Eurographics Symposium on Rendering (2026)
I. Gkioulekas and A. Jarabo (Editors)

Rethinking Color Precision: Real-Time Perceptual Spatial Quantization for Foveated Rendering

Z. Zhu G.D. Love and R.K. Dos Anjos

University of Leeds, United Kingdom

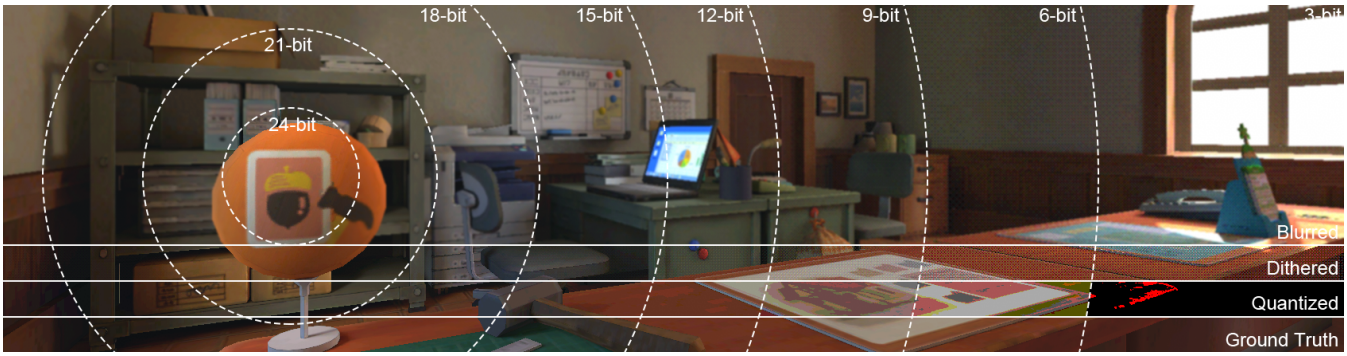

Figure 1: *Spatial quantization for bit depth reduction guided by retinal eccentricity. From fovea (left) to periphery (right), color precision is progressively reduced from 24 to 3 bits per pixel. From top to bottom: quantization with dithering and blur (ours), quantization with dithering, naïve quantization and ground truth. Our approach produces a close match to the ground truth.*

Abstract

Head-mounted displays demand extreme data bandwidth, yet traditional pipelines allocate uniform color precision across the entire frame, ignoring the rapid decline of chromatic sensitivity as retinal eccentricity increases. This paper proposes a real-time perceptual spatial quantization method that modulates bit depth based on retinal eccentricity rather than maintaining per-pixel precision. Through a psychophysical study, we model the detection threshold of our method across quantization levels ranging from 24 to 3 bits per pixel, the minimal level. Our method is the first to achieve this level of color reduction without detection, as it takes advantage of lowered peripheral color perception in tandem with the visual pooling effect. By integrating our model into the theoretical bandwidth measure, we demonstrate that our method can achieve a data reduction of up to 83.5% compared to uniform quantization. Our findings suggest potential for reducing bandwidth in wireless streaming to minimize latency and optimizing DRAM traffic to save power in mobile VR/AR architectures. In summary, our work provides a strategy for high-efficiency, perceptually lossless color reduction in foveated streaming systems.

CCS Concepts

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

1. Introduction

Virtual Reality (VR) platforms using head-mounted displays (HMDs) are continuing to increase in popularity and widespread adoption, moving from being solely a gaming platform to uses in education, engineering, manufacturing, and healthcare [LaV23]. One of the major contributors to this increased adoption is the shift toward standalone, cable-free form factors. The resulting portability allows for flexible usage scenarios where tethered solutions would be physically prohibitive.

However, this mobility comes at a steep computational cost. The demand for high-fidelity immersion and real-time interaction requires ultra-high rendering resolution and low latency, which often exceed the local processing capabilities of System-on-Chip (SoC) architectures on standalone devices. For instance, the required density for near-eye display to reach normal vision is approximately

2291.6 pixel-per-inch (ppi) [LaV23] and VR interfaces require total latency to remain under 16-60 ms to prevent impairing performance in demanding tasks [ARFK17]. Thus, wireless approaches with remote rendering are becoming increasingly vital [SLNC13].

To mitigate processing pressures, a popular approach to accelerate graphics pipelines in virtual reality has been foveated rendering [GFD*12; PSK*16; MDZV18], which relies on eye-tracking information from HMDs, allowing for reduced rendering quality or resolution in areas where the user is not currently fixated. This has been mainly driven by growing understanding of the Human Visual System (HVS) [SRJ11], and taking advantage of both loss of acuity [PSK*16] and pooling behavior of the visual cortex [WAF*21]. Physiologically, this degradation is mainly rooted in the non-uniform distribution of photoreceptors across the retina: while the fovea is densely packed with cones responsible for high acuity, cone density drops significantly with increasing eccentricity, giving way to a rod-dominated periphery. However, while current foveation techniques effectively minimize spatial resolution and geometric complexity, they largely focus on the density of samples.

Despite these advances, chromatic precision remains a largely under-exploited dimension. Although both psychophysical and physiological evidence show that color perception drops sharply outside the fovea [ACM25; Mul91], modern pipelines continue to allocate uniform 24-bit color depth across the entire frame. This "per-pixel" color accuracy represents a massive data redundancy in peripheral regions where the HVS cannot distinguish fine-grained chromatic gradients. A practical, real-time method for reducing color accuracy that remains compatible with existing streaming protocols has yet to be successfully proposed.

In this paper, we present a real-time method to drastically reduce the bit depth of rendered frames without perceived quality loss with a foveated streaming pipeline in mind. Our core insight is a shift from per-pixel precision to spatial color perception by introducing dithering. We take advantage of the pooling behavior of our visual cortex to recreate colors and mask visual artifacts introduced by the color precision reduction. By distributing color information spatially, we rely on the HVS's low-pass filtering nature in the retinal periphery to reconstruct smooth gradients from quantized data. This allows us to exploit our lowered color perception accuracy in peripheral vision further than previous work [LH08; ZAKL19]. We evaluate our method through a user study to derive the imperceptible thresholds of spatial quantization at all bit depths over eccentricities. We propose a step-mask model for color reduction and a safer zone balancing visual quality and compression rate. Our results suggest that the proposed method can achieve up to 83.5% in bandwidth reduction, demonstrating potential for reducing data transmission pressure in mobile VR/AR architectures.

2. Related Work

2.1. Foveated Rendering and Streaming

The HVS exhibits strong non-uniformities. Visual abilities decline significantly outside the fovea with increasing retinal eccentricity. This is related to both the sharp decrease in the density of photoreceptors and retinal ganglion cells outside the foveal region and the pooling behavior of our visual cortex. A large body of psychophysical, anatomical, and computational research has investigated the phenomenon of how visual acuity falls off [SRJ11].

Foveated rendering typically exploits the eccentricity-dependent vision acuity drop-off in real-time applications. Guenther et al. [GFD*12] developed the first well-known foveated rendering pipeline for 3D graphics that renders three layers of different resolutions and temporal rates based on a linear model for visual acuity over eccentricity. This line of research has been further explored, looking at factors such as contrast preservation [PSK*16; TAW*19], statistical behavior of oriented edges [WAF*21; TTD22], and minimizing artifacts from deferred pipelines [MDZV18], and others [MDV20; TAW*19].

