

Perceptually Calibrated Foveated Chroma Attenuation for Gaze-Contingent VR Rendering and Displays

Bipul Mohanto
University of Rostock
bipul.mohanto@uni-rostock.de

Martin Weier
RheinMain University of Applied
Sciences
martin.weier@hs-rm.de

Colin Groth
New York University
c.groth@nyu.edu

Jérôme Kudnick
RheinMain University of Applied
Sciences
jerome.kudnick@hs-rm.de

Alexander Marquardt
Nara Institute of Science and
Technology
marquardt@ieee.org

Stefan Bruckner
University of Rostock
stefan.bruckner@uni-rostock.de

Oliver Staadt
University of Rostock
oliver.staadt@uni-rostock.de



Figure 1: Amazon Bistro under uniform and foveated chroma attenuation (our) with degradation weights (0.70,0.35). Insets show the fovea (●), near periphery (●), and far periphery (●) from the gaze point, at the center marked with an eye.

Abstract

We present a perceptually calibrated foveated chroma attenuation method for gaze-contingent VR rendering. Operating in the YC_bC_r color space, the method preserves luma while progressively attenuating peripheral chroma with increasing eccentricity using perceptually motivated models for the blue-difference and red-difference channels. Through a series of calibration studies, we identify degradation weights associated with low perceptual detectability of foveated chroma under the tested conditions. In the main perceptual study, participants correctly identified the foveated chroma in only 17.7% of all trials. Among these correctly identified trials, it was judged to be no more distracting than uniform

color rendering in 62.7% of cases. Exploratory system-level measurements showed marginal display-power reductions, averaging 0.19% for static scenes and 0.53% for dynamic scenes. Moreover, video compression experiments yielded an average compression gain of 3.18%, although the magnitude varied substantially across scenes. These results demonstrate that peripheral chroma attenuation was difficult to localize under the employed sequential protocol in gaze-contingent VR, while providing modest and content-dependent secondary benefits for display power and video compression.

CCS Concepts

• Computing methodologies → Perception.

Keywords

Foveated color rendering, peripheral chroma attenuation

ACM Reference Format:

Bipul Mohanto, Martin Weier, Colin Groth, Jérôme Kudnick, Alexander Marquardt, Stefan Bruckner, and Oliver Staadt. 2026. Perceptually Calibrated Foveated Chroma Attenuation for Gaze-Contingent VR Rendering and Displays. In . ACM, New York, NY, USA, 11 pages. <https://doi.org/XXXXXXX>. XXXXXXXX

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org. VRST2026, Sendai, Japan

© 2026 Copyright held by the owner/author(s). Publication rights licensed to ACM. ACM ISBN 978-1-4503-XXXX-X/2026/06 <https://doi.org/XXXXXXX.XXXXXXX>

1 Introduction

Human color vision responds to a narrow portion ($< 1\%$) of the electromagnetic spectrum corresponding approximately to wavelengths between 380 and 780 nm [12]. Visual perception begins when light reflected from a scene passes by the eye optics, reaches the retina, is absorbed by the photoreceptors, is transmitted through the optic nerve, and is ultimately processed in the primary visual cortex [23, 52]. Accurate color reproduction plays an important role in near-eye displays (NEDs) used in VR, AR, and MR systems [19, 42]. However, color perception retinal eccentricity-dependent, varies across observers, and can be influenced by display characteristics, such as brightness and contrast. Therefore, a generic quantitative description of color perception is challenging. Accordingly, the mechanisms underlying color perception and detection remain complex and are not yet fully understood [4, 13, 14].

Foveated rendering is an optimization strategy that offers a compelling solution by exploiting the non-uniform sensitivity of human vision [20, 29, 39, 40, 51, 56]. This enables the dynamic allocation of rendering resources to the fovea while gradually reducing fidelity toward the periphery. Nevertheless, prior foveated rendering studies have primarily focused on spatial [20, 26, 28, 39–41, 55, 59] and temporal [29, 51, 56] degradation whereas the chroma attenuation has received comparatively less attention. One possible reason could be, chroma attenuation alone provides relatively small computational gain compared with spatial or temporal foveation. Nonetheless, color suppression has been well studied in optics [11], vision science [58], and psychology [12]. Neurophysiological evidence suggests that color perception deteriorates with increasing eccentricity. This is partly because the lower density of cones in the peripheral retina makes color perception less precise than in the fovea [37, 57]. These characteristics suggest that uniform color fidelity across the NEDs can be replaced by progressive, eccentricity-dependent color reduction (Figure 1). Such attenuation can provide several additional benefits; for example, shading simplification [13, 16], reduce power consumption [8, 9, 17], and improve compression efficiency [31]. Recent advances in wide-field VR displays with integrated eye tracking make gaze-contingent chroma attenuation increasingly practical. Moreover, compared with uniform color, foveated chroma attenuation more closely reflects the non-uniform nature of human vision and provides a suitable framework for studying peripheral color perception under immersive viewing conditions.

The main contribution of this study is an investigation of foveated chroma attenuation in a real-time, gaze-contingent VR rendering system. We propose respective chroma attenuation models for the blue-difference (C_b) and red-difference (C_r) channels in YC_bC_r color space and estimate the corresponding degradation weights through calibration studies (Section 5). On this basis, we evaluate the following research hypotheses in Section 6:

H1. Foveated color rendering is difficult to distinguish from uniform color rendering under the employed sequential viewing protocol.

H2. Foveated color rendering does not increase perceived distraction compared with uniform color rendering.

We further examine the effects of foveated chroma attenuation on display-power consumption (Section 6.2) and video compression (Section 6.3). The primary objective of this work is to address

a gap in foveated color rendering by investigating the relationships among the perceptual detectability of chroma attenuation, display-power consumption, and compression efficiency in a real-time stereoscopic VR rendering environment.

2 Related Work

Prior studies of peripheral chroma attenuation have followed primarily three directions: task-dependent color reduction, color-zone mapping, and just-noticeable difference (JND)-based degradation. In task-dependent studies, Watson et al. [50] reported that substantial reductions in spatiochromatic complexity can remain imperceptible in visual search tasks, while Geisler and Perry [18] showed that chromatic spatial resolution is approximately half that of luminance resolution. Mauderer et al. [32] used CIELab and ΔE_{00} to study peripheral color perception through color matching and sorting tasks. Nonetheless, the experiments relied on simplified stimuli and fixed CRT-based setups.

Color zone mapping is one of the earliest and most widely adopted approaches to peripheral color reduction. Sakurai et al. [44] introduced an early color zone map based on psychophysical measurements over a large visual field; though the study was limited to unique hues under photopic viewing conditions. Hamada and Yujiti [21] later extended this line of work to different illumination conditions, and Duchowski et al. [15] employed similar ideas in gaze-contingent displays. Duchowski and Bate [13] further examined peripheral color degradation for spatiotemporal content and reported that the spatial degradation can be tolerated farther into the periphery than color. Their model subsequently became influential in foveated rendering research [1, 4, 5]. For example, Bektas et al. [5] applied foveated color to hybrid Geographic Information System (GIS) rendering and reported substantial compression benefits, while follow-up work recommended a medium-sized attenuation mask for balancing performance and artifact visibility [4]. However, most of these studies were conducted in non-NED settings or in image-processing contexts.

A third line of work models peripheral color degradation using JND thresholds. Zhang et al. [58] proposed a peripheral color tolerance model based on chromaticity discrimination thresholds in CIELab, combined with eccentricity-dependent bit-depth reduction. Recently, Duinkharjav et al. [17] demonstrated that perceptually guided color modulation can substantially reduce VR power consumption while preserving visual fidelity and Chen et al. [10] compared several peripheral color deterioration functions and provided empirical guidance for selecting attenuation strategies. Notably, the display-power model proposed by Duinkharjav et al. [17] intentionally shifts peripheral blue hues toward a greenish tint, with power reduction as its primary objective. In contrast, our study focuses on developing and perceptually calibrating a gaze-contingent chroma attenuation model over an appropriate parameter range that maintains low perceptual detectability across the tested scenes and viewing conditions. Moreover, image-space evaluations are useful for characterizing the appearance of precomputed stimuli, but they do not capture temporal and implementation-dependent factors of an operational foveated renderer, such as gaze-tracking error and latency, frame-to-frame changes in attenuation, display-specific characteristics, or interactions with scene content. Mohanto

et al. [34] studied real-time foveated color rendering in HSV color space, applying degradation only to saturation via a piecewise linear fall-off. Our work, therefore, complements prior psychophysical and image-based investigations by evaluating perceptually calibrated chroma attenuation in a real-time, eye-tracked VR rendering under natural viewing conditions. In contrast to Mohanto et al. [34], our approach operates in the YC_bC_r color space and applies continuous chroma attenuation with eccentricity rather than using discrete attenuation regions. Foveated color rendering may also be applicable to data compression and cloud rendering, although explicit color treatment remains uncommon because most existing approaches (e.g., [2, 24, 27, 54]) emphasize spatial degradation and/or encoder quantization control.

