeating rasterization. Vaidyanathan et al. [61] proposed an architecture for the flexible control of shading rates in a GPU pipeline and tested their architecture for foveated rendering with a simplified visual acuity fall-off model. Weier et al. [62] combined the visual acuity fall-off model with the re-projection technique and applied it in the ray tracing algorithm for head mounted displays (HMDs). For each frame, if the re-projection technique cannot reuse the rendering result of the previous frame, the number of sampling rays required for the current pixel is determined by the corresponding visual acuity, with higher visual acuity requiring a larger number of sampling rays and lower visual acuity requiring a smaller number.

Binocular horopter Model. An empirical binocular horopter model was introduced by Panum [64], which reported that the sensory mechanism of the HVS fuses the images perceived by two eyes. This fusion leads to a single vision experience in the average visual direction and Panum's fusional area, as shown in Fig. 3. Mitchell [65] measured the upper limit of the parallax range, which was used to represent the upper disparity tolerance of the sensory mechanism for fusion.

In foveated rendering, Ohshima et al. [58] used fusion vision theory to control the geometric meshes level of detail. They reduced the complexity of geometry out of the fusional area to accelerate rendering. Based on the theory of fusion vision, many other studies focus on improving the depth-of-field blur effects [24, 66–72].
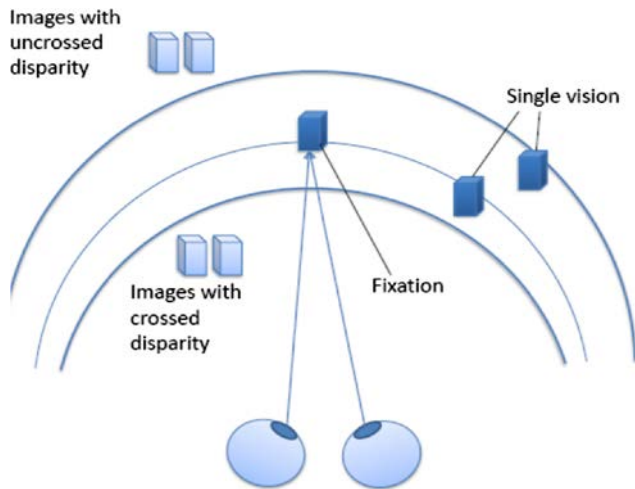
Fig. 3 Panum's fusional area: objects within the area are perceived as single images, objects further away are seen perceived with uncrossed disparity, and the objects closer to the viewer with crossed disparity. Reproduced with permission from Ref. [63], © ACM 2010.

Ocular dominance model. The ocular dominance model was proposed by Porac and Coren [46], which showed that the HVS tends to use one eye instead of both to perceive the scene. Shneor and Hochstein [73] evaluated the effect of ocular dominance under non-rivalry conditions and concluded that the dominant eye has priority in visual processing and may inhibit the performance of the non-dominant eye. Koçtekin et al. [74] evaluated the performance of the dominant eye for color vision discrimination ability among medical students with normal color vision and concluded that the dominant eye takes priority in the r/g color spectral region, probably including inhibition of the non-dominant eye.

Meng et al. [75] adopted the ocular dominance model into foveated rendering and rendered the non-dominant display with a more aggressive foveation to accelerate foveated rendering on HMDs.

2.2.2 Contrast sensitivity models

In foveated rendering, contrast sensitivity models mainly describe the HVS ability to distinguish objects from the background behind objects at different spatial frequencies [76], such as contrast sensitivity functions (CSFs); and the threshold at which an identical flickering stimulus varies in percept from flickering to stable, such as critical flicker fusions (CFFs). In CSF research, attention is paid not only to the influence of the most fundamental spatial frequencies, but also the influence of temporal frequencies, luminances, and colors [52, 77, 78]. In CFF research, attention is focused on measuring the threshold at which the HVS can perceive the stable flickering stimulus in the temporal domain [52, 79].

Spatial CSF. This was first proposed by Schade [84] and measured the contrast detection threshold of the most sensitive part of the range in a logarithmic scale range, and distributed evenly on the most sensitive part of this range, typically 1–16 cpd. Nowadays, the most commonly used spatial CSF is the threshold set measured by Watson [80] as a function of spatial frequency. Examples in Fig. 4(a) show that spatial CSF peaks between 4 and 5 cpd and falls rapidly at higher frequencies.

In early foveated rendering research, many researchers used spatial CSF to accelerate rendering by reducing the geometry complexity of the scene in the high static spatial frequency region [85–87]. Because the HVS is less sensitive to high-frequency patterns in the peripheral regions, the HVS can tolerate greater errors in the high-frequency regions of the rendered scene. Recently, Patney et al. [88] introduced a novel anti-aliasing algorithm to help recover peripheral details that are resolvable by our eyes. This algorithm provides details that the
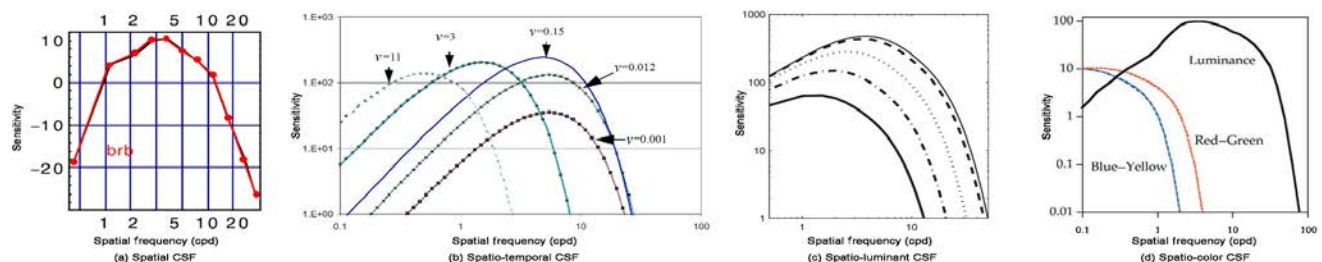

Fig. 4 CSFs over the range of spatial frequencies, temporal frequencies, luminances, and colors. (a) Spatial CSF with the measured data from Watson [80]. (b) Spatio-temporal CSF derived from sensitivity measurements in Yee et al. [81], where v is the velocities of the retinal images that measured in d/s. (c) Spatio-luminance CSF measured by Barten's model in Westland et al. [82] for stimulus of size 10 cpd and mean luminance 50 (thin solid line), 25 (dashed line), 2.5 (dotted line), 0.25 (dashdotted line), and 0.025 (thick solid line) cd/m². (d) Spatio-chromatic CSF for black–white, red–green, and yellow–blue contrast from Fairchild [83]. Reproduced with permission from Ref. [80], © Optical Society of America 2000; Ref. [81], © ACM 2001; Ref. [82], © Wiley Periodicals, Inc. 2006; Ref. [83], © John Wiley and Sons 2013.

periphery of the HVS can perceive. Koskela et al. [31] demonstrated that the smallest detail that humans can resolve is 60 cpd on average. If a rendering system could be built capable of showing 60 cpd, 95% of the rendered detail would be excessive. Then Koskela et al. [33] proposed a novel Visual-Polar coordinate space and distributed the samples according to the spatial CSF in the Visual-Polar coordinate space.

Spatio-temporal CSF. The HVS contrast sensitivity not only changes with spatial frequency but also with retinal velocities. The spatio-temporal CSF measures the HVS contrast sensitivity with spatial frequency and retinal-image motion. Kelly [49] measured the CSF by allowing the user to observe sine waves with different retinal velocities. Contrast sensitivity varies significantly with retinal velocity. Liu and Jiang [89] and Flipse et al. [90] reported that if the velocity of the retinal image is identical, the contrast sensitivity of the eye during fixation and pursuit will be equal, i.e., the motion of the retinal image, not the motion of the eye, determines contrast sensitivity. Figure 4(b) shows that the temporal CSF varies with different velocities of the retinal images.

In foveated rendering, Yee et al. [81] constructed a spatio-temporal error tolerance map based on a spatio-temporal CSF to accelerate rendering and achieved a significant improvement in speed. Stengel et al. [28] introduced a sampling scheme combined with a spatio-temporal CSF, which performs shading on regions of essential features in the image, and interpolates the remaining regions, to avoid affecting user perceived quality.

Spatio-luminance CSF. The HVS contrast sensitivity changes with luminance under the same spatial frequency. Van Meeteren and Vos [91] measured contrast sensitivity with a luminance ranging from 0.0001 to 10 cd/m² in the case of a spatial frequency ranging from 0.5 to 30 cpd. Kim et al. [77] extended the contrast sensitivity measure into higher luminance levels (150 cd/m²) with lower spatial frequencies, down to 0.125 cpd. Higher luminance levels are more relevant to photopic vision, and low frequencies are required to observe and model the CSF band-pass characteristic, especially for low luminance levels. Figure 4(c) shows that the luminance CSF varies with different mean luminances.

In foveated rendering, Stengel et al. [28] proposed a luminance map to adjust the sampling probability

such that the number of colored samples is further distributed in the image with essential features. Tursun et al. [29] proposed a new luminance-contrast-aware foveated rendering technique, which analyzed the local luminance contrast of the image to obtain a particular foveation to improve computational savings.

Spatio-chromatic CSF. The HVS contrast sensitivity changes significantly with sinusoidally changing colors at the same spatial frequency. Mullen [92] performed experiments that compare the decline in contrast sensitivity between the color-only (red–green) gratings and the monochromatic luminance gratings in the entire field of view when the spatial frequency is 2 cpd, at the center of the fovea and the eccentricity are 10° and 18° respectively. Anderson et al. [93] measured the CSF for eccentricities from 0° to 55° for chromatic red–green sinusoidal stimuli and reported that chromatic contrast declines more steeply than luminance contrast with eccentricity. Mullen and Kingdom [94] measured the cone contrast sensitivities for sine-wave grating stimuli (smoothly enveloped in space and time) for two colors (red–green and blue–yellow) and monochromatic luminance at a range of eccentricities in the nasal field (0°–25°). They identified that red–green cone opponency has a steep decline away from the fovea, while the loss in blue–yellow cone opponency is more gradual, showing a similar loss to that found for achromatic vision. Mullen et al. [95] measured the cone contrast for red–green and blue–yellow colors. The results showed that red–green cone opponency declines steeply across the human periphery and becomes behaviorally absent by 25°–30°. Chwesiuk and Mantiuk [78] reported that the color directions closer to the chromatic green-to-red axis show higher contrast sensitivity in comparison with achromatic stimuli, while for the yellow-to-blue axis, the sensitivity is lower. Figure 4(d) shows that the color CSF varies with black–white, red–green, and yellow–blue.

Duchowski and Çöltekin [96] introduced the possibility of developing a perceptually-based color degradation metric, which can be used to accelerate foveated rendering. They also investigated the peripheral color reduction with the color CSF; the results suggested that peripheral chromaticity cannot be reduced within the central 20° visual angle.

Critical flicker fusion. Besides CSFs that focus

on distinguishing objects from the background, researchers measured the threshold at which an identical flickering stimulus varies in percept from flickering to stable, i.e., critical flicker fusion (CFF) [52, 79]. Tyler and Hamer [79] introduced the Ferry–Porter law considering spatio-temporal frequency and luminance, which described that CFF increases linearly with log retinal luminance and log stimulus area, respectively. Tyler and Hamer [97] showed that the Ferry–Porter law also extends to higher eccentricities. Krajancich et al. [52] introduced a model to measure the eccentricity-dependent critical flicker fusion thresholds for space, time, and luminance. This showed that the CFF varies with spatial frequency and luminance and exhibited an anti-foveated effect, with the highest thresholds observed in the near–mid periphery of the visual field. Although no research directly applied the eccentricity-based CFF to foveated rendering algorithms, this provided a new model to improve foveated rendering efficiency.

3   3D foveated rendering taxonomies

Recent surveys proposed several taxonomies to classify existing foveated rendering techniques. In Weier et al.'s survey [39], the authors referred to foveated rendering methods as measurement-based perceptual approaches and classified them into two catalogs: one based on scene simplification, the other based on adaptive sampling. The methods in the first catalog are object-space methods. They use geometry techniques, such as LoD, to significantly reduce the scene's complexity using the visual acuity model or CSF according to the user's gaze position, thereby significantly improving time performance. The methods in the adaptive sampling class adaptively calculate the sampling rate in rendering paradigms, such as rasterization or ray tracing based on the visual acuity model or CSF.

Spjut et al. [98] proposed a two-dimensional taxonomy matrix of the foveated display. The first dimension is a resolution-contingent classification, and the second is a gaze-contingent classification. Resolution-contingent classification is based on the acuity distribution function of the human visual model. It describes how the non-linear fitting in which the angular resolution perceived by the user decreases as the gaze eccentricity or the angular displacement

from the center of gaze increases.

In the resolution classification, the foveated display can be divided into four categories according to the relationship between the visual acuity distribution function (ADF) and the display resolution distribution function (RDF) (Fig. 5). Class A is acuity matched. This is a conservative display method. The display resolution used in the foveal and peripheral region is higher than the perceptible resolution threshold in the visual acuity distribution function. This type of method ensures the user does not perceive the resolution drop. Class B is fovea matched, which means that the display resolution is higher than the user's perceptible threshold in the foveal region. In contrast, the display resolution is lower than the user's perceptible threshold in the peripheral region. To further improve efficiency, this type of method only focuses on the quality of the foveal region. Class C is periphery matched, i.e., the user does not perceive any artifacts in the peripheral region; however, the display resolution in the foveal region fails to meet or exceed user visual acuity. Class D is non-acuity matched, which means that in the foveal and peripheral region, the display resolution has not reached or exceeded the resolution threshold that can be perceived by human visual acuity. In using this type of method, the user is aware of artifacts in both regions. In the second dimension, the gaze-contingent classification, the foveated display can also be divided into four classes according to the gaze direction range in the
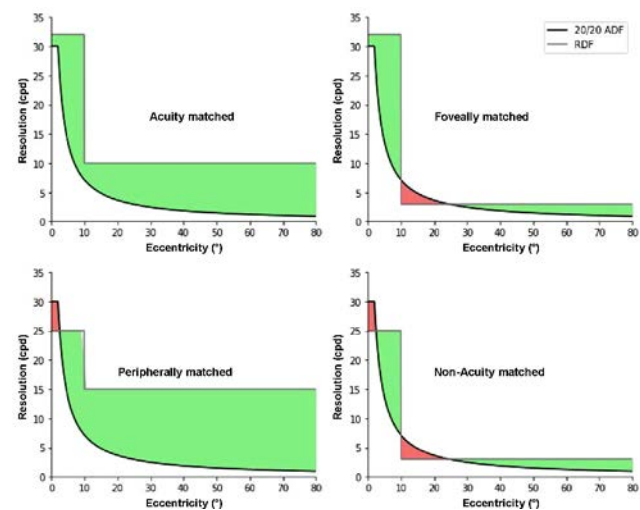

Fig. 5   Four possible comparisons of the user's visual acuity distribution function and the display's resolution distribution function. Reproduced with permission from Ref. [98], © IEEE 2020.

display. Class 1 is the fully foveated display, in which the gaze direction can be any direction in the display. Class 2 is the practically foveated display in which the gaze direction should be within ±15° from the center of the display. Class 3 is the partially foveated display in which the gaze direction should be much smaller than ±15°. Class 4 is the non-foveated display in which the single gaze direction is supported. Table 1 summarizes the relationship between resolution and gaze-contingent classifications, and provides further detailed descriptions for each.

Weier et al. [39] provided an overview of the HVS perceptual mechanisms and classified existing rendering techniques according to different perceptual mechanisms. Spjut et al. [98] focused on the classification of display effects, which is suitable for hardware display devices and measures the degree of support for foveated rendering by display devices.

