



High-fidelity color characterization in virtual reality across head mounted displays, game engines, and materials

FRANCISCO DÍAZ-BARRANCAS,^{†,*}  RAQUEL GIL RODRÍGUEZ,[†]
FLORIAN S. BAYER, AVI AIZENMAN,
AND KARL R. GEGENFURTNER 

Department of Psychology, Justus-Liebig-Universität, Alter Steinbacher Weg 38, Giessen 35394, Germany

[†]Co-first authors

*francisco.diaz-barrancas@psychol.uni-giessen.de

Abstract: We present a comprehensive colorimetric analysis of three head mounted displays (HMDs) - HTC Vive Pro Eye, Pimax 8K X DMAS, and Varjo Aero - focusing on their color calibration and uniformity across different game engines (Unity and Unreal) and for different materials/shaders. We developed a robust methodology combining hardware and software tools, including spectroradiometry and imaging colorimetry, to characterize and calibrate these HMDs for accurate color reproduction. The study showcases substantial advancements in colorimetric accuracy, with a reduction in the average ΔE_{00} of 90% or more across all tested HMDs and conditions. This level of color reproduction quality is below human discrimination thresholds, ensuring that any color inaccuracies remain imperceptible to the human eye. We also identified key areas for improvement, particularly in display uniformity, which could impact peripheral color reproduction. By making our tools and code publicly available, this study aims to facilitate future research and development in virtual reality (VR) technology, emphasizing the importance of color fidelity in virtual environments. The new insight enabled by our work is the extension and application of a traditional calibration method to currently available HMDs.

© 2024 Optica Publishing Group under the terms of the [Optica Open Access Publishing Agreement](#)

1. Introduction

Virtual reality (VR) has advanced considerably, leading to a diverse array of applications for VR systems [1–3]. Developments in head mounted display (HMD) technology and VR have significantly enhanced the immersive experience in virtual environments [4]. These advancements are largely attributable to progress in physics-based rendering [5] and computer graphics [6], which have enriched the visual experience. Despite these advancements, achieving accurate color reproduction under varying lighting conditions in VR remains a challenge [7]. Accurate color representation is crucial for the physical authenticity of a virtual scene [8]. Enhanced color rendering in VR content is especially important for applications demanding high realism [9].

Display technologies, such as those utilized in HMDs, commonly use red, green, and blue color channels (RGB) for managing color. The RGB model is inherently device-specific, meaning that the same RGB inputs may produce varied light emissions across different devices. Additionally, the choice of game engines in VR applications affects the RGB color model. Therefore, to ensure uniform color rendering among various devices and game engines, calibration is essential [10]. This process involves developing a model to accurately translate between the RGB input values and the corresponding visual stimuli presented on the HMD, and vice versa.

This paper presents an in-depth colorimetric analysis of three HMDs: HTC Vive Pro Eye, Pimax 8K X DMAS, and Varjo Aero. Throughout this study, we have thoroughly described our methodology in a step-by-step manner, ensuring it can be replicated with different HMDs, using various materials, shaders, or under different lighting conditions. Our results reveal a high level of

colorimetric accuracy in each HMD, evaluated across different materials/shaders and two popular game engines, Unity [11] and Unreal [12]. Additionally, we assessed the display uniformity of these devices. Our findings offer valuable insights for HMD manufacturers and VR application developers, guiding them towards improving color fidelity in virtual scene simulations.

2. Related work

Calibration and colorimetric characterization are crucial for accurate color control on display devices [13]. Calibration typically involves setting a display to known values, such as adjusting the white point, gain, and offset. In contrast, colorimetric characterization demands an understanding of the relationship between a device's input signals and its output. Given the vast range of chromatic stimuli that digital devices can display, it is impractical to measure all possible combinations. Therefore, it is commonly assumed that each color channel functions independently, allowing the reduction of necessary measurements to a feasible quantity. Various display characterization techniques have been systematically compared in recent literature [14]. Our research focuses on the colorimetric characterization of HMDs for psychophysical studies in color perception, employing a look-up table (LUT) model to achieve precise color reproduction across diverse display conditions.

Ongoing research in VR has been centered on enhancing chromatic characterization, with notable advancements in color reproduction. A study by [15] examines the application of INFITEC technology, a passive 3D technology for large screen projections, in immersive VR environments. This research underscores the significance of accurate chromatic characterization for creating realistic virtual scenarios and the future integration of perceptual principles. Conversely, [16] proposes a method for enhancing color fidelity in real-time 3D-rendered VR, focusing on color calibration, lighting, and hyperspectral textures. [17] introduced a framework that utilizes an imaging colorimeter for effective color calibration of VR HMDs, employing Unreal Engine for 3D rendering. Through the control of visual stimuli, this study facilitated immersive color vision research with high consistency, thereby advancing the application of VR technology in vision studies. [18] discusses the utility of VR in assessing indoor visual environments, emphasizing its advantages in controlling variables and reducing costs. The review also highlights the need for standardized investigation protocols in lighting research within VR. Additionally, [19] discusses adapting content production workflows for VR storytelling, specifically focusing on the challenges of color management. This work presents a workflow for characterizing and calibrating VR devices to ensure color accuracy. Finally, [20] discusses the control of stimulus luminance in VR perception experiments, offering linearization techniques and MATLAB code for application in Unity.

Although there have been studies on color calibration in VR, most have focused on a single HMD or graphics engine. [21] proposes a device-agnostic approach to color calibration in VR, ensuring precise control over visual stimuli. [22] presents color characterization of an OLED VR headset, demonstrating a reduction in error by approximately 44% post-calibration. Despite these advancements, there remains a gap in comprehensive studies quantifying the influence of various factors on HMD behavior, including different gaming engines and scene settings. Our work shows that methodologies for desktop monitor characterization can be adapted for HMDs [13,23–26], considering their unique optical designs and the indirect content control via game engines.

3. Methodology

This section is structured into subsections detailing the hardware and software utilized in our study, enhancing the replicability of our methods. Additionally, the code used in our research is publicly accessible at [27].

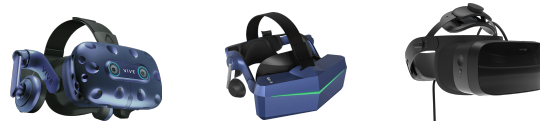
3.1. Hardware

Below, we describe the hardware components employed in our study, including HMDs, the computer system, and measurement devices.

3.1.1. Head mounted displays

We used three different HMDs for our measurements: HTC Vive Pro Eye, Varjo Aero, and Pimax 8K X DMAS. Detailed specifications of these HMDs are provided in Table 1. Please note that we used the default device settings for all configurations.

Table 1. Specifications of the different HMDs analyzed in this work (HTC Vive Pro Eye, Pimax 8K X DMAS, and Varjo Aero).



	HTC Vive Pro Eye	Pimax 8K X DMAS	Varjo Aero
Released Date	2018	2018	2021
Resolution (per eye)	1440 × 1660	3840 × 2160	2880 × 2720
Display	AMOLED	CLPL	Mini LED
Refresh Rate	90 Hz	90 Hz	90 Hz
Lenses	Fresnel	Fresnel	Aspheric
Field of View (FOV)	110°	159°	115°
Interpupillary distance (IPD)	61 – 72 mm	60 – 72 mm	57 – 73 mm

3.1.2. Computer

The measurements were conducted on a desktop computer running Windows 10, equipped with an Intel Core i9-9900k processor at 3.60GHz, 128GB RAM, and an Nvidia Titan RTX graphics card.

3.1.3. Spectroradiometer Konica Minolta CS-2000A

The spectral characterization of the HMDs was performed using the Konica-Minolta CS-2000A spectroradiometer [28], which offers a spectral resolution of 1 nm between 380 and 780 nm and using 1 degree of aperture. Our measurement approach aligns with recommendations from existing literature on color measurements in near-eye displays [29–32].

3.1.4. I29 imaging colorimeter

The I29 imaging colorimeter [33] allows for measurements of chromaticity CIE(x,y) and luminance values Y(cd/m²) with a resolution of 6576 × 4384 pixels and an accuracy of ±0.003 for x and y coordinates. This instrument was employed in conjunction with a specialized AR/VR lens [34], tailored for assessing HMD devices so that the colorimeter can be optimally positioned at the HMD's entrance pupil. The AR/VR lens offers a FOV of 120° (horizontal) × 80° (vertical).

3.2. Virtual scene

To simulate a controlled scenario in which both the position of the lights and the reproduction of the environment are faithful to a real environment, we have created a virtual twin of a light booth with a test disk (white reference) at the center (see Fig. 1). The image in the middle depicts the luminance of a real Just Normlicht when measured using the I29 colorimeter. The left and right images show simulations of a light booth using the Unreal Engine and Unity, respectively.