3 Human Visual System

Color perception varies across retinal eccentricity [52]. This non-uniformity is mainly governed by three factors: photoreceptor absorption, retinal signal transmission, and cortical processing. The first factor is light absorption by photoreceptors. Under photopic vision, color perception is mediated by three classes of cones, namely S, M, and L cones, with peak sensitivities near 440, 540, and 570 nm, respectively [18, 33]. Cone density is highest in the fovea and declines with eccentricity, with the steepest decrease occurring near the center of the visual field [43]. Nevertheless, this decline is not uniform across cone classes, because the fovea is dominated by densely packed L and M cones, whereas S cones are sparse and largely absent from the central fovea [7]. Outside the fovea, the relative proportion of S cones increases, while the densities of L and M cone decrease [3, 22, 36]. The second factor is retinal signal transmission. The cone generated signals are processed by networks of bipolar, horizontal, and amacrine cells before being transmitted by retinal ganglion cells (RGCs) through the optic nerve. RGC density is also strongly eccentricity-dependent, peaking in the central retina and decreasing substantially toward the periphery [48, 49]. As a result, central vision receives far denser neural sampling than peripheral vision. The third factor is cortical processing. Visual information is processed across multiple cortical areas, with distinct emphases on foveal and peripheral input [46]. Foveal input receives a disproportionately large cortical representation, whereas sensitivity to spatial and chromatic detail is lower in peripheral input [25, 35, 53]. Consequently, peripheral color perception increasingly diverges from fovea with increasing eccentricity. Red-green sensitivity falling faster than blue-yellow or luminance sensitivity [3, 22]. This asymmetry is highly relevant for chroma attenuation, since different chroma components may permit different degrees of peripheral attenuation. However, the exact limits of peripheral color perception remain debated, especially under naturalistic and immersive viewing conditions [12]. These observations motivate our methodology (Section 4) in which we employ separate chroma-attenuation models for the C_b and C_r channels.

4 Methodology

We introduce a perceptually-motivated and psychophysically calibrated foveated chroma attenuation method for gaze-contingent VR rendering. The method involves the following primary steps:

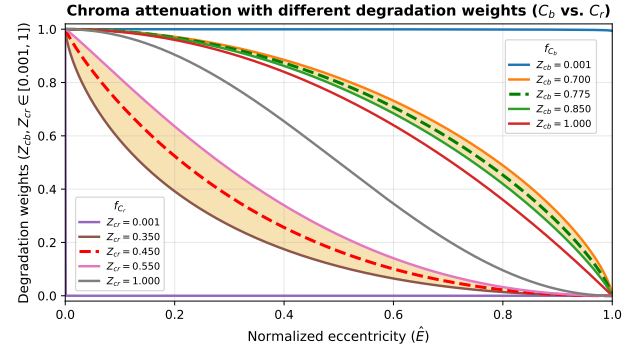


Figure 2: Foveated chroma attenuation models (Equation 1, Equation 2) with derived degradation weights $Z_{cb} \in [0.70, 0.85]$ and $Z_{cr} \in [0.35, 0.55]$ from Section 5.3.

4.1 Eccentricity Mapping

We used a vertical field of view ($vFoV$)-anchored pinhole projection to map a target eccentricity E_{deg} (in $^\circ$) from the gaze point to a radius in a fullscreen shader (uv space). Since Unity evaluates trigonometric functions in radians (E), additional degrees to radians conversion is necessary. The mapping is, $R(E) = 0.5(\tan(E)/(\tan(vFoV/2)))$; here, $R(E)$ denotes the radius in normalized shader coordinates under $vFoV$ normalization. The aspect-corrected radial distance has been computed as, $r_{uv} = \sqrt{(\alpha \Delta u)^2 + (\Delta v)^2}$, where, $\alpha = W/H$, is the framebuffer aspect ratio. This scales the u -axis so that radial distance remains isotropic in screen space. For the HTC Vive Pro Eye, the per-eye framebuffer resolution is 1440×1600 , resulting in $\alpha = 0.9$. The eccentricity is then normalized to $\hat{E} \in [0, 1]$ as $\hat{E} = \text{clamp}(r_{uv}/R(E_{max}), 0.0, 1.0)$; where, E_{max} defines the outer eccentricity limit of display, for HTC Vive Pro Eye that is 55° .

4.2 Foveated Chroma Attenuation

We use the YC_bC_r color space rather than a physiological opponent-color space such as DKL. YC_bC_r separates the luma component (Y) from the blue-difference (C_b) and red-difference (C_r) chroma components. This representation allows the chroma components to be attenuated independently while leaving the luma component unchanged. The channel-specific attenuation functions are therefore treated as empirically calibrated rendering functions rather than physiological models of opponent-color sensitivity. Prior vision research reports that chromatic sensitivity decreases with eccentricity and that different chromatic mechanisms exhibit different eccentricity-dependent sensitivity profiles [3, 6, 22, 45]. This evidence motivates the use of separate chroma attenuation models for the two chroma components. The C_b exhibits a relatively gradual and smooth sensitivity decline with eccentricity, consistent with the broader spatial distribution of S-cone-driven mechanisms [3, 6, 22, 45]. The Tukey biweight function (Equation 1) naturally captures this behavior. It produces a smooth, symmetric roll-off that approaches zero only gradually toward the periphery, thereby preserving a larger fraction of chroma energy in C_b at mid-eccentricities.

$$f_{C_b}(\hat{E}; Z_{cb}) = (1 - \hat{E}^2)^{Z_{cb}}, \quad 0 \leq \hat{E} \leq 1. \quad (1)$$

This function provides a smooth polynomial attenuation profile whose curvature is controlled by degradation weight, Z_{cb} . In contrast, C_r sensitivity declines more steeply and becomes behaviorally negligible around 25–30° [3, 22]. The raised cosine function (Equation 2) reflects this by providing a more aggressive, cosine-shaped attenuation that reaches near-zero more rapidly with eccentricity controlled by the degradation weight, Z_{cr} .

$$f_{C_r}(\hat{E}; Z_{cr}) = \frac{1}{2} \left[1 + \cos(\pi \hat{E}^{Z_{cr}}) \right], \quad 0 \leq \hat{E} \leq 1. \quad (2)$$

The calibration studies described in Section 5 estimate perceptually suitable operating ranges for these models. While these functional forms proved suitable for the present study, comparisons with alternative attenuation functions could provide further insight into their relative characteristics. The degradation weights Z_{cb} and Z_{cr} control the shapes of their respective attenuation functions. Larger Z_{cb} values produce stronger C_b attenuation, whereas smaller Z_{cr} values produce stronger C_r attenuation (Figure 2). Because the two degradation weights belong to different functional forms and have opposite relationships with attenuation strength, their numerical values are not directly comparable. They are therefore interpreted as function-specific operating degradation weights. Both Z_{cb} , Z_{cr} are constrained to [0.001, 1.0]. The lower bound avoids zero-exponent ambiguities at the endpoints of the normalized eccentricity domain. Finally, the modified YC_bC_r color space is subsequently converted back to the RGB color space for display. Both attenuation functions are continuous over the normalized eccentricity domain and computationally inexpensive to evaluate in a fullscreen fragment shader, making them suitable for real-time gaze-contingent rendering.

5 Calibration Study Design

We conducted two within-subject calibration studies (CSs) to determine the degradation weights for the Equation 1 and Equation 2. CS1 and CS2 were conducted in counterbalanced order, with a break in-between. Experimental conditions, e.g., lighting condition, sitting arrangement, display hardware, task instruction were kept constant throughout the studies. The participant numbers were limited during the CSs, which is not uncommon in foveated rendering research [29, 40]. After providing written consent, participants completed headset fitting, IPD adjustment, and eye calibration. We did not employ active task like searching or tracking during CSs and encouraged participants to explore the scenes naturally. After the CSs, participants rated the study design over, R1: Visual Comfort, R2: User Experience, and R3: Experiment Difficulty on a 10-point ordinal scale.

Software and Apparatus

All CSs including main study (Section 6) were conducted in Unity 6000.0.53f1 using an HTC VIVE Pro Eye headset. It integrates a Tobii eye tracker (120 Hz sampling rate and 0.50–1.1° spatial accuracy within an ~ 20° FoV), accessed through SRanipal Runtime v1.3.6.12 and Unity Plugin v1.3.6.8. The system was equipped with an Intel Core i9-12900KS CPU, 128 GB DDR5 RAM, and an NVIDIA GeForce RTX 3090 GPU with 24 GB VRAM.

