

# Realistic Luminance in VR

Nathan Matsuda\*

Meta  
USA

Alexandre Chapiro\*

Meta  
USA

Yang Zhao

Meta  
USA

Clinton Smith

Meta  
USA

Romain Bachy

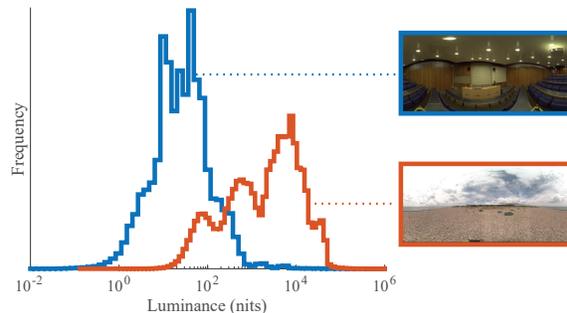
Meta  
USA

Douglas Lanman

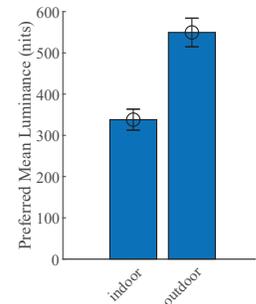
Meta  
USA



(a) Experimental Hardware



(b) Computed Real-World Luminances



(c) Preference Results

**Figure 1:** (a) Our high-dynamic-range VR prototype. (b) Recovered absolute luminance histograms of the SYN dataset. Outdoor scenes, in orange, are about two orders of magnitude brighter than indoor ones, in blue. Insets show sample scenes from the dataset. (c) Results of our perceptual study (Section 5.3). Users preferred significantly higher luminance values for outdoor scenes than for indoor ones.

## ABSTRACT

As virtual reality (VR) headsets continue to achieve ever more immersive visuals along the axes of resolution, field of view, focal cues, distortion mitigation, and so on, the luminance and dynamic range of these devices falls far short of widely available consumer televisions. While work remains to be done on the display architecture side, power and weight limitations in head-mounted displays pose a challenge for designs aiming for high luminance. In this paper, we seek to gain a basic understanding of VR user preferences for display luminance values in relation to known, real-world luminances for immersive, natural scenes. To do so, we analyze the luminance characteristics of an existing high-dynamic-range (HDR) panoramic image dataset, build an HDR VR headset capable of reproducing over 20,000 nits peak luminance, and conduct a first-of-its-kind study on user brightness preferences in VR. We conclude that current commercial VR headsets do not meet user preferences for display luminance, even for indoor scenes.

\*Denotes equal contribution.

Permission to make digital or hard copies of part or all 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 third-party components of this work must be honored. For all other uses, contact the owner/author(s).

SA '22 Conference Papers, December 6–9, 2022, Daegu, Republic of Korea

© 2022 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-9470-3/22/12.

<https://doi.org/10.1145/3550469.3555427>

## CCS CONCEPTS

• Computing methodologies → Perception.

## KEYWORDS

perception, virtual reality, high-dynamic-range

## ACM Reference Format:

Nathan Matsuda\*, Alexandre Chapiro\*, Yang Zhao, Clinton Smith, Romain Bachy, and Douglas Lanman. 2022. Realistic Luminance in VR. In *SIGGRAPH Asia 2022 Conference Papers (SA '22 Conference Papers)*, December 6–9, 2022, Daegu, Republic of Korea. ACM, New York, NY, USA, 8 pages. <https://doi.org/10.1145/3550469.3555427>

## 1 INTRODUCTION

The human visual system operating in natural conditions can resolve luminance values that range from over a million candelas per meter squared (nits) to near zero, and is able to simultaneously resolve over four orders of magnitude without adaptation [Kunkel and Reinhard 2010]. While most traditional displays can only replicate a fraction of the smaller simultaneous range, high-dynamic-range (HDR) displays aim to support luminance and contrast ranges closer to perceptual limits [Reinhard et al. 2010]. HDR displays have achieved widespread commercial success across cinema, home theater, and personal-use devices, but remain out of reach in the context of VR displays, which are typically limited to peak luminance values below 200 nits [Mehrfard et al. 2019]. The perceptual impact of HDR on this unique mix of controlled ambient illumination conditions in an enclosed headset, wide field-of-view (FOV), viewing

optics, and immersive presentation typical of VR displays is largely unexplored.

We begin to address this gap in understanding by analyzing an existing dataset of high-dynamic-range 360-degree photographs to identify trends in real-world luminance values for immersive scenes, creating a set of stimuli used to determine user preferences for VR headset luminance, relative to natural real-world values. To assess these preferences, we present a high-dynamic-range virtual reality demonstrator with a display system comprised entirely of off-the-shelf parts, capable of peak luminances over 20,000 nits. We achieve this without reducing the resolution or simultaneous contrast relative to commercially available VR headsets. Furthermore, stereo displays with high luminance used to study perceptual effects have typically been too bulky to be head mounted. Our head-tracked prototype has the potential to achieve a higher degree of perceptual realism than existing direct-view devices like HDR televisions, stereoscopes, and other high-luminance prototypes.

Finally, we conduct a user study to gain insight into display luminance requirements in a VR setting. Our study demonstrates that subjective brightness preference far exceeds what is available in today's headsets, and is affected by higher-order content-dependent factors, such as whether images represent indoor or outdoor scenes.

Our primary contributions are:

- A database of calibrated images that allow us to study real-world luminance distributions.
- A head-tracked VR prototype capable of producing 20,940 nits over a 62-degree field of view.
- A subjective study on brightness preferences for perceived realism in VR. Our results demonstrate that viewers have intrinsic preferences for brightness, which may not be satisfied by commercially available displays.