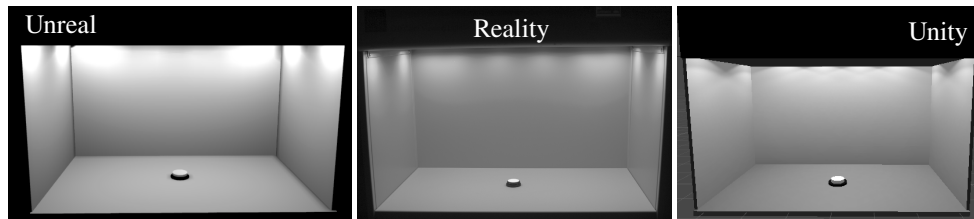


Fig. 1. Illustration of a real light booth from Just Normlicht (middle) vs the Virtual light booth for Unreal (left) and Unity (right). The RGB value of the test disk at the center of the light booth can be changed. In this example, the material of the disk is Standard for both engines.

3.3. Unity settings

For our study, Unity version 2019.1.5f1 was used. The color rendering behavior was set to linear mode, as Unity offers two rendering options (linear and gamma), which alter the internally used color space. We used linear mode to ensure the lighting calculations were mathematically correct [35]. This ensures our scenes are rendered with linear inputs and are not gamma-corrected. Deferred lighting behavior was applied, and HDR as well as any environmental post-processing by the graphics engine, were disabled. The deferred high definition render pipeline (HDRP) was chosen for lighting calculations over Forward HDRP, as this processes the lighting for every GameObject in the scene, rendering a high-quality virtual scene.

3.4. Unreal settings

We utilized Unreal Engine version 4.27.1. The light source in our light booth scene was set to stationary and positioned above the test disk. To maintain the original scene colors and prevent any unintended changes introduced by Unreal Engine's default post-processing, we disabled steps such as tone mapping or color grading. We set tone mapping to *none* prior to visualization.

3.5. Shaders/Materials

For both graphics engines, two shaders were selected for application: Unlit/Unlit and Standard/Lit, following Unity/Unreal terminologies, respectively. The subsequent paragraphs detail the differences between these materials, along with their primary characteristics.

In game engines, shaders are crucial for rendering the appearance of objects and scenes, where unlit and lit / standard (Unity/Unreal) materials are basic shader types. The color and luminance of unlit materials remain constant, irrespective of lighting conditions; e.g., unlit materials are displayed by the same RGB values in the HMD, regardless of whether the material is presented in a brightly lit or fully dark room in VR. The invariance of unlit materials across lighting conditions makes them ideal for controlled experimental designs. However, unlit materials look unrealistic in naturalistic scenes.

Lit/standard materials offer more realistic behavior by dynamically responding to scene lighting. The RGB values assigned to these materials are interpreted as the material's reflectance. Consequently, the RGB values ultimately displayed on the HMD are influenced by the illumination intensity and color, and local factors such as shadows and interreflections. Standard materials enable realistic light interactions and are suitable for rendering realistic scenes.

From now on, we will refer to the RGB value given to the engine as the *reflectance* value, regardless of the material/shader choice, and *reflected light* as the output XYZ value displayed in the headset.

The choice of shader significantly influences the visual outcome of the virtual environment. Figure 2 provides schematics comparing the visual effects of both the Unlit and Standard shaders

under identical lighting and rendering conditions in Unity and Unreal, respectively. For both shaders we introduce a reflectance value of $R = G = B = 1$ for the spheres presented in the VR scene.

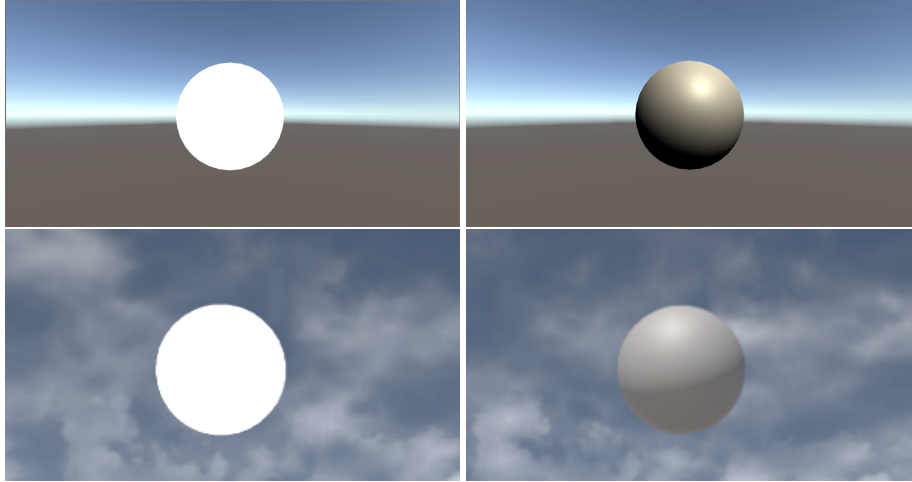


Fig. 2. Illustration of the Unlit shader (left) versus the Standard shader (right) as rendered in Unity (first row) and Unreal (second row). All the spheres have a reflectance value of $R = G = B = 1$ in the four different setups.

For more technical details on these two materials/shaders, please refer to *Shaders/Material* section in the [Supplement 1](#).

3.6. Display measurements setup

For the measurements, we utilized a custom-designed mechanical mount for the HMD with three degrees of freedom. This allows for the necessary adjustment of the headset's position and orientation relative to the measuring device. This arrangement was sufficient to achieve a satisfactory alignment of both components. Figure 3 depicts various configurations of this setup as employed in the measurement process.

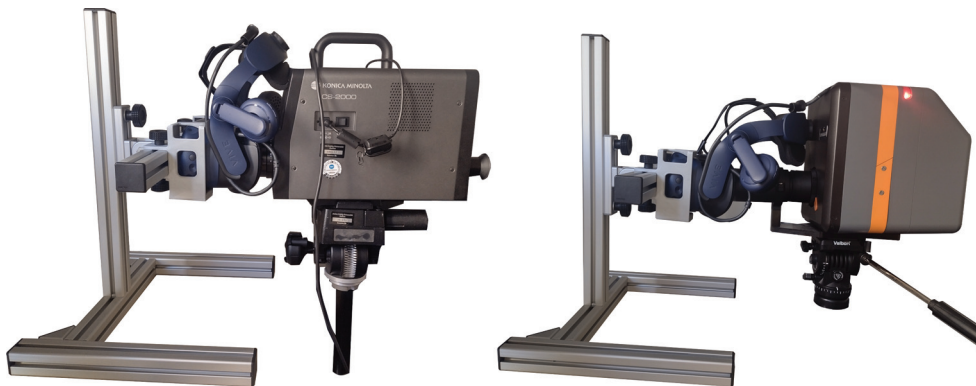


Fig. 3. Measurement setups using a custom support holder for the HMD: the spectroradiometer setup is shown on the left, and the colorimeter setup on the right.

HMDs incorporate a display and a lens for each eye. The lens is designed to maintain a compact form factor – typically only a few centimeters thick – while providing an optically comfortable viewing distance for the user, generally around 1 meter or more. This design, however, involves certain compromises: the quality of the resulting image is influenced by the relative position and orientation of the eye to the screen. To address this, the lens is engineered to optimize image quality within a defined, limited region known as the eye box. Even within this eye box, the image quality is contingent on the eye's orientation, specifically the user's gaze direction. It is presumed that the optical design is optimized for scenarios where the user is looking directly forward.

To ensure accurate measurement of the display and to faithfully replicate the optimal visual experience for users, positioning the measurement device within the eye box is crucial, with its orientation aligned to a forward gaze. To achieve this, we designed a scene in Unreal Engine where an image of concentric circles was projected onto a plane. By aligning the spectrometer and colorimeter with the center of these circles, we aim to measure the region that corresponds with a user's forward gaze direction. Figure 4 demonstrates this configuration, showing the alignment of the devices for precise measurements.

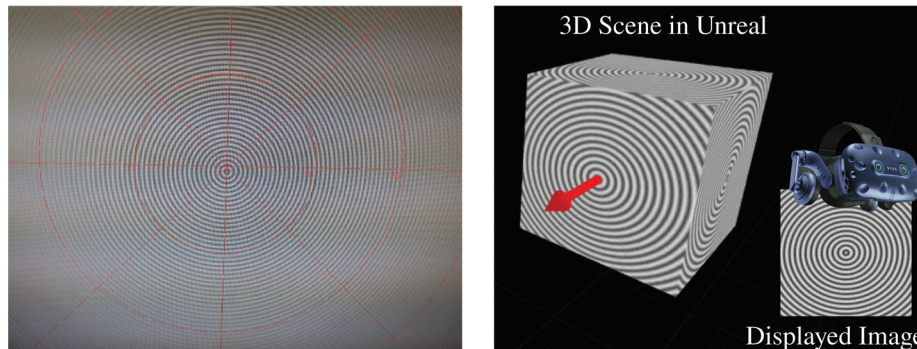


Fig. 4. Left: Image taken from the imaging colorimeter showing the alignment target (smallest circle). Right: Visualization of the alignment object in Unreal Engine. The red arrow marks the central position to which the headset will be aligned. The center of the concentric circles will be displayed in the projection center of the VR headset. This method ensures that all measurements are taken from the display's center and maintains consistency for repeated measurements.