Ethics

Participation in the study was voluntary, and all data were collected in anonymized form. Potential risks, including cybersickness, as well as the study procedures and tasks, were explained to each participant individually, both verbally and in the written consent documentation. Prior to taking part in the user study, all participants provided written informed consent. Given the short duration of the sessions and the nature of the tasks, the study was assessed as involving only minimal risk to participants. For these reasons, and in accordance with the legal framework of the institution, no ethical review board was involved in the review of the study. Participants were informed that they could withdraw from the study at any time without consequences. No conflicts of interest were declared.

5.1 First Calibration Study (CS1)

The primary goal of CS1 was to estimate perceptual operating points for the chroma degradation weights through participant-controlled adjustments. We therefore employed a Bidirectional Method of Adjustment (BMoA), in which a perceptual criterion is approached from two adjustment directions [30]. The descending series started from the full color state, whereas the ascending series started from the maximum attenuated state. This bidirectional procedure enabled us to assess possible starting-point or directional effects.

5.1.1 Participants and Stimuli. A total nine participants (6M, 3F, age range 23–47, mean 31.1, SD 7.8 years) took part. Five had prior VR experience and four had none. All participants had normal or corrected-to-normal visual acuity and self-reported normal color vision. To improve ecological validity, we evaluated six diverse scenes (Figure 6) under plausible indoor, mixed, and outdoor lighting conditions. Each scene was presented in both adjustment directions. Scene order was randomized per participant, while the descending series always preceded the ascending series for a given scene. A neutral gray screen was displayed for one second between trials.

5.1.2 Task and Procedure. The task followed a single-interval adjustment procedure rather than requiring a binary detection judgment or rating-scale response. Participants freely explored each scene using eye and head movements and adjusted Z_{cb} and Z_{cr} in steps of 0.05. In the descending series, they stopped when color degradation first became noticeable; in the ascending series, they stopped when further adjustment produced no perceptible color improvement. A button press confirmed one parameter pair (Z_{cb} , Z_{cr}) per trial. The nine participants completed 12 trials each, yielding 108 confirmed parameter pairs.

5.1.3 Results and Discussion. As summarized in Table 1, Z_{cb} had a mean of 0.589 and a median of 0.700 ($SD = 0.397$, $IQR = [0.289, 1.000]$), whereas Z_{cr} had a mean of 0.380 and a median of 0.351 ($SD = 0.344$, $IQR = [0.001, 0.613]$). The 95% confidence intervals, computed from participant-wise mean settings ($N = 9$), were [0.492, 0.687] for Z_{cb} and [0.293, 0.466] for Z_{cr} . The broad SDs and IQRs reflect combined variability across participants, scenes, and adjustment directions, including responses at the parameter bounds. Because Z_{cb} and Z_{cr} parameterize different nonlinear attenuation functions with opposite degradation directions, their numerical magnitudes were interpreted separately rather than compared directly. The combined medians, $Z_{cb} = 0.700$ and $Z_{cr} = 0.351$, were retained as

Table 1: Descriptive statistics of the CS1 degradation weights, combined across 9 participants. Mean, median, SD, and IQR describe the collective trial-level distributions; the 95% CIs are computed from participant-wise mean settings.

Weight	Mean	Median	SD	IQR [Q_1, Q_3]	95% CI (mean)
Z_{cb}	0.589	0.700	0.397	[0.289, 1.000]	[0.492, 0.687]
Z_{cr}	0.380	0.351	0.344	[0.001, 0.613]	[0.293, 0.466]

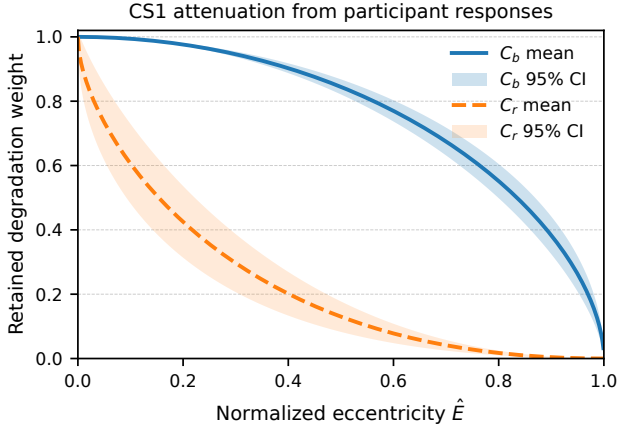


Figure 3: CS1 attenuation functions reconstructed from participant responses. The solid and dashed curves show the participant-wise mean C_b and C_r attenuation profiles, respectively; shaded regions denote the corresponding 95% confidence intervals across nine participants.

channel-specific scene-agnostic calibration estimates. To illustrate how variability in the estimated parameters influenced the attenuation profiles, the participant-wise mean settings were evaluated using the corresponding attenuation functions. Figure 3 presents the resulting C_b and C_r profiles together with their point-wise 95% confidence intervals. Moreover, paired Wilcoxon signed-rank tests on participant-wise mean settings found no significant difference between the ascending and descending series for either channel (Z_{cb} : $W = 17.0$, $p \approx 0.57$; Z_{cr} : $W = 19.0$, $p \approx 0.73$). Thus, no detectable directional hysteresis was observed, supporting the combination of both series.

5.2 Second Calibration Study (CS2)

The primary objective of CS2 was to evaluate the detectability of predefined degradation settings under foveated chroma attenuation. We employed Method of Constant Stimuli (MoCS), a non-adaptive psychophysical procedure in which predefined stimulus levels are repeatedly presented in randomized order and binary response proportions are used to characterize detection performance [38].

Because testing the full range of degradation weights at arbitrary step sizes in a user study is experimentally infeasible, we adopted a data-driven approach with BMoA procedure as described in Section 5.1. First, two experimenters with normal vision and color perception determined the relevant parameter ranges. For

Table 2: Summary statistics of the data-driven approach to determine initial degradation weight ranges for CS2.

Weight	Median	10% trimmed mean	IQR [Q_1, Q_3]
Z_{cb}	0.43	0.49	[0.25, 0.75]
Z_{cr}	0.35	0.38	[0.30, 0.50]

this, each experimenter completed three experimental repetitions with randomized scene (Figure 6) order, resulting in 720 trials per participant. Table 2 shows that the aggregated the degradation weights cluster in the mid-range of the admissible interval, avoiding both very strong and very weak chroma attenuation. For the C_b , the responses are centered at $Z_{cb} = 0.43$ (10% trimmed mean 0.49), with an interquartile range of [0.25, 0.75]. For the C_r , the corresponding values are $Z_{cr} = 0.35$ (10% trimmed mean 0.38), with a narrower interquartile range of [0.30, 0.50]. This pattern indicates that the viewers tolerated a broader range of C_b degradation, whereas the acceptable C_r settings were more concentrated.

5.2.1 Participants and Stimuli. A total nine participants (6M, 3F, age range 23–47, mean 31.1, SD 7.8 years) participated; five had prior, and the remaining four had no VR experience. All participants had normal or corrected-to-normal visual acuity with self-reported normal color perception. Each participant observed the same scenes as in CS1 (Figure 6). However, the study followed the MoCS procedure. For each scene, the full color (reference) was presented for three seconds, followed by a foveated color (target) created with a randomly assigned degradation-weight pairs from Table 2. Moreover, the scene order was randomized to mitigate order effects. Each of the six scenes was repeated 11 times, once for each randomly assigned degradation weight combination, resulting in 66 trials per participant, combining 594 binary observations. Between reference and target scene, we presented a neutral gray screen for one second.

5.2.2 Task and Procedure. The CS2 was a reference-anchored color perception test where a full color rendered scene followed by a foveated color scene was presented. After viewing each reference-target sequence, participants answered the question, “Can you perceive any color degradation in the pairwise sequence?”. A Yes response indicated that degradation was detected, whereas a No response indicated that it was not detected.