For 360° video and VR application streaming, tiling techniques are commonly used to stream only the visible portion, the viewport, at high quality and stream out-of-sight contents at a significantly lower quality or leave them out. More aggressively, streaming based on the fovea can further improve efficiency. The quality modulation can be roughly classified as multi-resolution techniques or quantization parameter (QP) adjustments in video encoding. Lungaro et al. [LSV*18] streamed high-quality (QP=22-27) HEVC-compliant tiles for the foveal region and lower-resolution background pre-encoded offline with two online video decoders for different qualities on the client side, saving 60% - 80% bandwidth compared to equirectangular 4K encoding. Illahi, Siekkinen, and Masala [ISM17] introduced real-time foveated video encoding (FVE) streaming for cloud gaming without modifying the gaming engine. Later, Illahi et al. [IGS*20] implemented gaze-adaptive encoding using the x264 encoder and evaluated QPs from 8 to 12, balancing saving and quality. Li et al. [LDB*21] introduced a log-rectilinear buffer for video encoding to enable foveated streaming. More recently, EyeNexus [WAT*25] combines real-time foveated spatial pixel amount compression and FVE, assigning video quantization non-uniformly. Industrial systems have also begun adopting foveated streaming techniques. For example, Steam [Val26] has introduced their foveated streaming, a streaming-first approach that dynamically allocates higher-quality pixels around the user's gaze point, reporting substantial improvements in effective bandwidth utilization and perceived image quality.

2.2. Color Perception and Reduction in Periphery

Several studies suggest that color perception also significantly declines from the fovea to the periphery, similar to acuity sensitivity [Mul91; CBR20]. Firstly, spatialchromatic contrast sensitivity, the ability to sense contrast details within colors, is weaker in the periphery. Mullen [Mul85] indicated that the perceptual cutoff for the achromatic channel is 34 cpd, while for chromatic channels it is 11 cpd. Mullen and Kingdom [Mul*02] measured the contrast sensitivities at eccentricities up to 25°, finding that red–green cone opponency has a steep decline away from the fovea, while the blue–yellow cone opponency is more gradual. Secondly, color discrimination, the ability to differentiate colors, also degrades significantly in the periphery [HPG09; DCT*22]. At a fixed luminance, the discrimination ellipses, the set of perceptually equivalent colors, are much larger at larger eccentricities than the parafoveal presentation, which means that there is a large fall-off of color dis-

criminative sensitivity. Additionally, the highest resolution limit for foveal vision is 94 pixels per degree (ppd) for achromaticity, 89 ppd for red-green and 53 ppd for yellow-violet [ACM25]. Both chromatic patterns have much larger drops than achromatic patterns: the achromatic resolution limit at 20° is 21 ppd, while it is 7 ppd for red-green and 5 ppd for yellow-violet.

There have been studies applying gaze-contingent color perception, mostly on perceptual enhancement [MM13; MFN16] and power reduction, shifting peripheral color for power saving in OLED display or during DRAM traffic [DCT*22; USC*24]. However, relatively limited work has applied color perception to data bandwidth reduction and real-time gaze-contingent streaming, as discussed in the following paragraphs.

Some studies explored gaze-contingent color reduction, reducing either the number of channels or bits per channel. One approach to reducing color is through saturation reduction, using only one channel in the periphery. Bektas et al. [BCKD15] implemented a gaze-contingent color degradation mask that progressively desaturates peripheral regions to reduce data transmission requirements, using a pixel-shader implementation from Duchowski and Coltekin [DÇ07]. Duchowski et al. [DBS*09] further explored color degradation maps by assigning grayscale values to pixels. However, they found that no desaturation should be applied within 20°. Hansen et al. [HPG09] also concluded that chromatic detection gets worse with increasing eccentricity but is still possible even up to 50°.

The two most relevant studies are gaze-contingent hue resolution reduction, or bit depth reduction. Liu and Hua [LH08] proposed a hue resolution foveation mask based on Barten's contrast sensitivity function model [Bar99], reducing the periphery color depth to 5 bits per channel (bpc), 15 bpp in total at eccentricity of 30°. However, this approach is purely calculated based on contrast sensitivity measured at the luminance level of the corresponding RGB color, implicitly assuming chromatic contrast sensitivity is the same as achromatic contrast sensitivity, while recent studies have shown that chromatic contrast sensitivity declines more steeply with retinal eccentricity [HPG09; ACM25]. Zhang et al. [ZAKL19] developed a peripheral color tolerance model based on the color discrimination model from Hansen et al. [HPG09] as a visual difference predictor for bit reduction. They used a look-up table for real-time conversion in LAB color space and evaluated the model under free-viewing conditions with a bit depth reduction of various bit levels at 10° eccentricity. But the perceptual tolerance was strongly content-dependent, showing up to a high probability of artifact detection.

In summary, while the HVS remains capable of detecting color information up to high eccentricity, the decline of chromatic sensitivity is steeper than that of achromatic sensitivity. Current foveated streaming solutions focus on spatial resolution or video compression parameters, yet most foveated graphics approaches largely ignore the potential of gaze-contingent color modulation. There remains a critical research gap for an effective approach capable of achieving radical color data savings.

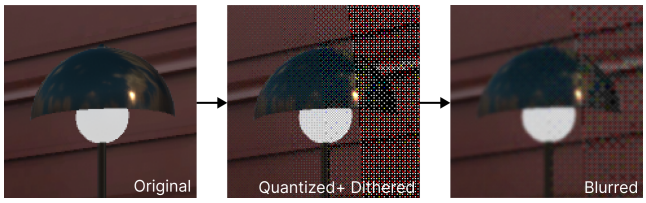

Figure 2: *Two-stage workflow (7x magnified) of the spatial quantization pipeline. The original color image (left) is quantized and dithered (middle), which introduces artifacts, including noise and lightness shift. The image is then recovered with a small spatial blur (right), reproducing the original image.*

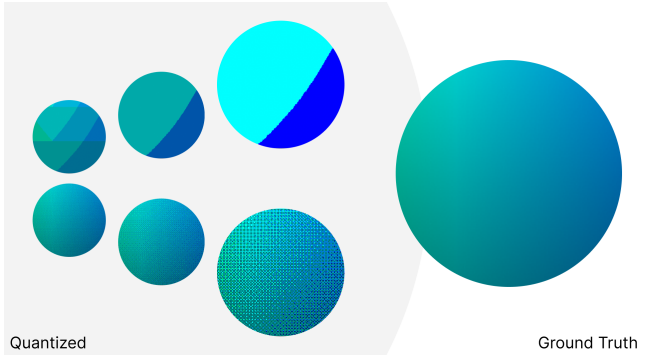

Figure 3: *Two different bit depth reduction approaches: the top three balls from left to right show the results of direct bit depth reduction at 3 to 1 bpc, which creates oriented edges clearly visible even in the far retinal periphery; the bottom three show the results of the same reduction with dithering and the results are smooth variations. Dithering breaks the spatial structure and allows the visual pooling effect to perceive the original shade. The large ball on the right is the referencing ground truth at 8 bpc.*

3. Method

3.1. Overview of the Proposed Pipeline

As discussed in Section 2.2, human color sensitivity drops significantly as eccentricity increases, which allows us to treat chromatic data as a prime candidate for aggressive bit reduction. Our pipeline (Figure 2) applies gaze-contingent data to perform a two-stage process: spatial quantization, which redistributes reduced color information into the spatial domain, and statistics recovery, which partially reconstructs the original contrast, lightness and color.

3.2. Acuity and Pooling

The core of our approach is rooted in the limitations of peripheral vision, with visual acuity being one of the primary considerations. While there are many visual acuity models used in foveated rendering [GFD*12; PSK*16] like the linear fall-off model [Wey58], the loss of visual acuity is often described as the feeling of "blurry".

Beyond simple blur, the periphery also fails to perceive individual features due to acuity loss, which is attributed to an effect called

Figure 4: *The same gray quantized and dithered from 7 bpc (left) to 1 bpc (right). At lower bit depths, a perceptual lightness shift emerges, appearing progressively "whiter". This effect is primarily caused by the sparse palette from quantization.*

pooling or visual crowding [SRJ11]. Pooling means that the spatial location of features is negligible, but the aggregate statistics (e.g., mean and variance [WAF*21]) matter perceptually. The size of the pooling regions increases with eccentricity at a rate significantly faster than the drop in acuity, which means that the patterns turn into statistics before acuity fails to resolve them.

