3-14-2022

Lightness, Brightness, and Transparency in Optical See-Through Augmented Reality

Lili Zhang
lxz6532@rit.edu

Lightness, Brightness, and Transparency in Optical See-Through Augmented Reality

Lili Zhang

March 14, 2022

Lili Zhang

Lightness, Brightness, and Transparency in Optical See-Through Augmented Reality

Ph.D. Dissertation, March 14, 2022

Advisor: Dr. Michael J. Murdoch

Committee Chair: Dr. Gabriel J. Diaz

Committee Members: Dr. Mark D. Fairchild and Dr. Elena A. Fedorovskaya

Rochester Institute of Technology

College of Science

Program of Color Science

Munsell Color Science Laboratory

1 Lomb Memorial Dr.

14623 Rochester, New York

MUNSELL COLOR SCIENCE LABORATORY
COLLEGE OF SCIENCE
ROCHESTER INSTITUTE OF TECHNOLOGY
ROCHESTER, NEW YORK

Ph.D. DEGREE DISSERTATION

The Ph.D. Degree Dissertation of Lili Zhang
has been examined and approved by the dissertation committee as satisfactory for the dissertation required for the Ph.D. degree in Color Science

Dr. Gabriel J. Diaz, Chair	Date
Dr. Michael J. Murdoch, Advisor	Date
Dr. Mark D. Fairchild, Committee Member	Date
Dr. Elena A. Fedorovskaya, Committee Member	Date

Lightness, Brightness, and Transparency in Optical See-Through Augmented Reality

Lili Zhang

Abstract

Augmented reality (AR), as a key component of the future metaverse, has leaped from the research labs to the consumer and enterprise markets. AR optical see-through (OST) devices utilize transparent optical combiners to provide visibility of the real environment as well as superimpose virtual content on top of it. OST displays distinct from existing media because of their optical additivity, meaning the light reaching the eyes is composed of both virtual content and real background. The composition results in the intended virtual colors being distorted and perceived transparent. When the luminance of the virtual content decreases, the perceived lightness and brightness decrease, and the perceived transparency increases. Lightness, brightness, and transparency are modulated by one physical dimension (luminance), and all interact with the background and each other. In this research, we aim to identify and quantify the three perceptual dimensions, as well as build mathematical models to predict them.

In the first part of the study, we focused on the perceived brightness and lightness with two experiments: a brightness partition scaling experiment to build brightness scales, and a diffuse white adjustment experiment to determine the absolute luminance level required for diffuse white appearances on 2D and 3D AR stimuli. The second part of the research targeted at the perceived transparency in the AR environment with three experiments. The transparency was modulated by the background Michelson contrast reduction in either average luminance or peak-to-peak luminance

difference to investigate, and later illustrated, the fundamental mechanism evoking transparency perception. The first experiment measured the transparency detection thresholds and confirmed that contrast sensitivity functions with contrast adaptation could model the thresholds. Subsequently, the transparency perception was investigated through direct anchored scaling experiment by building perceived transparency scales from the virtual content contrast ratio to the background. A contrast-ratio-based model was proposed predicting the perceived transparency scales. Finally, the transparency equivalency experiment between the two types of contrast modulation confirmed the mechanism difference and validated the proposed model.

Acknowledgments

I would like to express my deepest appreciation to my advisor Dr. Michael J Murdoch, for his invaluable advice, continuous support, and patience during my PhD journey. Without his guidance and persistent help, this dissertation would not have been possible.

I would like to thank my committee: Dr. Mark D. Fairchild, Dr. Elena A. Fedorovskaya, and Dr. Gabriel J. Diaz. Each member has provided me extensive professional guidance. I would like to extend my sincere thanks to the faculties and staff at the Color Science program, Dr. Roy S Berns, Dr. Susan Farnand, Dr. David Wyble, Stephanie Livingston-Heywood, and Valerie Hemink, for creating the most wonderful and supportive program.

This work would not have been possible without the financial support from Microsoft and Meta Reality Labs. I would also like to give special thanks to all observers participated in my research experiments.