5.2.3 Results and Discussion. The overall empirical detection rates (Figure 4) ranged from 0.204 at $(Z_{cb}, Z_{cr}) = (0.85, 0.55)$ to 0.389 at $(0.40, 0.35)$, with an overall Yes rate of 0.291. We fitted a binomial Generalized Estimating Equation (GEE) model, clustered by participant, with Z_{cb} and Z_{cr} as predictors. The coefficient estimates and robust standard errors were 0.23 ± 0.62 for the intercept, -0.74 ± 0.71 for Z_{cb} , and -1.60 ± 1.63 for Z_{cr} . Neither slope differed significantly from zero ($p = 0.29$ for Z_{cb} and $p = 0.32$ for Z_{cr}); consequently, the data do not provide statistical evidence that either weight independently affected detection probability within the tested domain. The model predicted $P(\text{Yes})$ values between approximately 0.217 and 0.375. Because neither the empirical observations nor the model predictions approached 0.50, the sampled data-derived settings did not bracket the perceptual transition. CS2 therefore did not estimate a 50% detection threshold or cross-validate the CS1 adjustment

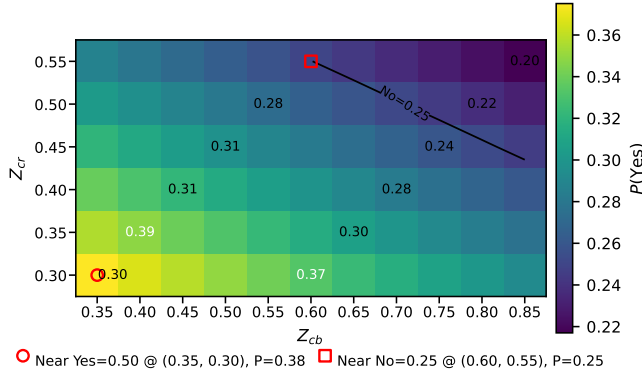


Figure 4: Scene-agnostic degradation-detection probabilities across the 11 tested degradation weight pairs. The background shows the participant-clustered GEE predictions over the displayed parameter domain, and the numerical labels show the empirical proportion of Yes responses at each tested pair (54 observations per pair). The $P(\text{Yes}) = 0.25$ contour is shown as a descriptive low-detection reference; no tested or predicted condition reached the conventional $P(\text{Yes}) = 0.50$ detection criterion.

points. Instead, it shows that the tested settings were detected in a minority of trials. The restricted detection range, joint variation of the two degradation weights, and between-participant response variability likely contributed to the wide coefficient uncertainties and limited sensitivity of the model.

5.3 Overall Discussion from CS1 and CS2

The two calibration studies provided complementary evidence for selecting degradation weights for the main study (Section 6). CS1 used participant-controlled bidirectional adjustments and identified clear scene-agnostic central tendencies (Table 1), with pooled medians of $Z_{cb} = 0.70$ and $Z_{cr} = 0.351$. CS2 evaluated 11 fixed, data-driven weight pairs using binary Yes–No degradation judgment. Detection probabilities remained low across the sampled domain, ranging from approximately 0.20 to 0.39, and neither Z_{cb} nor Z_{cr} significantly predicted detection probability in the participant-clustered GEE model (Figure 4). Therefore, CS2 did not identify or bracket a $P(\text{Yes}) = 0.50$ perceptual transition and should not be interpreted as a cross-validation of the CS1 settings. Instead, it showed that the tested data-derived combinations were detected in a minority of trials. For the main study, we consequently selected $Z_{cb} \in [0.70, 0.85]$ and $Z_{cr} \in [0.35, 0.55]$ as a pragmatic working range. The lower bounds were anchored near the CS1 median settings, while the upper bounds remained within the CS2-tested domain associated with generally low detection rates. This range represents an experimentally motivated parameter corridor rather than a statistically established JND interval.

Other than the degradation weights, across the CSs, participants reported (Figure 5) good R1: Visual Comfort and clearly positive R2: User Experience, with mean ratings around 7–8 on the 10-point ordinal scale (corresponding to slightly good to fairly good). For R3: Experiment Difficulty, responses were more heterogeneous. In

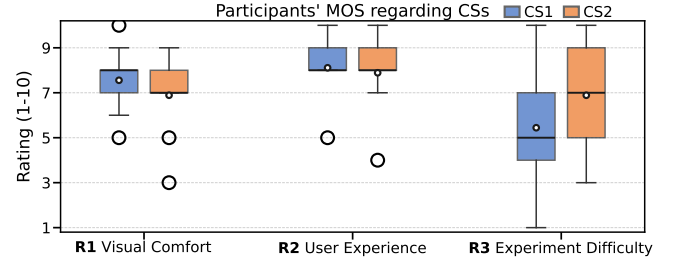


Figure 5: Distribution of questionnaire ratings for CS1 (●) and CS2 (●). Box plots show the IQR, median, data range excluding outliers, and mean rating for each questionnaire.

the CS1, ratings clustered around the middle of the scale, indicating an average to slightly easy task; however, with some participants experiencing it as hard and others as very easy. In the CS2, difficulty ratings shifted toward the easy side of the scale, suggesting that the CS2 was generally manageable to easy, while still showing considerable inter-individual variability.

6 Main Study

We conducted a within-subject repeated-measures study to evaluate the first two hypotheses introduced in Section 1.

Participants and Stimuli. An a priori G*Power analysis ($\alpha = 0.05$, $1 - \beta = 0.80$, medium effect size) indicated a minimum sample size of 26, which we adopted. Participants were aged 21–51 years (mean 31.9, SD 8.3), including 16 male (61.5%) and 10 female (38.5%). Fifteen participants (57.7%) reported prior VR experience, while 11 (42.3%) did not. All had normal or corrected-to-normal vision and self-reported normal color perception. The scenes are shown in Figure 6. We assigned no active task and emphasized on natural viewing. The scene order was randomized per participant. The participants were unaware about the specific objective of the study to reduce potential bias.

Task and Procedure. We employed a modified sequential triangulation procedure in which two uniform-color references and one foveated-color target were presented consecutively in randomized order. The three pairs of preselected degradation-weights were min (0.70, 0.35), avg (0.775, 0.45), and max (0.85, 0.55), based on Section 5.3; their assignment to the trials was randomized. Within each scene, the foveated target position was counterbalanced to occur once at each sequence position (A/B/C) across the three trials. Each representation was presented for eight seconds, during which they could freely move their gaze and head. Consecutive presentations were separated by a two-second neutral-gray interval to reduce immediate retinal afterimages.

After each sequence, participants answered RQ1, “Which representation looked different?”, using the response options A/B/C/Equal. The Equal option was included to avoid forcing an arbitrary positional response when participants perceived no reliable difference among the three representations, which was directly relevant to the indistinguishability hypothesis. Consequently, this task should be considered a three-position localization task with an additional

uncertainty response rather than a conventional forced-choice paradigm. Participants then answered RQ2, “Was the selected representation more or less distracting than the others?”, using the nominal response options *more/less/equal*. RQ2 was included in the analysis only when the foveated target had been correctly localized in RQ1. The main-study trials required 13.74 ± 1.48 minutes on average, while the complete session, including instructions, calibration, and debriefing, required 33.81 ± 4.56 minutes.

6.1 Results and Discussion

Across all 468 trials (26 participants $\times$ 6 scenes $\times$ 3 foveated chroma-attenuation settings), participants correctly localized the foveated target in 83 trials (17.7%). The nominal probability of selecting the correct target position from A/B/C is 33.3%; however, the additional *Equal* response means that the task was not a conventional triangulation procedure and that 33.3% does not constitute an exact chance model for the observed response process. Participants could select *Equal* rather than guess among A/B/C when they were uncertain. We therefore interpret the 17.7% rate descriptively as low target localization under the employed sequential protocol, rather than as statistically meaningful below-chance performance. Correct-localization rates were similarly low for the min (32/156, 20.5%), avg (26/156, 16.7%), and max (25/156, 16.0%) settings, with no reliable difference across settings (Friedman $\chi^2(2) = 2.60, p = 0.273$). These findings are consistent with H1, although they do not establish that stimulus indistinguishability alone caused the low localization rate. An exploratory scene-wise analysis showed rates between 12.8% and 16.7% for most scenes, whereas Emerald Square produced the highest localization rate (35.9%). Participant debriefings suggested that this scene-specific result may have been influenced by the nearby fence. Its high-frequency structure also produced a visible Moiré-like pattern, which may have increased discernibility independently of chroma attenuation. Nevertheless, the overall scene effect was not statistically significant (Friedman $\chi^2(5) = 9.08, p = 0.106$). The sequential presentation protocol constitutes an important limitation. Unlike simultaneous comparison, it required participants to retain each representation in visual working memory and may therefore have introduced temporal decay, adaptation, and primacy or recency effects. Randomized sequence order, counter-balanced target position, and neutral-gray intervals were intended to reduce systematic order and afterimage effects, but they could not eliminate the increased memory demand or measurement variability associated with sequential comparison. Consequently, the observed low localization rate may reflect a combination of perceptual similarity between the rendering conditions and uncertainty introduced by the presentation protocol. Additional future studies using simultaneous comparison and a strict forced-choice procedure without an *Equal* option would be required to completely disentangle these uncertainties, but would be considerably more demanding and laborious to conduct.

For H2, RQ2 was analyzed only for the 83 valid trials in which participants had correctly localized the foveated target in RQ1. In these trials, participants judged the selected representation as *less distracting* in 47 trials (56.6%), *equal* in 5 trials (6.0%), and *more distracting* in 31 trials (37.3%). Collapsing *less* and *equal* into a *not-more-distracting* category yielded 52/83 valid judgments (62.7%),

Table 3: Summary of target localization (RQ1) and valid distraction judgments (RQ2). RQ2 was analyzed only following correct target localization. L/E/M denotes less/equal/more distracting.