However, previous work has demonstrated that strongly oriented features are still resolved in the periphery when creating a ventral metamer [FS11; WAF*21]. While we may not discern similarly distributed patterns [ZS21], we will still perceive oriented lines or contrast shifts accurately. In Freeman and Simoncelli's work [FS11], this is captured by both marginal statistics of the steerable pyramid, and self and cross correlation between different bands and scales, and serves as a good approximation of how the HVS is perceiving images. Figure 3 demonstrates two different patterns using the same color precision. The patterns on the top row have a high mean in a specific band of the steerable pyramid, and high correlation between different scales on that band. The bottom pattern will have a low mean overall in different bands, and not easily discerned from a different one [ZS21].

3.3. Hue Resolution Reduction

Previous studies reduced per-pixel hue resolution directly, mapping the original pixel color to the nearest color in the new representation (the nearest element in the sparse color space). As shown in the top row of Figure 3, artifacts including hue shift, banding, detail loss, and false contours, are visible even in the far periphery.

Rather than treating color reduction as a purely per-pixel problem, we shift the focus to spatial color perception. We take inspiration from spatial quantization (SQ), the idea proposed by Puzicha et al. [PHK*00] to move beyond simple bit-truncation. It is the contextual image processing concept that considers chromatic and spatial information by performing CQ and digital halftoning simultaneously. The quantization errors are hence spatially distributed across neighboring pixels. Through visual pooling, the spatial average of these local color variations can produce the perception of intermediate colors that are not explicitly represented in the reduced color space. This mechanism becomes increasingly effective in peripheral vision. As eccentricity increases, visual acuity decreases and color sensitivity decreases, while the pooling effect increases, integrating information over larger spatial regions. In the fovea, this dithering pattern might still be perceived as noise; however, in the periphery, larger pooling "reconstructs" the intended colors, making imperceptible bit reduction possible.

3.4. Contrast Recovery

The previous strategy introduces several new artifacts, including spatial noise and local hue shift; however, the most notable is a significant lightness shift. Lightness is the subjective perception of the reflectance of an object, which is distinct from physical luminance or absolute brightness. Lightness shift occurs mainly because palette reduction forces smooth gradients into discrete steps, changing the color mean and significantly increasing local contrast. According to Wallach's Ratio Principle [Wal48], our perception of lightness is determined by the luminance ratios between adjacent areas rather than absolute values. Therefore, when quantization alters these ratios, the perceived lightness of the region increases as the bit depth decreases. Also, high contrast sparse pixels make visual pooling more difficult. This can be seen in Figure 4 at lower bit depths and in the histogram of Figure 5.

To mitigate these artifacts in the peripheral regions, we take inspiration from the "blurring of periphery" which is commonplace in foveated graphics [MDZV18]. Lower resolution in the periphery is typically not detected due to the loss of visual acuity.

In our framework, this low-pass filter serves as a reconstruction operator that aligns the quantized digital signal with the integration mechanisms of the peripheral HVS. By performing a weighted spatial average of the local neighborhood, it smooths discrete transitions caused by quantization. This reduces the spatial contrast ratios to a state closer to the original data, which reduces the lightness shift induced by palette sparsity. Simultaneously, the filter also suppresses high-frequency noise patterns that fall within the peripheral resolution limit. Statistically, this blur filter recovers the mean and variance collapsed during high-level quantization (Figure 5). In summary, the filter functions as an efficient spatial dequantization mechanism at last, as it "mixes" neighbor colors to create intermediate color values that do not exist in the reduced bit-stream.

3.5. Near-eye Displays

Finally, we consider the reality of these frames being viewed through a near-eye display. HMDs feature significantly higher physical pixel densities than traditional desktop monitors. The individual dithered pixels are spatially small enough to fall below the threshold of detection. Two specific features do support our proposed approach. First, using a screen-space high-frequency dithering pattern may introduce a "screen-door effect" on the displayed graphics. However, this is a recognized effect and common concern on HMDs [ZYX*20], which we argue cannot be distinguished clearly from the dithering approach. Second, optical systems in

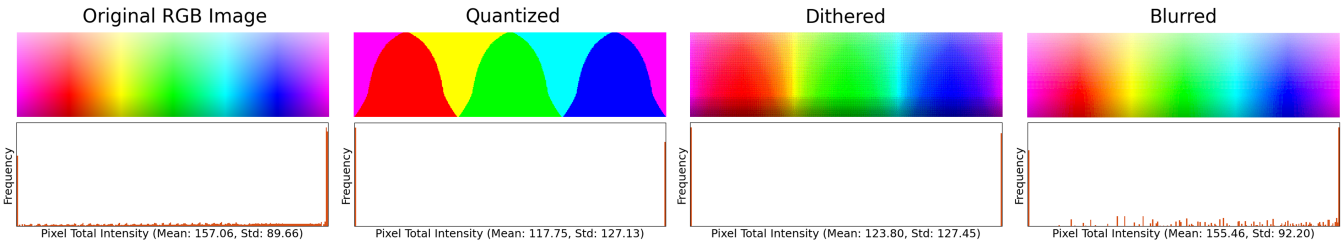

Figure 5: *Histograms of a full color gradient image, its naïve quantization (1 bpc), quantization (1 bpc) with dithering, and quantization with dithering and blur. The histograms illustrate the color intensity distributions for three channels, where the quantization changes the mean and increases the variance. Blur produces a closer approximation to the ground truth visually and statistically.*

HMDs exhibit both monochromatic aberrations (e.g., defocus) and chromatic aberrations (e.g., longitudinal and transverse) towards the lens periphery [BKB19]. This distortion reduces effective display acuity in the peripheral visual field, acting as a hardware-level low-pass filter. This allows our blurring operation not to introduce a noticeable loss of acuity, which is done at a small scale where details could already not be clearly distinguished under normal HMD viewing conditions.

4. Implementation

We implement the proposed foveated spatial quantization as a real-time post-processing effect within the Unity Built-In Rendering pipeline using HLSL fragment shaders. Real-time gaze coordinates are captured via the Tobii XR SDK and passed to the shader as a normalized screenspace vector. The implementation is designed mainly to aim for efficiency and real-time usage with minimal computational budgets.

Previous studies in foveated color reduction applied hue resolution reduction in RGB or LAB space, which fall under image-independent uniform scalar quantization (IUSQ) schemes. We also employ quantization in the RGB color space motivated by both computational efficiency and compatibility with the native rendering pipeline. Implementing IUSQ is straightforward and efficient in an orthotope-shaped color space such as RGB, wherein quantization cells are rectangular as stated in previous studies [Cel23]. Additionally, a vital consideration is the performance under extreme conditions as low as 1 bpc. As noted by Hill et al. [HRV97], while spaces like CIELAB provide perceptual uniformity, they are abstract representations that lack the boundaries of the related rendering and display gamut. Our test across several color spaces (Figure 6) shows significant hue shifts under extremely low bit depth, which is also not recoverable with our contrast recovery method. While image-dependent CQ might achieve higher compression, it evaluates a new palette every frame, which is computationally expensive and introduces temporal artifacts, which are very easy to see on mid-fovea and are still visible in the periphery [SWK25]. Given these considerations, RGB-based uniform quantization provides a favorable trade-off between efficiency, stability, and implementation simplicity for real-time foveated rendering applications.

A wide range of digital halftoning algorithms has been proposed, including dithering, error diffusion, and direct binary search (DBS).