## 2 RELATED WORK

**Real-World Luminance Statistics:** Capture of high dynamic range photographs, e.g., by using exposure bracketing, is well established [Debevec and Malik 2008; Froehlich et al. 2014; Mann and Picard 1995], resulting in widespread use in research and in the entertainment industry. Beyond artistic uses, these techniques may also be employed for the photometric evaluation of real-world luminances [Pierson et al. 2021]. In fact, before we can explore what luminance ranges users prefer in a VR headset, we need to understand the statistics of real-world luminances over wide fields of view and dynamic ranges. The vast majority of image content available in research datasets has been captured with conventional cameras — that is, the FOVs and dynamic ranges are limited. The SYNS dataset [Adams et al. 2016] offers 360-degree, HDR images of indoor and outdoor environments captured with a Spheron SceneCam. The entries in this dataset provide an accurate linear recording of scene luminances over a large dynamic range (26 EV). While the luminance values are correct up to a scale factor, metadata available for each image from the dataset, such as time exposure, ISO sensitivity and aperture, in conjunction with laboratory measurements we conducted with the same camera, allow us to calibrate the dataset content to recover physical luminance values (see Section 3.1 for more detail).

**High Luminance Displays and User Preference:** Computational display research over the past two decades sought to increase peak luminance and contrast reproducible by digital displays beyond the limits of the underlying components. Seetzen et al. [2003] overcame the low contrast and bit depth of LCD-based digital displays by cascading a projector and transmissive LCD panel to effectively double the contrast and bit depth. Subsequent work with this prototype explored users' preferred contrast for a given luminance [Seetzen et al. 2006]. Such display pipelines also allow to further explore appearance preferences — for example, Radonjic et al. [2011] evaluated perceived contrast over a large range of luminances. These display configurations increase the dynamic range by increasing overall contrast, but absolute luminance is independently controlled by the backlight (or projector lamp) power.

Daly and colleagues [2013] determined that when evaluating the reproduction of specular highlights on an HDR display, a peak luminance of  $4,000 \text{ cd/m}^2$  was needed on average to satisfy 90% of users. The resulting guidelines for direct-view display preference [Dolby 2016], however, do not necessarily translate to immersive head-mounted displays due to the differences in viewing conditions.

**High Luminance Virtual Reality and User Preference:** Due to the lack of high luminance VR headsets, which rarely exceed 200 nits [Mehrfard et al. 2019], much prior work on HDR in VR looks at the viability of tonemapping to improve perceptual fidelity and immersion [Chamilothori et al. 2019; Luidolt et al. 2020; Najaf-Zadeh et al. 2017; Proulx 2020; Regalbutto 2019; Rockcastle et al. 2021], rather than studying the effects of headset luminance directly.

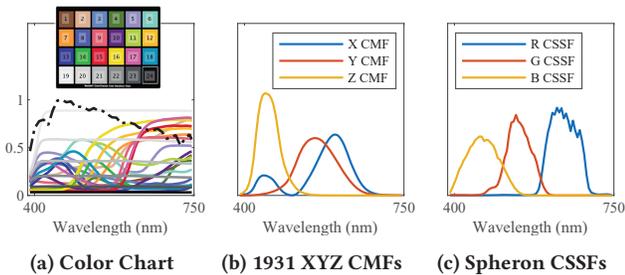
However, well before the current wave of VR research, Ledda et al. [2004] built a binocular HDR display and performed a series of detection tests with a focus on determining the effect of HDR displays on the visibility of peripheral stimuli. Recently, Zhong et al. [2021] built a multifocal stereoscope capable of reproducing HDR content. While not strictly a VR headset, and not specifically targetting users' preferences with respect to dynamic range, this work does get at the underlying question of realism that we aim to explore here. In both of these cases the testbeds were capable of producing binocular stimuli without head tracking. Our aim is to begin to understand the hardware requirements for immersive (i.e., head-tracked, 360-degree) content to feel real to users, beginning at a basic level: the peak luminance of the display.

## 3 REAL-WORLD LUMINANCE

To begin discussing realistic luminance values in the context of VR, we first need to understand real-world luminance statistics. This will help us contrast what is reproducible on a display with reality, and enable studies to explore whether being able to faithfully reproduce absolute luminance is necessary in order to provide users with a realistic experience.

### 3.1 Luminance Estimation

Measuring luminance from a scene directly can be challenging. A spectrometer or luminance meter can provide this measurement, but they are generally slow and most perform only spot measurements (as opposed to capturing an entire scene). It is desirable to estimate the luminance of a scene using commercially available



**Figure 2: (a) Reflectances of a Macbeth color chart, shown above, and an arbitrarily scaled D65 sunlight illuminant spectrum (dotted line). (b) CIE 1931 2-deg. XYZ color matching functions. (c) The RGB spectral responses of a Spheron HDR camera used to capture the SYNS dataset [Adams et al. 2016], measured using a monochromator.**

RGB cameras due to their ease of operation, but this leaves two questions: (1) how can we convert RGB values to luminance? (2) how precise are these estimates?

Beginning by assuming an ideal camera and well-exposed, linear RGB pixels (not over- or under-exposed – for instance, images taken using carefully set bracketed exposures), light reflected from a real-world object can be calculated based on its reflectance  $O(\lambda)$  and illuminant  $S(\lambda)$  as the product of the two:  $S(\lambda) * O(\lambda)$  [Fairchild 2013]. Sample reflectances of patches on a Macbeth color chart, and standard D65 sunlight spectral power distribution (SPD)<sup>1</sup> are shown in Figure 2a.

To calculate the perceived color of the object, we also need to model the human visual system’s (HVS) response. This is done using a set of standard trichromatic color matching functions (CMFs<sup>2</sup>, see Figure 2b), which can be defined in the XYZ color space. In this context, Y represents luminance, and is often defined in a standard 0-1 range in units relative to a reference white, but (like linear RGB values) remain proportional to the physical luminance in nits – the quantity we want to determine from an RGB image [Wyszecki and Stiles 1982]. A camera’s spectral sensitivity functions (CSSFs) can be analogously characterized for R, G, and B channels, as shown in Figure 2c.