Two types of measurements were conducted with the spectroradiometer:

1. **Characterization values:** To construct the characterization model, 52 intensity levels for each primary color were recorded. Another set of 52 intensities was also measured for each of the three color channels together (from black $RGB = (0, 0, 0)$ to white $RGB = (1, 1, 1)$), where all primary colors were combined at the same RGB intensity.
2. **Validation sets:**
 - A collection of 100 random RGB values was recorded for validating the calibration, as shown in the Results subsection 6.5.
 - Furthermore, 115 random values in the $CIE LCh_{ab}$ color space (LCH) were chosen, with their corresponding RGB values subsequently measured. This procedure is demonstrated in the Results subsection 6.5.

A total of 423 measurements were performed for each graphics engine and HMD combination. To streamline the color measurement process, a TCP/IP connection was established between

Matlab, which managed the spectroradiometer, and the graphics engines (Unreal or Unity). This setup enabled the automation of the measurement procedure.

4. Colorimetric characterization model

We have selected a display characterization model that emphasizes both simplicity and accuracy. This model aims to form a parameter-free and accurate relationship between the RGB values and the chromatic values of the stimulus across various reference color spaces.

Equation (1) outlines the characterization model utilized in our study. The input vector $(R, G, B)^T$ denotes the normalized intensity values provided to the engine, while the output vector $(X, Y, Z)^T$ represents the values measured using the spectroradiometer.

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = M \cdot \begin{pmatrix} R' \\ G' \\ B' \end{pmatrix}, \text{ where } \begin{pmatrix} R' \\ G' \\ B' \end{pmatrix} = f \left(\begin{pmatrix} R \\ G \\ B \end{pmatrix} \right) \text{ and } M = \begin{pmatrix} X_{R_{max}} & X_{G_{max}} & X_{B_{max}} \\ Y_{R_{max}} & Y_{G_{max}} & Y_{B_{max}} \\ Z_{R_{max}} & Z_{G_{max}} & Z_{B_{max}} \end{pmatrix}. \quad (1)$$

Our model contains two sequential transformations: firstly, a linearization function $f(\cdot)$ that converts the input $(R, G, B)^T$ into the linearized form $(R', G', B')^T$; and secondly, a linear transformation linking $(X, Y, Z)^T$ to the linearized values $(R', G', B')^T$.

The linear transformation in our model is represented by a matrix M , where each column corresponds to the $(X, Y, Z)^T$ values of the individual color channels at their maximum intensity. For example, setting $R = 1$, $G = 0$, and $B = 0$ denotes the maximum normalized intensity for the red channel. This matrix transforms the linearized $(R', G', B')^T$ values into $(X, Y, Z)^T$ values.

To establish the linearization function $f(\cdot)$, we chose to use a Lookup Table (LUT) to define the relationship between the nonlinear input $(R, G, B)^T$ and the linearized outputs $(R', G', B')^T$. Following the methods suggested by [23,36], we measured the $(X, Y, Z)^T$ for 52 equidistant RGB intensities where $R = G = B$, representing achromatic values in RGB terms. This equates to values ranging from 0 to 255 in 8-bit representation at intervals of 5. After normalizing the input and applying the inverse of matrix M to these measured $(X, Y, Z)^T$ values, the relationship between linear and nonlinear RGB values was determined. This led to the formulation of three 1D LUTs, one for each color channel. For unmeasured values, linear interpolation was employed. The rationale for using LUTs was their parameter-free nature, inherently providing greater generalizability and accuracy. This approach substantially improved calibration outcomes when contrasted with our earlier work [37], which utilized parametric functions.

From now on, we will define the *forward model* as the transformation from the non-linear $(R, G, B)^T$ input of the engine to the measured $(X, Y, Z)^T$ values on the display. Additionally, the *backward or inverse model* will be defined as the transformation from the measured $(X, Y, Z)^T$ values to the non-linear $(R, G, B)^T$ engine's input.

Figure 5 summarizes the color characterization model, irrespective of the game engine, material settings or HMD choice.

Each configuration (material, graphics engine, and HMD) follows the same characterization process. As a result, every setup is assigned a unique matrix and a set of 1D LUTs for each color channel. Figure 6 illustrates the LUT used for the red channel in HTC Vive Pro Eye and Varjo Aero for all their configurations. We refer the reader to Section 6.2 for a better understanding of this relation in the LUTs.

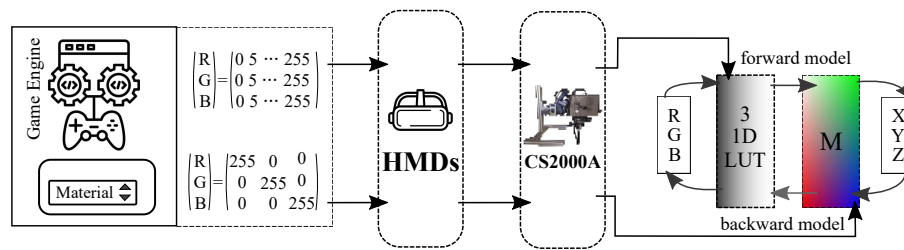


Fig. 5. Flowchart of the colorimetric characterization model.

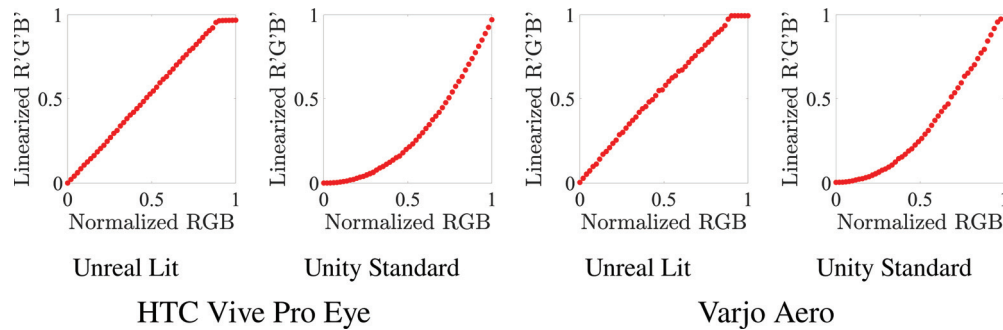


Fig. 6. LUT for the red channel, for HTC Vive Pro Eye and Varjo Aero, arranged from left to right, using Unreal and Unity standard configurations.

5. Step-by-Step procedure

This section details the systematic procedure for characterizing other VR devices. The corresponding code for this process is available publicly and can be accessed at [27].

- Aligning the device for measurement:** Position the VR device in front of the spectroradiometer as detailed in Section 3.6. Adjust the spectroradiometer to be approximately 5mm from the headset. Launch the Unreal project (**CenterHMD**) and align the spectroradiometer.
- Open Unity/Unreal project for calibration:** Open the Unity or Unreal project **CalibrationHMD** (based on the engine in use). Initiate the code to project the test disk at the center of the scene. Ensure the disk spans the entire width of the display and presents a uniform color over the visible field. The 'Display VR View' feature in Steam VR can be used to confirm display uniformity.
- Execute MATLAB script for spectroradiometer and Unity/Unreal control:** With the Unity or Unreal project active, run the MATLAB script for spectroradiometer measurements (**ConnectionUnity.m/ConnectionUnreal.m**). This automated process should be completed in approximately 10 minutes.
- Define the Colorimetric Characterization Model:** Post-measurement, process the data using the MATLAB script (**Characterization_HTCVive_Unity.m** for the HTC Vive Pro Eye in Unity, for example). This script generates the matrix M and three 1D LUTs, one for each color channel.
- Evaluate the calibration error of the VR device:** To assess the calibration error, execute the appropriate MATLAB script (**Characterization_HTCVive_Unity.m** for the HTC

Vive Pro Eye in Unity, for instance). This script computes the colorimetric error ΔE_{00} [38].

The above steps can be used to replicate our findings. The project settings in Unity or Unreal can be updated with the desired material or processing configurations. Note that the above pipeline describes a methodology that can be replicated with other shaders, settings, or HMDs.

6. Results

In the following sections, we will share our measurement results, discuss their implications, and explore how they enhance the community's efforts to improve existing methodologies.

6.1. Primaries spectra

Figure 7 presents the spectral composition for each channel across different HMDs for the Unlit material in Unity. The primary colors for the HTC Vive Pro Eye and Pimax HMDs display a high degree of independence with minimal spectral overlap in their power distributions. Conversely, the Varjo HMD demonstrates multiple peaks in the spectral power distributions for all primary colors, with peak wavelengths exhibiting similarities across these primaries. This implies that multiple LEDs of this HMD are combined to generate each of the individual primaries.

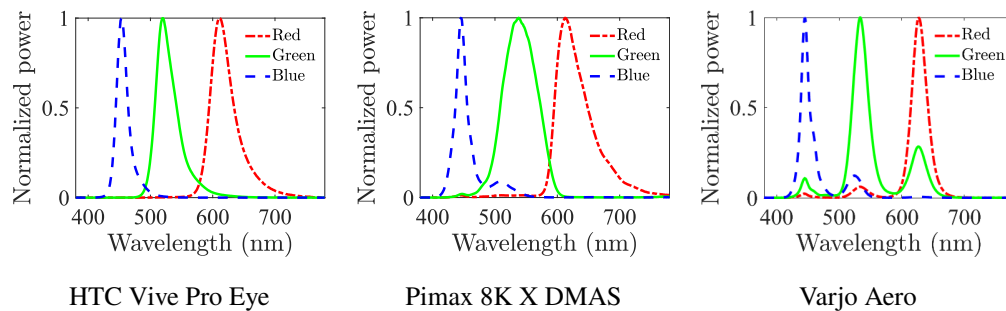


Fig. 7. Normalized power spectra for R, G, and B (red, green, and blue lines) under the Unity Unlit configuration across different headsets (from left to right, HTC Vive Pro Eye, Pimax 8K X DMAS, and Varjo Aero).