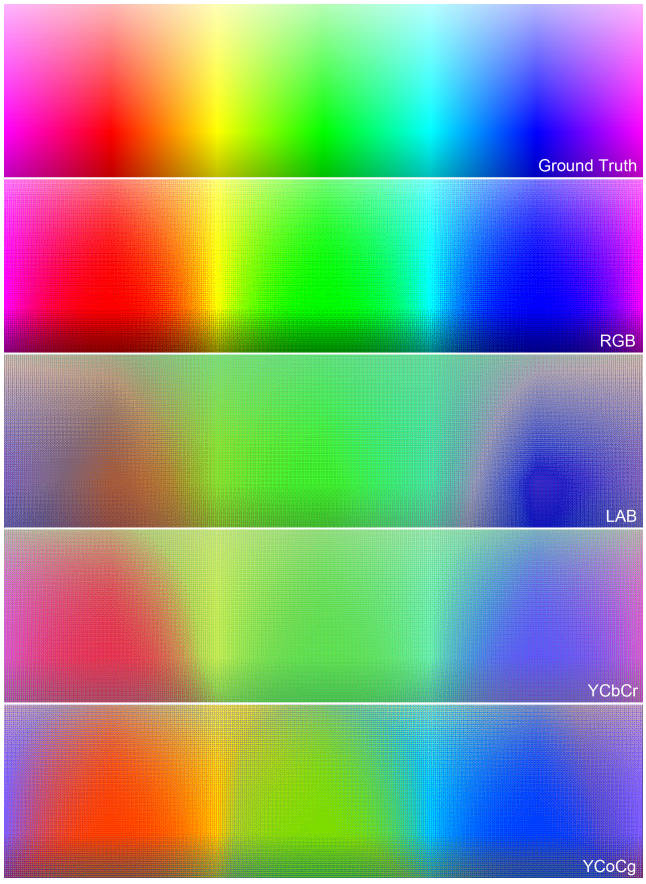

Figure 6: *Quantized and dithered to 1 bpc in RGB, LAB, YCbCr and YCoCg color spaces.*

Though error diffusion, such as Floyd-Steinberg, is widely known for its quality [Flo76], diffusion relies on sequential pixel dependencies calculated from neighboring pixels, so it is poorly suited for parallel execution and thus difficult to integrate into real-time rendering pipelines. Similarly, DBS [AA92] is very powerful in producing high-quality results but time-consuming because of its iterative procedure. Therefore, we select dithering methods for their ef-

ficiency. Dithering can be processed in parallel per pixel and hence applied in real time, including clustered-dot and dispersed-dot.

We implement quantization using ordered dithering on each RGB channel independently using an 8×8 Bayer matrix [Bay76]. It is a practical compromise between computational efficiency, temporal stability, and implementation simplicity. For each pixel, the original value in each channel is compared against a spatially varying threshold determined by the Bayer matrix. The threshold comparison assigns the pixel to one of the available quantization levels, producing a low-bit-depth representation while spatially distributing quantization error across neighboring pixels. Though Bayer ordered dithering may not outperform other noise patterns, the grid-aligned structure interacts favorably with the display and the subsequent dequantization step. We selected the 8×8 matrix as a practical compromise between quality and computational efficiency: the 4×4 matrix provided insufficient spatial distribution of quantization error and resulted in more noticeable artifacts, whereas the 8×8, 16×16 and larger matrices performed similarly well.

We employ a minimal blur for the purpose of recovering the statistics and contrast artifacts with minimal computational cost and minimal possibility of being noticed (Figure 5). We implement a modified Kawase blur filter [Kaw03] for its efficiency in real-time rendering. In our implementation, we utilize 2 iterations of ping-ponging with a constant blur offset of 0.2 pixels, which provides a very small but sufficiently low-pass cutoff to neutralize the quantization.

To maintain high visual acuity where the user is looking, we define a circular region to keep the fovea ($\sim 5^\circ$) at full fidelity. To prevent visible discontinuities at the boundary of the fovea, we implement a smooth transition in the parafoveal region ($\sim 5^\circ$–9°). This transition is implemented via linear interpolation based on the acuity linear fall-off model, with blending weights defined by a `smoothstep` function.

5. User Study

5.1. Experiment

We designed a psychophysical user study to estimate perceptual thresholds for foveated spatial quantization under varying eccentricities and quantization levels.

In practice, the main goal is to minimize the total amount of bits to represent pixels while preserving lossless perceptual quality. To quantify the theoretical benefit of our findings, we define a perceptual bit allocation model. Assuming a radially symmetric display, the inner regions of the area are a circular disk, and the outermost region is a rectangular region. The contribution of bit weight, w_i, is proportional to its area in zone i:

$$w_i \propto \begin{cases} \pi r_i^2, & i = 0, \\ \pi(r_{i+1}^2 - r_i^2), & i = 1, \ldots, 6, \\ A_{screen} - \pi r_i^2, & i = 7. \end{cases} \tag{1}$$

where r_i and r_{i+1} denote the inner and outer normalized radius of the circular zone i, and A_{screen} denotes the total image area. Assuming the bit depth per channel decreases linearly from 8 bpc (fovea) to 1 bpc (periphery), such that the bit cost b per pixel in zone i is

$$b_i = 3 \times (8 - i) \tag{2}$$

Therefore, the total weighted bit cost B (bpp) is:

$$B = \sum_{i=0}^{7} b_i w_i \tag{3}$$

$$= 3A_{screen} + 3\pi(r_1^2 + \ldots + r_7^2) \tag{4}$$

The conversion from normalized radius r to eccentricity θ (in degrees) is given by:

$$\theta = \frac{180}{\pi} \arctan\left(2r \cdot \tan\left(\frac{\Phi \cdot \pi}{360}\right)\right) \tag{5}$$

where Φ denotes the field-of-view (FOV) corresponding to the axis along which r is normalized. This means for a given bit depth, we seek to determine the smallest eccentricity beyond which artifacts fall below perceptual detection thresholds. We fix the spatial blur parameters at the minimum level as stated in the previous section, as blur is introduced solely to attenuate high-frequency artifacts and should not introduce too much additional computational overhead.

5.1.1. Stimulus

Visual sensitivity is very content-dependent and attention-sensitive [ZAKL19; CBR20]. To avoid biasing measurements toward a specific stimulus configuration, we design a controlled and diverse scene that spans a range of perceptually relevant signal characteristics for the user study.

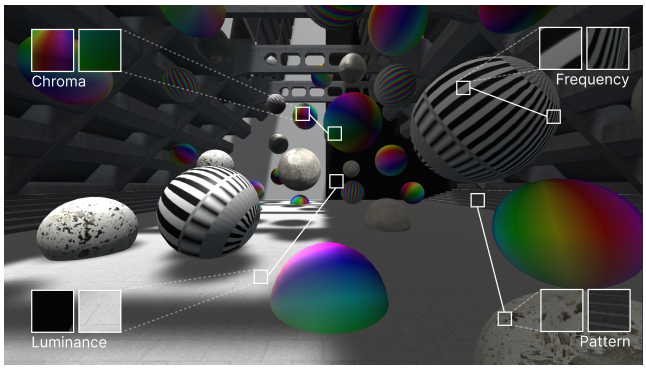

Figure 7: *Controlled 3D stimulus scene used in the experiment. The environment contains spherical primitives textured with various visual dimensions, including luminance, chromatic gradients, high-frequency patterns, and pattern statistics.*

As shown in Figure 7, the scene is designed to probe the perceptual visibility of artifacts under varying visual conditions. It comprises two global luminance regimes and a set of parametrically defined spherical primitives. Spheres are textured with key stimulus dimensions, including various chroma, spatial frequencies, and pattern statistics (i.e., structured or random). Spheres with Granger Rainbow texture provide smooth color gradients at low spatial frequency, making them sensitive to hue distortions and banding artifacts. Black-white-striped and color-striped spheres are structured patterns with various spatial frequencies at high contrast, which

are sensitive to blur and aliasing. We also include natural textures with randomized noise, such as stone and cement, enabling evaluation under more realistic conditions. These elements are distributed across different depths and under two luminance regions.