The proposed taxonomy focuses on enabling researchers to easily understand the actual functions, basic ideas, technical framework of current methods and the fundamental design factors that support designers in considering and making technical decisions when designing new methods. We classify the current foveated rendering methods according to three dimensions: (1) required input data type;

(2) foveation principle; (3) rendering paradigm. Table 2 shows the elements in each dimension.

Foveated rendering works for different input data. Before understanding or designing a foveated rendering method, it is necessary to consider the processed data type. The input data type is taken as the first dimension of our foveated rendering taxonomy. Present foveated rendering methods can process the data types: image/video, volume data, geometric meshes, point cloud, hologram data, and light field.

The foveation principle is used as the second dimension to classify the previous methods. Foveated rendering provides high-quality rendering for the HVS fovea and provides unnoticeably lower-quality rendering for the periphery. Its core principle is multi-resolution rendering. The present methods use one or several different types of multi-resolution rendering under this ideology, including multi-spatial, multi-temporal, multi-luminance, multi-color, and multi geometry resolution, which is typically referred to as the LoD.

Multi-spatial resolution reduces rendering quality in the output image according to the visual acuity models and the spatial CSFs from the foveal to the peripheral region. Multi-temporal resolution based

Table 1 The classification matrix is produced by combining RDF classification (letters) with motion classification (numbers) from Spjut et al. [98]

	Class A Acuity Matched	Class B Foveally Matched	Class C Peripherally Matched	Class D Non-Acuity Matched
Class 1 Fully Foveated	For any gaze direction, the display meets or exceeds the user's visual acuity without any peripheral artifacts	For any gaze direction the foveal inset matches user acuity, but peripheral artifacts are present	The foveal inset fails to match user acuity, but achieves equal resolution over all gaze directions with no peripheral artifacts	Neither the foveal inset nor periphery matches user acuity, but the display achieves equal resolution over all gaze directions
Class 2 Practically Foveated	For a practical sub-set of gaze directions the display meets or exceeds the user's visual acuity without any peripheral artifacts	For a practical sub-set of gaze directions the foveal inset matches user acuity w peripheral artifacts present	The foveal inset fails to match user acuity, but achieves equal resolution over a practical sub-set of gaze directions with no peripheral artifacts	Neither the foveal inset nor periphery matches user acuity, but the display achieves equal resolution over a practical sub-set of gaze directions
Class 3 Partially Foveated	For a small sub-set of gaze directions the display meets or exceeds the user's visual acuity without any peripheral artifacts	For a small sub-set of gaze directions the foveal inset matches user acuity w peripheral artifacts present	The foveal inset fails to match user acuity, but achieves equal resolution over a small sub-set of gaze directions with no peripheral artifacts present	Neither the foveal inset nor periphery matches user acuity, but the display achieves equal resolution over a small subset of gaze directions
Class 4 Non-Foveated	For a single gaze direction the display meets or exceeds the user's visual acuity without any peripheral artifacts	For a single gaze direction the foveal inset matches user acuity w peripheral artifacts present	The foveal inset fails to match user acuity and foveal acuity changes with gaze, but no peripheral artifacts are ever present	Neither the foveal inset nor periphery matches user acuity, and the RDF appears to change for any given gaze direction

清华大学出版社 Tsinghua University Press ⌖ Springer

Table 2   Taxonomy vocabulary

	a	Image/Video
	b	Volume Data
1. Data Type	c	Geometric Meshes
	d	Point Cloud
	e	Hologram Data
	f	Light Field
	a	Multi-spatial Resolution
	b	Multi-temporal Resolution
2. Foveation	c	Multi-luminance Resolution
Principle	d	Multi-color Resolution
	e	Level of Detail
	a	Rasterization
	b	Ray Tracing
	c	Ray Casting
3. Rendering	d	Instant Radiosity
Paradigm	e	Shadow Mapping
	f	Online/Offline Simplification
	g	Neural Rendering
	h	Photon Mapping
	i	Phase Retrieval
	j	Data Transmission

methods render one image with multiple resolutions based on spatio-temporal CSFs. Researchers not only consider the HVS spatial error tolerance but also the spatio-temporal error tolerance of dynamic objects and take advantage of the HVS to ensure greater spatio-temporal error tolerance of dynamic objects to effectively perform foveated rendering, which achieves significant improvement in speed [81, 88]. Multi-luminance resolution based methods render one image with multiple resolutions according to spatio-luminance CSFs. Based on the HVS luminance-contrast-awareness, researchers reduce the resolution of peripheral regions with low luminance-contrast sensitivity more aggressively to further improve foveated rendering performance [29]. Foveated rendering alongside the concept of multi-color resolution [99] takes advantage of peripheral chromatic degradation, i.e., acceptable peripheral chromatic LoD, and renders one image with multiple color resolutions based on spatio-chromatic CSFs. Multi-luminance resolution and multi-color resolution based methods are also spatially multi-resolution; however, a particular difference remains in the foveation principle used. To assist readers in more clearly understanding these methods, in this survey, we separated multi-luminance and multi-color resolution based methods from the traditional multi-spatial resolution based methods. LoD reduces the complexity of the scene geometry in the periphery through visual acuity models and CSFs to reduce computing resources required to render the virtual environment.

The third classification dimension is the rendering paradigm used by existing methods to achieve multi-resolution rendering, which includes: rasterization, ray tracing, ray casting, instant radiosity, shadow mapping, online/offline simplification, photon mapping, neural rendering, and phase retrieval for holographic data. For the 360° video which is extremely popular in VR applications recently, the encoding, decoding, and transmission mode combined with foveal information directly affect foveated rendering, Thus, we also introduce the data transmission of the 360° video as an element of the rendering paradigm.

Table 3 shows the classification of 90 published reports on the foveated rendering methods from

Table 3   Summary of foveated rendering technique implementations (*: survey, +: patent, -: cutting-edge equipment)

Implementation	Data Type	Foveation Principle	Rendering Paradigm
Levoy 1990 I3D [35]	b	a	c
Funkhouser 1993 SIGGRAPH [100]	c	e	f
Ohshima 2002 VR [58]	c	e	f
Luebke 2000 Tech.Rep [59]	c	a, b, e	f
Luebke 2001 EG [19]	c	a, e	f
Parkhurst 2001 ETRA [60]	c	e	f
Loschky 2001 ARL [18]	a	a	a
Reddy 2001 CGA [101]	c	b, e	f
Yee 2001 TOG [81]	c	b	a
Murphy 2001 EG [102]	c	e	f
Parkhurst 2002 HF* [23]	/	/	/
Cheng 2003 SPA [103]	c	e	f

(Continued)

Implementation	Data Type	Foveation Principle	Rendering Paradigm
Duchowski 2004 Citeseer* [104]	/	/	/
Reingold 2003 SAGE* [105]	/	/	/
Zhou 2004 BDM [106]	b	a	c
Yu 2005 VC [107]	b	a	c
Lu 2006 EG [108]	b	a	c
Duchowski 2007 TOMM* [96]	/	/	/
Hillaire 2008 CG&A [109]	c	a	a
Hillaire 2008 VR [66]	c	a	a
Duchowski 2009 TAP [99]	a	d	a
Murphy 2009 SAP [110]	c	a	c
Mantiuk 2011 SGDA [67]	c	a	a
Guenter 2012 TOG [26]	c	a	a
Gallo 2013 ISRN [111]	b	a	c
Duchowski 2014 SAP [24]	c	a	a
Fujita 2014 SIGGRAPHAsia [112]	c	a	b
Vaidyanathan 2014 Eurographics [61]	c	a	a
Mauderer 2014 CHI [69]	c	a	a
Patney 2016 SIGGRAPH⁻ [113]	/	/	/
Patney 2016 TOG [88]	c	b	a
Stengel 2016 CGF [28]	c	b, c	a
Swafford 2016 SAP [114]	c	a, e	a, f
Pai 2016 SIGGRAPH⁻ [115]	/	/	/
Lindeberg 2016 [116]	c	e	f
Koskela 2016 ISVC [31]	c	a	b
Weier 2016 CGF [62]	c	a	b
Weier 2017 EG [39]	/	/	/
Albert 2017 TAP [117]	a	a	a
Blackmon 2017 USPatent⁺ [118]	c	a, e	a, b, f
Koskela 2017 SIGGRAPH [119]	c	a	b
Hsu 2017 MM [120]	a	a	a
Sun 2017 TOG [121]	f	a	b
Lungaro 2018 TVCG [122]	a	a	j
Meng 2018 TOG [123]	c	a	a
Koskela 2018 CVM [124]	c	a	b
Turner 2018 VR [25]	c	a	a
Molenaar 2018 [32]	c	a	b
Weier 2018 TAP [71]	c	a	b
Zheng 2018 VRST [20]	c	e	f
Tan 2018 Opt.Express⁻ [125]	/	/	/
Wilson 2018 USPatent⁺ [126]	a	a	a
Young 2019 USPatent⁺ [127]	d	a	a
Kaplanyan 2019 TOG [38]	a	a	g
Wei 2019 Appl.Opt. [128]	e	a	b
Young 2019 USPatent⁺ [129]	c	e	f
Tavakoli 2019 USPatent⁺ [30]	/	/	/
Stafford 2019 USPatent⁺ [130]	c	e	f
Tursun 2019 TOG [29]	c	c	b
Koskela 2019 EG [33]	c	a	b
Friston 2019 TOG [131]	c	a	a
Schütz 2019 VR [21]	d	e	f

(Continued)

Implementation	Data Type	Foveation Principle	Rendering Paradigm
Bruder 2019 EuroVis [37]	b	a	c
Radkowski 2019 HCII [132]	c	a	a
Siekawa 2018 MMM [133]	c	a	b
Kim 2019 TOG[-] [134]	/	/	/
Lee 2019 Opt.Express[-] [135]	/	/	/
Bastani 2020 USPatent[+] [27]	c	a	a
Spjut 2020 TVCG* [98]	/	/	/
Young 2020 USPatent[+] [136]	c	a	a, e
Koskela 2020 [34]	c	a	b
Kang 2020 ACCESS [72]	b	a	c
Ananpiriyakul 2020 EI [137]	b	a	c
Wang 2020 ISMAR [36]	c	c	d
Konrad 2020 TOG [138]	c	a	a
Joshi 2020 Access [139]	c	a	a
Meng 2020 TVCG [75]	c	a	a
Meng 2021 TVCG [140]	f	a	a
Frieß 2021 TVCG[-] [141]	/	/	/
Yoo 2020 OpEx[-] [142]	/	/	/
Bitterli 2020 SIGGRAPH [143]	c	b	c
Deza 2020 [144]	a	a	g
Yang 2021 C&G [145]	c	c	d
Franke 2021 CGF [146]	c	a, b	a
Surace 2021 [147]	a	a	g
Kim 2021 ISMAR [148]	c	a	b
Liu 2021 ISMAR [149]	c	a	b
Walton 2021 TOG [150]	a	a	a
Li 2021 TVCG [151]	a	a	j
Shi 2021 TVCG [152]	c	c	h
Chakravarthula 2021 TVCG [153]	e	a	i
Jindal 2021 TOG [154]	c	b, c	a

1990 to 2021 according to the three dimensions. In addition, the table also lists publications on new devices for foveated rendering (marked with "-"), the related surveys (*), and the related patent (+).

4   Early research from 1990 to 2011

As foveated rendering research is plentiful spanning a period greater than 30 years, it is divided into two parts organized by chronological order: early research from 1990 to 2011 and recent research over the last 10 years from 2012 to 2021.

One reason for this is that, with the development of technology, the focus of the recent research has changed compared with early research.

Firstly, the early research in this topic area focused on developing LoD techniques to reduce the complexity of geometric meshes, simulating visual blur effects to enhance the visual appearance of the rendering results by rasterization on geometric meshes and accelerating the ray casting process for rendering volume data. With the emergence of Ray Tracing Texel eXtreme (RTX), a high-end professional visual computing platform created by Nvidia that supports real-time ray tracing [155], recent research in foveated rendering paid more attention to accelerating ray tracing for geometric meshes. Prior to the emergence of RTX, previous ray tracing was only available for non-real-time applications, such as offline rendering for cinematic visual effects or photo-level realism [156].

Secondly, recent developments in other technologies have also led to a change in foveated rendering focus. For example, cloud rendering became a trend

with the development of communication technologies such as 5G, which enables content providers to render 3D programs using a remote server and send back rendered images to user terminals interactively [157]. Cloud rendering has revolutionized foveated rendering-based data transmission. The development of deep learning techniques has also used foveated images or videos to improve the accuracy of deep learning models for specific computer vision tasks. These key developments have initiated vital research hotspots in the field.

Thirdly, researchers proposed new data types, such as point cloud, hologram data, and light fields, to meet the requirements of different applications. Incorporating the rendering paradigm for these new data types into the foveated rendering framework has also become a critical research area.

Another reason for this division is that readers may have different requirements for early and recent research. For the former, typically readers solely require understanding of the methods function and fundamental ideas. While for the latter, because it is the state of the art, it may be necessary to reproduce and compare recent research, such that readers can establish a deeper understanding of contemporary foveated rendering.

Foveated 3D graphics [26] proposed in 2012 is an essential milestone for dividing research on the topic into two parts. This introduced a rasterization-based foveated rendering system to improve rasterization rendering performance, demonstrating that users cannot perceive the degradation of rendering quality from foveated rendering in this system because of the publication of extremely detailed perceptual experiments. Prior to this, foveated rendering primarily mimicked HVS visual effects to improve the visual appearance of images. In subsequent research, foveated rendering focused on improving rendering performance without perceptual loss.

In this section, early foveated rendering research is reviewed. Figure 6 visualizes the frequency of various research on the topic from 1990 to 2011 according to the proposed taxonomy. We initially summarized discussions in review papers from 1990 to 2011 (Section 4.1). Subsequently, we introduced the methods according to their frequency of occurrence from high to low: (1) foveated rendering based on LoD (Section 4.2); (2) foveated rendering based on multi-
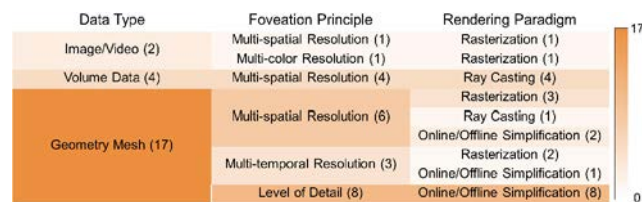

Fig. 6 Frequency of research in foveated rendering from 1990 to 2011. Numbers in parentheses indicate the number of studies using the specific data type, and foveated principle and rendering paradigm are listed in front of parentheses.

spatial resolution for volume data (Section 4.3); (3) foveated rendering based on multi-spatial resolution for geometric meshes (Section 4.4). Additionally, some techniques closely related to foveated rendering in early research are introduced (Section 4.5). In early research, foveated rendering was also referred to as gaze-directed rendering (GDR), gaze-contingent rendering (GCR), or gaze-contingent display (GCD).

4.1 Reviews

Several reviews discuss and summarize the early foveated rendering research.

For example, Reingold et al. [105] discussed gaze-contingent multi-resolution displays (GCMRD) in different areas, including engineering design research on the development of GCMRDs, multi-resolution image processing, multi-resolution sensors, human factors research on multi-resolution displays, gaze-contingent displays, and human–computer interaction. Focus was placed on reviewing methods to solve two questions regarding gaze-contingent multi-resolution displays: (1) image degradation owing to the characteristics of multi-resolution images, vision model based multi-resolution images generation methods, discrete/continuous-resolution drop-off, and color resolution drop-off were reviewed; (2) for perceptible image motion caused by image updating, gaze/head/hand-contingent displayed area of interest (D-AOI) movement-based methods and predictive D-AOI movement-based methods were analyzed. Parkhurst and Niebur [23] reviewed variable-resolution displays from the three aspects: (1) potential computational savings achieved with variable-resolution displays; (2) practical constraints in implementing variable-resolution displays; (3) the behavioral consequences of using variable-resolution displays, such as perceptual quality, task performance, and eye movement

measures. The authors also explained that gaze-related rendering in virtual reality is only one variable-resolution display application. Variable-resolution displays could also be used in low-vision enhancement and internet image transmission applications. Duchowski et al. [104] divided gaze-contingent display methods into two categories: model-based graphical displays and screen-based displays. Model-based methods used the objects' LoD to generate the image matching the resolvability of the human retina, while the screen-based methods adjusted the image quality at the pixel level. Focus plus context methods were also discussed, which were extremely similar to foveal and peripheral displays. Duchowski and Çöltekin [96] reviewed perceptually loss-less gaze-contingent displays, space-variant imaging based on the pyramidal idea, and gaze-contingent displays for stereoscopic imaging. The authors also summarized GPU-based gaze-contingent displays before 2007 and introduced related technologies including mipmapping, multitexturing, and fragment programming.