6.2. Luminance and additivity

As outlined in Section 3.5, Unlit materials in Unreal Engine are unaffected by lighting conditions in the final material calculation, which is a unique property. Figures 8, 9, and 10 display the relationship between input intensity and luminance for different game engines and materials across each HMD. Within Unity, for all materials and shaders examined, this input-luminance relationship conforms to a standard gamma function, expressed as:

$$L = l \cdot b^\gamma. \quad (2)$$

Here, L denotes luminance (measured in cd/m^2), l is a constant, b represents the normalized input intensity, and γ is the exponent in the power law function.

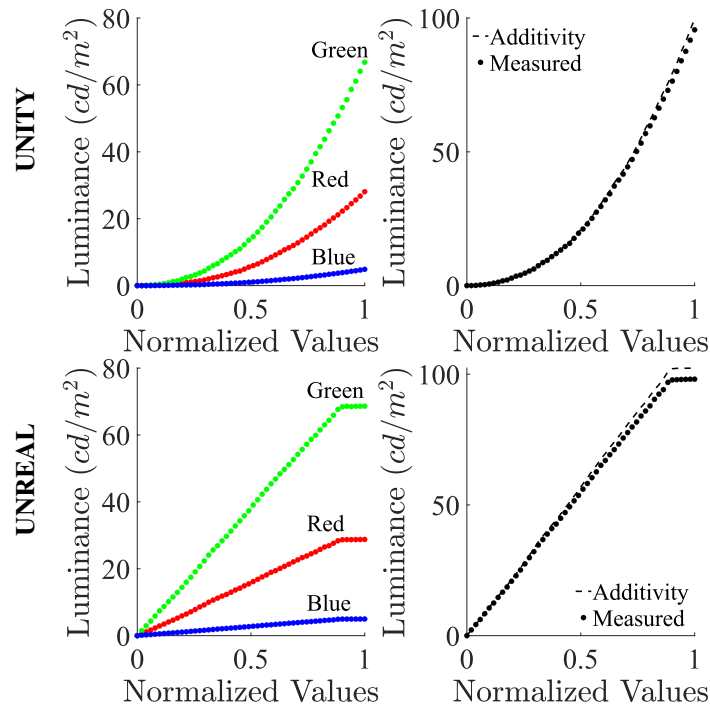


Fig. 8. Relationship between luminance and normalized intensity values in the HTC Vive Pro Eye for Unity and Unreal, using Standard material/shader. Red, green, and blue points indicate the R, G, and B channels, respectively, while black points depict the combined luminance of all three channels (achromatic). The solid lines represent the luminance predicted by the fitted function.

In Unreal, both unlit and standard shaders/materials exhibit a relationship between input intensity and luminance describable by

$$L = \begin{cases} m \cdot b + k, & m \cdot b + k < c, \\ c, & m \cdot b + k \geq c, \end{cases} \quad (3)$$

where L and b retain the units described above, while m , k , and c are constants representing the slope, offset, and clipping value, respectively.

In Figs. 8, 9, and 10, black circles illustrate the relation between measured luminance and reflectance for achromatic values for Standard material/shader. The dashed lines represent the cumulative luminance of the R, G, and B channels for each reflectance level. Across all evaluated conditions, the measured luminance values align closely with the sum of the individual R, G, and B luminance values. However, the display demonstrates sub-additive behavior in the HTC Vive Pro Eye, while the Pimax 8K X DMAS and Varjo Aero exhibit super-additive characteristics. Please refer to Figs. S1, S2, and S3 in the Supplement 1 for Unlit material/shader.

Regarding the HTC Vive Pro Eye using Unity, the luminance behavior parallels that of a standard monitor, as characterized by Eq. (2). Notably, a subadditivity of -2.95% is observed, indicating that the luminance in achromatic measurements is lower than the cumulative luminance of the isolated primaries. In contrast, when using Unreal, the display initially follows a linear pattern at lower intensities. Upon reaching saturation (around 50% intensity for Unlit and 90% for Standard material), it transitions to a constant output as described by Eq. (3).

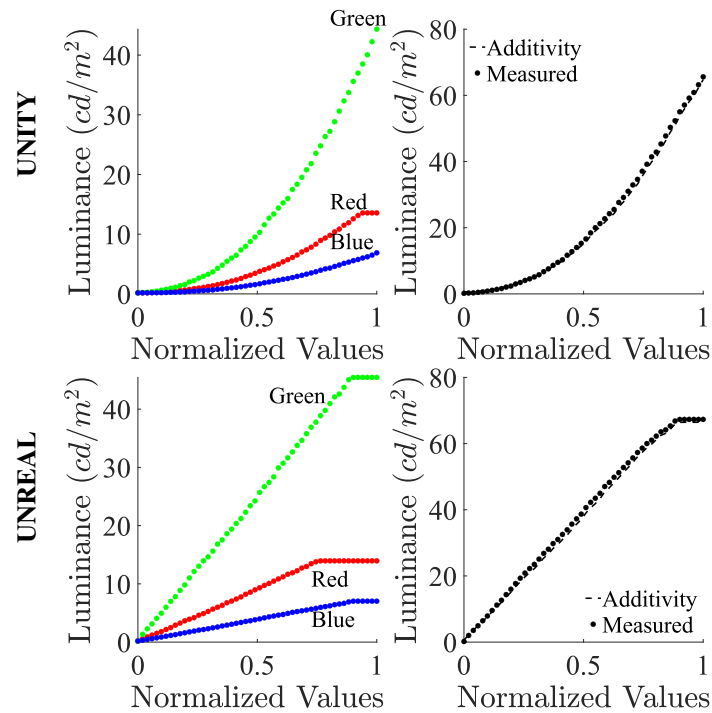


Fig. 9. Relationship between luminance and normalized intensity values in the Pimax 8K X DMAS for Unity and Unreal, using Standard material/shader. Red, green, and blue points indicate the R, G, and B channels, respectively, while black points depict the combined luminance of all three channels (achromatic). The solid lines represent the luminance predicted by the fitted function.

The luminance performance of Pimax is akin to that of the HTC Vive Pro Eye, albeit with a lower maximum luminance. In Unity, saturation was noted in the red channel, suggesting a luminance limit for this channel. Differing from HTC, Pimax shows superadditivity of +1.18%, where the achromatic luminance exceeds the sum of the luminances of the isolated primaries.

The Varjo Aero exhibits the highest luminance values across the R, G, B channels and when all channels are combined, demonstrating a superadditivity of +2.87%.

6.3. Primaries chromaticity

When measuring isolated primaries, we observed a variation in chroma with intensity that was particularly pronounced for Unity. At lower intensities, primaries tend to converge towards the achromatic axis. This trend is consistent across all headsets, with the least marked effect seen in the HTC Vive Pro, as shown in Fig. 11. The effect is more notable in Pimax 8K X DMAS and Varjo Aero when using Unity (refer to Figs. 12 and 13).

An obvious explanation for the decrease of chromaticity at low intensities is that the black point of the HMDs is not truly black. Adding a constant term to compensate for the black point did not improve the overall accuracy of our model. This indicates that the cause is not a constant backlight imposed on the output of the primaries. Furthermore, when operating with Unity, this variation of chromaticity is more pronounced for all HMDs. We interpret this as evidence that the game engine modifies the RGB input in a way that introduces a primary interdependence. As a result, higher chromaticity errors at lower intensities are anticipated, particularly when using Unity.

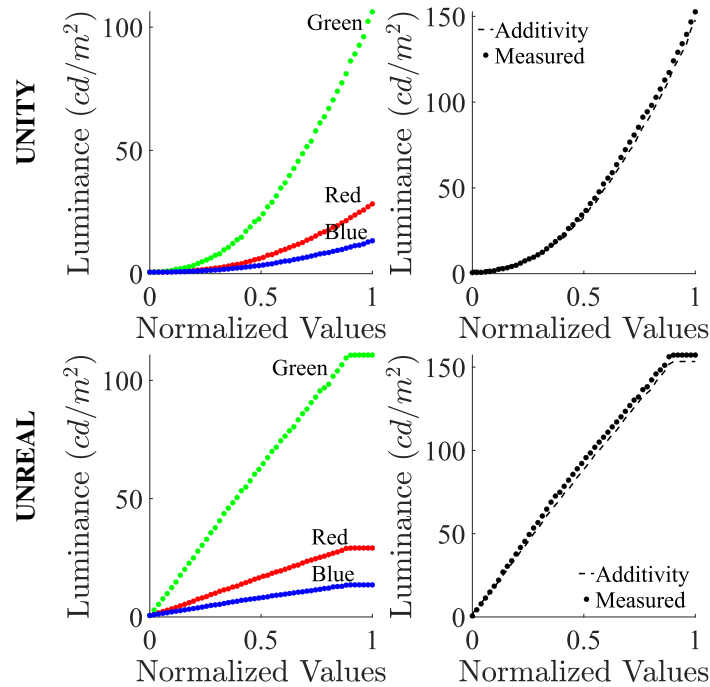


Fig. 10. Relationship between luminance and normalized intensity values in the Varjo Aero for Unity and Unreal, using Standard material/shader. Red, green, and blue points indicate the R, G, and B channels, respectively, while black points depict the combined luminance of all three channels (achromatic). The solid lines represent the luminance predicted by the fitted function.