By jointly manipulating chroma, luminance, contrast, spatial frequency, and pattern properties within a unified scene, this design provides a comprehensive testbed for capturing and assessing when and where artifacts become perceptually detectable.

5.1.2. Procedure

We employ a psychophysical paradigm with a built-in verification step to determine the perceptual thresholds at which artifacts are not noticeable. The study is structured as a nested loop iterating through two primary independent variables:

- **Bit depth** L, corresponding to progressively reduced chromatic precision. Specifically, L controls the per-channel bit depth, ranging from 7 bpc to 1 bpc.
- **Foveal radius** R, the radius defined in normalized screen space. We adopt R directly instead of eccentricity as the primary control variable to obtain efficiency and consistency. Assuming a unit square render target, $R = 0.5$ corresponds to a circular region that reaches the screen boundaries. We discretize R from the parafoveal boundary ($\sim 9^\circ$) ($R \approx 0.064$) to the maximum eccentricity extent ($R = 0.5$) with a step size of 0.05.

To ensure accurate real-time eye tracking, each session begins with a standard eye-tracking calibration provided by the headset. This ensures the foveal mask remains accurately centered on the participant's gaze.

Prior to the formal session, participants are instructed to complete a practice trial where the non-foveated ground truth (GT) stimulus is displayed to establish a baseline "artifact-free" rendering and to get familiarized with the interface. In the formal session, each trial consists of a pair of stimuli, including the reference GT stimulus and the foveated stimulus at the current L, R combination shown in a randomized order. The first stimulus is displayed for 10 seconds and then it gradually fades to black. The participant is then prompted with a panel asking if there is any artifact and is asked to provide a response using the left and right arrow keys as input for "yes" and "no". After the selection, the scene gradually fades back to normal, presenting the other stimulus in the pair. Then the participant is prompted again. This is designed as a validation to prevent guessing or being influenced by other potential rendering artifacts. If the participant reports an artifact in the GT, the current trial is flagged as an error and reset. A notification is displayed to the participant to inform them that the previous render is GT. A successful detection of an artifact is recorded only if the participant identifies an artifact in the foveated stimulus while reporting no artifact in the GT. We would increase R to expand the high-quality foveal region, testing if the artifact remains visible at a higher eccentricity with the same bit depth at the current L. If the participant fails to detect an artifact in both GT and the foveated stimulus, we decrease the bit depth L at the current R. The session continues until either R reaches the maximum extent or the bit depth at L reaches the most aggressive level. To ensure data integrity, we excluded sessions where a participant exceeded 10 errors in identifying GT, requiring a re-completion of the session.

Following the session, a semi-structured interview is conducted, asking what kind of artifact the participant sees, where they identify the artifact, and what specific scene regions, stimuli, or viewing conditions trigger artifact visibility.

5.1.3. Apparatus

The study was conducted on a Pimax Crystal headset. The FOV of Pimax Crystal is approximately 103° horizontally and vertically [Mus26]. We set the display mode to the highest image quality, without GPU upscaling or sharpness adjustments, and displayed at 120 Hz. The application was developed in Unity Version 2022.3.19 using the Built-In Rendering pipeline.

5.1.4. Participants

We recruited 29 participants (aged 18-34, with half aged 22-25; 12 female, 17 male). All candidates were required to have normal or corrected-to-normal visual acuity and normal color vision. Adherence to these criteria was verified via pre-screening: color vision was assessed using Ishihara pseudo-isochromatic plates, and visual acuity was self-reported or verified through corrected lenses. Following the screening, 7 participants were excluded from the primary analysis: 3 exhibited color vision deficiencies, and 4 failed to meet the corrected-to-normal vision requirement at the time of testing. The final dataset comprises 22 valid participants.

All experimental procedures were conducted in accordance with institutional ethical guidelines. Informed consent was obtained from all participants prior to the session, and all data were strictly anonymized.

5.2. Results

To ensure the reliability of our psychophysical thresholds, we apply a strict exclusion threshold based on GT reporting. We exclude data if a participant reported artifacts in the GT stimuli more than 10 times from the analysis. Following this criterion and visual requirements, 21 valid datasets from valid participants were retained for statistical evaluation, and an additional 6 valid datasets were from participants who reported visual deficiency.

Prior to the core analysis, we convert the recorded foveal radius R from normalized screen space back to angular eccentricity θ (in degrees) to ensure a perceptually relevant metric using Equation 5.

By taking the mean for each bit level in the population, we obtain the average eccentricity at which a person cannot see any artifacts. As illustrated in Figure 8, the mean thresholds consistently increase from 7 bpc to 1 bpc. The thresholds for 7 to 5 bpc conditions remain relatively stable (9.44°-11.24°), staying close to the parafoveal baseline. While lower bit depths (below 4 bpc) exhibit a sharp increase. 1 bpc (mean: 27.15; std: 15.37) shows high inter-participant variance and requires significantly larger eccentricity to hide artifacts compared to higher bit depth in general.

We summarize the threshold at each bit depth with the mean value as a step-mask in Table 1. While individual responses exhibit variability, the mean provides a convenient approximation for constructing a staircase of bit depth reduction. We also propose a transition zone for safer reduction that balances perceptual quality

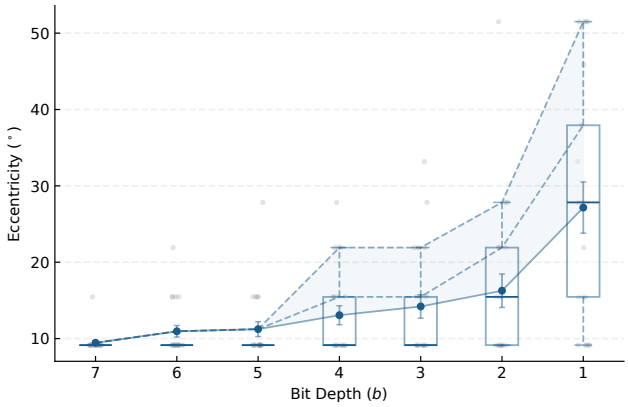

Figure 8: *Perceptual eccentricity thresholds across bit depths (7 to 1 bpc). The bottom line illustrates the mean threshold, which is used in our step-mask model. The top dashed line illustrates the max threshold (4 to 1 bpc), and the middle dashed line illustrates the Q3 threshold (4 to 1 bpc). The blue zone is the transition zone that balances perceptual quality and compression rate.*

Figure 9: *Foveated mask with mean thresholds from the study. Settings match the bottom line in Figure 8.*

and compression rate, bounded by the mean and conservative upper cap (Figure 8). For lower bit depth (< 5 bpc), we characterize the detection distribution by examining the third quartile (Q_3) and maximum threshold values across the bit depth conditions as shown in Table 1. These thresholds are synthesized into the foveated spatial quantization mask as illustrated in Figure 9.

Based on the semi-structured interview (N=21), participants reported that artifacts were most noticeable in regions with high spatial frequency (N=2), edges (N=7), and brighter regions (N=3), where they appeared as shimmering patterns or as blurring. Only one participant reported detecting artifacts related to color. This suggests that chromatic artifacts are effectively attenuated by the small spatial blurring, while some high-frequency patterns are more visible. Artifacts were reported to be more salient during rapid gaze shifts or saccades (N=3), which is likely attributable to temporal misalignment caused by eye-tracking latency.