4.2 LoD

Clark [158] introduced the concept of discrete LoD, which defined several versions of the model at different levels, using a detailed grid when the object is close to the observer and replacing it with a coarser approximation when the object is far from the observer. The LoD technique can be combined with foveated rendering to reduce the complexity of scenes according to the user's gaze position and the perceptual models, which significantly improves time performance [159]. Funkhouser and Séquin [100] proposed a gaze-directed dynamic LoD selection system that considers motion blur and visual acuity. The motion blur value is expressed by the speed at which the object image moves on the retina. The visual acuity value is expressed by the distance from the object to the center of the user's gaze. Owing to the lack of an accurate perceptual model, the effect of motion blur is controlled by a slider set by the user. As there is no eye-tracking system, the user's gaze is assumed to be at the center of the screen. This research firstly introduced the concept of gaze-directed perceptual LoD. Ohshima et al. [58] used the ultrasonic sensors built into the eye-trackers to measure head direction, which is used as a substitute

for gaze direction. The authors introduced a visual acuity fall-off model, a binocular horopter model, and a kinetic vision model, respectively, to calculate visual acuity according to eye direction, and subsequently mapped the minimum visual acuity calculated by the three models to control the LoD for rendering.

As the discrete LoD technique cannot locally change details, for example, the side of a large object near the view cannot be rendered in great detail while simultaneously reducing its distant details. Rather than calculating a series of static LoDs in the pre-process, Hoppe [160] introduced the concept of continuous LoD. They built a data structure from which the desired LoD can be extracted at runtime. In foveated rendering, Luebke et al. [59] proposed a gaze-directed continuous LoD framework. They employed a commercial eye tracker to measure the user's gaze over a desktop display in real time, and then introduced a perceptual metric to measure the level of geometric meshes based on the visual acuity fall-off model proposed in Ref. [161] and the spatio-temporal contrast sensitivity function proposed in Ref. [49]. Murphy and Duchowski [102] employed a binocular eye-tracked VR system to obtain the gaze in VR, and then modeled visual acuity fall-off for both eyes based on the gaze. Subsequently, they proposed the gaze-contingent continuous LoD to degrade the resolution of meshes based on visual acuity. Luebke and Hallen [19] provided a perception based node fold system for the vertex tree, which is a hierarchical clustering of vertices. They identified that the perceptible result of a change induced by simplification can be conservatively equal to the change of its lowest spatial frequency and maximum contrast. Thus the perception-based node expansion system visits each node in the vertex tree top–down. If the lowest spatial frequency and maximum contrast induced by folding the node are less than the pre-defined threshold contrast, the system folds the node. Otherwise, the node will remain unfolded, and traversal continues. Reddy [101] noted that previous perceptually based LoD research used pre-simplified versions of an object that can be selected for rendering in a view-dependent manner. They performed a per-pixel calculation of the pixel's spatial frequency by employing the GPU, and then used the spatial frequency to determine the LoD based on an eccentricity-based spatio-temporal CSF [49].

Parkhurst et al. [60] conducted virtual search tasks to evaluate straight-forward gaze-contingent continuous LoD rendering, in which the LoD decreases linearly as the distance from the rendered object to the point of gaze increases. The results demonstrated that the behavioral performance gains could offset behavioral performance costs of gaze-contingent LoD techniques owing to increased rendering performance. Cheng [103] used surface information obtained from a 3D scanner and allowed a user to select a foveal point, and then proposed an interactive LoD update with foveation.

4.3 Multi-spatial resolution for volume data

Rendering volume data inherently consumes massive computing resources owing to large data size. Thus real-time rendering of large volume datasets was infeasible using desktop personal computers in earlier years. One solution is to use ray casting to render volume data based on the concept of multi-spatial resolution, i.e., to render objects in the foveal region at full resolution and ignore details of objects in the peripheral region, which can reduce calculation and communication requirements.

Levoy and Whitaker [35] first explored the method for incorporating foveated rendering into volume rendering. They used the Eye-Mark eye tracker to obtain the user's gaze direction and directed this at an object by rotating the user's eyes or head until the object's projection falls on the fovea. Subsequently, they distributed the number of casting rays per unit area and the number of samples taken along the unit length of each ray based on a visual acuity fall-off model. For weakening unnecessary objects in the peripheral region, Zhou et al. [106] adjusted the opacity of the sample according to the distance from the sample point to the center of the foveal region for volume feature enhancement, which assisted users in focusing more on objects in the foveal region. To further accelerate foveated volume rendering, Yu et al. [107] remapped the mask which was used to sample the rays and the length of each ray into a small number of wavelet coefficients in the wavelet domain according to the visual acuity fall-off model. Figure 7 visualizes the rendering results of full-resolution ray casting and the proposed method. Lu et al. [108] used a camera to focus on one eye and record eye movements as the user observes the volume, and employed the eccentricity-based spatio-temporal
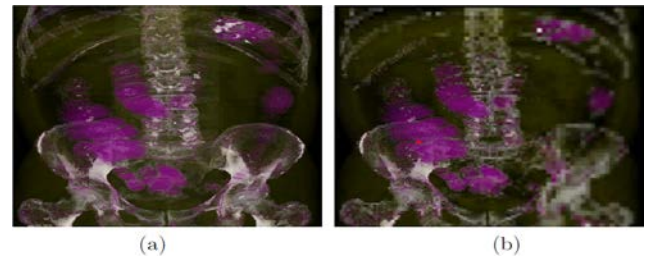

Fig. 7 Fast rendering of foveated volumes in wavelet-based representation proposed by Yu et al. [107]. (a) is rendered with a full resolution, (b) is rendered with this method, and the fovea is situated at the red dot. This method achieved a 1.3–8× improvement in speed compared with the full resolution ray casting. Reproduced with permission from Ref. [107], © Springer-Verlag 2005.

CSF [49] to acquit the HVS importance information, subsequently, they used this importance information to fix object shapes, positions and to tune opacity transfer functions automatically.

4.4 Multi-spatial resolution for geometric meshes

In addition to volume data, the concept of multiple spatial resolution is also used to accelerate geometric mesh rendering.

Murphy et al. [110] proposed a hybrid technique based on the visual acuity fall-off model and the spatial CSF, which used ray casting to sample the scene's geometry. This technique enables non-isotropic degradation within meshes without directly manipulating mesh geometry.

As early geometry models were coarse, geometric mesh performance rendering in the entire image at high resolution is acceptable. Thus, researchers in foveated rendering focused more on simulating the HVS visual appearance, i.e., gaze-contingent depth-of-field (DoF) rendering, rather than accelerating geometric mesh rendering. The traditional pinhole camera model in computer graphics can sharply present objects at all distances. However, in the eyes and real cameras, only objects within the focal range can be sharply displayed, while objects far away or close to the viewpoint are blurred. To simulate the fact that humans only perceive sharp objects within a certain distance range near the focal length and to improve the user's immersion, gaze-contingent DoF rendering was introduced in Hillaire et al. [66] and Mantiuk et al. [67]. In this review, we regard gaze-contingent DoF as a specific type of foveated rendering, which pays more attention to the scene

depth range in the foveal region.

For improving gaze-contingent DoF perception during first-person navigation in virtual environments (VE), Hillaire et al. [109] proposed a gaze-contingent DoF blur filter which simulates the blurring of objects located in front of or behind the focus point of the eyes, and a peripheral blur filter which simulates the blurring of objects situated in the periphery of the field of vision. Hillaire et al. [66] subsequently described an algorithm for calculating the focal length and point in the 3D virtual environment and used the gaze-contingent DoF blur and peripheral blur filters proposed in Hillaire et al. [109] to render the DoF blur effects to simulate the fact that humans only perceive sharp objects within a certain distance range near the focal length. Mantiuk et al. [67] evaluated human impression regarding the existence of the DoF phenomenon in the 3D virtual environment. The results demonstrated that people noticed and preferred the DoF visualization controlled by the eye tracker. The best impression was achieved with the medium blurriness level (the lens aperture diameter was 7 cm).

In early foveated rendering research, researchers also adopted concepts of multi-color and multi-temporal resolution in foveated rendering. Duchowski et al. [99] investigated the color reduction in the peripheral region. The results demonstrated that peripheral chromaticity could not be reduced within the central 20° visual angle, i.e., the color reduction should be maintained isotropically across the central 20° visual field.

4.5　Other related work

From 1990 to 2011, some other related foveated rendering research emerged, such as *perception based rendering*, *focus+context visualization*, *selective rendering*, and *multi-resolution display*.

Perception-based rendering refers to use of the HVS features and associated perceptual models to improve rendering performance and to enhance the perceptual quality of rendering results. For example, Yee et al. [81] constructed a spatio-temporal error tolerance map based on a spatio-temporal CSF that accepts low-quality rendering in highly error-tolerant regions without degrading perceptual quality, thus improving rendering speed. Unlike perception-based rendering, all HVS features and perceptual models

in foveated rendering are highly related to the HVS foveal features.

In early research involving perception-based rendering, researchers conducted user studies of perceptual models to obtain useful parameters and error metrics that have a direct impact on foveated rendering. For example, Ramasubramanian et al. [162] introduced an error metric considering the spatial-luminance CSF, which predicted the perceptual threshold to detect artifacts in 3D scenes. Myszkowski et al. [163] presented a perceptual error metric based on a spatio-temporal CSF, which retained inherent noise in the animation generated using stochastic methods below human observer sensitivity.

Focus+context visualization is a rendering technique that visualizes more critical information by removing or suppressing less critical parts of the scene. Critical information typically has semantic integrity. Focus+context visualization typically uses distortion and highlighting to visualize interested objects in focus and nearby related objects in context [164–171], while foveated rendering is based on HVS perception theories to allocate further computing sources to the foveal region.

Carpendale et al. [167] highlighted data by dedicating additional space to this and applied distortions to abstract graphs to observe interested graph nodes clearly. Viola et al. [171] proposed a view-dependent model for automatic focus+context volume visualization. This model enables interested objects to be displayed more accurately to view further details, while occluded objects are displayed with low accuracy or completely suppressed.

Selective rendering is task-dependent rendering, which uses HVS knowledge to select the objects in scenes that require rendering based on application tasks [172–175], i.e., different tasks require different objects to be drawn. For example, if the task is to count the number of pencils in a mug on a table in a room, only the image in the visual angle of the fovea centered around the pencils is rendered with high quality. Cater et al. [172] designed perceptual experiments to prove that users would ignore parts of the scene that were not related to a specific task, which can be used to reduce rendering time without affecting visual quality in interactive tasks. Sundstedt et al. [174, 175] investigated the extent to which

image resolution, edge anti-aliasing and reflection, and shadow parameters can be reduced between non-task-related and task-related regions when viewers cannot perceive image quality degradation.

Multi-resolution display focused on a more general pipeline of multi-resolution rendering [51, 176–179]. In addition to foveated rendering, the multi-resolution display can also be used for perception-based and selective rendering, etc. Duchowski and McCormick [176] introduced a multi-resolution display method based on mipmap texture mapping. They retained the original image resolution in multiple regions of interest (ROIs) selected by users and gradually reduced the periphery around each ROI according to the specified resolution mapping function. Geisler and Perry [51] developed a foveated multi-resolution pyramid video coding/decoding system that uses a foveated multi-resolution pyramid to encode each image into five or six regions of different resolutions and eliminated spatial edge artifacts between the regions generated by foveation through raised-cosine blending across levels of the pyramid and "foveation point interpolation" within pyramid levels. Geisler and Perry [177] described a multi-resolution pyramid method that used a pyramid encoder to divide the image into 2–6 layers, and used a pyramid decoder to sample each layer at different rates. Parkhurst et al. [178] introduced a two-region gaze-contingent display and investigated behavioral effects on the display based on a visual search task. They identified that reaction time and accuracy co-vary as a function of the foveal region size. For the small foveal region, slow reaction time is accompanied by high accuracy. Conversely, for the large foveal region, fast reaction time is accompanied by low accuracy. Geisler and Perry [179] proposed a method to generate completely arbitrary variable-resolution displays based on image pyramidal pre-processing [51].

5 Foveated rendering over the past decade (2012–2021)

This section reviews foveated rendering research published most recently over the past decade. Figure 8 visualizes the frequency of various foveated rendering research from 2012 to 2021 according to the proposed classification method. LoD or multi-
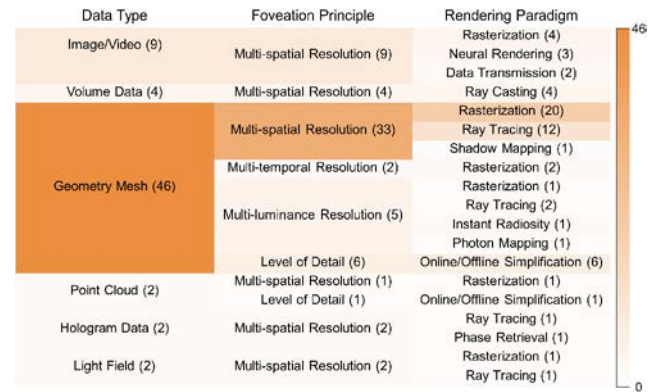

Fig. 8 Frequency of research in foveated rendering from 2012 to 2021. Numbers in parentheses indicate the number of studies using the specific data type, and foveated principle and rendering paradigm are listed in front of parentheses.

spatial resolution rasterization methods for geometric meshes, and multi-spatial resolution methods for volume data remain research hotspots. Furthermore, methods such as multi-spatial resolution ray tracing for geometric meshes, and multi-spatial resolution methods for images or videos have also attracted keen attention from researchers. In the following subsections, we introduce methods in these classes according to the frequency of occurrence from high to low: (1) foveated rendering based on multi-spatial resolution, rendering geometric meshes with rasterization (Section 5.1); (2) foveated rendering based on multi-spatial resolution, rendering geometric meshes with ray tracing (Section 5.2); (3) foveated rendering based on multi-spatial resolution, rendering image/video data (Section 5.3); (4) foveated rendering based on LoD (Section 5.4); (5) multi-spatial resolution for volume data (Section 5.5); (6) multi-luminance resolution method for geometric meshes (Section 5.6); and (7) foveated rendering for nascent data types (Section 5.7).

5.1 Multi-spatial resolution rasterization for geometric meshes

In recent years, with the development of modeling technology, the complexity of 3D models and the scale of virtual scenes have increased. In multiple virtual reality applications, using high-resolution and high-quality rasterization of the scene cannot achieve real-time frame rates. Therefore, many researchers focused on the foveated rendering method alongside improving geometric mesh rasterization performance based on the concept of multi-spatial resolution.