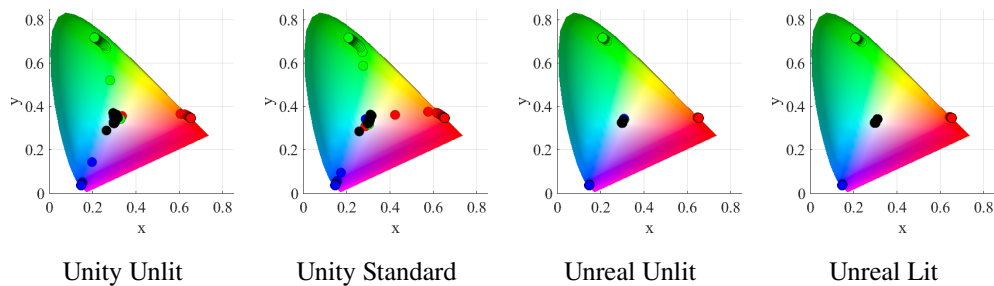


Fig. 11. The chromaticity of isolated (red, green, blue dots) and combined primaries (black dots) at different intensities for the HTC Vive Pro Eye.

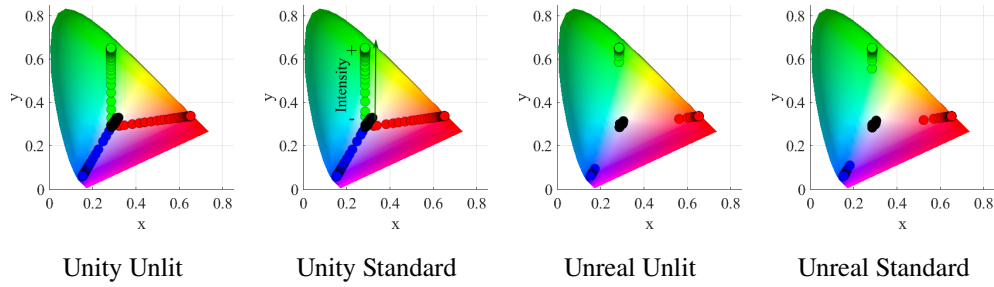


Fig. 12. The chromaticity of isolated (red, green, blue dots) and combined primaries (black dots) at different intensities for the Pimax 8K X DMAS.

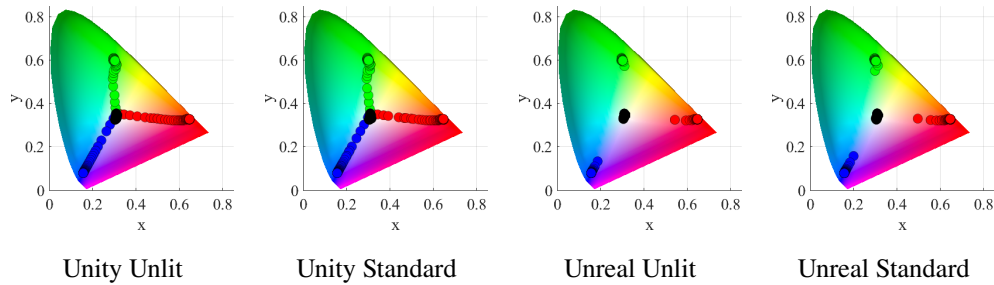


Fig. 13. The chromaticity of isolated (red, green, blue dots) and combined primaries (black dots) at different intensities for the Varjo Aero.

We cannot exclude the possibility that there is also an HMD internal mechanism that introduces primary interdependence, which is a common procedure on consumer displays to mitigate artifacts or enhance the visual experience.

6.4. Gamut and white points

The HTC Vive Pro Eye is capable of displaying the widest range of colors, as depicted in Fig. 14. In terms of the color gamut ratio compared to the sRGB color space, the HTC Vive Pro Eye achieves 145.20%, followed by the Pimax 8K X DMAS at 114.89%, and the Varjo Aero at 97.09% (for the configuration Unreal Lit). Please note that the gamut and the white point for a specific HMD remain consistent across different configurations. Table 2 presents the specific white points for each configuration.

Table 2. (x, y, Y) coordinates of the measured white points in each configuration.

		HTC Vive Pro Eye			Varjo Aero			Pimax 8K X DMAS		
		x	y	Y	x	y	Y	x	y	Y
Unity	Unlit	0.298	0.323	96	0.314	0.346	156	0.284	0.296	66
	Standard	0.299	0.323	96	0.313	0.346	153	0.285	0.296	66
Unreal	Unlit	0.299	0.325	99	0.314	0.347	159	0.289	0.301	68
	Lit	0.299	0.324	98	0.313	0.346	157	0.287	0.298	67

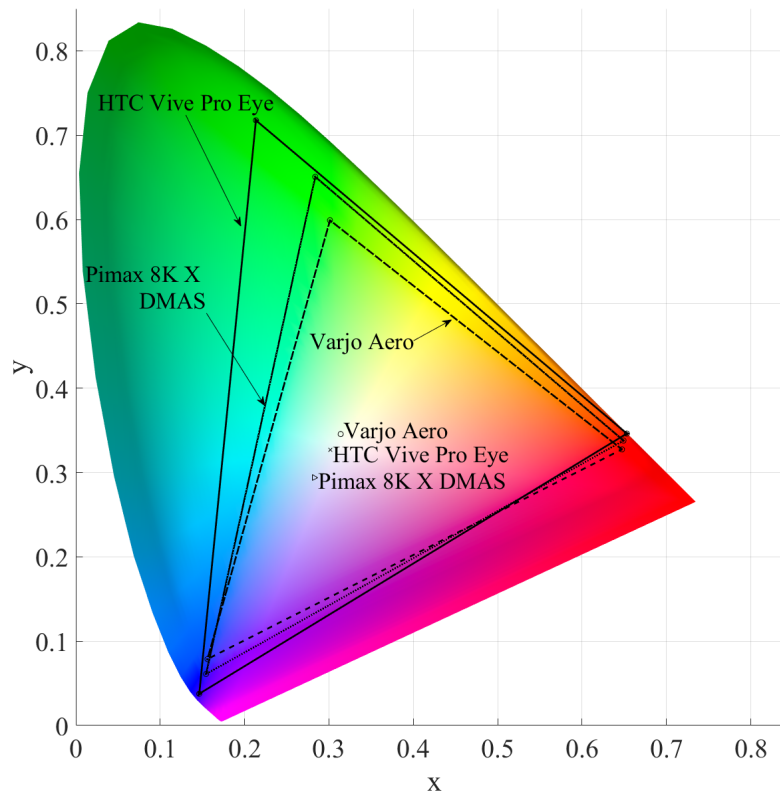


Fig. 14. Comparison of each display's color gamut for the configuration Unreal Lit with its reference white (located at the diagram's center).

6.5. Validation of the characterization model

For model validation, we compared measured and predicted xyY values corresponding to 1) a unique set of RGB values across all configurations (RGB validation set: same RGB, different xyY), and 2) using configuration-specific RGB values, each tailored to yield the same predicted xyY values (LCH validation set: different RGB, same xyY).

We use the colorimetric error ΔE_{00} [38] for quantitative results. This computation requires a white point for the transformation to $L^*a^*b^*$ color space. In our results, we selected the specific white point for each configuration.

6.5.1. RGB validation set

To validate the calibration across both game engines and materials, we assessed 100 random RGB values, as illustrated in Fig. 15 (first row). These values are the same for all different configurations. The deviation from the predictions of the calibration model remained below 1 ΔE_{00} , in terms of the mean, for the HTC Vive Pro Eye and Pimax 8K X DMAS, a discrepancy that is typically imperceptible to the human eye. For the Varjo Aero, this deviation was maintained below 1.5 ΔE_{00} , as depicted in Fig. 16. Please refer to Tables 3, 4, and 5 for more detailed statistics before and after considering the color characterization model.