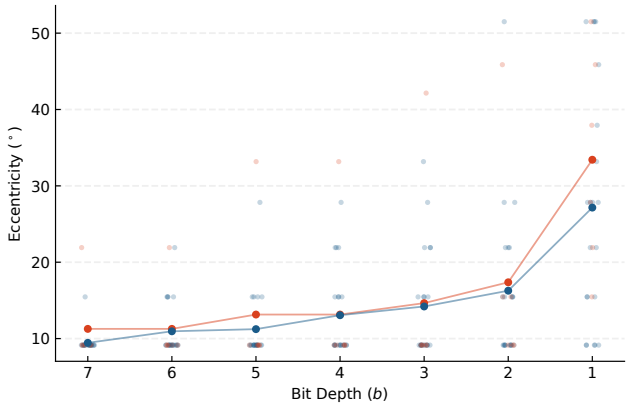

Figure 10: *Perceptual eccentricity thresholds across bit depths (7 to 1 bpc) with two groups. Blue indicates the results with participants without visual deficiencies (N=21); red indicates the results with participants with self-reported visual deficiencies (N=6).*

Surprisingly, participants with self-reported visual deficiencies (N=6) did not show statistically significant differences in either quantitative thresholds or qualitative reports compared to the main cohort, as shown in Figure 10.

6. Evaluation

6.1. Bandwidth

To evaluate the impact of our findings regarding bandwidth saving, we compute the area-weighted bit depth across eccentricity. Given the Pimax Crystal display configuration and Equation 4, when viewing at the center, the resulting area-weighted bandwidth based on the mean values is approximately 3.97 bpp, compared to the original 24 bpp for a standard true RGB color representation. This corresponds to a bandwidth reduction of 83.5%. The upper bound weighted bandwidth of the transition zone area is approximately 7.15 bpp, which is 70.2% cut.

Our current implementation in Unity costs 30 μs for Pimax Crystal, with 10 μs for quantization with dithering and 20 μs for blur. As both these postprocessing implementations scale linearly with image size, it maintains high-performance real-time throughput at the high resolutions required for modern HMDs.

6.2. Application

We explain the potential of the practical utility of our foveated spatial quantization framework for a standalone mobile VR/AR headset. In the setup of client-server streaming architecture [USC*24], a

Zone	Bit Depth (*bpc*)	Mean Threshold (°)	Q_3 Threshold (°)	Max Threshold (°)
0	8	0 - 9.44	0 - 9.44	0 - 9.44
1	7	9.44 - 10.95	9.44 - 10.95	9.44 - 10.95
2	6	10.95 - 11.24	10.95 - 11.24	10.95 - 11.24
3	5	11.24 - 13.06	11.24 - 15.46	11.24 - 21.91
4	4	13.06 - 14.20	-	-
5	3	14.20 - 16.27	15.46 - 21.91	21.91 - 27.83
6	2	16.27 - 27.15	21.91 - 37.93	27.83 - max
7	1	27.15 - max	37.93 - max	-

Table 1: *Eccentricity-based bit allocation zones derived from mean, Q_3, and max perceptual thresholds. The Q_3 and max values are used for safer compression in lower bit regimes (< 5 bpc).*

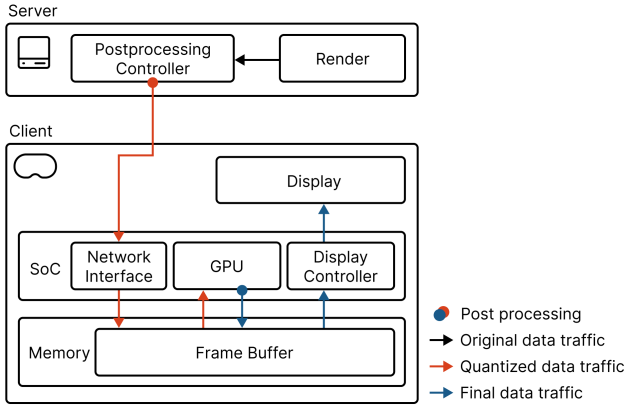

Figure 11: *Data communication traffic in a client-server streaming architecture for standalone VR/AR. Black arrows indicate original data traffic, red arrows denote quantized data traffic, and blue arrows represent the final display traffic.*

powerful server (e.g., a cloud instance or a nearby base station) processes the rendered GT content with a single-pass post-processing spatial quantization, while the client (e.g., a mobile HMD with gaze tracking) is responsible for a minimal blur recovery post-processing (Figure 11). By applying our foveated spatial quantization on the server side, we can largely reduce the data bandwidth for transmission from server to client, and also lower computational load on the client device.

The most immediate application is the reduction of wireless network traffic from the server to the client. This reduction directly lowers the data required for transmitting each frame, enabling smaller data packets and consequently reducing transmission latency and buffering delays. In bandwidth-constrained or latency-sensitive scenarios such as wireless HMD streaming or cloud gaming, this can substantially improve responsiveness and user-perceived quality of experience. Unlike previous approaches based on tiling or video encoding QP modification, our approach specifically exploits color representation drop-off and operates directly in the spatial domain as a lightweight post-process, making it a potential complement to existing optimizations. Modern codecs (e.g., HEVC or H.264/x264) are designed for the visual fidelity in the fovea, whereas human chromatic acuity drops more rapidly. The interaction between our method and codec behavior (e.g., quantization parameter adaptation, motion compensation, and temporal prediction) remains an open question.

Another application is the reduction of on-device DRAM traffic on the client side. DRAM traffic, the constant movement of frame data between the frame buffer and SoC components, drives a significant portion of an SoC's power consumption. Reducing DRAM traffic is paramount for mobile HMDs because data transfer energy far outweighs computational energy [USC*24]. Transferring a single Byte of information through DRAM consumes approximately 800× more energy than a single Multiply-Accumulate operation on 1-Byte fixed-point data. Compared to the analysis for the popular Base + Delta compression at around 9 bpp on average and Ujjainkar et al's [USC*24] improvement at around 7 bpp on average, our method can potentially store and fetch as low as 4 bpp data without considering metadata. This compression is also less content-dependent than previous compression, hence providing a more stable compression rate.

7. Limitations and Future Work

Our current framework applies a consistent blur-recovery level across all reduced bit depths. However, for moderate quantization, for example, reducing to 5 bpc, this process might not be necessary. At these higher precision levels, the differential thresholds for color discrimination and chromatic contrast sensitivity may be sufficiently high that the quantization error falls below the acuity drop even without recovery. Additionally, increasing blur progressively with eccentricity could further mask quantization artifacts, potentially enabling more aggressive bit reduction in the periphery. However, for computational efficiency and implementation simplicity, we still keep it at only one level.

Based on the qualitative results, more artifacts appear during saccades. This can be mainly due to the latency in eye-tracking and rendering updates, which leads to temporal misalignment between the gaze location and the applied quantization pattern. In our study, participants were explicitly instructed to search for artifacts, which may amplify their visibility compared to natural viewing conditions. In practical scenarios, where gaze behavior is task-driven and not focused on artifact detection, such artifacts may be less perceptually salient.

We use image-independent CQ for its efficiency and suitability for real-time rendering. While image-dependent CQ methods that adapt to image content can achieve higher reconstruction fidelity, they are generally computationally expensive and thus unsuitable for real-time applications. Among the fastest methods are Median Cut and Wu's algorithm, as demonstrated in comparative studies [Pér21]. We implemented a modified Median Cut algorithm as an image-dependent CQ baseline, using histogram-based initialization to accelerate palette generation. However, the method runs below 30 fps when targeting palettes larger than 512 colors (9 bpp), while palette size smaller than this introduces noticeable temporal flickering due to palette instability across frames. These results highlight a trade-off between reconstruction fidelity and temporal stability in image-dependent CQ methods.

Our current strategy does not fully exploit the spatial characteristics of human vision. We use a uniform spatial resolution across

the visual field. In practice, foveated rendering techniques often reduce spatial resolution in the periphery, which could potentially be jointly optimized with chromatic quantization for additional bandwidth savings. Future work could explore integrated models that reduce both spatial and hue resolution for more bandwidth saving.