Condition	Q1 Right/N (%)	Valid Q2	L/E/M	Not-more (%)
Overall	83/468 (17.7)	83	47/5/31	62.7
min	32/156 (20.5)	32	17/1/14	56.2
avg	26/156 (16.7)	26	14/2/10	61.5
max	25/156 (16.0)	25	16/2/7	72.0

Table 4: Summary of questionnaire responses (N = 26).

Question	Median [Q1–Q3]	Mode	Range
R1: Visual Comfort (↑)	8.0 [6–8]	8	2–10
R2: User Experience (↑)	8.5 [8–9]	8	7–10
R3: Experiment Difficulty (↑)	7.5 [5–9]	10	1–10

compared with 31/83 (37.3%) *more distracting* judgments. This descriptive pattern is consistent with H2: when the foveated rendering was noticed, it was more often judged as not more distracting than as more distracting. However, because valid RQ2 responses were conditional on correct target localization and were unevenly distributed across repeated observations from participants, the result should be interpreted as qualified rather than confirmatory evidence. The small number of explicit *equal* responses also indicates that correctly detected differences were rarely experienced as identical in perceived distraction. Participant debriefings further suggested that the limited distraction was primarily associated with high-frequency scene textures rather than peripheral chroma attenuation. Overall, the evidence for H2 is therefore weaker than the evidence supporting H1.

Beyond the hypotheses, the post-experiment subjective evaluations (Table 4) were favorable. R1: Visual Comfort was likewise rated positively, with a median of 8 and nearly three-quarters of responses falling within the 7–10 range. R2: User Experience emerged as the strongest dimension, with all participants assigning ratings in the upper portion of the scale and a median score of 8.5, corresponding to assessments between fairly good and good. R3: Experiment Difficulty was generally evaluated as being on the easier side of the scale; however, it exhibited the greatest variability among the three measures, indicating comparatively larger inter-participant differences in perceived task demands.

6.2 Power Measurement

Due to the proprietary nature of the VR headset’s power management, on device power measurement was impracticable. Hence, we built a custom setup to measure the power consumption. The setup has a Wisecoco 3.81” Micro OLED stereo display (1080 $\times$ 1200 resolution per-eye and 90 Hz refresh rate), driven by an H381DLN 01.2 MIPI controller and monitored with an Arduino UNO R3 and INA219 current sensor (Figure 7). The characteristics were selected to approximate the HTC Vive Pro Eye used in the preceding experiment. Six scenes in Figure 6 were evaluated under static and dynamic viewing conditions. The dynamic condition used identical

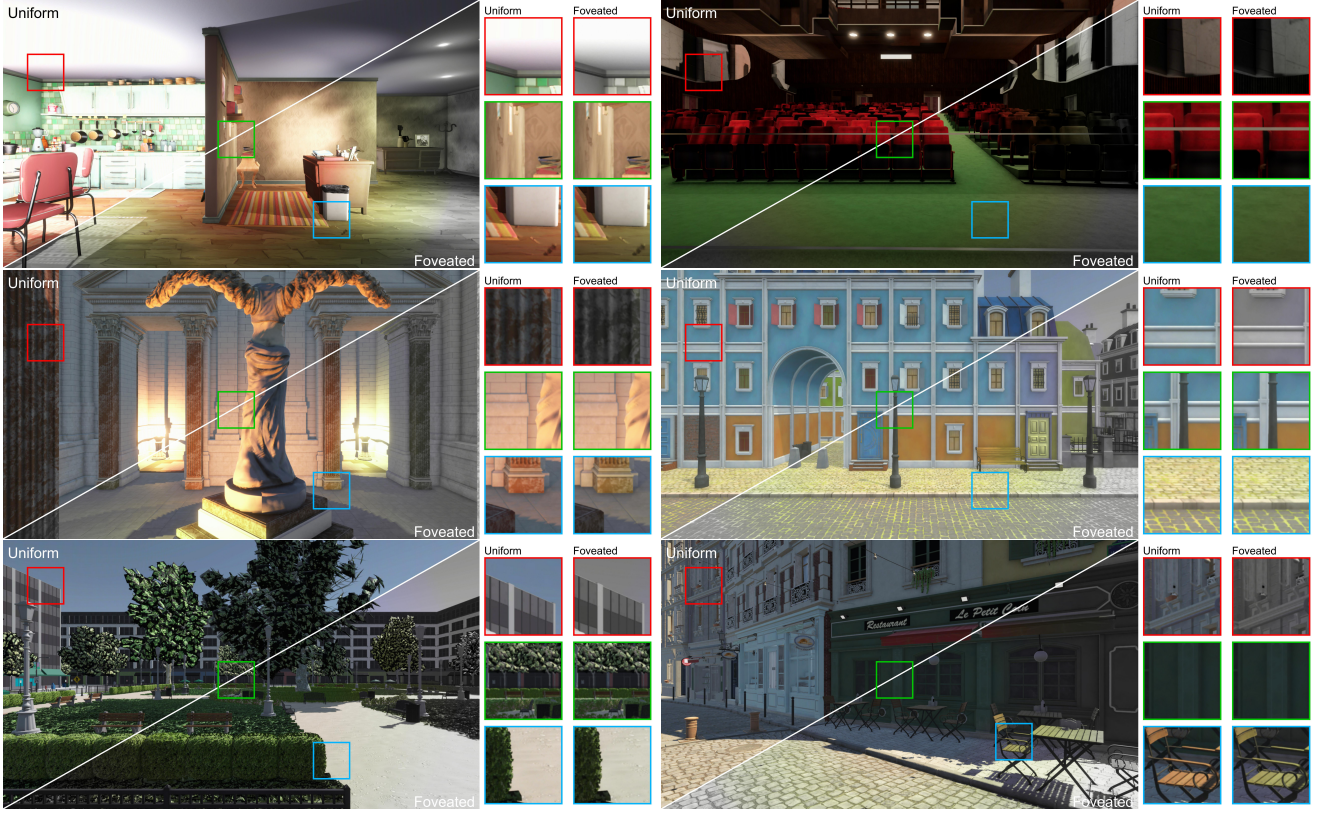


Figure 6: The uniform and foveated (degradation weights (0.70, 0.35)) color rendering. The inset view represents fovea (●), near periphery (●), and far periphery (●) over six scenes. Top row, Stylized House Interior and Gwangju Theater; mid row, Sun Temple and Modular Victorian City; bottom row, Emerald Square and Amazon Bistro. Gaze point is at the center.

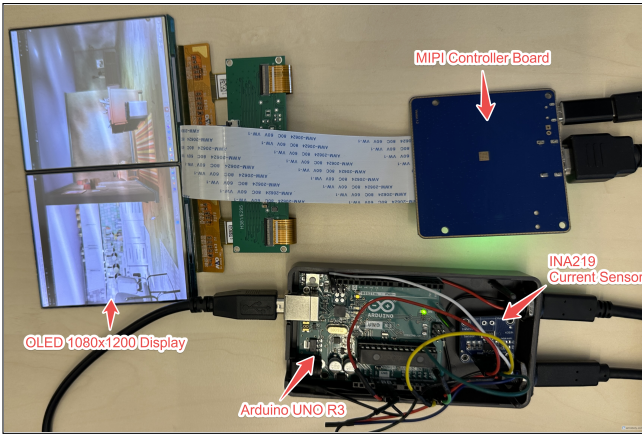


Figure 7: Display power measurement setup with Wisecoco Micro OLED, INA219 current sensor, and Arduino UNO R3.

predefined scanpaths and scene motion across foveated chroma attenuation settings. We sent the video signal at fullscreen 4K resolution from the Unity to inspect the maximum power consumption during the test. Under each conditions, we analyzed approximately

Table 5: Power reduction relative to the corresponding uniform color, reported as min/avg/max degradation weights.

Scene	Dynamic (%)	Static (%)
Amazon Bistro	1.10 / 1.10 / 0.47	0.29 / 0.22 / 0.24
Emerald Square	0.37 / 0.33 / 0.27	0.15 / 0.16 / 0.15
Stylized House Interior	0.19 / 0.14 / 0.03	0.19 / 0.15 / 0.12
Sun Temple	0.66 / 0.67 / 0.65	0.20 / 0.18 / 0.19
Gwangju Theater	0.60 / 0.59 / 0.54	0.12 / 0.09 / 0.10
Modular Victorian City	0.62 / 0.60 / 0.59	0.29 / 0.30 / 0.27
Overall	0.59 / 0.57 / 0.43	0.21 / 0.18 / 0.18

180-seconds of power data with 500-ms intervals. As all runs had equal duration, we report percentage reduction in mean power relative to the corresponding uniform color baseline.