Guenter et al. [26] took advantage of the visual acuity fall-off model and rendered three nested layers by rasterization. The pipeline for this method is described in Fig. 9. These nested layers are rasterized as the angular diameter decreases in resolution to achieve improved rendering performance. Finally, three layers are mixed to form the final image. The results demonstrate that the rendering speed of this method is 5–6× that of the traditional method. The quality users visually perceive is comparable to traditional rendering. Vaidyanathan et al. [61] presented a novel architecture to flexibly control shading rates in a rasterization pipeline named Coarse
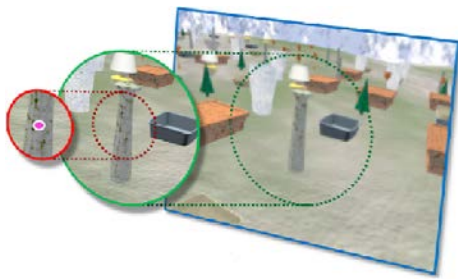
Fig. 9 Foveated 3D graphics proposed by Guenter et al. [26]. Three nested layers were rendered (red, green, and blue) at three different resolutions through rasterization based on a visual acuity fall-off model. The three nested layers are combined to generate the final image. This method could achieve comparable perceptual quality with reference to traditional full-resolution rendering, but at a 4–6.2× speed improvement. Reproduced with permission from Ref. [26], © ACM 2012.

Pixel Shading (CPS) and tested the architecture for foveated rendering with a visual acuity fall-off model. As CPS pipelines require adaptive shading features not yet commonly available on commodity GPUs, Meng et al. [123] presented a simple two-pass kernel foveated rendering (KFR) pipeline that maps well onto modern GPUs. In the first pass, they computed the kernel log-polar transformation and rendered it to a reduced-resolution buffer. The second pass carried out the inverse-log-polar transformation with anti-aliasing to map reduced-resolution rendering to the full-resolution screen. The results showed that KFR could achieve a 2.8–3.2× speed improvement in rendering on 4K UHD (2160 p) displays with less perceptual LoD.

In addition to considering the spatial factor, much research considered the temporal factor, based on the concept of multi-temporal resolution to further accelerate geometric mesh rasterization. Stengel et al. [28] introduced a sampling method based on the visual acuity fall-off model, the spatio-temporal and the spatio-luminance CSFs, and subsequently integrated the sampling method into the deferred shading pipeline. Only important image features were shaded while interpolating the remaining features without affecting perceived quality. The visualization results are shown in Fig. 10. Patney et al. [88]
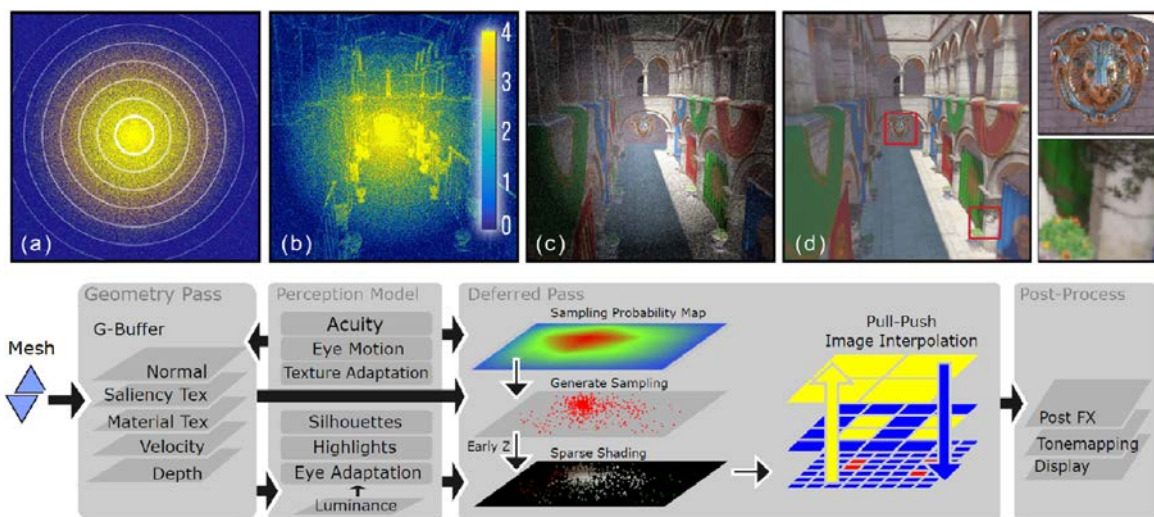

Fig. 10 Adaptive image-space sampling method for foveated rendering proposed by Stengel et al. [28]. A perceptual adaptive sampling pattern (b) was constructed for sparse shading (c), which combined visual cues such as visual acuity (a), spatial, spatio-temporal, and spatio-luminance CSFs. Fast image interpolation was performed in the periphery (d) to achieve the same perceptual quality with less shading cost. Row 2 shows the pipeline of the proposed method: in the geometry pass, this generates the G-Buffer; in the deferred pass, it first generates the sampling pattern, then performs sparse shading based on the sampling pattern, and finally uses a pull-push operation to complete the missing image parts by interpolation; in the post-processing pass, it applies post-processing operations similarly to tone mapping and grading before displaying the final image. The final image contains high details in the fovea and low details in the periphery. Reproduced with permission from Ref. [28], © The Eurographics Association and John Wiley & Sons Ltd. 2016.

designed a foveated rendering system that reduces the number of shadings by up to 70%, and the authors subsequently introduced a novel anti-aliasing algorithm based on a visual acuity fall-off model and a spatio-temporal CSF. This anti-aliasing algorithm assists in recovering peripheral region details that are resolvable by human eyes albeit degraded by filtering. Franke et al. [146] presented a foveated rendering method that comprised recycling pixels in the periphery by spatio-temporally reprojecting them from previous frames to accelerate rendering performance. This reprojection detected and re-evaluated artifacts and disocclusions according to a confidence value determined by a perception-based metric. Jindal et al. [154] proposed the variable-rate shading pipeline to accelerate rasterization rendering performance. This approach divides the output image into a number of 16×16 image tiles, and subsequently adaptively adjusts the shading accuracy and refresh rate of each image tile based on spatio-temporal and the spatio-luminance CSFs.

To further improve calculation process speed, Turner et al. [25] aligned the rendered pixel grid to virtual scene content during rasterization and upsampling, which reduced the detectability of motion artifacts in the periphery without complex interpolation or anti-aliasing algorithms. Bastani et al. [27] rendered an intermediary image of the 3D scene in the intermediary compressed space and unwarped the image to generate the foveated image. Young et al. [136] adopted foveated rendering to accelerate shadow rendering. Shadow mapping was used to obtain two shadow maps of different resolutions and geometric meshes in the foveal region were rendered with the high-resolution shadow map, while that of the peripheral region were rendered using the low-resolution shadow map.

Towards HMDs with latency and field-of-view requirements, Friston et al. [131] presented a rasterization pipeline that achieved foveated rendering in one rasterization pass with per-fragment ray-casting. Meng et al. [75] accelerated foveated rendering on HMDs with more aggressive foveation based on the theory of ocular dominance.

Foveated rendering improves the frame rate and quality of foveal vision by reducing peripheral vision resolution. However, foveated rendering optimization is a difficult task. This requires careful selection of multiple parameters, such as the number of layers, eccentricity, resolution of the peripheral region, and foveated rendering perceptibility must be evaluated. Therefore, many researchers designed perceptual studies to optimize and evaluate the task. Patney et al. [88] designed a user study to evaluate users' perceptual abilities of peripheral vision when viewing today's displays. The results demonstrated: (1) filtering peripheral regions would reduce contrast, thereby creating a sense of tunnel vision; (2) when applying the post-processing contrast enhancement function, the object could tolerate a $2\times$ larger blur radius before detecting the difference from the non-foveated ground truth. Swafford et al. [114] applied foveated rendering to the multi-resolution, screen-space ambient occlusion, and tessellation methods. Practical rules for each method were proposed to achieve significant performance gains with user studies and the newly proposed rendering quality metrics.

Recent research also concentrated on gaze-contingent DoF rendering based on the concept of multi-spatial resolution. Mauderer et al. [69] designed a user study to demonstrate that gaze-contingent DoF increased subjective perceived realism and depth and could contribute to the perception of ordinal depth and distance between objects; however, it was limited in accuracy. Duchowski et al. [24] used gaze-contingent DoF to reduce users' visual discomfort when viewing stereoscopic displays. However, similar to earlier attempts, participants disliked gaze-contingent DoF, which may be attributed to eye tracker spatial inaccuracy and the DoF simulation's noticeable temporal lag. Konrad et al. [138] extended gaze-contingent DoF rendering to ocular parallax rendering, which described the small amounts of depth-dependent image shifts on the retina created as the eye rotates. They introduced ocular parallax rendering technology that accurately rendered small amounts of gaze-contingent parallax capable of improving depth perception and realism in VR. The results demonstrated that ocular parallax rendering provided an effective ordinal depth cue and improved the impression of realistic depth in VR. Walton et al. [150] believed that the HVS perceives that the periphery is more than just blurry, and proposed a real-time method to compute images identical to ground truth images in terms of peripheral perception.

In addition, researchers applied foveated rendering to VR interaction. Joshi and Poullis [139] presented foveated rendering-based redirected walking in VR, which capitalized on naturally occurring saccades and blinks to completely refresh the framebuffer. Radkowski and Raul [132] conducted a user study to demonstrate whether the foveated rendering technique would distract users and reduce their training effect in VE. The results demonstrated that the user noticed the technology but was not negatively affected by it, and the performance difference was insignificant, except for some outliers caused by technical eye-tracking limitations.

In addition to geometric meshes, multi-spatial resolution rasterization can also be used for foveated rendering on point clouds [127].

5.2 Multi-spatial resolution based ray tracing

Ray tracing is capable of controlling the number of rays emitted from each pixel. The more rays emitted from a single pixel, the higher the rendering quality of that pixel. Therefore, the ray tracing framework naturally supports spatial multi-resolution rendering. Koskela et al. [31] provided a theoretical estimation that 94% of the rays could be omitted by integrating foveated rendering with ray tracing. Thus many researchers focused on ray tracing based on foveated rendering with the concept of multi-spatial resolution.

Fujita and Harada [112] first implemented the foveated rendering system based on ray tracing. A pre-computed sampling pattern was used with a kNN scheme to reconstruct images from sparse samples. Their system showed artifacts, without considering the eye sensitivity to contrasts and lacked pertinent input from relevant user studies. To address these challenges, Weier et al. [62] combined ray tracing based foveated rendering with reprojection rendering, using information from the previous frame to reduce the sampling rays for new frames. Subsequently, the authors applied a temporal caching and resampling scheme to improve reconstruction quality for regions that expose high contrasts and silhouettes. The results of user studies conducted demonstrated that the method achieved a real-time frame rate and compared with the fully rendered image, the visual difference was difficult to detect. Blackmon et al. [118] combined ray tracing and rasterization in a single pipeline. Ray tracing was used to render

the foveal region and rasterization to render the peripheral region. To speed up previewing the artist's points of interest, Koskela et al. [119, 124] applied foveated rendering to progressive Monte Carlo rendering, which omits more than 90% of rays that must be traced in real time. Their user study demonstrated that the perceived convergence of the proposed method was 10× faster than that of a conventional preview, and participants rated the method to have only marginally more artifacts in areas where it had to start rendering from scratch. Molenaar [32] traced rays based on the visual acuity fall-off model, and reconstructed images based on a spatial CSF. Experimental results demonstrated that this method provided a basic speed improvement of 4.3×.

Willberger et al. [180] introduced a hybrid path tracing approach to accelerate the global illumination calculation in foveated rendering. The method uses screen space path tracing to render objects with diffuse, specular, and glossy materials, using multi-bounced path tracing to render objects with the transparent material. To render direct lighting from millions of dynamic light sources interactively with ray tracing, Bitterli et al. [143] introduced the spatiotemporal reservoir resampling method to resample a set of candidate light samples based on the spatio-temporal feature, and subsequently traced rays from sampled lights to illuminate the scene. Kim et al. [148] proposed a perceptually efficient pixel sampling method suitable for HMD ray tracing, which combined the Jin et al. [181] selective oversampling technique with the foveated rendering scheme.

As linear fall-off still requires many rays in the periphery [32, 62, 112], Koskela et al. [33] traced rays and denoised in Visual-Polar space, and subsequently mapped the results to the screen space. In this method, when perceived quality is similar, rendering and denoising speed will increase by 2.5×, and ray traversal speed will increase by 1.3–1.5×. This is because primary rays maintain high coherence, and GPU resource utilization is improved. The pipeline of this method is shown in Fig. 11. Koskela [34] proposed a working prototype of a foveated ray tracing system that combined the novel Visual-Polar coordinate space proposed in Koskela et al. [33] and the regression-based reconstruction filter proposed in Koskela et al. [182] for ray tracing that runs in
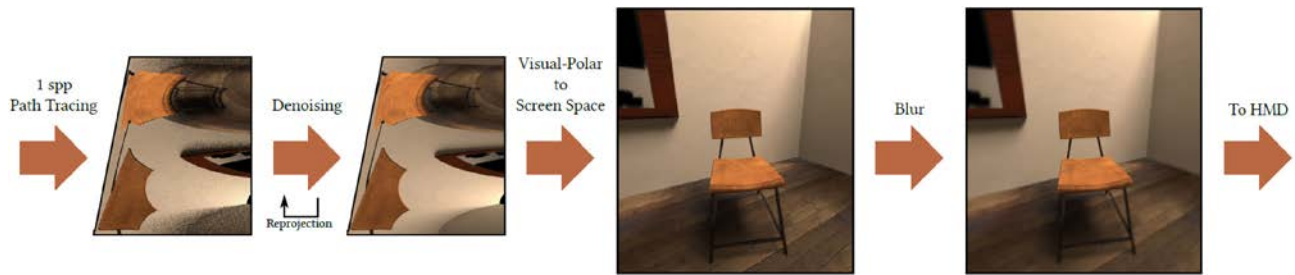
Fig. 11 Foveated real-time path tracing in visual-polar space proposed by Koskela et al. [33]. Rays were traced and rendering results denoised in a Visual-Polar space, the results were then mapped to the screen space, and finally the Guassian blur was performed to generate the final HMD rendering result. Ray tracing and denoising in Visual-Polar space increase both by 2.5× faster. Reproduced with permission from Ref. [33], © The Author(s) 2019.

real time.

Most previous methods model the sensitivity as a function of eccentricity and control the number of rays emitted according to these functions, without considering that displayed content also strongly influenced sensitivity. Tursun et al. [29] proposed a new luminance-contrast-aware foveated ray tracing technique. This technique showed that if the spatio-luminance CSF is considered in foveated rendering, the number of tracing rays can be significantly reduced. The disadvantage is that a low-quality image must be generated for each frame, indicating areas with different luminances.

For applying DoF effects in foveated ray tracing, Weier et al. [71] proposed a foveated rendering system that integrates DoF filters to hide potential visual artifacts. Results of the perceptual study showed that tracing rays reduced by more than 69% while rendering quality of this system was rated almost on par with full rendering. Liu et al. [149] developed a mathematical model to simulate the DoF effects of human eyes in VR and subsequently performed DoF-based stochastic sampling to simulate retinal blur according to this mathematical model.

5.3 Muti-spatial resolution for image/video

Muti-spatial resolution for image/video research can be divided into three categories: (1) conducting perceptual research on foveated images or videos; (2) neural rendering on foveated images or videos; (3) accelerating the encoding and transmission of 360° video streaming.

In the first category, some researchers used high-quality images/videos taken by cameras or rendered with 3D models to generate foveated images/videos by filtering or down-sampling high-quality images/videos in the peripheral region and designed user studies

to evaluate foveated rendering performance and quality parameters. Albert et al. [117] explored the effect of foveated rendering latency in VR applications. The results showed that larger foveal regions allow for more aggressive foveation, which is further pronounced for temporally stable foveation techniques. The results also demonstrated that increasing eye-tracking latency by 80–150 ms causes a significant reduction in the acceptable amount of foveation; however, a similar decrease in acceptable foveation was not identified for shorter eye-tracking latencies of 20–40 ms, suggesting that a total system latency of 50–70 ms could be tolerated. Hsu et al. [120] proposed a regression model to demonstrate the relationship between human perceived quality and foveated rendering parameters, such as the number of layers, the eccentricity degrees, and resolution of the peripheral region. The results demonstrated that (1) no absolute superior subjective assessment method exists, (2) subjects must complete further observations to confirm that foveated rendering is more imperceptible than perceptible, (3) when the eccentric angle is 7.5°+, and the peripheral region resolution is 540 p+, subjects barely notice foveated rendering, and (4) the quality of experiments level is highly dependent on the individuals and scenes.