8. Conclusion

In this paper, we present a real-time perceptual spatial quantization method for the extreme bandwidth demands of gaze-contingent HMD-based systems. By shifting the focus from uniform per-pixel bit depth to eccentricity-dependent spatial color perception, we exploit the pooling behavior and color perception in the periphery. Our approach utilizes ordered dithering to maintain perceived chromatic gradients even at extreme quantization levels and utilizes blur to recover artifacts from quantization. Through a user study, we model the detection thresholds over eccentricity across quantization levels from 8 bpc to 1 bpc. Based on the results, we propose a step-mask model and a zone that achieves a theoretical data reduction of 70.2% - 83.5% on Pimax Crystal configurations, which suggests a potential for reducing wireless streaming bandwidth to minimize wireless and on-chip traffic in mobile VR/AR architectures. Future research will explore frame-dependent adaptive color quantization and the interplay between spatial and hue resolution.

Acknowledgments

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) under Grant No. 3102516.

References

[AA92] ANALOUI, MOSTAFA and ALLEBACH, JAN P. "Model-Based Halftoning Using Direct Binary Search". *Human Vision, Visual Processing, and Digital Display III*. Vol. 1666. SPIE, Aug. 1992, 96–108. DOI: 10.1117/12.135959 5.

[ACM25] ASHRAF, MALIHA, CHAPIRO, ALEXANDRE, and MANTIUK, RAFAŁ K. "Resolution Limit of the Eye — How Many Pixels Can We See?": *Nature Communications* 16.1 (Oct. 2025), 9086. DOI: 10.1038/s41467-025-64679-2 2, 3.

[ARFK17] ATTIG, CHRISTIANE, RAUH, NADINE, FRANKE, THOMAS, and KREMS, JOSEF F. "System Latency Guidelines Then and Now – Is Zero Latency Really Considered Necessary?": *Engineering Psychology and Cognitive Ergonomics: Cognition and Design*. Ed. by HARRIS, DON. Cham: Springer International Publishing, 2017, 3–14. DOI: 10.1007/978-3-319-58475-1_1 2.

[Bar99] BARTEN, PETER G. J. *Contrast Sensitivity of the Human Eye and Its Effects on Image Quality*. 1000 20th Street, Bellingham, WA 98227-0010 USA: SPIE, Dec. 1999. DOI: 10.1117/3.353254 3.

[Bay76] BAYER, BRYCE E. "An optimum method for two-level rendition of continuous-tone pictures". *Ineternl. Conf. on Comm.* Vol. 50. 1976, 69–77 6.

[BCKD15] BEKTAS, KENAN, CÖLTEKIN, ARZU, KRÜGER, JENS, and DUCHOWSKI, ANDREW T. "A Testbed Combining Visual Perception Models for Geographic Gaze Contingent Displays." *EuroVis (Short Papers)*. 2015, 67–71. DOI: 10.2312/eurovisshort.20151127 3.

[BKB19] BEAMS, RYAN, KIM, ANDREA S, and BADANO, ALDO. "Transverse chromatic aberration in virtual reality head-mounted displays". *Optics express* 27.18 (2019), 24877–24884. DOI: 10.1364/OE.27.024877 5.

[CBR20] COHEN, MICHAEL A., BOTCH, THOMAS L., and ROBERTSON, CAROLINE E. "The Limits of Color Awareness during Active, Real-World Vision". *Proceedings of the National Academy of Sciences* 117.24 (June 2020), 13821–13827. DOI: 10.1073/pnas.1922294117 2, 6.

[Cel23] CELEBI, M. EMRE. "Forty Years of Color Quantization: A Modern, Algorithmic Survey". *Artificial Intelligence Review* 56.12 (Dec. 2023), 13953–14034. DOI: 10.1007/s10462-023-10406-6 5.

[DBS*09] DUCHOWSKI, ANDREW T., BATE, DAVID, STRINGFELLOW, PARIS, et al. "On Spatiochromatic Visual Sensitivity and Peripheral Color LOD Management". *ACM Trans. Appl. Percept.* 6.2 (2009), 9:1–9:18. DOI: 10.1145/1498700.1498703 3.

[DÇ07] DUCHOWSKI, ANDREW T. and ÇÖLTEKIN, ARZU. "Foveated Gaze-Contingent Displays for Peripheral LOD Management, 3D Visualization, and Stereo Imaging". *ACM Trans. Multimedia Comput. Commun. Appl.* 3.4 (2007), 6:1–6:18. DOI: 10.1145/1314303.1314309 3.

[DCT*22] DUINKHARJAV, BUDMONDE, CHEN, KENNETH, TYAGI, ABHISHEK, et al. "Color-Perception-Guided Display Power Reduction for Virtual Reality". *ACM Trans. Graph.* 41.6 (Dec. 2022), 1–16. DOI: 10.1145/3550454.3555473 2, 3.

[Flo76] FLOYD, ROBERT W. "An adaptive algorithm for spatial gray-scale". *Proc. Soc. Inf. Disp.* Vol. 17. 1976, 75–77 5.

[FS11] FREEMAN, JEREMY and SIMONCELLI, EERO P. "Metamers of the ventral stream". *Nature Neuroscience* 14.9 (2011), 1195–1201. DOI: 10.1038/nn.2889 4.

[GFD*12] GUENTER, BRIAN, FINCH, MARK, DRUCKER, STEVEN, et al. "Foveated 3D Graphics". *ACM Trans. Graph.* 31.6 (2012), 164:1–164:10. DOI: 10.1145/2366145.2366183 2, 3.

[HPG09] HANSEN, THORSTEN, PRACEJUS, LARS, and GEGENFURTNER, KARL R. "Color Perception in the Intermediate Periphery of the Visual Field". *Journal of Vision* 9.4 (Apr. 2009), 26. DOI: 10.1167/9.4.26 2, 3.

[HRV97] HILL, B., ROGER, TH., and VORHAGEN, F. W. "Comparative Analysis of the Quantization of Color Spaces on the Basis of the CIELAB Color-Difference Formula". *ACM Trans. Graph.* 16.2 (Apr. 1997), 109–154. DOI: 10.1145/248210.248212 5.

[IGS*20] ILLAHI, GAZI KARAM, GEMERT, THOMAS VAN, SIEKKINEN, MATTI, et al. "Cloud Gaming with Foveated Video Encoding". *ACM Trans. Multimedia Comput. Commun. Appl.* 16.1 (Feb. 2020), 7:1–7:24. DOI: 10.1145/3369110 2.

[ISM17] ILLAHI, GAZI, SIEKKINEN, MATTI, and MASALA, ENRICO. "Foveated Video Streaming for Cloud Gaming". *2017 IEEE 19th International Workshop on Multimedia Signal Processing (MMSP)*. Oct. 2017, 1–6. DOI: 10.1109/MMSP.2017.8122235 2.

[Kaw03] KAWASE, MASAKI. "Frame Buffer Postprocessing Effects in DOUBLE-S.T.E.A.L". *Game Developers Conference (GDC)*. 2003. URL: https://www.gdcvault.com/view/public/1012351 6.

[LaV23] LAVALLE, STEVEN M. *Virtual reality*. Cambridge university press, 2023 1, 2.

[LDB*21] LI, DAVID, DU, RUOFEI, BABU, ADHARSH, et al. "A Log-Rectilinear Transformation for Foveated 360-Degree Video Streaming". *IEEE Transactions on Visualization and Computer Graphics* 27.5 (May 2021), 2638–2647. DOI: 10.1109/TVCG.2021.3067762 2.

[LH08] LIU, SHENG and HUA, HONG. "Spatialchromatic Foveation for Gaze Contingent Displays". *Proceedings of the 2008 Symposium on Eye Tracking Research & Applications*. ETRA '08. New York, NY, USA: Association for Computing Machinery, Mar. 2008, 139–142. DOI: 10.1145/1344471.1344507 2, 3.