6.5.2. LCH validation set

In addition to the standard validation procedures previously described, we further assessed the model's accuracy by examining its performance with a uniform set of xyY outputs across all

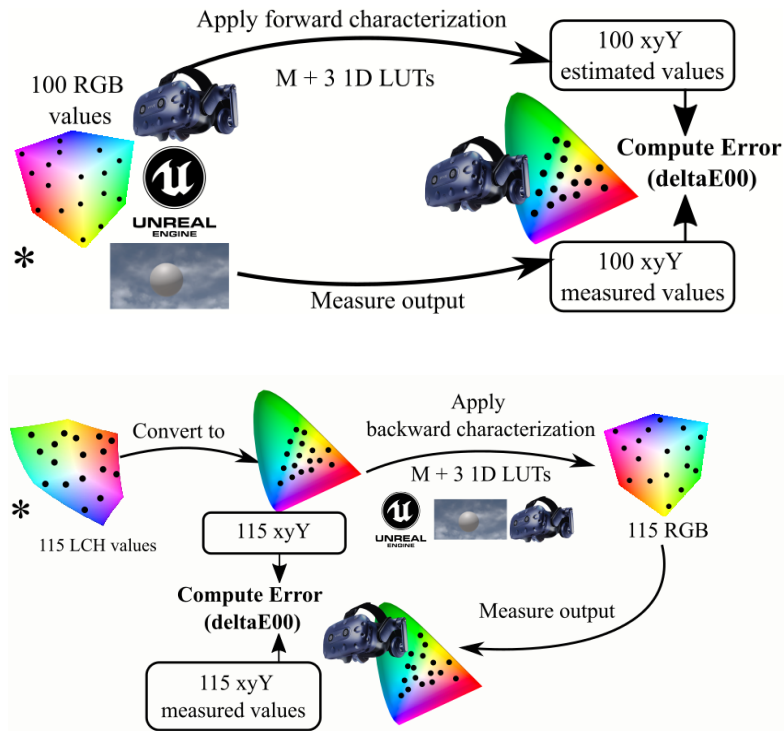


Fig. 15. Both validation sets for the characterization model, with (*) denoting the starting point.

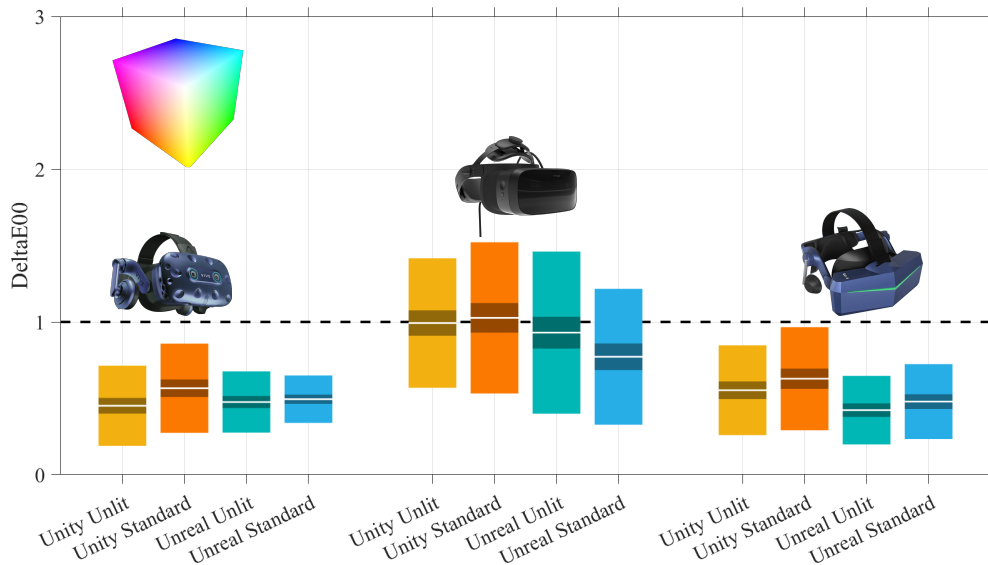


Fig. 16. ΔE_{00} error across all configurations and HMDs for the RGB validation set, from left to right: HTC Vive Pro Eye, Varjo Aero, and Pimax 8K X DMAS. In each bar, the larger outer box represents the standard deviation of the measurements, while the shaded inner box indicates the 95% confidence interval for the mean. The central white line denotes the mean value of the measurements.

Table 3. Comparison of deltaE00 statistics with and without utilizing estimated color characterization for HTC Vive Pro Eye headset on the RGB validation set.

		HTC Vive Pro Eye									
		Without characterization					After characterization				
		Mean	Med	Std	Min	Max	Mean	Med	Std	Min	Max
Unity	Unlit	17.15	15.72	5.48	7.31	30.09	0.45	0.40	0.26	0.08	2.31
	Standard	17.13	15.74	5.48	7.14	30.07	0.57	0.54	0.29	0.11	2.28
Unreal	Unlit	13.96	14.11	4.23	3.60	23.94	0.48	0.49	0.20	0.02	0.98
	Lit	5.99	5.94	1.72	1.89	10.62	0.49	0.48	0.16	0.12	1.03

Table 4. Comparison of deltaE00 statistics with and without utilizing estimated color characterization for Varjo Aero headset on the RGB validation set.

		Varjo Aero									
		Without characterization					After characterization				
		Mean	Med	Std	Min	Max	Mean	Med	Std	Min	Max
Unity	Unlit	13.96	12.94	6.62	3.75	27.67	0.99	0.89	0.42	0.24	2.64
	Standard	13.91	12.90	6.63	3.77	27.40	1.03	0.93	0.49	0.16	2.51
Unreal	Unlit	13.47	13.98	3.91	2.70	23.46	0.93	0.84	0.53	0.22	3.04
	Lit	3.88	3.89	0.91	1.52	6.23	0.77	0.73	0.44	0.11	2.44

Table 5. Comparison of deltaE00 statistics with and without utilizing estimated color characterization for Pimax 8K X DMAS headset on the RGB validation set.

		Pimax 8K X DMAS									
		Without characterization					After characterization				
		Mean	Med	Std	Min	Max	Mean	Med	Std	Min	Max
Unity	Unlit	13.72	12.63	5.52	3.11	25.89	0.55	0.52	0.29	0.09	1.50
	Standard	13.88	12.72	5.55	3.08	26.17	0.63	0.59	0.33	0.10	1.71
Unreal	Unlit	13.92	14.62	4.34	2.95	23.80	0.42	0.39	0.22	0.03	1.10
	Lit	5.19	4.56	1.96	1.53	9.48	0.48	0.45	0.25	0.04	1.24

configurations. For this purpose, 115 points were selected within the LCH color space, ensuring their placement within the common gamut of all HMDs and configurations. These LCH values were then transformed into the XYZ color space using a common white point, specifically the HTC Vive Pro Eye under the Unreal Engine Unlit configuration, resulting in a single set of xyY values for all HMDs and configurations. Subsequently, we calculated the specific RGB values for each configuration using the corresponding characterization model, measured the headset's output, and compared these measurements with the predetermined xyY values, as depicted in Fig. 15 (second row). This method of validation is particularly relevant as it mirrors the practical application of the model, which is to produce specific colors defined within a non-RGB color space.

Figure 17 illustrates the chromaticity error associated with the 115 predefined xyY values. Generally, the chromaticity errors observed here are greater than those noted in the characterization error section (Section 6.5). This increase in error can be attributed to the cumulative inaccuracies inherent in converting values from an LCH device-independent color space to an RGB device-dependent color space.

In line with the results observed in the RGB validation set (Section 6.5.1), both HTC and Pimax demonstrate the lowest chromaticity errors, with an average below 1 deltaE00 unit. The

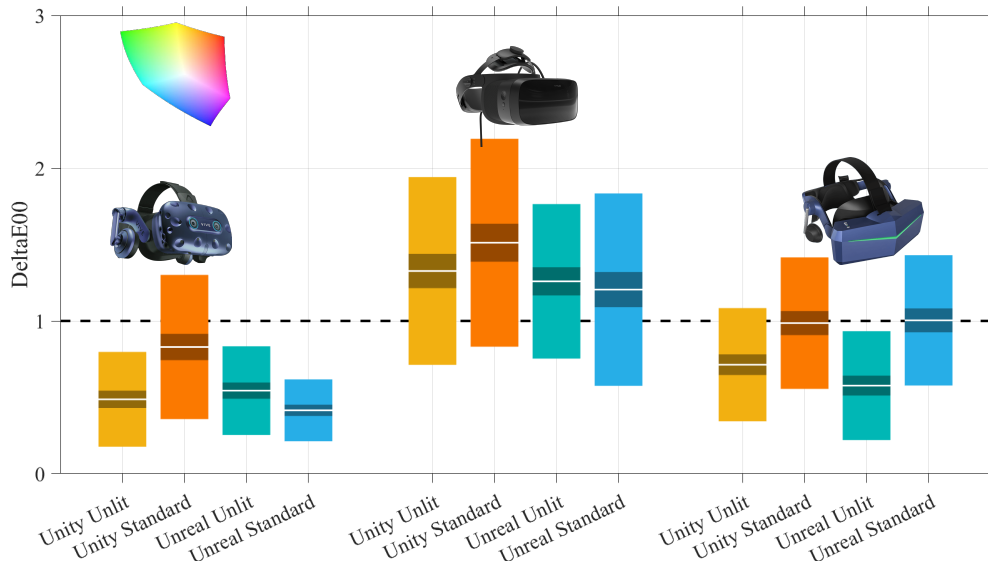


Fig. 17. DeltaE00 error for the xyY validation set across four engine configurations and three headsets: HTC Vive Pro Eye, Varjo Aero, and Pimax 8K X DMAS. The bar plots are the same as in Fig. 16.

Varjo, meanwhile, shows chromaticity errors remaining under 1.5 units in deltaE00. Please refer to Table 6 for more detailed statistics

Table 6. DeltaE00 statistics on the LCH validation set for all VR headsets.