To further improve foveated rendering speed, a small fraction of pixels are provided in the peripheral region for each frame, and hence, the image quality of the peripheral region is unacceptable. A neural rendering model was introduced to solve this problem. Kaplanyan et al. [38] proposed a generative adversarial neural network to improve the quality of images/videos in the peripheral region. The method can achieve real-time frame rates with gaze-contingent head-mounted displays on modern hardware. Figure 12 compares the results among the
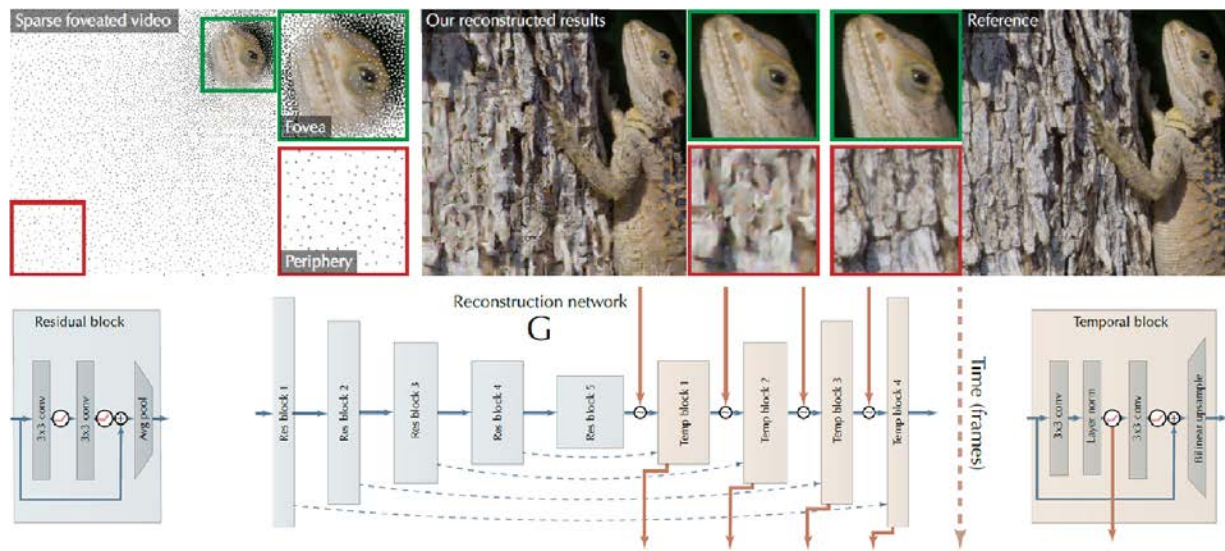
Fig. 12 Neural reconstruction for foveated rendering and video compression proposed by Kaplanyan et al. [38]. The authors reconstructed the foveated video through a generative adversarial neural network from the sparse foveated video frames with 10% of pixels (top left). This method reconstructed the video compressed by more than 14× of the original video, and the reconstructed result (top middle) had no significant reduction in perceptual quality compared with the reference (top right). The recurrent video encoder–decoder network architecture is visualized in the bottom. Reproduced with permission from Ref. [38], © Owner/Author 2019.

compressed video, reconstructed video, and reference video frames.

Some research focuses on improving the accuracy of deep learning models for specific computer vision tasks based on foveated images or videos. Deza and Konkle [144] explored the visual representation of the human foveated perceptual system, encoded the feature, and trained a convolutional neural network named Foveation-Nets to perform scene categorization. The results demonstrated that the visual representation of Fovation-Nets learning was different from the network without foveated input, and Fovation-Nets had an impact on generalization, robustness, and perceptual sensitivity. This provided computational support for the idea that the HVS foveated nature may confer a functional advantage for scene representation. Surace et al. [147] proposed a procedure to train a generative network for foveated image reconstruction. This procedure penalized perceptually significant deviations in the output to maintain perceived rather than natural image statistics.

The immersive experience offered in VR via 360° video is becoming increasingly popular. However, current bandwidth can barely accommodate the 360° video streaming solution that delivers the entire HD 360° video frame in real time. As most of the pixels in 360° video are invisible or located in peripheral regions, streaming 360° video based on the fovea is a more efficient solution. Therefore, encoding and transmission of 360° video based on the fovea constitutes important foveated rendering research. Li et al. [151] proposed a log-linear transformation method to encode original HD 360° video frames based on the fovea and to transmit them to HMDs, which maintain full-resolution fidelity in the fovea and have improved perceptual blurring effects in the periphery. Figure 13 compares the final rendering results to the client, encoded by the traditional log-polar transformation and the log-rectilinear transformation in the server, respectively. To increase the transmission speed of the 360° video stream from the server to head-mounted displays, Lungaro et al. [122] proposed a gaze-aware transmission approach for 360° video streaming services, which delivered high visual quality images around the users' gaze points in real time while lowering quality elsewhere. The results of user studies demonstrated that compared with traditional solutions, the bandwidth required to provide users with a high quality of experience level, was reduced by up to 83%.

5.4 LoD

In recent years, some research focused on the foveated rendering method based on the LoD technique.
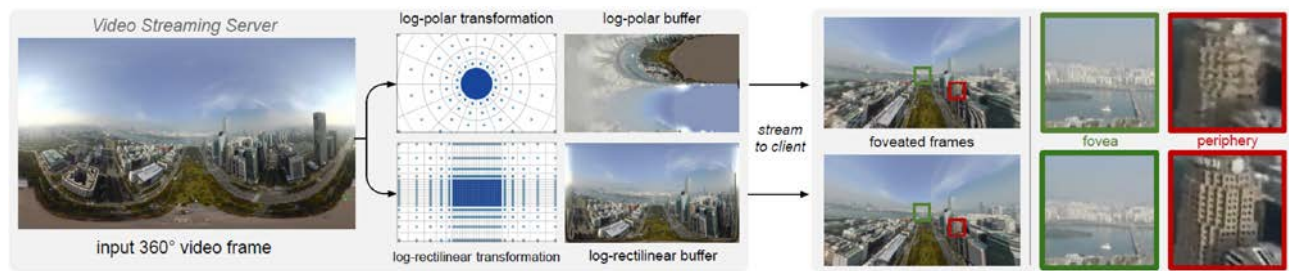
Fig. 13 Log-rectilinear transformation for foveated 360° video streaming proposed by Li et al. [151]. The upper and lower rows present the workflows with prior log-polar transformation and the proposed log-rectilinear transformation respectively. Both foveated methods convert the equirectangular video frames into down-sampled buffers, and subsequently encode and stream buffers to the client. On the client side, buffers are decoded to the screen space to generate the final results. The log-rectilinear transformation reduces flickering and aliasing artifacts in both the foveal and peripheral regions more significantly than that of the prior log-polar transformation. Reproduced with permission from Ref. [151], © IEEE 2021.

Different from previous years, researchers focused on designing user studies to optimize or select various parameters involved in the previous method or refine previous methods instead of proposing new LoD methods.

Swafford et al. [114] designed a user study that compares a foveated rendered image with an eccentric angle of 9° and a reference image at full resolution in random order. Three LoDs are generated on the scene geometry: high, medium, and low. A lower level means that there is a less tessellated grid for each tile. The results demonstrated that users had a similar visual experience to the foveated LoD rendered image with the medium level in the peripheral region and the full-resolution reference image. However, time performance could be improved by 3×. As Swafford et al. [114] only applied the tessellation method to fixed-size triangles, the results of tessellation of much larger or smaller triangles do not match the visual perceptual size. Zheng et al. [20] adaptively adjusted the tessellation levels and culling region based on visual sensitivity. Young and Stafford [129] adjusted the foveal region size and shape to correct the gaze tracing error or state parameters and combined this technique with LoD to render foveated images. Stafford and Young [130] selectively filtered the images in the peripheral region to reduce visual artifacts owing to contrast resulting from the lower LoD before compositing foveated images for presentation. Lindeberg [116] proposed a gaze-contingent depth of field tessellation that applies tessellation to all objects within the focal plane, gradually decreasing tessellation levels as applied blur increases. User studies demonstrated that this technique helps reduce the number of primitives

rendered by approximately 70% and frame time by approximately 9% compared with using fully adaptive tessellation.

Researchers not only applied LoD-based foveated rendering to scenes with geometric meshes, but also to point clouds to improve time performance. Schütz et al. [21] proposed a continuous LoD method for rendering large point clouds in real time. This method continuously recreated a down-sampled vertex buffer from the full point cloud, based on camera orientation, position, and distance to the camera, in a point-wise fashion and at a speed of 17 million points per millisecond.

5.5 Multi-spatial resolution for volume data

In recent years, with the increase in GPU computing power, researchers have further proposed more complex techniques to improve the efficiency of volume data foveated rendering.

Gallo and Placitelli [111] introduced a hybrid CPU–GPU volume ray-casting system for interactive, medical-quality visualization using an ordinary desktop PC. The system combined three parts: a gaze-directed volume rendering tool that renders the foveal region in maximum resolution, an inner structure tool that enables interactive inspection of data using two different transfer functions simultaneously, and a localized oversampling tool that allows users to interactively execute oversampling and antialiasing techniques in the foveal region. Bruder et al. [37] accelerated volume rendering through the Linde–Buzo–Gray sampling method based on the visual acuity fall-off model and natural neighbor interpolation. Ananpiriyakul et al. [137] smoothly transited the resolution from the foveal to the

peripheral region with the use of face-tracking to drive adaptive-resolution volume data visualization. The results demonstrated a 2–2.5× frame rate improvement on interactive explorations. Kang et al. [72] proposed a thin lens camera model to simulate rays passing through different parts of the lens for volume data visualizations. The model is implemented in the GPU pipeline with no pre-processing. The results demonstrated that the method could generate volume data visualizations with better depth perception than existing DoF methods, and the speed was 9× faster.

5.6 Multi-luminance resolution

The concept of multi-luminance resolution has only been used in foveated rendering in the past 5 years.

Stengel et al. [28] presented a luminance map to adjust the sampling probability of the periphery to obtain shading samples that can effectively shade important features of the image. Tursun et al. [29] proposed a novel luminance-contrast-aware foveated rendering technique that improves computational savings by analyzing the local luminance contrast of the image, this method pipeline is demonstrated in Fig. 14. Wang et al. [36] proposed the foveated

Fig. 14 Luminance-contrast-aware foveated rendering proposed by Tursun et al. [29]. A low-resolution image was first rendered, and then divided into multiple small patches, and subsequently, the standard deviation ρ was calculated to obtain the maximum acceptable resolution reduction for each patch. Finally, the luminance-contrast-aware adaptive resolution rendering was performed through real-time ray tracing. Compared with standard foveated rendering, this method achieved a 0.8–2.6× acceleration and improved perceptual quality. Reproduced with permission from Ref. [29], © ACM 2019.

instant radiosity method that casts more VPLs to illuminate the foveal region such that more accurate global illumination effects in the foveal region and less accurate global illumination in the peripheral region can be rendered. Yang et al. [145] improved the method proposed by Wang et al. [36] and created a CMF-based perceptual probability map to manage virtual point lights more accurately to further improve rendering quality in the fovea. Because the method of Wang et al. [36] and Yang et al. [145] only supports diffuse scenes, Shi et al. [152] adopted the photon mapping method to foveated rendering, which renders high-quality global illumination effects in the foveal region at interactive frame rates for the scenes that include diffuse, specular, glossy, and transparent materials.

5.7 Foveated rendering for nascent data types

With the rise of 3D display technologies, new data types appear, such as hologram data and light fields. However, current hardware and graphic algorithms cannot enable high quality and low latency for 3D displays. Researchers extend the foveated rendering algorithms to support these nascent data types.

Researchers extended foveated rendering methods from 3D geometry scenes to 4D light fields based on the concept of multi-spatial resolution. Sun et al. [121] proposed a 4D light field foveated rendering method with importance sampling and a sparse reconstruction scheme based on the spectral bounds and depth perception measurements. The results demonstrated that the technique traced only 16%–30% rays without compromising perceptual quality. Meng et al. [140] introduced a 3D-kernel foveated rendering method to observe light fields, which provided similar visual results as the original light fields. However, this achieves a speed improvement of up to 7.28× for the light fields with a resolution of 25×25×1024×1024 p with minimal perceptual loss of detail.

Foveated rendering research has also been published based on the concept of multi-spatial resolution to improve the rendering of holograms. Wei and Sakamoto [128] proposed an angle-changeable foveated ray tracing method for rendering the computer-generated hologram (CGH) with better performance and almost no observable artifacts for the user. Chakravarthula et al. [153] reduced the
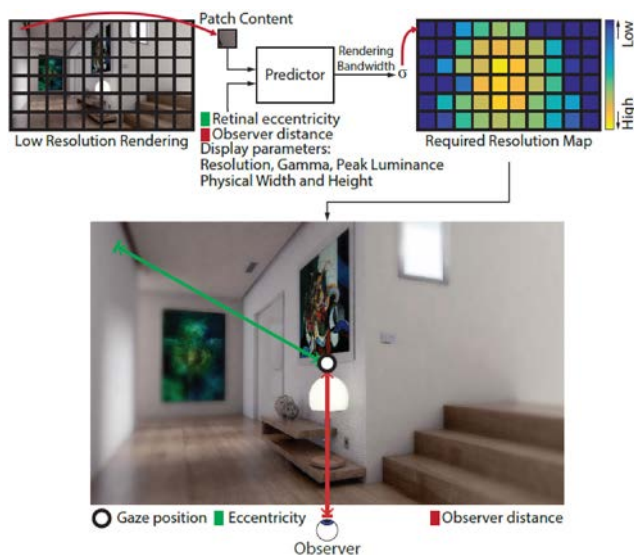
perceived speckle noise by integrating two factors into the phase hologram computation: (1) foveal and peripheral vision HVS characteristics; (2) the retinal point spread function. With this new method, the perceived speckle noise can be pushed from the fovea to the periphery.

6 Discussion

Although foveated rendering has been a focus area in research and industry for more than two decades, there are still many opportunities and open questions to be solved.

One potential opportunity is to take full advantage of the human visual features for foveated rendering. The current foveated rendering method only uses parts of the HVS features, including visual acuity and contrast sensitivity, and other features that may be beneficial in this context are not reflected in existing research; therefore, further research is required to investigate this. For example, visual masking may be utilized for accelerating foveated rendering. This explains that the visibility of one image, called a target, can be reduced by the presence of another image, called a mask [183]. For example, as the luminance or scene changes sharply, the HVS sensitivity will decrease when a new scene suddenly appears. Therefore, decreasing rendering quality of the foveal image in the subsequent frames will not cause the user to notice the difference. We believe that the next important step towards foveated rendering is effectively capitalizing of human visual features to achieve more aggressive foveated rendering without compromising perceptual awareness.

Another potential opportunity is to apply computer vision and artificial intelligence technologies to address some issues for current foveated rendering methods. Some explorations on this aspect have been completed. To further improve user gaze tracking accuracy, Arabadzhiyska et al. [184] proposed a method to predict the landing position of the gaze position during saccades in foveated rendering preprocessing. Kaplanyan et al. [38] employed a generated adversarial neural network in the foveated rendering post-processing stage, which reconstructed details in the fovea and generated temporally stable peripheral content. Other technologies, for example, the attention model, could also be considered for

integration into the foveated rendering paradigm to improve quality and performance.