Table 5 shows that foveated chroma attenuation consistently reduced display power under dynamic viewing across all scenes. Averaged over scenes, the reduction relative to uniform rendering was 0.59%, 0.57%, and 0.43% for the min, avg, and max chroma attenuation settings, respectively. The strongest dynamic savings were observed for Amazon Bistro (1.10% for both min and avg), followed

Table 6: Scene-wise compression results over six scenes, sorted from best to worst overall compression gain. For each scene, uniform color rendering is the baseline, and the foveated chroma attenuation is the mean over the three settings. Negative Δ values indicate that the foveated videos were *smaller* than the uniform baseline. Positive ratio-change values (last two columns) indicate stronger compression for the foveated videos. The overall compression gain is the mean of Δ H.264 and Δ HEVC.

Scene	Uniform H.264 (MB)	Foveated mean H.264 (MB)	Δ H.264 (%)	Uniform HEVC (MB)	Foveated mean HEVC (MB)	Δ HEVC (%)	Δ ratio H.264 (%)	Δ ratio HEVC (%)	Overall gain (%)
Emerald Square	37.451	34.028	-9.14	10.195	9.171	-10.04	10.07	11.16	-9.59
Modular Victorian City	46.662	45.168	-3.20	9.773	9.507	-2.72	3.31	2.80	-2.96
Stylized House Interior	30.192	29.337	-2.83	7.026	6.843	-2.60	2.92	2.67	-2.72
Amazon Bistro	27.158	26.677	-1.77	8.290	8.084	-2.48	1.81	2.56	-2.13
Sun Temple	10.697	10.673	-0.22	3.858	3.787	-1.83	0.38	1.99	-1.03
Gwangju Theater	237.168	234.241	-1.23	60.163	60.145	-0.03	1.28	0.08	-0.63
Mean over scenes	64.888	63.354	-3.07	16.551	16.256	-3.28	3.29	3.54	-3.18

by Modular Victorian City (0.59–0.62%), Gwangju Theater (0.54–0.60%), and Sun Temple (0.65–0.67%). Even the weakest dynamic case, Stylized House Interior max, remained below the uniform baseline (0.03%), indicating that the dynamic condition yields stable; however, scene-dependent power benefits. Under static viewing condition which is also view-dependent, the savings were smaller but remained positive for all scenes and all chroma attenuation settings. Averaged over scenes, the static reductions were 0.21%, 0.18%, and 0.18% for the min, avg, and max attenuation settings, individually. The largest static gains occurred in Modular Victorian City (0.27–0.30%) and Amazon Bistro (0.22–0.29%), both of which contain relatively large blue regions. Emerald Square and Gwangju Theater showed more modest but still consistent reductions. Averaged over all foveated configurations, the mean reduction was 0.19% for static conditions and 0.53% for dynamic conditions, corresponding to approximate average energy savings of 0.48 J and 1.39 J per run, respectively.

These results should not be interpreted as evidence that the proposed transformation is an energy-optimal foveated color rendering strategy. The method preserves the luminance component while attenuating C_b and C_r , progressively moving peripheral colors toward less chromatic values. Because our attenuation model does not incorporate primary-specific power coefficients, it cannot guarantee a monotonic relationship between attenuation strength and display-power reduction. Moreover, our objective is different from power-aware color optimization methods such as that of Duinkharjav et al. [17], which explicitly modify peripheral colors according to a display-power model and can consequently achieve substantially larger reductions. In contrast, our model is constrained by luminance preservation and perceptually calibrated chroma attenuation, without deliberately shifting colors toward lower-power primaries. The present measurements quantify the incidental energy effect of that perceptual constraint rather than demonstrating a competitive power-optimization method. From a practical perspective, reductions below (1%) are unlikely to produce a meaningful extension of battery life when used as a standalone technique. The findings instead establish that luminance-preserving peripheral desaturation offers only a marginal direct power benefit on the tested hardware. Such reductions could potentially complement spatial foveation, refresh-rate adaptation, brightness control, or primary-aware color optimization, but their combined effect was not evaluated in this study. Accordingly, display-power reduction should be regarded

as a secondary system-level observation rather than a primary advantage of the proposed method.

6.3 Video Compression

We captured 4K video under uniform and all three foveated chroma attenuation settings (min, avg, max) for static scenes and encoded them using FFmpeg. Table 6 reports the scene-wise compression results where the uniform rendering served as the baseline. The reported foveated value is the mean over the three foveated chroma attenuation settings.

Averaged across the six scenes, foveated chroma attenuation reduced compressed file size by 3.07% for H.264 and 3.28% for HEVC, corresponding to an overall mean reduction of 3.18%. However, the magnitude of this effect varied substantially across scenes. Emerald Square produced the largest reductions, amounting to 9.14% for H.264 and 10.04% for HEVC. For the remaining scenes, the overall reductions ranged from approximately 0.63% to 2.96%. The compression benefit should therefore be characterized as modest and strongly content dependent rather than as a uniform consequence of the attenuation model. Although the attenuation function itself is not content adaptive, its effect on an encoded bitstream necessarily depends on the input content. The number, magnitude, and spatial distribution of chroma differences removed by the method vary according to the chromatic composition of the scene, the amount of chromatic texture and edge detail, the position of those features relative to the gaze point, and the chroma-processing behavior of the codec. Scenes containing substantial peripheral chroma variation may therefore benefit more than scenes that are already weakly chromatic or whose bitrate is dominated by luminance structure, geometric detail, or temporal motion. The comparatively large reduction for Emerald Square and the much smaller reductions for the other scenes illustrate this dependency. The observed mean reduction of approximately 3% is unlikely to constitute a substantial bandwidth saving for an individual short sequence. Unlike approaches that explicitly exploit perceptual color-discrimination limits during image representation or encoding [47], our experiment applies the same luminance-preserving chroma attenuation used in the perceptual study and subsequently encodes the resulting frames using conventional codecs. Consequently, the present experiment does not optimize quantization, chroma representation, or memory traffic according to perceptual thresholds. The smaller compression gains should therefore not be interpreted as

contradicting prior perceptually optimized encoding approaches, but rather as the incidental compressibility benefit of the proposed attenuation model. It may nevertheless provide an auxiliary benefit in bandwidth-constrained or large-scale streaming scenarios, particularly when combined with spatial foveation, adaptive quantization, and gaze-contingent encoder. Such practical benefits cannot be established from the current experiment, which evaluates a limited set of scenes under static gaze and reports compressed file sizes rather than a complete rate-distortion analysis. Accordingly, the present results demonstrate that peripheral chroma attenuation can improve compressibility for some content, but they do not support a general claim of substantial or content-independent bandwidth reduction. Future evaluation should include dynamic gaze, a broader range of chromatic and motion characteristics, controlled bitrate comparisons, and rate-distortion measurements at matched perceptual quality.

6.4 Limitations and Future Work

The current study has several limitations. Our work is not intended to compete with established foveated color models; rather, it aims to define chroma attenuation functions and identify suitable degradation weight ranges. The CSs used a small participant sample and a limited set of discrete degradation weights. The experiments were conducted on a single VR headset with specific brightness-contrast and only short-term exposure was evaluated. As a result, transferability across devices and long-term effects on visual comfort or fatigue remain unclear. Future work should include more participants, additional VR platforms, and longer exposure time. Separating Z_{cb} and Z_{cr} would allow a broader exploration of the parameter space. Since headset color calibration was limited, the study relied on default device settings. The observed power savings were also limited, possibly due to the unchanged luminance channel. The video compression has been evaluated under only static conditions. Future work should evaluate the effect of dynamic gaze and motion on video compression. Finally, the evaluation focused mainly on subjective measures; adding objective metrics such as task completion time and accuracy would provide a more complete assessment.

7 Conclusion

This study demonstrates that peripheral chroma attenuation can be integrated into gaze-contingent VR rendering with limited perceptual cost. Across the calibration studies, perceptual evaluation, power analysis, and video-compression experiments, the results indicate that reducing peripheral chroma while leaving the luma component unchanged offers a practical trade-off among visual fidelity, energy consumption, and compression efficiency. Although the measured power and compression savings are modest, they were consistent and achieved without substantially increasing perceptual salience or distraction. These findings support foveated color as a promising direction for perception-driven VR rendering and a useful component in broader energy-aware display pipelines.

Acknowledgments

This research was partially supported by project THEMIS (01IS22077), JSPS KAKENHI (JP25KF0113), and a JSPS Postdoctoral Fellowship

for Research in Japan. The authors specially thank Martin Becker, University of Marburg, Germany.