The perceived XYZ color coordinates (e.g., luminance Y), are derived as:

$$Y = \int_{\lambda} S(\lambda) O(\lambda) \text{CMF}_Y(\lambda) \quad (1)$$

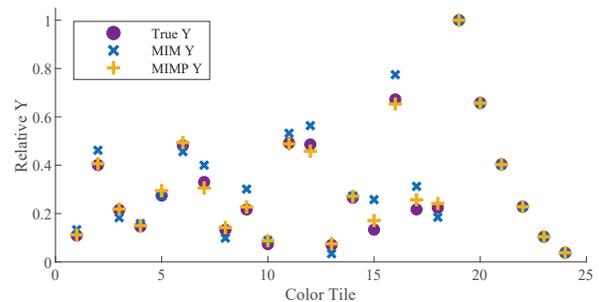
Similarly, the linear RGB response is derived from the CSSFs as, e.g.:

$$R = \int_{\lambda} S(\lambda) O(\lambda) \text{CSSF}_R(\lambda) \quad (2)$$

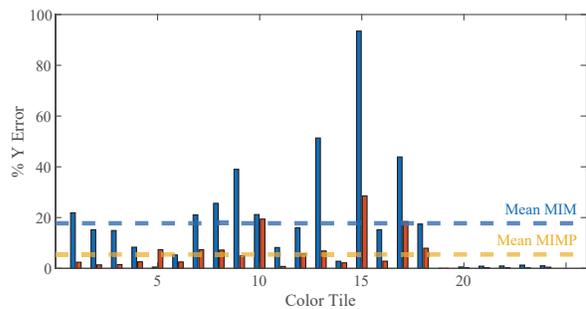
Computationally, these functions are sampled as N-long arrays at some wavelength intervals, and the integral turns to a summation. For simplicity, we can represent the CMFs as an  $N \times 3$  array A and

<sup>1</sup>Color checker and D65 distribution from <https://www.rit.edu/science/munsell-color-science-lab-educational-resources>

<sup>2</sup>More information on CMFs can be found at <http://www.cvrl.org/>



**(a) Luminance Estimation Comparison**



**(b) Luminance Estimation Relative Error**

**Figure 3: (a) Scaled comparison between ground truth  $Y$  and  $Y_{\text{RGB}}$  obtained using MIM and MIMP. (b) Percentage error per sample. Notably, the gray-scale patches at the bottom of the color chart show the best accuracy. Patch 15 is a saturated red color. The average error is approximately 17.7% for MIM and 5.5% for MIMP. X-axes are patches from the color chart in Figure 2, sorted top-left to bottom-right.**

the RGB CSSFs in an  $N \times 3$  array B. The conversion between RGB and XYZ color spaces is done through a  $3 \times 3$  matrix T, whose values are determined based on the target responses (CMFs and CSSFs):

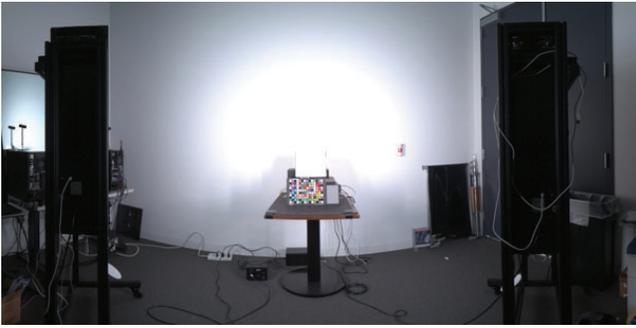
$$B \approx A \times T \quad (3)$$

A simple (but not optimal) approximation for T can be obtained by finding a least-squares solution to the linear system above (termed the “maximum ignorance method” (MIM) [Fang et al. 2017; Finlayson and Drew 1996]. This result can be further improved through the use of a support  $N \times N$  matrix L which enforces physically-plausible non-negative values for radiant energy (termed maximum ignorance method with positivity, or MIMP).

We use the data shown in Figure 2 to run a synthetic experiment, comparing the ground-truth Y obtained from the CMF as in Eq. (1) to an estimated  $Y_{\text{RGB}}$ , obtained by applying the matrices T obtained using MIM and MIMP to RGB values drawn using the CSSFs as in Eq. (2). A comparison between Y and  $Y_{\text{RGB}}$  for the Macbeth chart SPDs from Figure 2a illuminated by D65 is shown in Figure 3.

## 3.2 Pipeline Validation

To test the MIMP computational pipeline presented in Section 3 in practice, we must address three salient points:



**Figure 4: Our lab capture setting using controlled illumination and standard reflectance targets.**

- (1) Obtain linear RGB values, which typically requires recording camera parameters: exposure, ISO and aperture, and disabling or inverting gamma encoding and white-balance
- (2) Ensure linear RGB pixel values are reliable and not under- or over-exposed. Sensor noise and optics must not significantly impact the result.
- (3) To obtain a result in an absolute scale, at least one known patch must be measured (e.g., a standard reflectance patch with known value).

*Hardware.* We acquired a Spheron HDR camera, which is able to take high-resolution 360° images and has highly controllable software allowing for fine control of capture parameters like exposure, ISO, aperture, and white balance. These metadata parameters are output in a text file format and loaded into our software suite.