I am grateful for my fellow students with whom I have had the pleasure to work, Dr. Anku Manderna, Dr. Fu Jiang, Dr. Matthew Ronnenberg, Dr. AdiRobinson, Dr. Nargess Hassani, Dr. Katherine Carpenter, Dr. Yongmin Park, Hao Xie, Luke Hellwig, Olivia Kuzio, Jenibel Paray, Yue Yuan, Chen Shen, Tucker Downs, Minyao Li, Zilong Li.

I would like to thank my parents, whose love is with me regardless of the physical distance. I would like to thank Dawei Liu and my cat, Bamboo, for whose companionship during the pandemic time.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The ultimate display would, of course, be a room within which the computer can control the existence of matter.

- Ivan E Sutherland, *The Ultimate Display*[1]

Augmented reality (AR), sometimes referred to as mixed reality (MR), supplements the real world with virtual objects, and modulates reality with virtual characteristics. The technology can be traced back to flight simulators and military head-up displays. Now its value is carried to both enterprise and consumer domains improving efficiency and experiences in training, manufacturing, retail, gaming, social networking, etc. AR is not restricted to particular display technologies [2]. One method of achieving the blending between the virtual objects and the real world is through optical see-through (OST) displays. OST displays, also known as transparent displays or additive displays, utilize optical combiners to superimpose virtual rendering on the real world. The combiners are transparent or semi-transparent, ensuring the visibility of the real environment. The ”additivity” depicts that the physical light reaching the eyes is the additive blend of the display itself and the light from the real world modulated by the optical elements. Sometimes, OST displays are coupled with head and/or eye tracking, resulting in anchored content from the see-through display

in the real-world coordinate rather than the display coordinate.

The technology for AR OST displays is still evolving while facing lots of challenges for broad adoption of this novel media into everyday life. There have been studies improving display efficiency, uniformity, gamut, blur, field-of-view (FoV), registration, form-factor, etc. A few of them directly relate to visual qualities. OST displays are uniquely distinctive from conventional display media due to their transparency and additivity. The interaction between the reality and the virtual object, as the core of AR, is the most intriguing part of it.

One example is given in Figure 1.1 capturing the additivity of AR OST displays. On the panel labeled as ”background”, a sinusoidal grating represents the real world background visible through the AR optical elements. The grating is surrounded by uniform medium gray color. Repetitive 6-step neutral scales on the medium gray surround are for reference, representing the full luminance range in the background. On the OST AR display shows 8 virtual square patches of different luminance with black surround. The bottom photograph illustrated when the virtual content is superimposed on the background, the black surround on the AR display is invisible because no additional light is added to the background. The 8 virtual patches change their appearances due to the patterned background blending. Here we mainly focus on the lightness, brightness, and transparency in AR content appearance. In CIE International Lighting Vocabulary, [3], brightness is defined as "a visual sensation according to which an area appears to emit, or reflect, more or less light". Lightness is defined as "the brightness of an area judged relative to the brightness of a similarly illuminated area that appear to be white or highly transmitting". The appearance of the virtual patches not only change in apparent brightness, but also the in transparency. Both the AR overlay and the backgrounds contributed to the appearance changes.

There are some appearance problems in OST AR that do not present on other media. Virtual