			Mean	Median	Std	Min	Max
HTC Vive Pro	Unity	Unlit	0.49	0.42	0.31	0.06	1.88
		Standard	0.83	0.68	0.47	0.21	2.66
	Unreal	Unlit	0.54	0.49	0.29	0.12	1.72
		Lit	0.42	0.38	0.20	0.09	0.94
Varjo Aero	Unity	Unlit	1.33	1.25	0.61	0.21	3.57
		Standard	1.51	1.39	0.68	0.35	3.65
	Unreal	Unlit	1.26	1.20	0.51	0.22	3.68
		Lit	1.21	1.13	0.63	0.13	3.55
Pimax 8K	Unity	Unlit	0.71	0.62	0.37	0.15	2.06
		Standard	0.99	0.91	0.43	0.37	2.20
	Unreal	Unlit	0.58	0.44	0.36	0.13	2.11
		Lit	1.00	0.91	0.43	0.24	2.39

6.6. Uniformity

The uniformity of the displays was assessed using the I29 imaging colorimeter. This instrument captures images at a resolution of 6576×4384 pixels, providing chromaticity (x, y) and luminance Y values for each pixel. The alignment of the colorimeter was conducted following the same procedure as for the spectroradiometer, see Fig. 4. In addition, we placed the colorimeter at the same distance from all display lenses ($\sim 5\text{mm}$).

Please note that each VR headset has an optimal position relative to the eye, which is not publicly disclosed. Viewing from this position can be considered as the upper limit for visual performance. However, our main interest lies in capturing what the user actually experiences. Therefore, any measurement within the range where individuals' eyes may be located is adequate for us, even if the uniformity results vary when measured from different positions. By manually adjusting the measurement devices to a location where the display's visual quality is judged to be satisfactory and by aligning it with the center of the alignment pattern, we essentially mimic the adjustment process a user would follow.

Using Unity, a uniformly white image covering the entire display was projected, akin to the calibration process. We then recorded the chromaticity and luminance values of this image. Prior to analysis, the image was resized to 10% of its original size, and a Gaussian filter was applied to eliminate any Moiré patterns. Subsequently, we employed two binary masks for further analysis:

1. *Region of Interest*: This mask excluded areas outside the display's boundaries.
2. *One-Degree Center*: A circular mask with a diameter of one degree was centered on the display, corresponding to the measurement aperture of the spectroradiometer. This central region served as the reference for subsequent analyses.

We separately examined the uniformity of luminance and chromaticity.

6.6.1. Luminance uniformity

As Fig. 18 shows, the average luminance at the display's center is 103.19 cd/m^2 for the HTC Vive Pro Eye, 172.72 cd/m^2 for Varjo, and 71.20 cd/m^2 for Pimax. According to our findings, within the entire image, the percentage of the area where luminance exceeds 50% relative to the central 1° region is 67% for the HTC Vive Pro Eye, 21% for the Varjo Aero, and 27% for the Pimax. This implies that the luminance drops more quickly in Varjo Aero and Pimax 8K X DMAS as compared to HTC Vive Pro Eye.

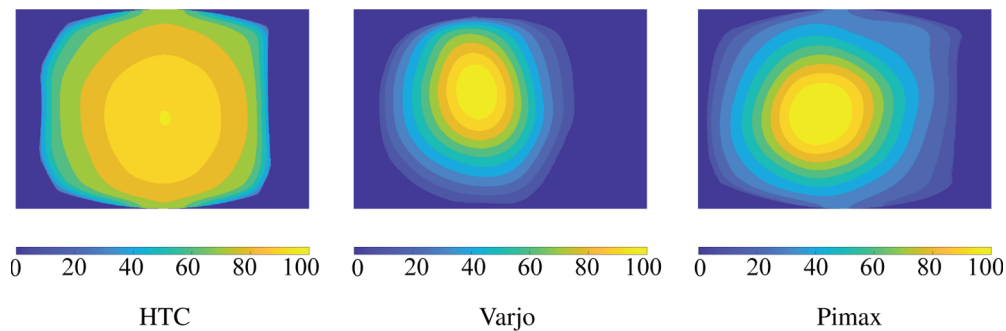


Fig. 18. Relative luminance measurements for white in HTC Vive Pro Eye, Varjo Aero, and Pimax 8K X DMAS. The scale represents the percentage with respect to the mean value from the center of the display spanning 1° .

6.6.2. Chromaticity uniformity

We calculated the average (x, y) chromaticity (computed in XYZ coordinates) in the one-degree center region to evaluate chromaticity variation across the display. We then determined the chromaticity error (ΔE_{00} , excluding luminance) for each pixel in the colorimeter image relative to the averaged center region. Figure 19 illustrates these chromaticity errors for each display. Unlike the luminance channel, the chromaticity error does not exhibit a consistent

increase or decrease relative to the display center but rather shows an irregular pattern of variation. The maximum error for all three HMDs is 10.5, 12.49 and 12.78 for HTC Vive Pro Eye, Varjo Aero and Pimax 8K X DMAS, respectively. See Table 7 for more detailed statistics.

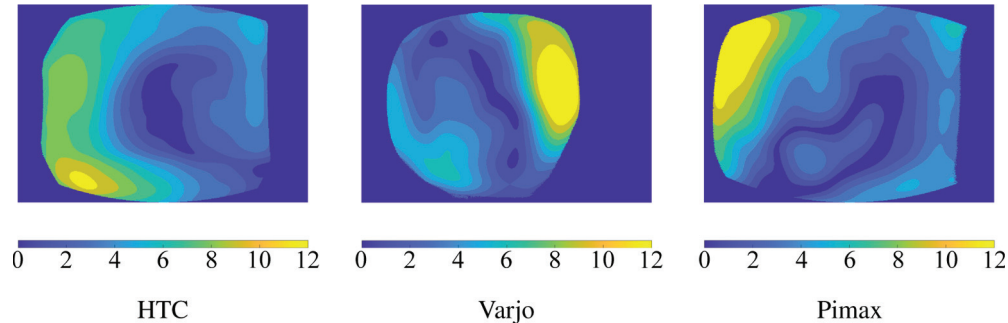


Fig. 19. Computed chromaticity error $\Delta E_{00}[a, b]$ based on the mean value in the CIE Lab color space from the center of the display spanning 1° .

Table 7. Statistics chromaticity error $\Delta E_{00}[a, b]$ based on the mean value in the CIE Lab color space from the center of the display spanning 1° . See Fig. 19 for chromaticity uniformity.

	Mean	Median	Std	Min	Max
HTC Vive Pro Eye	4.65	4.21	2.60	0.01	10.47
Varjo Aero	4.545	3.68	3.02	0.01	12.49
Pimax 8K X DMAS	4.03	3.52	2.83	0.01	12.78

7. Discussions

This study addresses the challenge of achieving accurate color reproduction in VR devices through a detailed color calibration process. Our research offers a methodical guide for precise color calibration across a variety of VR devices currently in use. To facilitate this, we have developed a consistent characterization approach applicable to both Unity and Unreal engines, which serves as a basis for comparative evaluations between these two graphics engines.

An important finding in our study is the effectiveness of the AMOLED panel in the HTC Vive Pro Eye, which showed high color reproduction accuracy and low chromaticity errors across all configurations. However, using a Fresnel lens in this model diminishes immersion in virtual environments as there is a stark contrast between the high-quality resolution in the center of the display and the low-resolution periphery. In contrast, the Varjo headset, despite a slightly higher error margin, provides a more immersive experience due to the use of aspheric lens design. The Pimax headset closely matches the HTC in color accuracy and additionally offers a wider field of view (FOV), making it suitable for specific applications. This highlights the need for more focused research to assess the factors influencing user immersion in VR environments directly.

Our findings indicate that combining an AMOLED panel with an aspheric lens could offer a promising solution for achieving authentic color reproduction in VR headsets.

In our assessment of display uniformity, we noted a significant reduction in luminance toward the periphery of the display (similar to conventional CRTs monitors), with decreases surpassing 50% in all tested headsets. Additionally, we documented variations in chromaticity across the FOV, leading to considerable increases in error outside the central measurement region. Our present model does not consider such discrepancies resulting from uniformity variations.

Although previous research indicates that fixation points in HMDs are predominantly biased centrally [39], where these errors are less pronounced, this aspect may pose challenges in scenarios where accurate color reproduction in peripheral areas is essential.

We would like to mention as well the potential value of psychophysical studies. In fact, the colorimetric characterization we describe serves as a necessary step for future projects studying color perception and appearance in VR. In a related study by Gil Rodriguez et al. [17], a color constancy experiment in VR was conducted, requiring a prior characterization process. Currently, we're engaged in several projects investigating different mechanisms of color constancy in VR.

Please note that previous works in color perception [40–42] feature color stimuli presented over the visual field and the periphery, and they consistently show that the peripheral representation of color is poorer than in the fovea. Our study focus on the foveal region (central 1° area) and deltaE00 is a metric design for this specific area. Evaluating peripheral viewing would require measurements of human color perception across the whole visual field, and comparing them to a colorimetric evaluation of the headset across the visual field.