*Measurement.* We used our camera to acquire captures of a static scene in a lab setting containing a real-world MacBeth chart, as shown in Figure 4. In addition to the images and metadata, a Konica Minolta CS-160 chroma-meter was employed to measure real luminance values of several neutral reference surfaces under a large range of luminance, giving us the opportunity to test the algorithm across a wide dynamic range. We use these ground-truth measurements to compare against the performance of our RGB-to-luminance conversion using the MIMP method described in Section 3.1. Each surface was captured using one of five sets of camera parameters, as shown in Table 1.

**Table 1: Camera parameters**

	A	B	C	D	E
Exposure	1/15	1/15	1/15	1/8	1/30
Aperture	f4	f5.6	f8	f5.6	f5.6
ISO	200	200	200	200	200

Note that, for a given luminance value, we expect these meta-variables to affect the RGB result in predictable ways. In particular, exposure and ISO speed are approximately linear factors, and aperture is approximately quadratic for well-exposed pixels (i.e., pixels that are neither over- nor under-exposed). The relationships between these values can be expressed through the Exposure Value

(EV) formula [Jacobson et al. 2013], which we will adapt here as:

$$2^{EV} = \frac{N^2}{t} = \frac{LS}{K} \quad (4)$$

where  $N$  is the relative aperture,  $t$  is the exposure time,  $L$  is the luminance value,  $S$  is the ISO speed, and  $K$  is a calibration constant. Following this, we can obtain:

$$L = \frac{N^2 K}{tS} \quad (5)$$

The values Table 1 are chosen so that the predicted luminance follows an approximately known pattern. Reducing the aperture from f4 to f5.6 and f8 should result in approximately half the response per step, similar to halving the exposure time from, e.g., 1/15 to 1/30. As such, we would expect the RGB responses of cases A and D, and C and E to be similar. Each group A/D, B, and C/E should have approximately half the response of the previous group.

Using our measurements, we can estimate the unknown in the EV equation: the calibration constant  $K$ . We use this single parameter to convert our Spheron measurements from RGB to luminance in nits by using the matrix  $T$  derived in Eq. (3). We expect that the Y component of XYZ, obtained as  $XYZ = T^*RGB$ , is linearly correlated to luminance in nits and can be corrected using Eq. (5).

Our practical test, shown in Figure 5, reveals that recovered luminance  $Y_{RGB}$  lies within 20% of the ground truth for values above 3 nits, which we deem acceptable for our application.

### 3.3 Luminance Dataset

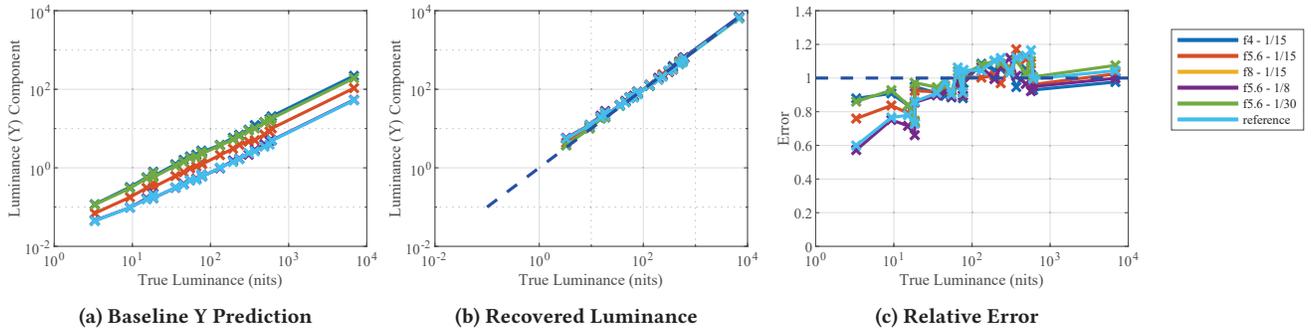
As our main goal is the study of luminance requirements for realistic VR presentation, we aim to create a dataset of stimuli that can be used to conduct a subjective study. To do this, we apply the pipeline presented here to The Southampton-York Natural Scenes (SYNS<sup>3</sup>) dataset, which consists of panoramic HDR data measured from a diverse set of rural and urban locations [Adams et al. 2016]. The authors generated these samples with the aim of producing a representative array of indoor and outdoor environments.

This SYNS dataset was collected using a Spheron HDR camera identical to the one used in this experiment. We selected images with available camera parameters (9 indoor and 10 outdoor), which were processed using the pipeline described in Section 3.1 resulting in recovered estimates of per-pixel luminance (Figure 1b).

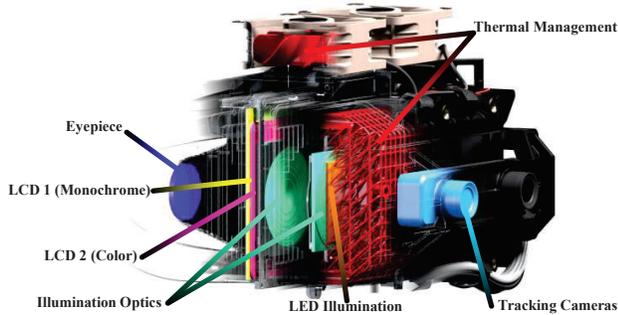
## 4 EXPERIMENTAL HARDWARE

In order to conduct a study on luminance preference for realistic VR, we designed and built a custom headset that can reproduce a wide range of luminance values. We closely followed existing VR architectures while relaxing power, thermal, and weight limits. We use one Lighten Phoenix LTHX1212-060-005 chip-on-board phosphor LED as the backlight for each eye, driven by a Thorlabs DC2200 LED power supply. Each LED is shaped by an Edmund Optics 13-457 f=10mm Fresnel lens, a 28mm air gap, and finally an Edmund Optics 43-024 f=38.1mm Fresnel lens to steer the emission cone toward the display eyebox. Following the work of Seetzen et al. [2004], we employ a dual-modulation approach using liquid crystal displays (LCD). Due to the low efficiency and potential for color moiré patterns when using two color LCDs, we use one color

<sup>3</sup><https://syns.soton.ac.uk/>



**Figure 5:** (a)  $Y_{RGB}$  predictions are compared to luminance measurements. Lines represent cases shown in Table 1. (b) Luminances are corrected as described in Section 3.2. As camera parameters are accurately compensated for, all lines converge to the true luminance values (c) The error ratio between ground truth and recovered luminance per patch.