[LSV*18] LUNGARO, PIETRO, SJÖBERG, RICKARD, VALERO, ALFREDO JOSÉ FANGHELLA, et al. "Gaze-Aware Streaming Solutions for the Next Generation of Mobile VR Experiences". *IEEE Transactions on Visualization and Computer Graphics* 24.4 (Apr. 2018), 1535–1544. DOI: 10.1109/TVCG.2018.2794119 2.

[MDV20] MENG, XIAOXU, DU, RUOFEI, and VARSHNEY, AMITABH. "Eye-dominance-guided foveated rendering". *IEEE transactions on visualization and computer graphics* 26.5 (2020), 1972–1980. DOI: 10.1109/TVCG.2020.2973442 2.

[MDZV18] MENG, XIAOXU, DU, RUOFEI, ZWICKER, MATTHIAS, and VARSHNEY, AMITABH. "Kernel foveated rendering". *Proceedings of the ACM on Computer Graphics and Interactive Techniques* 1.1 (2018), 1–20. DOI: 10.1145/3203199 2, 4.

[MFN16] MAUDERER, MICHAEL, FLATLA, DAVID R., and NACENTA, MIGUEL A. "Gaze-Contingent Manipulation of Color Perception". *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI '16. New York, NY, USA: Association for Computing Machinery, 2016, 5191–5202. DOI: 10.1145/2858036.2858320 3.

[MM13] MANTIUK, RADOSŁAW and MARKOWSKI, MATEUSZ. "Gaze-Dependent Tone Mapping". *Image Analysis and Recognition*. Ed. by KAMEL, MOHAMED and CAMPILHO, AURÉLIO. Berlin, Heidelberg: Springer, 2013, 426–433. DOI: 10.1007/978-3-642-39094-4_48 3.

[Mul*02] MULLEN, KATHY T et al. "Differential distributions of red–green and blue–yellow cone opponency across the visual field". *Visual neuroscience* 19.1 (2002), 109–118. DOI: 10.1017/S0952523802191103 2.

[Mul85] MULLEN, K T. "The Contrast Sensitivity of Human Colour Vision to Red-Green and Blue-Yellow Chromatic Gratings." *The Journal of Physiology* 359 (Feb. 1985), 381–400. DOI: 10.1113/jphysiol.1985.sp015591 2.

[Mul91] MULLEN, K. T. "Colour Vision as a Post-Receptoral Specialization of the Central Visual Field". *Vision Research* 31.1 (Jan. 1991), 119–130. DOI: 10.1016/0042-6989(91)90079-K 2.

[Mus26] MUSIL, RICHARD. *HMD Geometry Database*. https://risa2000.github.io/hmdgdb/. Feb. 2026 7.

[Pér21] PÉREZ-DELGADO, MARÍA-LUISA. "Revisiting the Iterative Anttree for Color Quantization Algorithm". *Journal of Visual Communication and Image Representation* 78 (July 2021), 103180. DOI: 10.1016/j.jvcir.2021.103180 9.

[PHK*00] PUZICHA, J., HELD, M., KETTERER, J., et al. "On Spatial Quantization of Color Images". *IEEE Transactions on Image Processing* 9.4 (Apr. 2000), 666–682. DOI: 10.1109/83.841942 4.

[PSK*16] PATNEY, ANJUL, SALVI, MARCO, KIM, JOOHWAN, et al. "Towards Foveated Rendering for Gaze-Tracked Virtual Reality". *ACM Trans. Graph.* 35.6 (2016), 179:1–179:12. DOI: 10.1145/2980179.2980246 2, 3.

[SLNC13] SHEA, RYAN, LIU, JIANGCHUAN, NGAI, EDITH C.-H., and CUI, YONG. "Cloud Gaming: Architecture and Performance". *IEEE Network* 27.4 (July 2013), 16–21. DOI: 10.1109/MNET.2013.6574660 2.

[SRJ11] STRASBURGER, HANS, RENTSCHLER, INGO, and JÜTTNER, MARTIN. "Peripheral Vision and Pattern Recognition: A Review". *Journal of Vision* 11.5 (Dec. 2011), 13. DOI: 10.1167/11.5.13 2, 4.

[SWK25] STANCU, FLORIN-VLADIMIR, WEISS, TOMER, and KUFFNER DOS ANJOS, RAFAEL. "Foveated Animations for Efficient Crowd Simulation". *Proc. ACM Comput. Graph. Interact. Tech.* 8.1 (May 2025), 9:1–9:14. DOI: 10.1145/3728306 5.

[TAW*19] TURSUN, OKAN TARHAN, ARABADZHIYSKA-KOLEVA, ELENA, WERNIKOWSKI, MAREK, et al. "Luminance-contrast-aware foveated rendering". *ACM Trans. Graph.* 38.4 (July 2019). DOI: 10.1145/3306346.3322985 2.

[TTD22] TARIQ, TAIMOOR, TURSUN, CARA, and DIDYK, PIOTR. "Noise-based enhancement for foveated rendering". *ACM Trans. Graph.* 41.4 (2022), 1–14. DOI: 10.1145/3528223.3530101 2.

[USC*24] UJJAINKAR, NISARG, SHAHAN, ETHAN, CHEN, KENNETH, et al. "Exploiting Human Color Discrimination for Memory- and Energy-Efficient Image Encoding in Virtual Reality". *Proceedings of the 29th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 1*. Vol. 1. ASPLOS '24. New York, NY, USA: Association for Computing Machinery, Apr. 2024, 166–180. DOI: 10.1145/3617232.3624860 3, 8, 9.

[Val26] VALVE CORPORATION. *Steam Frame*. https://store.steampowered.com/sale/steamframe. Accessed: 2026-04-13. 2026 2.

[WAF*21] WALTON, DAVID R., ANJOS, RAFAEL KUFFNER DOS, FRISTON, SEBASTIAN, et al. "Beyond blur: real-time ventral metamers for foveated rendering". *ACM Trans. Graph.* 40.4 (July 2021). DOI: 10.1145/3450626.3459943 2, 4.

[Wal48] WALLACH, HANS. "Brightness Constancy and the Nature of Achromatic Colors". *Journal of Experimental Psychology* 38.3 (1948), 310–324. DOI: 10.1037/h0053804 4.

[WAT*25] WU, ZE, ALHILAL, AHMAD, TSUI, YUK HANG, et al. "EyeNexus: Adaptive Gaze-Driven Quality and Bitrate Streaming for Seamless VR Cloud Gaming Experiences". *Proc. ACM Netw.* 3.CoNEXT4 (Nov. 2025), 42:1–42:25. DOI: 10.1145/3768989 2.

[Wey58] WEYMOUTH, FRANK W. "Visual Sensory Units and the Minimal Angle of Resolution". *American Journal of Ophthalmology* 46.1, Part II (1958), 102–113. DOI: 10.1016/0002-9394(58)90042-4 3.

[ZAKL19] ZHANG, LILI, ALBERT, RACHEL, KIM, JOOHWAN, and LUEBKE, DAVID. "Developing a Peripheral Color Tolerance Model for Gaze-Contingent Rendering". *Journal of Vision* 19.10 (Sept. 2019), 298c. DOI: 10.1167/19.10.298c 2, 3, 6.

[ZS21] ZIEMBA, COREY M and SIMONCELLI, EERO P. "Opposing effects of selectivity and invariance in peripheral vision". *Nature communications* 12.1 (2021), 4597. DOI: 10.1038/s41467-021-24880-5 4.

[ZYX*20] ZHAN, TAO, YIN, KUN, XIONG, JIANGHAO, et al. "Augmented reality and virtual reality displays: perspectives and challenges". *Iscience* 23.8 (2020). DOI: 10.1016/j.isci.2020.101397 4.