References

- [1] Kaan Akşit, Praneeth Chakravarthula, Kishore Rathinavel, Youngmo Jeong, Rachel Albert, Henry Fuchs, and David Luebke. 2019. Manufacturing application-driven foveated near-eye displays. *IEEE transactions on visualization and computer graphics* 25, 5 (2019), 1928–1939.
- [2] Ahmad Alhailal, Ze Wu, Yuk Hang Tsui, and Pan Hui. 2024. FovOptix: Human Vision-Compatible Video Encoding and Adaptive Streaming in VR Cloud Gaming. In *Proceedings of the 15th ACM Multimedia Systems Conference (Bari, Italy) (MMSys '24)*. Association for Computing Machinery, New York, NY, USA, 67–77. doi:10.1145/3625468.3647612
- [3] Miyoshi Ayama, Masato Sakurai, Otto Carlander, Gunilla Derefeldt, and L. Eriksson. 2004. Color appearance in peripheral vision. In *Human Vision and Electronic Imaging IX*, Bernice E. Rogowitz and Thrasyvoulos N. Pappas (Eds.), Vol. 5292. International Society for Optics and Photonics, SPIE, 260 – 271. doi:10.1117/12.522240
- [4] Kenan Bektaş, Arzu Çöltekin, Jens Krüger, Andrew T. Duchowski, and Sara Irina Fabrikant. 2019. GeoGCD: improved visual search via gaze-contingent display. In *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications (Denver, Colorado) (ETRA '19)*. Association for Computing Machinery, New York, NY, USA, Article 84, 10 pages. doi:10.1145/3317959.3321488
- [5] Kenan Bektaş, Arzu Çöltekin, Jens Krüger, and Andrew T. Duchowski. 2015. A Testbed Combining Visual Perception Models for Geographic Gaze Contingent Displays. In *Eurographics Conference on Visualization (EuroVis) - Short Papers*, E. Bertini, J. Kennedy, and E. Puppo (Eds.). The Eurographics Association. doi:10.2312/eurovisshort.20151127
- [6] Norick R Bowers, Karl Gegenfurtner, and Alexander Goettker. 2025. Chromatic and Achromatic Contrast Sensitivity in the far Periphery. *bioRxiv* (2025), 2025–03.
- [7] David J. Calkins. 2001. Seeing with S cones. *Progress in Retinal and Eye Research* 20, 3 (2001), 255–287. doi:10.1016/S1350-9462(00)00026-4
- [8] Kenneth Chen, Budmonde Duinkharjav, Nisarg Ujjainkar, Ethan Shahan, Abhishek Tyagi, Jiayi He, Yuhao Zhu, and Qi Sun. 2023. Imperceptible Color Modulation for Power Saving in VR/AR. In *ACM SIGGRAPH 2023 Emerging Technologies* (Los Angeles, CA, USA) (SIGGRAPH '23). Association for Computing Machinery, New York, NY, USA, Article 8, 2 pages. doi:10.1145/3588037.3595388
- [9] Kenneth Chen, Nathan Matsuda, Thomas Wan, Ajit Ninan, Alexandre Chapiro, and Qi Sun. [n. d.]. ML-PEA: Machine Learning-Based Perceptual Algorithms for Display Power Optimization. *Computer Graphics Forum* n/a, n/a ([n. d.]), e70369. arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.70369 doi:10.1111/cgf.70369
- [10] Kenneth Chen, Thomas Wan, Nathan Matsuda, Ajit Ninan, Alexandre Chapiro, and Qi Sun. 2024. PEA-PODs: Perceptual Evaluation of Algorithms for Power Optimization in XR Displays. *ACM Trans. Graph.* 43, 4, Article 67 (jul 2024), 17 pages. doi:10.1145/3658126
- [11] Zaiqing Chen, Hui Liu, Qi Xiong, Xiaoqiao Huang, Yonghang Tai, and Junsheng Shi. 2023. A quantitative measurement of binocular fusion for achromatic and chromatic colors. *Optical Review* 30, 1 (2023), 50–60.
- [12] Michael A. Cohen, Thomas L. Botch, and Caroline E. Robertson. 2020. The limits of color awareness during active, real-world vision. *Proceedings of the National Academy of Sciences* 117, 24 (2020), 13821–13827. arXiv:https://www.pnas.org/doi/pdf/10.1073/pnas.1922294117 doi:10.1073/pnas.1922294117
- [13] Andrew T. Duchowski, David Bate, Paris Stringfellow, Kaveri Thakur, Brian J. Melloy, and Anand K. Gramopadhye. 2009. On spatiochromatic visual sensitivity and peripheral color LOD management. *ACM Trans. Appl. Percept.* 6, 2, Article 9 (mar 2009), 18 pages. doi:10.1145/1498700.1498703
- [14] Andrew T. Duchowski and Arzu Çöltekin. 2007. Foveated Gaze-Contingent Displays for Peripheral LOD Management, 3D Visualization, and Stereo Imaging. *ACM Trans. Multimedia Comput. Commun. Appl.* 3, 4, Article 6 (Dec. 2007), 18 pages. doi:10.1145/1314303.1314309
- [15] Andrew T. Duchowski, Nathan Cournia, and Hunter Murphy. 2004. Gaze-Contingent Displays: A Review. *CyberPsychology & Behavior* 7, 6 (2004), 621–634. arXiv:https://doi.org/10.1089/cpb.2004.7.621 doi:10.1089/cpb.2004.7.621 PMID: 15687796
- [16] Andrew T. Duchowski and Tiras D Eaddy. 2009. A Gaze-Contingent Display Compensating for Scotomata.. In *Eurographics (Short Papers)*, 9–12.
- [17] Budmonde Duinkharjav, Kenneth Chen, Abhishek Tyagi, Jiayi He, Yuhao Zhu, and Qi Sun. 2022. Color-Perception-Guided Display Power Reduction for Virtual Reality. *ACM Trans. Graph.* 41, 6, Article 210 (nov 2022), 16 pages. doi:10.1145/3550454.3555473
- [18] Wilson S. Geisler and Jeffrey S. Perry. 1998. Real-time foveated multiresolution system for low-bandwidth video communication. In *Human Vision and Electronic Imaging III*, Bernice E. Rogowitz and Thrasyvoulos N. Pappas (Eds.), Vol. 3299. International Society for Optics and Photonics, SPIE, 294 – 305. doi:10.1117/12.