**Figure 6:** A cutaway illustration showing our hardware configuration. To achieve higher luminance and contrast than existing VR headsets, light emitted by a chip-on-board LED is concentrated by Fresnel lenses to illuminate a color LCD and a monochrome LCD, stacked for dual modulation, which are imaged by an eyepiece.

and one monochrome LCD. Until recently, transmissive LCDs were uncommon, requiring the disassembly of displays with integrated backlights [Rhodes et al. 2019]. The rise in popularity of desktop resin 3D printers has led to the widespread availability of high resolution, high contrast transmissive displays. We use a Wisecoco 6-inch 1620×2560 transmissive color display, and a Sharp LS060R1SX02 6-inch 1440×2560 transmissive color display. We used Thorlabs 50mm achromatic doublets as viewing optics, which achieve a  $\sim 62$  degree field of view. The monochrome LCD is placed at the eyepiece’s focal length, with a virtual image conjugate at infinity. The color display is imaged beyond infinity to help reduce moiré effects induced by the two pixel grids.

The primary challenge for such a configuration is heat. Both LCDs have an operating temperature below 60 degrees Celsius, above which the liquid crystals remain in an isotropic state regardless of the electric field and cannot display an image. We employ a pair of 60mm fans to draw cool air from below the headset across a the front and back of the LCD stack and a sintered metal heat sink

for the LED backlights before exiting the top of the headset. The relatively heavy prototype (2.5kg) is suspended from a tool balancer mounted to an overhead truss. Users move the headset by holding a pair of DJI RS 2 control handles, which we used for their convenient mounting features and button layout that matched our user study needs. To read the button input in our study software, we wired the handle’s built-in buttons to a microcontroller, which presents user input to the host PC as mouse events. Zed Mini camera is affixed to the front of the headset to provide head tracking.

See Figure 1 for a photograph of the headset and Figure 6 for a cutaway view of the key headset components.

#### 4.1 Software

Dual-layer HDR decompositions often require a deconvolution step to correct the mismatch in optical resolution of the two layers, which is particularly important for the local dimming arrays commonly found in HDR televisions. In our case, the circle of confusion for a point on the rear color LCD has a footprint of 1.75 pixels on the front monochrome LCD. Because of this small size and the HVS’ lower sensitivity to color contrast [Kim et al. 2013], we use a simple color-mono factorization without any deconvolution. The target color image  $[I_r^*, I_g^*, I_b^*]$  is normalized to the floating point intensity range  $[0,1]$ . A square-root monochrome image  $I_m$  is calculated using sRGB color coefficients  $C_r$ ,  $C_g$ , and  $C_b$ :

$$I_m = \sqrt{C_r I_r^* + C_g I_g^* + C_b I_b^*} \quad (6)$$

The color component  $I_c$  is extracted using a naive factorization:

$$I_c = \begin{cases} \left[ \frac{I_r^*}{I_m}, \frac{I_g^*}{I_m}, \frac{I_b^*}{I_m} \right], & \text{if } I_m > 0 \\ [0, 0, 0], & \text{otherwise} \end{cases} \quad (7)$$

This process is implemented as a post-processing shader in Unity.

The headset is connected via two HDMI cables to an NVIDIA RTX 3090 GPU in a desktop PC. The headset runs in real time at the 50Hz native refresh rate of the two LCDs.

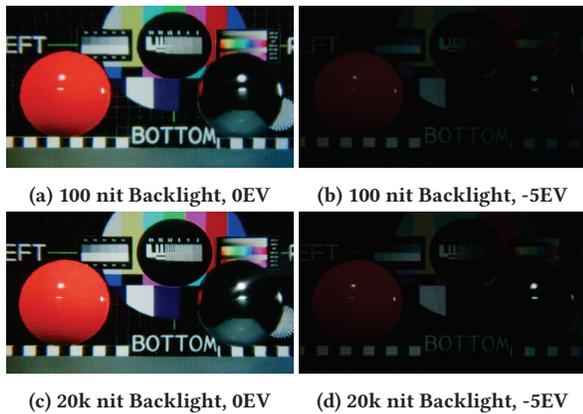


Figure 7: *Top*: A through-the-lens capture of our prototype headset with luminance clipped at 100 nits to emulating current VR devices. *Bottom*: Equivalent capture at full 20,940 nit luminance. The right image in both rows is adjusted down 5 EV to show how specular highlights are preserved in high luminance mode. Note that an HDR image cannot be fully reproduced in a traditional medium such as this figure.

## 4.2 Headset Performance

We measured the peak luminance and dynamic range of the headset using a Konica Minolta CS-2000A spectroradiometer with a VR lens attachment. Peak luminance, sampled in the center, was measured at 20,940 nits. The dark state was measured at 0.05 nits, for a sequential contrast ratio of 418,800:1. Instantaneous contrast was estimated by capturing a checkerboard pattern with 2.5 degree squares and its complement. This contrast ratio was measured at 78:1.

Figure 7 depicts the headset in operation using a target HDR virtual scene created in Unity using the real-time decomposition described in Sec. 4.1. The top row shows a low luminance mode where the backlight is set to 100 nits and the digital image is normalized such that content is clipped at 100 nits. The bottom row shows the full dynamic range of the headset. These images were captured using a Red Komodo and Sigma EX DG fisheye lens at  $f/3.5$ , 320 ISO, and 41 millisecond exposure time. The right hand column in Figure 7 shows the view through the lens with a 5 stop digital reduction in exposure. Note how the specular highlights in the low dynamic range mode are clipped, whereas they are preserved in the HDR mode. This additional visual information (which is best observed in person) supports the wider range of luminance that we target with this demonstrator. Refer to our supplementary video to view this example in motion.