The development of cutting-edge foveated displays is another potential avenue for foveated rendering. In recent years, Tan et al. [125] used beam splitters with different magnifications to combine two identical displays to demonstrate a dynamic foveal VR display. Lee et al. [135] introduced a time-multiplexed see-through fixed foveated holographic display using a beam splitter and tunable lens, with a foveal field of view of 1.04° and a peripheral field of view of 22.6°. Kim et al. [134] presented a foveated display with resolution and focal depth dynamically driven by gaze tracking for AR. The display combines a traveling micro-display for the high-resolution foveal region with a wide field-of-view peripheral display that follows the viewer's pupil during eye movement. However, current foveated displays for VR and AR have high mechanical complexities and drawbacks for responsiveness and power draw. Focus depth estimation of current displays is not robust; although previous research supports the feasibility of estimating focal depth based on binocular astigmatism alone, it has also been reported that half diopters or more are inaccurate [185]. The combination of foveated displays and prescription corrective optics also presents a challenge.

Based on the analysis and summary of existing foveated rendering methods, some open questions require urgent solutions.

Currently, many studies have been published on foveated rendering methods for volume data and geometric meshes, and concepts are relatively mature. Only in recent years foveated rendering research of hologram data and the light fields is nascent. Generally, foveated rendering methods involving volume data and geometric meshes are used for reference, such as the ray tracing method. Thus, further research is required to identify a more suitable foveated rendering method for these new data types.

Although the ray tracing framework can be adopted into foveated rendering in a straightforward manner, this is inefficient for some special effects in 3D rendering, such as global illumination for the scene containing point light sources, and high-detailed caustics. Some rendering paradigms render these special effects more efficiently; however, they cannot be directly integrated into foveated rendering.

Adopting these rendering paradigms to support foveated rendering is therefore a challenge. Methods proposed in Refs. [36, 145, 152] are interesting attempts. Based on the concept of multi-luminance resolution, they adopt instant radiosity and photon mapping to foveated rendering. Based on different foveation principles, many other efficient real-time rendering paradigms, such as bidirectional path tracing [186] and vertex connection and merging [187], etc., can be applied to foveated rendering for improved performance.

To evaluate foveated images/video quality, the straightforward method is to design perceptual experiments to collect user's perception information. As perceptual experiments are typically time-consuming and costly, they should be performed for methods with a greater chance of success. Therefore, some objective metrics based on the biological and physical theories involving foveated rendering must be proposed to quickly evaluate the feasibility of the tested foveated rendering methods. Currently, some metrics exist to evaluate foveated image quality, for example: (1) The foveal signal-to-noise ratio (FSNR) [188] valued the distortion between foveated images and reference images with a weighted signal-to-noise ratio. FSNR failed to consider user perception of foveated images quality, which may cause perceptual deviations in evaluating foveated images. (2) The foveated wavelet image quality index (FWQI) [189] calculated the wavelet coefficient difference between foveated and reference images with the integration of spatial CSF. FWQI did not consider spatio-temporal CSF while it was reported that the contrast sensitivity of the HVS can be significantly influenced by the retinal velocity [190]. (3) The foveated mean squared error (FMSE) [191] evaluated foveated video quality with the consideration of both spatial and spatio-temporal CSF. FMSE assumed that eye fixation points are always located at the center of images. This assumption potentially introduces biases in evaluating visual quality. (4) The window-based structural similarity index (WSSIM) [192] used different rules to evaluate foveated image quality for different windows on the foveated images, the scoring rules for the window closer to the fovea will be more stringent. WSSIM relies on selecting an appropriate saliency model. However, this may bias foveated image evaluation results. Thus far,

the lack of a more general, comprehensive, and widely accepted metric has significantly complicated the evaluation of foveated images/video quality. In addition, constructing datasets to evaluate different foveated images/video aspects could ensure improved comparability of evaluation results.

In recent years, most foveated rendering methods designed are for VR applications, and few methods aim toward AR applications. Kim et al. [134] investigated foveated rendering under AR. The focus was predominantly on the design of a dynamically-foveated augmented reality display. For AR applications that require virtual and real fusion, the degree of fusion will directly affect the quality of rendering results; therefore, the question of how to control the degree of fusion to generate images of different qualities in different regions remains an open challenge. For information-enhanced AR applications, it is also worth exploring whether relevant content such as scene semantic and task target information can be added to foveated rendering.

In addition to improving rendering speed, foveated rendering can also be used to complete specific tasks. For example, Joshi and Poullis [139] presented foveated rendering-based redirected walking in VR, which rendered a high-quality region to guide the spatially-varying rotation and updated peripheral framebuffer during inattentional blindness. Whether foveated rendering can assist or improve other VR and AR tasks is yet to be explored.

7   Conclusions

This paper surveys research and development involving foveated rendering over the past 31 years. Visual perception theories and taxonomies regarding foveated rendering are discussed in-depth. We respectively review early foveated rendering technologies (from 1990 to 2011) and those that have more recently emerged over the past decade (from 2012 to 2021). Finally, we discuss potential opportunities and open questions for future research in this field.

Declaration of competing interest

The authors have no competing interests to declare that are relevant to the content of this article.

References

[1] Corrêa, C. G.; Nunes, F. L. S.; Bezerra, A.; Carvalho, P. M. Evaluation of VR medical training applications under the focus of professionals of the health area. In: Proceedings of the ACM Symposium on Applied Computing, 821–825, 2009.

[2] Hsieh, M. C.; Lin, Y. H. VR and AR applications in medical practice and education. *Hu Li Za Zhi* Vol. 64, No. 6, 12–18, 2017.

[3] Hsieh, M. C.; Lee, J.-J. Preliminary study of VR and AR applications in medical and healthcare education. *Journal of Nursing and Health Studies* Vol. 3, No. 1, 1, 2018.

[4] Rizzo, A.; Morie, J. F.; Williams, J.; Pair, J.; Buckwalter, J. G. Human emotional state and its relevance for military VR training. In: Proceedings of the 11th International Conference on Human Computer Interaction, 2005.

[5] Lele, A. Virtual reality and its military utility. *Journal of Ambient Intelligence and Humanized Computing* Vol. 4, No. 1, 17–26, 2013.

[6] Ahir, K.; Govani, K.; Gajera, R.; Shah, M. Application on virtual reality for enhanced education learning, military training and sports. *Augmented Human Research* Vol. 5, No. 1, 7, 2020.

[7] Ong, S. K.; Nee, A. Y. C. *Virtual and Augmented Reality Applications in Manufacturing.* London: Springer London, 2004.

[8] Choi, S.; Jung, K.; Do Noh, S. Virtual reality applications in manufacturing industries: Past research, present findings, and future directions. *Concurrent Engineering* Vol. 23, No. 1, 40–63, 2015.

[9] Doolani, S.; Wessels, C.; Kanal, V.; Sevastopoulos, C.; Jaiswal, A.; Nambiappan, H.; Makedon, F. A review of extended reality (XR) technologies for manufacturing training. *Technologies* Vol. 8, No. 4, 77, 2020.

[10] Avila, L.; Bailey, M. Virtual reality for the masses. *IEEE Computer Graphics and Applications* Vol. 34, No. 5, 103–104, 2014.

[11] Bialkova, S.; Van Gisbergen, M. S. When sound modulates vision: VR applications for art and entertainment. In: Proceedings of the IEEE 3rd Workshop on Everyday Virtual Reality, 1–6, 2017.

[12] Saint-Louis, C.; Hamam, A. Survey of haptic technology and entertainment applications. In: Proceedings of the SoutheastCon, 1–7, 2021.

[13] Puggioni, M. P.; Frontoni, E.; Paolanti, M.; Pierdicca, R.; Malinverni, E. S.; Sasso, M. A content creation tool for AR/VR applications in education: The ScoolAR framework. In: *Augmented Reality, Virtual Reality, and Computer Graphics. Lecture Notes in Computer Science, Vol. 12243.* De Paolis, L.; Bourdot, P. Eds. Springer Cham, 205–219, 2020.

[14] Ferdani, D.; Fanini, B.; Piccioli, M. C.; Carboni, F.; Vigliarolo, P. 3D reconstruction and validation of historical background for immersive VR applications and games: The case study of the Forum of Augustus in Rome. *Journal of Cultural Heritage* Vol. 43, 129–143, 2020.

[15] Tanenbaum, T. J.; Hartoonian, N.; Bryan, J. "How do I make this thing smile?": An inventory of expressive nonverbal communication in commercial social virtual reality platforms. In: Proceedings of the CHI Conference on Human Factors in Computing Systems, 1–13, 2020.

[16] Potter, M. C.; Wyble, B.; Hagmann, C. E.; McCourt, E. S. Detecting meaning in RSVP at 13 ms per picture. *Attention, Perception, & Psychophysics* Vol. 76, No. 2, 270–279, 2014.

[17] Hendrickson, A. E.; Yuodelis, C. The morphological development of the human fovea. *Ophthalmology* Vol. 91, No. 6, 603–612, 1984.

[18] Loschky, L. C.; McConkie, G. W.; Yang, J.; Miller, M. E. Perceptual effects of a gaze-contingent multi-resolution display based on a model of visual sensitivity. In: Proceedings of the ARL Federated Laboratory 5th Annual Symposium-ADID Consortium, 53–58, 2001.

[19] Luebke, D.; Hallen, B. Perceptually driven simplification for interactive rendering. In: *Rendering Techniques 2001. Eurographics.* Gortler, S. J.; Myszkowski, K. Eds. Springer Vienna, 223–234, 2001.

[20] Zheng, Z. P.; Yang, Z.; Zhan, Y. W.; Li, Y. Q.; Yu, W. X. Perceptual model optimized efficient foveated rendering. In: Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology, 1–2, 2018.

[21] Schütz, M.; Krösl, K.; Wimmer, M. Real-time continuous level of detail rendering of point clouds. In: Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces, 103–110, 2019.

[22] Loschky, L. C.; McConkie, G. W. User performance with gaze contingent multiresolutional displays. In: Proceedings of the Symposium on Eye Tracking Research & Applications, 97–103, 2000.

[23] Parkhurst, D. J.; Niebur, E. Variable-resolution displays: A theoretical, practical, and behavioral evaluation. *Human Factors* Vol. 44, No. 4, 611–629, 2002.

[24] Duchowski, A. T.; House, D. H.; Gestring, J.; Wang, R. I.; Krejtz, K.; Krejtz, I.; Mantiuk, R.; Bazyluk, B. Reducing visual discomfort of 3D stereoscopic displays with gaze-contingent depth-of-field. In: Proceedings of the ACM Symposium on Applied Perception, 39–46, 2014.

[25] Turner, E.; Jiang, H. M.; Saint-Macary, D.; Bastani, B. Phase-aligned foveated rendering for virtual reality headsets. In: Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces, 1–2, 2018.

[26] Guenter, B.; Finch, M.; Drucker, S.; Tan, D.; Snyder, J. Foveated 3D graphics. *ACM Transactions on Graphics* Vol. 31, No. 6, Article No. 164, 2012.

[27] Bastani, B.; Funt, B.; Vignaud, S.; Jiang, H. Smoothly varying foveated rendering. US Patent 10,546,364, 2020.

[28] Stengel, M.; Grogorick, S.; Eisemann, M.; Magnor, M. Adaptive image-space sampling for gaze-contingent real-time rendering. *Computer Graphics Forum* Vol. 35, No. 4, 129–139, 2016.

[29] Tursun, O. T.; Arabadzhiyska-Koleva, E.; Wernikowski, M.; Mantiuk, R.; Seidel, H. P.; Myszkowski, K.; Didyk, P. Luminance-contrast-aware foveated rendering. *ACM Transactions on Graphics* Vol. 38, No. 4, Article No. 98, 2019.

[30] Tavakoli, M.; Khan, M.; Renschler, M.; Mondal, M. Scene-based foveated rendering of graphics content. US Patent 10,482,648, 2019.

[31] Koskela, M.; Viitanen, T.; Jääskeläinen, P.; Takala, J. Foveated path tracing. In: *Advances in Visual Computing. Lecture Notes in Computer Science, Vol. 10072.* Springer Cham, 723–732, 2016.

[32] Molenaar, E. N. Towards real-time ray tracing through foveated rendering. Master Thesis. University of Utrecht, 2018.

[33] Koskela, M.; Lotvonen, A.; Mäkitalo, M.; Kivi, P.; Viitanen, T.; Jääskeläinen, P. Foveated real-time path tracing in visual-polar space. In: *Eurographics Symposium on Rendering - DL-only and Industry Track.* Boubekeur, T.; Sen, P. Eds. The Eurographics Association, 2019.

[34] Koskela, M. Foveated path tracing with fast reconstruction and efficient sample distribution. Dissertation. Dissertation. Tampere University, 2020.

[35] Levoy, M.; Whitaker, R. Gaze-directed volume rendering. In: Proceedings of the Symposium on Interactive 3D Graphics, 217–223, 1990.

[36] Wang, L. L.; Li, R. Z.; Shi, X. H.; Yan, L. Q.; Li, Z. C. Foveated instant radiosity. In: Proceedings of the IEEE International Symposium on Mixed and Augmented Reality, 1–11, 2020.

[37] Bruder, V.; Schulz, C.; Bauer, R.; Frey, S.; Weiskopf, D.; Ertl, T. Voronoi-based foveated volume rendering. In: *EuroVis 2019 - Short Papers.* Johansson, J.; Sadlo, F.; Marai, G. E. Eds. The Eurographics Association, 2019.

[38] Kaplanyan, A. S.; Sochenov, A.; Leimkühler, T.; Okunev, M.; Goodall, T.; Rufo, G. DeepFovea: Neural reconstruction for foveated rendering and video compression using learned statistics of natural videos. *ACM Transactions on Graphics* Vol. 38, No. 6, Article No. 212, 2019.

[39] Weier, M.; Stengel, M.; Roth, T.; Didyk, P.; Eisemann, E.; Eisemann, M.; Grogorick, S.; Hinkenjann, A.; Kruijff, E.; Magnor, M.; et al. Perception-driven accelerated rendering. *Computer Graphics Forum* Vol. 36, No. 2, 611–643, 2017.

[40] Cline, D. *Dictionary of Visual Science.* Chilton Book Company, 1980.

[41] Ivančić Valenko, S.; Cviljušac, V.; Modrić, D. The impact of physical parameters on the perception of the moving elements in peripheral part of the screen. *Tehnički vjesnik* Vol. 26 No. 5, 1444–1450, 2019.

[42] Schaadt, A. K. Disorders of binocular convergent fusion and stereoscopic space perception following acquired brain damage: Treatment and neuro-anatomical implications. Dissertation. Universität des Saarlandes, 2015.

[43] Strasburger, H.; Rentschler, I.; Jüttner, M. Peripheral vision and pattern recognition: A review. *Journal of Vision* Vol. 11, No. 5, 13, 2011.

[44] Fender, D.; Julesz, B. Extension of panum's fusional area in binocularly stabilized vision. *Journal of the Optical Society of America* Vol. 57, No. 6, 819–830, 1967.

[45] Georgeson, M. A.; Wallis, S. A. Binocular fusion, suppression and diplopia for blurred edges. *Ophthalmic and Physiological Optics* Vol. 34, No. 2, 163–185, 2014.

[46] Porac, C.; Coren, S. The dominant eye. *Psychological Bulletin* Vol. 83, No. 5, 880–897, 1976.

[47] Robson, J. G. Spatial and temporal contrast-sensitivity functions of the visual system. *Journal of the Optical Society of America* Vol. 56, No. 8, 1141–1142, 1966.

[48] Campbell, F. W.; Robson, J. G. Application of Fourier analysis to the visibility of gratings. *The Journal of Physiology* Vol. 197, No. 3, 551–566, 1968.

[49] Kelly, D. H. Motion and vision. II. Stabilized spatio-temporal threshold surface. *Journal of the Optical Society of America* Vol. 69, No. 10, 1340–1349, 1979.