8. Conclusions

The new insight enabled by our work is the extension and application of a traditional method for color characterization to currently available HMDs, responding to the growing relevance of VR technology across diverse fields. To facilitate wider accessibility and usage, we have made our tools and code publicly available and open source on GitHub (27). By following all the steps detailed in this work, new experiments can be reproduced in different HMDs or setup conditions such as materials or illumination.

Our approach, while straightforward and efficient, reduces color reproduction errors by 90% or more to levels below the threshold of human perceptibility. In the Unity engine, we observed the largest decrease in error, from up to 17 deltaE00 units to as low as 0.45 deltaE00. In the Unreal engine, the largest reductions in error were from around 14 deltaE00 units (Unlit) and 6 deltaE00 units (Standard) to below 0.5 deltaE00. These results are specific to the HTC Vive; with other HMDs, we achieved similar, albeit slightly smaller, error reductions. For all HMDs tested here, the decrease in error was less pronounced in the case of the Unreal Standard mode, which is attributed to Unreal's superior color reproduction management compared to Unity, even before applying a calibration. Nonetheless, this reduction signifies a substantial improvement.

While our method successfully attains color reproduction quality that surpasses human discrimination standards in the most critical area of the field of view (FOV), it does not consider the non-uniformities resulting from the optical design of contemporary VR HMDs. Addressing this limitation presents an opportunity for future research to refine and extend the calibration procedures outlined in our current study.

Funding. European Research Council (884116); Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) (222641018, SFB/TRR 135 Project C2); Alexander von Humboldt-Stiftung.

Acknowledgment. Alexander von Humboldt fellowship to author AA.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are available in Ref. [27].

Supplemental document. See [Supplement 1](#) for supporting content.

References

1. W. Huang and R. D. Roscoe, "Head-mounted display-based virtual reality systems in engineering education: A review of recent research," *Comp. Appli. In Engineering* **29**(5), 1420–1435 (2021).
2. H. Cwierz, F. Díaz-Barrancas, J. G. Llinás, *et al.*, "On the validity of virtual reality applications for professional use: A case study on color vision research and diagnosis," *IEEE Access* **9**, 138215–138224 (2021).
3. L. Li, F. Yu, D. Shi, *et al.*, "Application of virtual reality technology in clinical medicine," *American Journal of Translational Research* **9**, 3867 (2017).

4. L. Fan, H. Li, and M. Shi, "Redirected walking for exploring immersive virtual spaces with HMD: a comprehensive review and recent advances," *IEEE Trans. Visual. Comput. Graphics* **29**(10), 4104–4123 (2022).
5. M. Höll, M. Oberweger, C. Arth, *et al.*, "Efficient physics-based implementation for realistic hand-object interaction in virtual reality," in *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, (IEEE, 2018), pp. 175–182.
6. Y. M. Tang and H. L. Ho, "3D modeling and computer graphics in virtual reality," in *Mixed Reality and Three-Dimensional Computer Graphics*, (IntechOpen, 2020).
7. F. Diaz-Barrancas, H. Cwierz, P. Pardo, *et al.*, "Improvement of realism sensation in virtual reality scenes applying spectral and colour management techniques," in *Proceedings of the 25th Symposium of the International Colour Vision Society (ICVS 2019)*, Riga, Latvia, (2019), p. 79.
8. P. J. Pardo, M. I. Suero, and Á. L. Pérez, "Correlation between perception of color, shadows, and surface textures and the realism of a scene in virtual reality," *J. Opt. Soc. Am. A* **35**(4), B130–B135 (2018).
9. A. S. Kim, W.-C. Cheng, R. Beams, *et al.*, "Color rendering in medical extended-reality applications," *J. Digit. Imaging* **34**(1), 16–26 (2021).
10. D. H. Brainard, D. G. Pelli, and T. Robson, *Display Characterization* (American Cancer Society, 2002).
11. Unity, Unity, (2022).
12. Unreal, Unreal Engine, (2022).
13. M. I. Suero, P. J. Pardo, and A. Pérez, "Colour characterization of handheld game console displays," *Displays* **31**(4-5), 205–209 (2010).
14. U. Bhaumik, F. B. Leloup, and K. Smet, "Systematic comparison of head mounted display colorimetric performance using various color characterization models," *Optics Continuum* **2**(6), 1490–1504 (2023).
15. D. Gadia, C. Bonanomi, M. Rossi, *et al.*, "Color management and color perception issues in a virtual reality theater," in *Stereoscopic Displays and Applications XIX*, vol. 6803 (SPIE, 2008), pp. 269–280.
16. F. Díaz-Barrancas, H. Cwierz, P. J. Pardo, *et al.*, "Spectral color management in virtual reality scenes," *Sensors* **20**(19), 5658 (2020).
17. R. Gil Rodríguez, F. Bayer, M. Toscani, *et al.*, "Colour calibration of a head mounted display for colour vision research using virtual reality," *SN Computer Science* **3**(1), 1–10 (2022).
18. A. Bellazzi, L. Bellia, G. Chinazzo, *et al.*, "Virtual reality for assessing visual quality and lighting perception: A systematic review," *Build. Sci.* **209**, 108674 (2021).
19. T. Pouli, P. Morvan, S. Thiebaud, *et al.*, *Color Management for VR Production*, (Association for Computing Machinery, New York, NY, USA, 2018), VRIC '18.
20. R. F. Murray, K. Y. Patel, and E. S. Wiedenmann, "Luminance calibration of virtual reality displays in unity," *Journal of Vision* **22**(13), 1 (2022).
21. N. Zaman, P. Sarker, and A. Tavakkoli, "Calibration of head mounted displays for vision research with virtual reality," *Journal of Vision* **23**(6), 7 (2023).
22. W.-C. Cheng, "82-4: Color characterization of virtual reality devices using professional-and consumer-grade instruments," in *SID Symposium Digest of Technical Papers*, vol. 54 (Wiley Online Library, 2023), pp. 1158–1161.
23. R. S. Berns, "Methods for characterizing CRT displays," *Displays* **16**(4), 173–182 (1996).
24. T. Johnson, "Methods for characterizing colour printers," *Displays* **16**(4), 193–202 (1996).
25. M. Xu and H. Hua, "Systematic method for modeling and characterizing multilayer light field displays," *Opt. Express* **28**(2), 1014–1036 (2020).
26. R. Salama and M. Elsayed, "A live comparison between unity and unreal game engines," *Global Journal of Information Technology: Emerging Technologies* **11**(01), 1–7 (2021).
27. GitHub repository, <https://github.com/rkl-gilro/ColorCharacterization>, (2024).
28. Konica Minolta, Spectroradiometer CS-2000, (2022).
29. J. Penczek, R. L. Austin, S. Obheroi, *et al.*, "54-5: Measuring direct retinal projection displays," in *SID Symposium Digest of Technical Papers*, vol. 51 (Wiley Online Library, 2020), pp. 807–810.
30. H. Hong, "A measurement method of the focal distance of the virtual image in an augmented reality or virtual reality device," *J. Soc. Inf. Disp.* **29**(4), 230–236 (2021).
31. R. S. Draper, J. Penczek, R. Varshneya, *et al.*, "72-2: Standardizing fundamental criteria for near eye display optical measurements: determining eye point position," in *SID Symposium Digest of Technical Papers*, vol. 49 (Wiley Online Library, 2018), pp. 961–964.
32. R. Varshneya, R. S. Draper, J. Penczek, *et al.*, "50-4: Standardizing fundamental criteria for near eye display optical measurements: determining the eye-box," in *SID Symposium Digest of Technical Papers*, vol. 51 (Wiley Online Library, 2020), pp. 742–745.
33. Radiant Vision Systems, Radiant ProMetric I, (2020).
34. Radiant Vision Systems, Radiant AR/VR Lens, (2020).
35. A. Aldandarawy, "[Unity]Always Be Linear:Shader-Based Gamma Correction," (2020). <https://medium.com/@abdulla.aldandarawy/unity-always-be-linear-1a30db4765db> [Accessed: December 2023].
36. M. Fairchild and D. Wyble, "Colorimetric characterization of the apple studio display (flat panel LCD)," *Munsell Color Science Laboratory Technical Report* (1998).

37. F. Díaz-Barrancas, R. Gil-Rodríguez, A. Aizenman, *et al.*, “Color calibration in virtual reality for unity and unreal,” in *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, (2023), pp. 733–734.
38. G. Sharma, W. Wu, and E. N. Dalal, “The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations,” *Color Res. Appl.* **30**(1), 21–30 (2004).
39. A. M. Aizenman, G. A. Koulieris, A. Gibaldi, *et al.*, “The statistics of eye movements and binocular disparities during VR gaming: implications for headset design,” *ACM Trans. Graph.* **42**(1), 1–15 (2023).
40. M. A. Cohen, T. L. Botch, and C. E. Robertson, “The limits of color awareness during active, real-world vision,” *Proc. Natl. Acad. Sci.* **117**(24), 13821–13827 (2020).
41. B. Duinkharjav, K. Chen, A. Tyagi, *et al.*, “Color-perception-guided display power reduction for virtual reality,” *ACM Trans. Graph.* **41**(6), 1–16 (2022).
42. T. Hansen, L. Pracejus, and K. R. Gegenfurtner, “Color perception in the intermediate periphery of the visual field,” *Journal of Vision* **9**(4), 26 (2009).