## 5 PERCEPTUAL EXPERIMENT

Having obtained our baseline dataset of real-world luminances (Section 3.3) and built our experimental hardware (Section 4), we proceed to a subjective study of brightness preferences in VR. Our goal is to gain insight into display luminance requirements in VR, setting the stage for further exploration of HDR VR. In particular, we want to understand whether users focusing on immersive, realistic presentation of live-action scenes will have non-trivial preferences

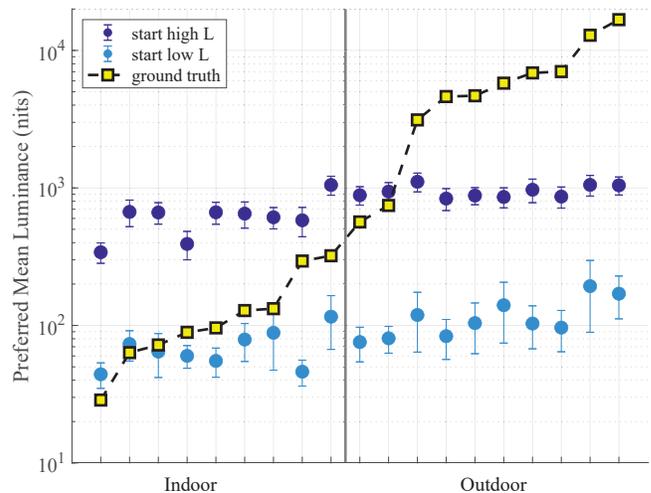


Figure 8: The results of our subjective study. Along the horizontal axis, scenes are sorted in order of increasing ground truth mean luminance (yellow squares). Purple and blue circles show the “start high” and “start low” conditions, respectively. Note that the means of many outdoor scenes were not reproducible on our device. A comparison between preferences for indoor and outdoor scenes is shown in Figure 1c.

for brightness, and how their VR preferences are affected by learned experience of luminance in the real world scenarios.

### 5.1 Stimuli

Our goal is to study subjective preference of luminance variation without affecting contrast and eliminating as many sources of external influence as possible. To achieve this, our stimulus set was generated based on a luminance reconstruction of the SYNS dataset (Section 3.3). Following Dolby’s HDR study [Daly et al. 2013], stimuli were presented in grayscale to avoid conflicting signals from the Hunt Effect (luminance increases perceived colorfulness of chromaticity and saturation) and artifacts resulting from color clipping.

Given an image  $I$  with mean value  $\mu_I$ , a target mean luminance value  $\mu_n$  can be obtained without modifying Michelson contrast by applying a log-luminance offset [Chapiro et al. 2018]:

$$I_n = 10^{\log(I) - \log(\mu_I) + \log(\mu_n)} \quad (8)$$

Mean values are calculated using values in the image’s 5th to 95th percentile to avoid undue influence from sparse outliers [Daly et al. 2013]. Evenly spaced logarithmic means  $\mu_n$  were used, with a maximum permissible mean of 2766 nits. This value was chosen by calculating the largest value for which no more than 5% of pixels are clipped for any image in the dataset after applying Eq. (8).

### 5.2 Experimental Procedure

20 participants with normal or corrected-to-normal vision took part in our study (12F, 8M, aged 26-51). This study was approved in an internal health hazard review. Prior to starting, an assistant read through an explanation of the task and oversaw a training phase helping familiarize participants to the task, headset, and controls.

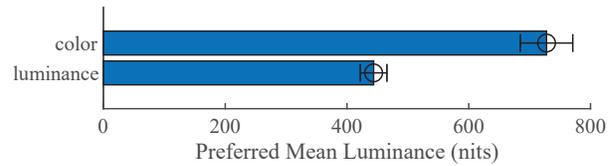
The study consisted of 19 trials (9 indoor and 10 outdoor scenes, presented in random order). For each trial, participants were tasked with selecting the luminance level that feels most *realistic* and *immersive*. Participants were instructed to evaluate the scene as a whole, and spend enough time on each scene to consider the entire setting, including dark and bright regions, shadows and highlights. They were also instructed to disregard any image artifacts such as moiré patterns or over-exposed regions present in the data.

Selection of a preferred brightness level was done using a Method-of-Adjustment (MOA) paradigm, with users changing mean luminance by rotating a dial up or down. Each time luminance was changed, a mask with the same luminance as the mean of the users' new frustrum was displayed for 250 ms to avoid direct comparison and help users adapt [Pattanaik et al. 2000]. As our study values are well within the photopic range and increment over small logarithmic steps, this mask duration was deemed qualitatively appropriate to drive adaptation [Davson 1990]. In order to gauge the magnitude of the effect of anchoring on lightness perception [Gilchrist et al. 1999], the study was run twice in random order, with all trials starting at either the lowest or highest permissible mean level. We hypothesized that hysteresis would influence trials starting from a bright value to result in a brighter preference, and the reverse for trials starting from a dim setting.

In addition to numeric results, qualitative feedback for selection strategy was gathered, with most users mentioning using their memory of real-life luminance, focusing on details such as highlights or shadows, and focusing on illuminated areas such as windows or clouds to make their decision. Some users reported minor difficulties which were not judged to affect their performance, such as simulator sickness, or physical strain from using the device. This was mitigated by allowing for breaks and keeping trial length short (average of 36 minutes). Users selected the maximum luminance value in approximately 1% of trials.