[50] Mullen, K. T. The contrast sensitivity of human colour vision to red-green and blue-yellow chromatic gratings. *The Journal of Physiology* Vol. 359, No. 1, 381–400, 1985.

[51] Geisler, W.; Perry, J. Real-time foveated multiresolution system for low-bandwidth video communication. In: Proceedings of the SPIE 3299, Human Vision and Electronic Imaging III, 294–305, 1998.

[52] Krajancich, B.; Kellnhofer, P.; Wetzstein, G. A perceptual model for eccentricity-dependent spatio-temporal flicker fusion and its applications to foveated graphics. *ACM Transactions on Graphics* Vol. 40, No. 4, Article No. 47, 2021.

[53] Weymouth, F. W. Visual sensory units and the minimal angle of resolution. *American Journal of Ophthalmology* Vol. 46, No. 1, 102–113, 1958.

[54] Weymouth, F. W. Visual sensory units and the minimum angle of resolution. *Optometry and Vision Science* Vol. 40, No. 9, 550–568, 1963.

[55] Daniel, P. M.; Whitteridge, D. The representation of the visual field on the cerebral cortex in monkeys. *The Journal of Physiology* Vol. 159, No. 2, 203–221, 1961.

[56] Levi, D. M.; Klein, S. A.; Aitsebaomo, A. P. Vernier acuity, crowding and cortical magnification. *Vision Research* Vol. 25, No. 7, 963–977, 1985.

[57] Nakayama, K. Properties of early motion processing: Implications for the sensing of egomotion. In: *Perception and Control of Self-motion.* Psychology Press, 93–104, 1990.

[58] Ohshima, T.; Yamamoto, H.; Tamura, H. Gaze-directed adaptive rendering for interacting with virtual space. In: Proceedings of the IEEE Virtual Reality Annual International Symposium, 103–110, 2002.

[59] Luebke, D.; Hallen, B.; Newfield, D.; Watson, B. Perceptually driven simplification using gaze-directed rendering. Technical Report CS-2000-04. Department of Computer Science, University of Virginia, 2000.

[60] Parkhurst, D.; Law, I.; Niebur, E. Evaluating gaze-contingent level of detail rendering of virtual environments using visual search. 2001.

[61] Vaidyanathan, K.; Salvi, M.; Toth, R.; Foley, T.; Akenine-Möller, T.; Nilsson, J.; Munkberg, J.; Hasselgren, J.; Sugihara, M.; Clarberg, P.; et al. Coarse pixel shading. In: Proceedings of the High Performance Graphics, 9–18, 2014.

[62] Weier, M.; Roth, T.; Kruijff, E.; Hinkenjann, A.; Pérard-Gayot, A.; Slusallek, P.; Li, Y. Foveated real-time ray tracing for head-mounted displays. *Computer Graphics Forum* Vol. 35, No. 7, 289–298, 2016.

[63] Mikkola, M.; Boev, A.; Gotchev, A. Relative importance of depth cues on portable autostereoscopic display. In: Proceedings of the 3rd Workshop on Mobile Video Delivery, 63–68, 2010.

[64] Panum, P. L. *Physiologische Untersuchungen über das Sehen mit zwei Augen.* Schwer, 1858.

[65] Mitchell, D. E. A review of the concept of "panum's fusional areas". *Optometry and Vision Science* Vol. 43, No. 6, 387–401, 1966.

[66] Hillaire, S.; Lecuyer, A.; Cozot, R.; Casiez, G. Using an eye-tracking system to improve camera motions and depth-of-field blur effects in virtual environments. In: Proceedings of the IEEE Virtual Reality Conference, 47–50, 2008.

[67] Mantiuk, R.; Bazyluk, B.; Tomaszewska, A. Gaze-dependent depth-of-field effect rendering in virtual environments. In: *Serious Games Development and Applications. Lecture Notes in Computer Science, Vol. 6944.* Ma, M.; Fradinho Oliveira, M.; Madeiras Pereira, J. Eds. Springer Berlin Heidelberg, 1–12, 2011.

[68] Vinnikov, M.; Allison, R. S. Gaze-contingent depth of field in realistic scenes: The user experience. In: Proceedings of the Symposium on Eye Tracking Research and Applications, 119–126, 2014.

[69] Mauderer, M.; Conte, S.; Nacenta, M. A.; Vishwanath, D. Depth perception with gaze-contingent depth of field. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 217–226, 2014.

[70] Gupta, K.; Kazi, S. Gaze contingent depth of field display. 2016. Available at http://stanford.edu/class/ee367/Winter2016/Gupta_Kazi_Report.pdf.

[71] Weier, M.; Roth, T.; Hinkenjann, A.; Slusallek, P. Foveated depth-of-field filtering in head-mounted displays. *ACM Transactions on Applied Perception* Vol. 15, No. 4, Article No. 26, 2018.

[72] Kang, J.; Lee, J.; Shin, Y. G.; Kim, B. Depth-of-field rendering using progressive lens sampling in direct volume rendering. *IEEE Access* Vol. 8, 93335–93345, 2020.

[73] Shneor, E.; Hochstein, S. Eye dominance effects in feature search. *Vision Research* Vol. 46, No. 25, 4258–4269, 2006.

[74] Koçtekin B.; Gündoğan, N. Ü.; Altıntaş, A. G. K.; Yazıcı, A. C. Relation of eye dominance with color vision discrimination performance ability in normal subjects. *International Journal of Ophthalmology* Vol. 6, No. 5, 733–738, 2013.

[75] Meng, X. X.; Du, R. F.; Varshney, A. Eye-dominance-guided foveated rendering. *IEEE Transactions on Visualization and Computer Graphics* Vol. 26, No. 5, 1972–1980, 2020.

[76] Owsley, C. Contrast sensitivity. *Ophthalmology Clinics of North America* Vol. 16, No. 2, 171–177, 2003.

[77] Kim, K. J.; Mantiuk, R.; Lee, K. H. Measurements of achromatic and chromatic contrast sensitivity functions for an extended range of adaptation luminance. In: Proceedings of the SPIE 8651, Human Vision and Electronic Imaging XVIII, 86511A, 2013.

[78] Chwesiuk, M.; Mantiuk, R. Measurements of contrast sensitivity for peripheral vision. In: Proceedings of the ACM Symposium on Applied Perception, 1–9, 2019.

[79] Tyler, C. W.; Hamer, R. D. Analysis of visual modulation sensitivity. IV. Validity of the Ferry-Porter law. *Journal of the Optical Society of America A* Vol. 7, No. 4, 743–758, 1990.

[80] Watson, A. B. Visual detection of spatial contrast patterns: Evaluation of five simple models. *Optics Express* Vol. 6, No. 1, 12–33, 2000.

[81] Yee, H.; Pattanaik, S.; Greenberg, D. P. Spatiotemporal sensitivity and visual attention for efficient rendering of dynamic environments. *ACM Transactions on Graphics* Vol. 20, No. 1, 39–65, 2001.

[82] Westland, S.; Owens, H.; Cheung, V.; Paterson-Stephens, I. Model of luminance contrast-sensitivity function for application to image assessment. *Color Research & Application* Vol. 31, No. 4, 315–319, 2006.

[83] Fairchild, M. D. *Color Appearance Models.* John Wiley & Sons, 2013.

[84] Schade, O. H. Optical and photoelectric analog of the eye. *Journal of the Optical Society of America* Vol. 46, No. 9, 721–739, 1956.

[85] Xia, J. C.; Varshney, A. Dynamic view-dependent simplification for polygonal models. In: Proceedings of the 7th Annual IEEE Visualization, 327–334, 2009.

[86] Hoppe, H. View-dependent refinement of progressive meshes. In: Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques, 189–198, 1997.

[87] Luebke, D.; Erikson, C. View-dependent simplification of arbitrary polygonal environments. In: Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques, 199–208, 1997.

[88] Patney, A.; Salvi, M.; Kim, J.; Kaplanyan, A.; Wyman, C.; Benty, N.; Luebke, D.; Lefohn, A. Towards foveated rendering for gaze-tracked virtual reality. *ACM Transactions on Graphics* Vol. 35, No. 6, Article No. 179, 2016.

[89] Liu, Y. M.; Jiang, B. C. Contrast sensitivity measured during smooth pursuit movement. *Science China Chemistry* Vol. 27, No. 7, 710–721, 1984.

[90] Flipse, J. P.; Wildt, G. J. V. D.; Rodenburg, M.; Keemink, C. J.; Knol, P. G. M. Contrast sensitivity for oscillating sine wave gratings during ocular fixation and pursuit. *Vision Research* Vol. 28, No. 7, 819–826, 1988.

[91] Van Meeteren, A.; Vos, J. J. Resolution and contrast sensitivity at low luminances. *Vision Research* Vol. 12, No. 5, 825–833, 1972.

[92] Mullen, K. T. Colour vision as a post-receptoral specialization of the central visual field. *Vision Research* Vol. 31, No. 1, 119–130, 1991.

[93] Anderson, S. J.; Mullen, K. T.; Hess, R. F. Human peripheral spatial resolution for achromatic and chromatic stimuli: Limits imposed by optical and retinal factors. *The Journal of Physiology* Vol. 442, No. 1, 47–64, 1991.

[94] Mullen, K. T.; Kingdom, F. A. A. Differential distributions of red-green and blue-yellow cone opponency across the visual field. *Visual Neuroscience* Vol. 19, No. 1, 109–118, 2002.

[95] Mullen, K. T.; Sakurai, M.; Chu, W. Does L/M cone opponency disappear in human periphery? *Perception* Vol. 34, No. 8, 951–959, 2005.

[96] Duchowski, A. T.; Çöltekin, A. Foveated gaze-contingent displays for peripheral LOD management, 3D visualization, and stereo imaging. *ACM Transactions on Multimedia Computing, Communications, and Applications* Vol. 3, No. 4, Article No. 6, 2007.

[97] Tyler, C. W.; Hamer, R. D. Eccentricity and the ferry-porter law. *Journal of the Optical Society of America A* Vol. 10, No. 9, 2084–2087, 1993.

[98] Spjut, J.; Boudaoud, B.; Kim, J.; Greer, T.; Albert, R.; Stengel, M.; Aksit, K.; Luebke, D. Toward standardized classification of foveated displays. *IEEE Transactions on Visualization and Computer Graphics* Vol. 26, No. 5, 2126–2134, 2020.

[99] Duchowski, A. T.; Bate, D.; Stringfellow, P.; Thakur, K.; Melloy, B. J.; Gramopadhye, A. K. On spatiochromatic visual sensitivity and peripheral color LOD management. *ACM Transactions on Applied Perception* Vol. 6, No. 2, Article No. 9, 2009.

[100] Funkhouser, T. A.; Séquin, C. H. Adaptive display algorithm for interactive frame rates during visualization of complex virtual environments. In: Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques, 247–254, 1993.

[101] Reddy, M. Perceptually optimized 3D graphics. *IEEE Computer Graphics and Applications* Vol. 21, No. 5, 68–75, 2001.

[102] Murphy, H.; Duchowski, A. T. Gaze-contingent level of detail rendering. In: Proceedings of the Eurographics 2001 - Short Presentations, 2001.

[103] Cheng, I. Foveated 3D model simplification. In: Proceedings of the 7th International Symposium on Signal Processing and Its Applications, 241–244, 2003.

[104] Duchowski, A. T.; Cournia, N.; Murphy, H. Gaze-contingent displays: Review and current trends. 2004. Available at http://andrewd.ces.clemson.edu/gcd/adc04.pdf.

[105] Reingold, E. M.; Loschky, L. C.; McConkie, G. W.; Stampe, D. M. Gaze-contingent multiresolutional displays: An integrative review. *Human Factors* Vol. 45, No. 2, 307–328, 2003.

[106] Zhou, J. L.; Döring, A.; Tönnies, K. D. Distance based enhancement for focal region based volume rendering. In: *Bildverarbeitung für die Medizin 2004. Informatik aktuell.* Tolxdorff, T.; Braun, J.; Handels, H.; Horsch, A.; Meinzer, H. P. Eds. Springer Berlin Heidelberg, 199–203, 2004.

[107] Yu, H.; Chang, E. C.; Huang, Z. Y.; Zheng, Z. J. Fast rendering of foveated volumes in wavelet-based representation. *The Visual Computer* Vol. 21, No. 8, 735–744, 2005.

[108] Lu, A. D.; Maciejewski, R.; Ebert, D. Volume composition using eye tracking data. In: Proceedings of the 8th Joint Eurographics/IEEE VGTC Conference on Visualization, 115–122, 2006.

[109] Hillaire, S.; Lécuyer, A.; Cozot, R.; Casiez, G. Depth-of-field blur effects for first-person navigation in virtual environments. *IEEE Computer Graphics and Applications* Vol. 28, No. 6, 47–55, 2008.

[110] Murphy, H. A.; Duchowski, A. T.; Tyrrell, R. A. Hybrid image/model-based gaze-contingent rendering. *ACM Transactions on Applied Perception* Vol. 5, No. 4, Article No. 22, 2009.

[111] Gallo, L.; Placitelli, A. P. High-fidelity visualization of large medical datasets on commodity hardware. *ISRN Biomedical Engineering* Vol. 2013, Article ID 892967, 2013.

[112] Fujita, M.; Harada, T. Foveated real-time ray tracing for virtual reality headset. In: Proceedings of the SIGGRAPH Asia, 2014.

[113] Patney, A.; Kim, J.; Salvi, M.; Kaplanyan, A.; Wyman, C.; Benty, N.; Lefohn, A.; Luebke, D. Perceptually-based foveated virtual reality. In: Proceedings of the ACM SIGGRAPH Emerging Technologies, 1–2, 2016.

[114] Swafford, N. T.; Iglesias-Guitian, J. A.; Koniaris, C.; Moon, B.; Cosker, D.; Mitchell, K. User, metric, and computational evaluation of foveated rendering methods. In: Proceedings of the ACM Symposium on Applied Perception, 7–14, 2016.

[115] Pai, Y. S.; Tag, B.; Outram, B.; Vontin, N.; Sugiura, K.; Kai, K. Z. GazeSim: Simulating foveated rendering using depth in eye gaze for VR. In: Proceedings of the ACM SIGGRAPH Posters, 1–2, 2016.

[116] Lindeberg, T. Concealing rendering simplifications using gaze contingent depth of field. Master Thesis. KTH Royal Institute of Technology, 2016.

[117] Albert, R.; Patney, A.; Luebke, D.; Kim, J. Latency requirements for foveated rendering in virtual reality. *ACM Transactions on Applied Perception* Vol. 14, No. 4, Article No. 25, 2017.

[118] Blackmon, S.; Peterson, L. T.; Ozdas, C.; Clohset, S. J. Foveated rendering. US Patent App. 15/372,589, 2017.

[119] Koskela, M.; Immonen, K.; Viitanen, T.; Jääskeläinen, P.; Multanen, J.; Takala, J. Foveated instant preview for progressive rendering. In: Proceedings of the SIGGRAPH Asia Technical Briefs, 1–4, 2017.

[120] Hsu, C. F.; Chen, A.; Hsu, C. H.; Huang, C. Y.; Lei, C. L.; Chen, K. T. Is foveated rendering perceivable in virtual reality? Exploring the efficiency and consistency of quality assessment methods. In: Proceedings of the 25th ACM International Conference on Multimedia, 55–63, 2017.

[121] Sun, Q.; Huang, F. C.; Kim, J.; Wei, L. Y.; Luebke, D.; Kaufman, A. Perceptually-guided foveation for light field displays. *ACM Transactions on Graphics* Vol. 36, No. 6, Article No. 192, 2017.

[122] Lungaro, P.; Sjoberg, R.; Valero, A. J. F.; Mittal, A.; Tollmar, K. Gaze-aware streaming solutions for the next generation of mobile VR experiences. *IEEE Transactions on Visualization and Computer Graphics* Vol. 24, No. 4, 1535–1544, 2018.