- 320120
- [19] Firdevs Gökmenoğlu and Saadet Akbay. 2023. Virtual interior environment: Influence of colour on the sense of immersion. *Cultura e Scienza del Colore-Color Culture and Science* 15, 01 (2023), 13–18.
 - [20] Brian Guenter, Mark Finch, Steven Drucker, Desney Tan, and John Snyder. 2012. Foveated 3D Graphics. *ACM Transactions on Graphics* 31, 6, Article 164 (Nov. 2012), 10 pages. doi:10.1145/2366145.2366183
 - [21] Hiroshi Hamada and Akira Yujiri. 2004. Color Zone Map in peripheral vision at various illuminance levels. *Optical review* 11 (2004), 240–248.
 - [22] Thorsten Hansen, Lars Pracejus, and Karl R. Gegenfurtner. 2009. Color perception in the intermediate periphery of the visual field. *Journal of Vision* 9, 4 (04 2009), 26–26. arXiv:https://arxiv.org/abs/0904.2666
 - [23] Tomomi Ichinose and Samar Habib. 2022. ON and OFF Signaling Pathways in the Retina and the Visual System. *Frontiers in ophthalmology* 2 (2022), 989002.
 - [24] Gazi Karam Illahi, Thomas Van Gemert, Matti Siekkinen, Enrico Masala, Antti Oulasvirta, and Antti Ylä-Jääski. 2020. Cloud gaming with foveated video encoding. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)* 16, 1 (2020), 1–24.
 - [25] Daniel Jiménez-Navarro, Ana Serrano, and Sandra Malpica. 2024. Minimally disruptive auditory cues: their impact on visual performance in virtual reality. *The Visual Computer* (2024), 1–15.
 - [26] Henry Kam, Colin Groth, Jenna Kang, Pratham Saraf, Qi Sun, and Kenneth Chen. 2026. Dichoptic Foveation. In *Proceedings of the Special Interest Group on Computer Graphics and Interactive Techniques Conference Papers (SIGGRAPH Conference Papers '26)*. Association for Computing Machinery, New York, NY, USA, Article 139, 11 pages. doi:10.1145/3799902.3811108
 - [27] Anton S. Kaplanyan, Anton Sochenov, Thomas Leimkühler, Mikhail Okunev, Todd Goodall, and Gizem Rufo. 2019. DeepFovea: neural reconstruction for foveated rendering and video compression using learned statistics of natural videos. *ACM Trans. Graph.* 38, 6, Article 212 (Nov. 2019), 13 pages. doi:10.1145/3355089.3356557
 - [28] Jonghyun Kim, Youngmo Jeong, Michael Stengel, Kaan Akşit, Rachel Albert, Ben Boudaoud, Trey Greer, Joohwan Kim, Ward Lopes, Zander Majercik, Peter Shirley, Josef Spjut, Morgan McGuire, and David Luebke. 2019. Foveated AR: dynamically-foveated augmented reality display. *ACM Trans. Graph.* 38, 4, Article 99 (July 2019), 15 pages. doi:10.1145/3306346.3322987
 - [29] Brooke Krajancich, Petr Kellnhofer, and Gordon Wetzstein. 2021. A perceptual model for eccentricity-dependent spatio-temporal flicker fusion and its applications to foveated graphics. *ACM Trans. Graph.* 40, 4, Article 47 (jul 2021), 11 pages. doi:10.1145/3450626.3459784
 - [30] Tsuyoshi Kuroda and Emi Hasuo. 2014. The very first step to start psychophysical experiments. *Acoustical Science and Technology* 35, 1 (2014), 1–9. doi:10.1250/ast.35.1
 - [31] Sheng Liu and Hong Hua. 2008. Spatialchromatic foveation for gaze contingent displays. In *Proceedings of the 2008 symposium on Eye tracking research & applications*. 139–142.
 - [32] Michael Mauderer, David R. Flatla, and Miguel A. Nacenta. 2016. Gaze-Contingent Manipulation of Color Perception. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) (CHI '16). Association for Computing Machinery, New York, NY, USA, 5191–5202. doi:10.1145/2858036.2858320
 - [33] Victor Medina. 2016. *Visuo-perceptual validation methods for physically based image synthesis*. Theses. Université Paris sciences et lettres. https://pastel.hal.science/tel-01585882
 - [34] Bipul Mohanto, Martin Weier, Colin Groth, and Oliver Staadt. 2025. A Unified Real-Time Spatial and Chromatic Foveated Rendering in VR. In *GI Workshop 2025 on Virtual and Augmented Reality*. Gesellschaft für Informatik, Bonn, 10–18420.
 - [35] Micah M Murray and Mark T Wallace. 2011. *The neural bases of multisensory processes*. CRC Press.
 - [36] Debarshi Mustafi, Andreas H Engel, and Krzysztof Palczewski. 2009. Structure of cone photoreceptors. *Progress in retinal and eye research* 28, 4 (2009), 289–302.
 - [37] Cornelis Noorlander, Jan J Koenderink, Ron J Den Olden, and B Wigbold Edens. 1983. Sensitivity to spatiotemporal colour contrast in the peripheral visual field. *Vision research* 23, 1 (1983), 1–11.
 - [38] Adrien Paire, Anne Hillairet de Boisferon, and Céline Paeye. 2023. Empirical validation of QUEST+ in PSE and JND estimations in visual discrimination tasks. *Behavior Research Methods* 55, 8 (2023), 3984–4001.
 - [39] Anjul Patney, Marco Salvi, Joohwan Kim, Anton Kaplanyan, Chris Wyman, Nir Benty, David Luebke, and Aaron Lefohn. 2016. Towards foveated rendering for gaze-tracked virtual reality. *ACM Trans. Graph.* 35, 6, Article 179 (dec 2016), 12 pages. doi:10.1145/2980179.2980246
 - [40] Anjul Patney, Marco Salvi, Joohwan Kim, Anton Kaplanyan, Chris Wyman, Nir Benty, David Luebke, and Aaron Lefohn. 2016. Towards Foveated Rendering for Gaze-Tracked Virtual Reality. *ACM Trans. Graph.* 35, 6, Article 179 (dec 2016), 12 pages. doi:10.1145/2980179.2980246
 - [41] Lifeng Qiu Lin, Henry Kam, Qi Sun, and Kaan Akşit. 2025. Foveation Improves Payload Capacity in Steganography. In *Proceedings of the SIGGRAPH Asia 2025 Posters (SA Posters '25)*. Association for Computing Machinery, New York, NY, USA, Article 44, 2 pages. doi:10.1145/3757374.3771423
 - [42] Samad Roohi and Aynaz Forouzandeh. 2019. Regarding color psychology principles in adventure games to enhance the sense of immersion. *Entertainment Computing* 30 (2019), 100298.
 - [43] Galina I. Rozhkova, Alexander V. Belokopytov, Maria A. Gracheva, Egor I. Ershov, and Petr P. Nikolaev. 2021. A simple method for comparing peripheral and central color vision by means of two smartphones. *bioRxiv* (2021). arXiv:https://www.biorxiv.org/content/early/2021/01/13/2021.01.12.426150.full.pdf doi:10.1101/2021.01.12.426150
 - [44] Masato Sakurai, Miyoshi Ayama, and Takeshi Kumagai. 2003. Color appearance in the entire visual field: color zone map based on the unique hue component. *J. Opt. Soc. Am. A* 20, 11 (Nov 2003), 1997–2009. doi:10.1364/JOSAA.20.001997
 - [45] Samuel G Solomon, Barry B Lee, Andrew JR White, Lukas Rüttiger, and Paul R Martin. 2005. Chromatic organization of ganglion cell receptive fields in the peripheral retina. *Journal of Neuroscience* 25, 18 (2005), 4527–4539.
 - [46] Nicholas T. Swofford, José A. Iglesias-Guitián, Charalampos Koniaris, Bochang Moon, Darren Cosker, and Kenny Mitchell. 2016. User, Metric, and Computational Evaluation of Foveated Rendering Methods. In *Proceedings of the ACM Symposium on Applied Perception (Anaheim, California) (SAP '16)*. Association for Computing Machinery, New York, NY, USA, 7–14. doi:10.1145/2931002.2931011
 - [47] Nisarg Ujjainkar, Ethan Shahan, Kenneth Chen, Budmonde Duinkharjav, Qi Sun, and Yuhao Zhu. 2024. Exploiting Human Color Discrimination for Memory- and Energy-Efficient Image Encoding in Virtual Reality. In *Proceedings of the 29th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 1 (<conf-loc>, <city>La Jolla-</city>, <state>CA-</state>, <country>USA-</country>, </conf-loc>)* (ASPLOS '24). Association for Computing Machinery, New York, NY, USA, 166–180. doi:10.1145/3617232.3624860
 - [48] Brian A Wandell. 1996. *Foundations of vision*. SPIE.
 - [49] Andrew Watson. 2014. A formula for human retinal ganglion cell receptive field density as a function of visual field location. *Journal of vision* 14 (06 2014). doi:10.1167/14.7.15
 - [50] Benjamin Watson, Neff Walker, Larry F. Hodges, and Aileen Worden. 1997. Managing Level of Detail through Peripheral Degradation: Effects on Search Performance with a Head-Mounted Display. *ACM Trans. Comput.-Hum. Interact.* 4, 4 (Dec. 1997), 323–346. doi:10.1145/267135.267137
 - [51] Martin Weier, Thorsten Roth, Ernst Kruijff, André Hinkenjann, Arsène Pérard-Gayot, Philipp Slusallek, and Yongmin Li. 2016. Foveated Real-Time Ray Tracing for Head-Mounted Displays. *Comput. Graph. Forum* 35, 7 (Oct. 2016), 289–298.
 - [52] Christoph Witzel and Karl R Gegenfurtner. 2018. Color perception: Objects, constancy, and categories. *Annual review of vision science* 4, 1 (2018), 475–499.
 - [53] BR Wooten and George Wald. 1973. Color-vision mechanisms in the peripheral retinas of normal and dichromatic observers. *The Journal of general physiology* 61, 2 (1973), 125–145.
 - [54] Ze Wu, Ahmad Alhilal, Yuk Hang Tsui, Matti Siekkinen, and Pan Hui. 2025. EyeNexus: Adaptive Gaze-Driven Quality and Bitrate Streaming for Seamless VR Cloud Gaming Experiences. *Proc. ACM Netw.* 3, CoNEXT4, Article 42 (Nov. 2025), 25 pages. doi:10.1145/3768989
 - [55] Howard Xiao, Jan Ackermann, Boyang Deng, and Gordon Wetzstein. 2026. Policy-based Foveated Imaging and Perception. arXiv:2606.02565 [cs.CV] https://arxiv.org/abs/2606.02565
 - [56] Kai Xiao, Gabor Liptor, and Karthik Vaidyanathan. 2018. Coarse Pixel Shading with Temporal Supersampling. In *Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games (Montreal, Quebec, Canada) (I3D '18)*. Association for Computing Machinery, New York, NY, USA, Article 1, 7 pages. doi:10.1145/3190834.3190850
 - [57] Ariel Zeleznikow-Johnston, Yasunori Aizawa, Makiko Yamada, and Naotsugu Tsuchiya. 2023. Are color experiences the same across the visual field? *Journal of cognitive neuroscience* 35, 4 (2023), 509–542.
 - [58] Lili Zhang, Rachel Albert, Joohwan Kim, and David Luebke. 2019. Developing a peripheral color tolerance model for gaze-contingent rendering. *Journal of Vision* 19, 10 (2019), 298c–298c.
 - [59] Yan Zhang, Keyao You, Xiaodan Hu, Hangyu Zhou, Kiyoshi Kiyokawa, and Xubo Yang. 2024. Retinotopic Foveated Rendering. In *2024 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*. 903–912. doi:10.1109/VR58804.2024.00109