### 5.3 Results

The results of our experiment can be seen in Figure 8. N-way analysis of variance (ANOVA) indicated that individual differences were significant ( $p \ll 0.01$ ,  $F = 11.4$ ), which is expected for a study on preference. As predicted, a significant effect of hysteresis was observed. Preferences for trials starting from a bright setting were significantly higher than those starting dim (773 nits vs. 90.8 nits,  $p \ll 0.01$ ,  $F = 510.8$ ). Although no explicit distinction was made between indoor and outdoor scenes during the study, indoor scenes were set by users to significantly lower luminance levels than outdoor counterparts (333.4 nits vs. 530.3 nits,  $p \ll 0.01$ ,  $F = 29.34$ ). This follows the trend between real-life mean-value distributions of indoor and outdoor scenes (with means of 136 nits vs. 6,303 nits for our recovered luminance dataset). To further test whether the physical luminance values of each scene, which were unknown to participants, affected their responses, a Pearson's Linear Correlation Coefficient (PLCC) was computed separately for indoor and outdoor scenes as compared to brightness preference in our study. This resulted in values of 0.61 and 0.59, respectively. These values can be generally interpreted as large effect sizes [Cohen 1992], indicating the possibility that participants drew on an intrinsic estimate of the scene luminance to make their selection.



**Figure 9: Pilot data for 12 users showing luminance preference changes between grayscale and color conditions.**

## 6 CONCLUSIONS AND FUTURE WORK

We presented an investigation of luminance as a factor to realistic and immersive presentation in VR. We began by generating a dataset of real-world luminance values for representative scenes. Next, we built a novel HDR headset, which is able to reach values above 20,000 nits while maintaining the desirable immersiveness of VR. Finally, we ran a subjective study exploring user preferences for luminance in a VR environment. We found that users had distinct preferences when selecting luminance for different scenes, with outdoor scenes requiring significantly more luminance to be considered realistic and immersive when compared to indoor ones.

Our exploration of luminance preference also highlights that users consistently chose values higher than what can be achieved with commercial displays. This motivates the development of brighter headsets to bring more realistic and immersive experiences to users.

We believe the findings in this work merit further exploration into HDR VR. Following the template of direct-view HDR research [Daly et al. 2013], an interesting next step would be to study user preferences for contrast by manipulating highlights and shadows in HDR images. This is especially challenging in VR, because contrast is reduced by the lenses, but can be addressed with improved optics or careful control of the content being shown.

There are several additional aspects of HDR rendering which can be studied in a VR context, such as color. In our study, we used grayscale images, which is similar to the work of Daly et al. [2013] who avoided saturated colors in their stimuli. We ran a pilot experiment with 12 participants identical to the study described in Section 5, but with images presented in full color. Our preliminary results suggest a significant preference for higher luminance in the color condition ( $p \ll 0.01$ ,  $F = 43.3$ , see Figure 9). We theorize that this could be due to color helping disambiguate between time-of-day in the scenes, which are largely composed of day captures. Further study will be necessary to disambiguate this preference from other influences, such as the Hunt effect.

Finally, an exciting avenue for future work is tone mapping. Traditional tone mappers are generally unable to operate in VR without distracting flicker due to the head motion inherent in this type of presentation. However, the joint development of rendering algorithms and computational display hardware to achieve perceptually realistic luminance in VR may reduce the overall power requirements and open the door to practical HDR VR.

## ACKNOWLEDGMENTS

We thank Robert Birch, Saturnia Bocast, Ken Koh, Yu-Jen Lin, Chris Neugebauer, Ka Yan Wat, and Cameron Wood for their contributions toward conducting the user study described in this publication.