[123] Meng, X. X.; Du, R. F.; Zwicker, M.; Varshney, A. Kernel foveated rendering. *Proceedings of the ACM on Computer Graphics and Interactive Techniques* Vol. 1, No. 1, Article No. 5, 2018.

[124] Koskela, M. K.; Immonen, K. V.; Viitanen, T. T.; Jääskeläinen, P. O.; Multanen, J. I.; Takala, J. H. Instantaneous foveated preview for progressive Monte Carlo rendering. *Computational Visual Media* Vol. 4, No. 3, 267–276, 2018.

[125] Tan, G. J.; Lee, Y. H.; Zhan, T.; Yang, J. L.; Liu, S.; Zhao, D. F.; Wu, S. T. Foveated imaging for near-eye displays. *Optics Express* Vol. 26, No. 19, 25076–25085, 2018.

[126] Wilson, A.; Lanman, D. R.; Trail, N. D.; McEldowney, S. C.; McNally, S. J.; Sulai, Y. N. B. Rendering composite content on a head mounted display including a high resolution inset. US Patent 9,972,071, 2018.

[127] Young, A.; Ho, C.; Stafford, J. R. Foveal adaptation of particles and simulation models in a foveated rendering system. US Patent 10,339,692, 2019.

[128] Wei, L. J.; Sakamoto, Y. Fast calculation method with foveated rendering for computer-generated holograms using an angle-changeable ray-tracing method. *Applied Optics* Vol. 58, No. 5, A258–A266, 2019.

[129] Young, A.; Stafford, J. R. Real-time user adaptive foveated rendering. US Patent 10,192,528, 2019.

[130] Stafford, J. R.; Young, A. Selective peripheral vision filtering in a foveated rendering system. US Patent 10,169,846, 2019.

[131] Friston, S.; Ritschel, T.; Steed, A. Perceptual rasterization for head-mounted display image synthesis. *ACM Transactions on Graphics* Vol. 38, No. 4, Article No. 97, 2019.

[132] Radkowski, R.; Raul, S. Impact of foveated rendering on procedural task training. In: *Virtual, Augmented and Mixed Reality. Multimodal Interaction. Lecture Notes in Computer Science, Vol. 11574.* Chen, J.; Fragomeni, G. Eds. Springer Cham, 258–267, 2019.

[133] Siekawa, A.; Chwesiuk, M.; Mantiuk, R.; Piórkowski, R. Foveated ray tracing for VR headsets. In: *MultiMedia Modeling. Lecture Notes in Computer Science, Vol. 11295.* Kompatsiaris, I.; Huet, B.; Mezaris, V.; Gurrin, C.; Cheng, W. H.; Vrochidis, S. Eds. Springer Cham, 106–117, 2018.

[134] Kim, J.; Jeong, Y.; Stengel, M.; Akşit, K.; Albert, R.; Boudaoud, B.; Greer, T.; Kim, J.; Lopes, W.; Majercik, Z.; et al. Foveated AR. *ACM Transactions on Graphics* Vol. 38, No. 4, Article No. 99, 2019.

[135] Lee, J. S.; Kim, Y. K.; Lee, M. Y.; Won, Y. H. Enhanced see-through near-eye display using time-division multiplexing of a Maxwellian-view and holographic display. *Optics Express* Vol. 27, No. 2, 689–701, 2019.

[136] Young, A.; Ho, C.; Stafford, J. R. Optimized shadows in a foveated rendering system. US Patent 10,650,544, 2020.

[137] Ananpiriyakul, T.; Anghel, J.; Potter, K.; Joshi, A. A gaze-contingent system for foveated multiresolution visualization of vector and volumetric data. *Electronic Imaging* Vol. 32, No. 1, 374, 2020.

[138] Konrad, R.; Angelopoulos, A.; Wetzstein, G. Gaze-contingent ocular parallax rendering for virtual reality.

[139] Joshi, Y.; Poullis, C. Inattentional blindness for redirected walking using dynamic foveated rendering. *IEEE Access* Vol. 8, 39013–39024, 2020.

[140] Meng, X. X.; Du, R. F.; JaJa, J. F.; Varshney, A. 3D-kernel foveated rendering for light fields. *IEEE Transactions on Visualization and Computer Graphics* Vol. 27, No. 8, 3350–3360, 2021.

[141] Frieß, F.; Braun, M.; Bruder, V.; Frey, S.; Reina, G.; Ertl, T. Foveated encoding for large high-resolution displays. *IEEE Transactions on Visualization and Computer Graphics* Vol. 27, No. 2, 1850–1859, 2021.

[142] Yoo, C.; Xiong, J. H.; Moon, S.; Yoo, D.; Lee, C. K.; Wu, S. T.; Lee, B. Foveated display system based on a doublet geometric phase lens. *Optics Express* Vol. 28, No. 16, 23690–23702, 2020.

[143] Bitterli, B.; Wyman, C.; Pharr, M.; Shirley, P.; Lefohn, A.; Jarosz, W. Spatiotemporal reservoir resampling for real-time ray tracing with dynamic direct lighting. *ACM Transactions on Graphics* Vol. 39, No. 4, Article No. 148, 2020.

[144] Deza, A.; Konkle, T. Emergent properties of foveated perceptual systems. *arXiv preprint* arXiv:2006.07991, 2020.

[145] Yang, Q. Q.; Chen, Z. X.; Liu, Y. L.; Xing, G. Y.; Zhang, Y. C. Foveated light culling. *Computers & Graphics* Vol. 97, 200–207, 2021.

[146] Franke, L.; Fink, L.; Martschinke, J.; Selgrad, K.; Stamminger, M. Time-warped foveated rendering for virtual reality headsets. *Computer Graphics Forum* Vol. 40, No. 1, 110–123, 2021.

[147] Surace, L.; Wernikowski, M.; Tursun, C.; Myszkowski, K.; Mantiuk, R.; Didyk, P. Learning foveated reconstruction to preserve perceived image statistics. *arXiv preprint* arXiv:2108.03499, 2021.

[148] Kim, Y.; Ko, Y.; Ihm, I. Selective foveated ray tracing for head-mounted displays. In: Proceedings of the IEEE International Symposium on Mixed and Augmented Reality, 413–421, 2021.

[149] Liu, J. Y.; Mantel, C.; Forchhammer, S. Perception-driven hybrid foveated depth of field rendering for head-mounted displays. In: Proceedings of the IEEE International Symposium on Mixed and Augmented Reality, 1–10, 2021.

[150] Walton, D. R.; Dos Anjos, R. K.; Friston, S.; Swapp, D.; Akşit, K.; Steed, A.; Ritschel, T. Beyond blur: Real-time ventral metamers for foveated rendering. *ACM Transactions on Graphics* Vol. 40, No. 4, Article No. 48, 2021.

[151] Li, D.; Du, R. F.; Babu, A.; Brumar, C. D.; Varshney,

A. A log-rectilinear transformation for foveated 360-degree video streaming. *IEEE Transactions on Visualization and Computer Graphics* Vol. 27, No. 5, 2638–2647, 2021.

[152] Shi, X. H.; Wang, L. L.; Wei, X. H.; Yan, L. Q. Foveated photon mapping. *IEEE Transactions on Visualization and Computer Graphics* Vol. 27, No. 11, 4183–4193, 2021.

[153] Chakravarthula, P.; Zhang, Z.; Tursun, O.; Didyk, P.; Sun, Q.; Fuchs, H. Gaze-contingent retinal speckle suppression for perceptually-matched foveated holographic displays. *IEEE Transactions on Visualization and Computer Graphics* Vol. 27, No. 11, 4194–4203, 2021.

[154] Jindal, A.; Wolski, K.; Myszkowski, K.; Mantiuk, R. K. Perceptual model for adaptive local shading and refresh rate. *ACM Transactions on Graphics* Vol. 40, No. 6, Article No. 281, 2021.

[155] Alwani, R. Microsoft and Nvidia tech to bring photorealistic games with ray tracing. 2018. Available at https://www.gadgets360.com/laptops/news/microsoft-dxr-nvidia-rtx-ray-tracing-volta-gpu-metro-exodus-1826988.

[156] Sanzharov, V. V.; Frolov, V. A.; Galaktionov, V. A. Survey of nvidia RTX technology. *Programming and Computing Software* Vol. 46, No. 4, 297–304, 2020.

[157] Mukhina, K.; Bezgodov, A. The method for real-time cloud rendering. *Procedia Computer Science* Vol. 66, 697–704, 2015.

[158] Clark, J. H. Hierarchical geometric models for visible surface algorithms. *Communications of the ACM* Vol. 19, No. 10, 547–554, 1976.

[159] Luebke, D.; Reddy, M.; Cohen, J. D.; Varshney, A.; Watson, B.; Huebner, R. Temporal detail. In: *Level of Detail for 3D Graphics.* Amsterdam: Elsevier, 301–329, xvii, 2003.

[160] Hoppe, H. Progressive meshes. In: Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques, 99–108, 1996.

[161] Rovamo, J.; Virsu, V. An estimation and application of the human cortical magnification factor. *Experimental Brain Research* Vol. 37, No. 3, 495–510, 1979.

[162] Ramasubramanian, M.; Pattanaik, S. N.; Greenberg, D. P. A perceptually based physical error metric for realistic image synthesis. In: Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques, 73–82, 1999.

[163] Myszkowski, K.; Tawara, T.; Akamine, H.; Seidel, H. P. Perception-guided global illumination solution for animation rendering. In: Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques, 221–230, 2001.

[164] Furnas, G. W. Generalized fisheye views. *ACM SIGCHI Bulletin* Vol. 17, No. 4, 16–23, 1986.

[165] Sarkar, M.; Brown, M. H. Graphical fisheye views of graphs. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 83–91, 1992.

[166] Lamping, J.; Rao, R.; Pirolli, P. A focus+context technique based on hyperbolic geometry for visualizing large hierarchies. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 401–408, 1995.

[167] Carpendale, T.; Cowperthwaite, D. J.; Fracchia, F. D. Distortion viewing techniques for 3-dimensional data. In: Proceedings of the IEEE Symposium on Information Visualization, 46–53, 1996.

[168] Plaisant, C.; Grosjean, J.; Bederson, B. B. SpaceTree: Supporting exploration in large node link tree, design evolution and empirical evaluation. In: Proceedings of the IEEE Symposium on Information Visualization, 57–64, 2002.

[169] Kosara, R.; Miksch, S.; Hauser, H. Focus context taken literally. *IEEE Computer Graphics and Applications* Vol. 22, No. 1, 22–29, 2002.

[170] Munzner, T.; Guimbretière, F.; Tasiran, S.; Zhang, L.; Zhou, Y. H. TreeJuxtaposer: Scalable tree comparison using Focus+Context with guaranteed visibility. *ACM Transactions on Graphics* Vol. 22, No. 3, 453–462, 2003.

[171] Viola, I.; Kanitsar, A.; Groller, M. E. Importance-driven volume rendering. In: Proceedings of the IEEE Visualization, 139–145, 2004.

[172] Cater, K.; Chalmers, A.; Ledda, P. Selective quality rendering by exploiting human inattentional blindness: Looking but not seeing. In: Proceedings of the ACM Symposium on Virtual Reality Software and Technology, 17–24, 2002.

[173] Cater, K.; Chalmers, A.; Ward, G. Detail to attention: Exploiting visual tasks for selective rendering. In: Proceedings of the 14th Eurographics Workshop on Rendering, 270–280, 2003.

[174] Sundstedt, V.; Chalmers, A.; Cater, K. Selective rendering of task related scenes. In: Proceedings of the Symposium on Applied Perception in Graphics and Visualization, 174, 2004.

[175] Sundstedt, V.; Chalmers, A.; Cater, K.; Debattista, K. Top-down visual attention for efficient rendering of task related scenes. In: Proceedings of the Vision, Modeling and Visualization, 209–216, 2004.

[176] Duchowski, A. T.; McCormick, B. H. Simple multiresolution approach for representing multiple regions of interest (ROIs). In: Proceedings of the SPIE 2501, Visual Communications and Image Processing, 175–186, 1995.

[177] Geisler, W. S.; Perry, J. S. Variable-resolution displays for visual communication and simulation.

SID Symposium Digest of Technical Papers Vol. 30, No. 1, 420–423, 1999.

[178] Parkhurst, D.; Culurciello, E.; Niebur, E. Evaluating variable resolution displays with visual search: Task performance and eye movements. In: Proceedings of the Symposium on Eye Tracking Research & Applications, 105–109, 2000.

[179] Geisler, W. S.; Perry, J. S. Real-time simulation of arbitrary visual fields. In: Proceedings of the Symposium on Eye Tracking Research & Applications, 83–87, 2002.

[180] Willberger, T.; Musterle, C.; Bergmann, S. Deferred hybrid path tracing. In: *Ray Tracing Gems.* Berkeley: Apress, 475–492, 2019.

[181] Jin, B.; Ihm, I.; Chang, B.; Park, C.; Lee, W.; Jung, S. Selective and adaptive supersampling for real-time ray tracing. In: Proceedings of the Conference on High Performance Graphics, 117–125, 2009.

[182] Koskela, M.; Immonen, K.; Mäkitalo, M.; Foi, A.; Viitanen, T.; Jääskeläinen, P.; Kultala, H.; Takala, J. Blockwise multi-order feature regression for real-time path-tracing reconstruction. *ACM Transactions on Graphics* Vol. 38, No. 5, Article No. 138, 2019.

[183] Sherrington, C. S. On Reciprocal Action in the Retina as studied by means of some Rotating Discs. *The Journal of Physiology* Vol. 21, No. 1, 33–54, 1897.

[184] Arabadzhiyska, E.; Tursun, O. T.; Myszkowski, K.; Seidel, H. P.; Didyk, P. Saccade landing position prediction for gaze-contingent rendering. *ACM Transactions on Graphics* Vol. 36, No. 4, Article No. 50, 2017.

[185] Mlot, E. G.; Bahmani, H.; Wahl, S.; Kasneci, E. 3D gaze estimation using eye vergence. In: Proceedings of the International Joint Conference on Biomedical Engineering Systems and Technologies, 125–131, 2016.

[186] Veach, E. Robust Monte Carlo methods for light transport simulation. Ph.D. Thesis. Stanford University, 1997.

[187] Georgiev, I.; Křivánek, J.; Davidovič, T.; Slusallek, P. Light transport simulation with vertex connection and merging. *ACM Transactions on Graphics* Vol. 31, No. 6, Article No. 192, 2012.

[188] Lee, S.; Pattichis, M. S.; Bovik, A. C. Foveated video quality assessment. *IEEE Transactions on Multimedia* Vol. 4, No. 1, 129–132, 2002.

[189] Wang, Z.; Conrad Bovik, A.; Lu, L.; Kouloheris, J. L. Foveated wavelet image quality index. In: Proceedings of the SPIE 4472, Applications of Digital Image Processing XXIV, 2001.

[190] You, J. Y.; Ebrahimi, T.; Perkis, A. Attention driven foveated video quality assessment. *IEEE Transactions on Image Processing* Vol. 23, No. 1, 200–213, 2014.

[191] Rimac-Drlje, S.; Vranješ, M.; Žagar, D. Foveated mean squared error—A novel video quality metric. *Multimedia Tools and Applications* Vol. 49, No. 3, 425–445, 2010.

[192] Tsai, W. J.; Liu, Y. S. Foveation-based image quality assessment. In: Proceedings of the IEEE Visual Communications and Image Processing Conference, 25–28, 2014.

Lili Wang is a professor at the School of Computer Science and Engineering, Beihang University, and a researcher of the State Key Laboratory of Virtual Reality Technology and Systems. Her research interests include virtual reality, augmented reality, and rendering.

Xuehuai Shi is a Ph.D. student at the School of Computer Science and Engineering, Beihang University, China. His current research focuses on virtual reality and foveated rendering.

Yi Liu is a master student at the School of Computer Science and Engineering, Beihang University, China. His current research focuses on virutal reailty, augumented reality, and rendering.