## REFERENCES

- Wendy J Adams, James H Elder, Erich W Graf, Julian Leyland, Arthur J Lugtigheid, and Alexander Murry. 2016. The southampton-york natural scenes (SYNS) dataset: Statistics of surface attitude. *Scientific Reports* 6, 1 (2016), 1–17.
- Kynthia Chamilothoni, Jan Wienold, and Marilyne Andersen. 2019. Adequacy of immersive virtual reality for the perception of daylight spaces: comparison of real and virtual environments. *Leukos* 15, 2-3 (2019), 203–226.
- Alexandre Chapiro, Timo Kunkel, Robin Atkins, and Scott Daly. 2018. Influence of screen size and field of view on perceived brightness. *ACM Transactions on Applied Perception (TAP)* 15, 3 (2018), 1–13.
- Jacob Cohen. 1992. A power primer. *Psychological bulletin* 112, 1 (1992), 155.
- Scott Daly, Timo Kunkel, Xing Sun, Suzanne Farrell, and Poppy Crum. 2013. Preference limits of the visual dynamic range for ultra high quality and aesthetic conveyance. 86510J. <https://doi.org/10.1117/12.2013161>
- Hugh Davson. 1990. *Physiology of the Eye*. Bloomsbury Publishing.
- Paul E Debevec and Jitendra Malik. 2008. Recovering high dynamic range radiance maps from photographs. In *ACM SIGGRAPH 2008 classes*. 1–10.
- Dolby. 2016. *Dolby Vision™ for the Home*. Technical Report. Dolby Laboratories, Inc. [https://connect.avid.com/rs/149-WFZ-676/images/Avid\\_HDRSeminar\\_dolby-vision-white-paper.pdf](https://connect.avid.com/rs/149-WFZ-676/images/Avid_HDRSeminar_dolby-vision-white-paper.pdf)
- Mark D Fairchild. 2013. *Color appearance models*. John Wiley & Sons.
- Fufu Fang, Han Gong, Michal Mackiewicz, and Graham Finlayson. 2017. Colour correction toolbox. (2017).
- Graham D Finlayson and Mark S Drew. 1996. The maximum ignorance assumption with positivity. In *Color and Imaging Conference*. Society for Imaging Science and Technology, 202–205.
- Jan Froehlich, Stefan Grandinetti, Bernd Eberhardt, Simon Walter, Andreas Schilling, and Harald Brendel. 2014. Creating cinematic wide gamut HDR-video for the evaluation of tone mapping operators and HDR-displays. In *Digital photography X*, Vol. 9023. SPIE, 279–288.
- Alan Gilchrist, Christos Kossyfidis, Frederick Bonato, Tiziano Agostini, Joseph Cataliotti, Xiaojun Li, Branka Spehar, Vidal Annan, and Elias Economou. 1999. An anchoring theory of lightness perception. *Psychological review* 106, 4 (1999), 795.
- Ralph Jacobson, Sidney Ray, Geoffrey G Attridge, and Norman Axford. 2013. *Manual of Photography*. Routledge.
- Kil Joong Kim, Rafal Mantiuk, and Kyoung Ho Lee. 2013. Measurements of achromatic and chromatic contrast sensitivity functions for an extended range of adaptation luminance. In *Human vision and electronic imaging XVIII*. SPIE.
- Timo Kunkel and Erik Reinhard. 2010. A reassessment of the simultaneous dynamic range of the human visual system. In *Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization*.
- Patrick Ledda, Alan Chalmers, and Helge Seetzen. 2004. HDR displays: a validation against reality. In *2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No. 04CH37583)*, Vol. 3. IEEE, 2777–2782.
- Laura R Luidolt, Michael Wimmer, and Katharina Krösl. 2020. Gaze-dependent simulation of light perception in virtual reality. *IEEE Transactions on Visualization and Computer Graphics* 26, 12 (2020), 3557–3567.
- Steve Mann and Rosalind W. Picard. 1995. On being ‘undigital’ with digital cameras: Extending Dynamic Range by Combining Differently Exposed Pictures. In *Proceedings of IS&T*. 442–448.
- Arian Mehrfard, Javad Fotouhi, Giacomo Taylor, Tess Forster, Nassir Navab, and Bernhard Fuerst. 2019. A comparative analysis of virtual reality head-mounted display systems. *arXiv preprint arXiv:1912.02913* (2019).
- Hossein Najaf-Zadeh, Madhukar Budagavi, and Esmaeil Faramarzi. 2017. VR+ HDR: A system for view-dependent rendering of HDR video in virtual reality. In *2017 IEEE International Conference on Image Processing (ICIP)*. IEEE, 1032–1036.
- Sumanta N Pattanaik, Jack Tumblin, Hector Yee, and Donald P Greenberg. 2000. Time-dependent visual adaptation for fast realistic image display. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*. 47–54.
- Clotilde Pierson, Coralie Cauwerts, Magali Bodart, and Jan Wienold. 2021. Tutorial: luminance maps for daylighting studies from high dynamic range photography. *Leukos* 17, 2 (2021), 140–169.
- Josiane McGinn Proulx. 2020. *A Comparative Study of Lighting Perceptions between Tone-Mapped Virtual Reality Environments and Physical Environments*. Ph. D. Dissertation. University of Colorado at Boulder.
- Ana Radonjić, Sarah R. Allred, Alan L. Gilchrist, and David H. Brainard. 2011. The Dynamic Range of Human Lightness Perception. *Current Biology* 21, 22 (Nov. 2011), 1931–1936. <https://doi.org/10.1016/j.cub.2011.10.013>
- Alessio Regalbuto. 2019. Remote Visual Observation of Real Places through Virtual Reality Headsets. (2019).
- Erik Reinhard, Wolfgang Heidrich, Paul Debevec, Sumanta Pattanaik, Greg Ward, and Karol Myszkowski. 2010. *High dynamic range imaging: acquisition, display, and image-based lighting*. Morgan Kaufmann.
- TJ Rhodes, Gavin Miller, Qi Sun, Daichi Ito, and Li-Yi Wei. 2019. A transparent display with per-pixel color and opacity control. In *ACM SIGGRAPH Emerging Technologies*.
- S Rockcastle, M Danell, E Calabrese, G Sollom-Brotherton, A Mahic, K Van Den Wymelenberg, and R Davis. 2021. Comparing perceptions of a dimmable LED lighting system between a real space and a virtual reality display. *Lighting Research & Technology* 53, 8 (2021), 701–725.
- Helge Seetzen, Wolfgang Heidrich, Wolfgang Stuerzlinger, Greg Ward, Lorne Whitehead, Matthew Trentacoste, Abhijeet Ghosh, and Andrejs Vorozcovs. 2004. High dynamic range display systems. *ACM Transactions on Graphics (TOG)* (2004).
- Helge Seetzen, Hiroe Li, Linton Ye, Wolfgang Heidrich, Lorne Whitehead, and Greg Ward. 2006. Observations of Luminance, Contrast and Amplitude Resolution of Displays. *SID Symposium Digest of Technical Papers* 37, 1 (2006), 1229. <https://doi.org/10.1889/1.2433199>
- Helge Seetzen, Lorne A. Whitehead, and Greg Ward. 2003. A High Dynamic Range Display Using Low and High Resolution Modulators. *SID Symposium Digest of Technical Papers* 34, 1 (2003), 1450. <https://doi.org/10.1889/1.1832558>
- Gunter Wyszecki and Walter Stanley Stiles. 1982. *Color science*. Vol. 8. Wiley New York.
- Fangcheng Zhong, Akshay Jindal, Ali Özgür Yöntem, Param Hanji, Simon J. Watt, and Rafal K. Mantiuk. 2021. Reproducing reality with a high-dynamic-range multi-focal stereo display. *ACM Transactions on Graphics (TOG)* 40, 6 (Dec. 2021), 1–14. <https://doi.org/10.1145/3478513.3480513>