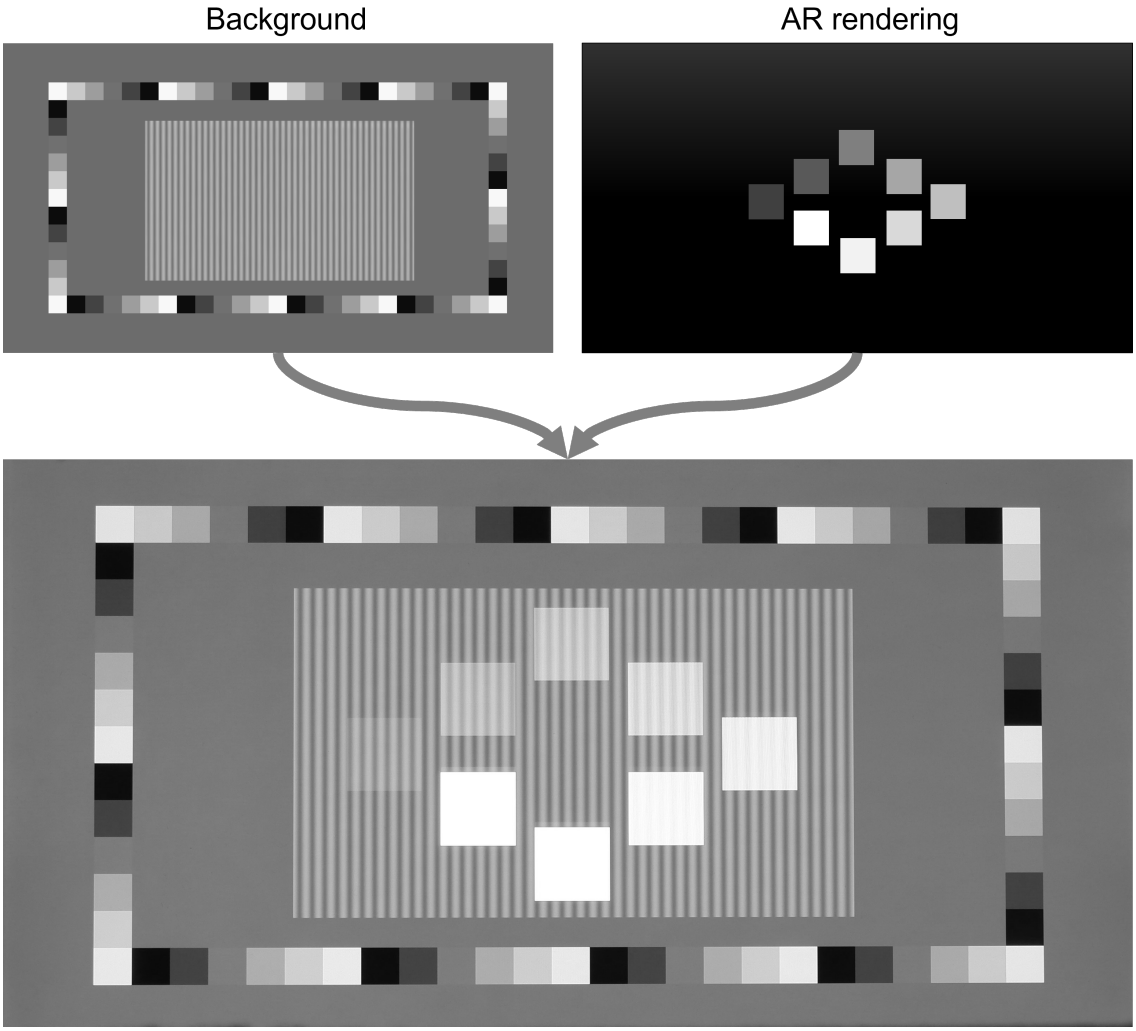


Figure 1.1: Background blending in AR OST display distorting the virtual color appearance: The background presents the sine pattern with neutral scales surrounded by medium gray color. The virtual image is comprised of 8 patches with different luminance surrounded by black. The overlaying process produced brightness changes and transparency perception on the virtual patches compared to the source image.

content colors are distorted by their according background colors if the rendering is not regulated by the environment information. Real world color is also distorted by the virtual content overlay. This effect is stronger if the virtual overlay is precisely confined and registered to the real world. It has been used to create virtual color swatch on the real objects. Virtual content can be perceived as transparent, but in a way that is different from physical transparent objects. Since the real world is always visible through the displays, the virtual content is physically transparent but with higher luminance than the backgrounds. Virtual content contrast is reduced because of the existence of background. An effect of this is the lack of black in AR OST displays. When the OST display emits minimum light, the virtual content disappears, leaving only the background visible. Rendering black virtual content is rather difficult without a dark real world background, which is impractical considering the common user cases of AR. White balance on the virtual content can be inconsistent with the real world. New attributes introduced by 3D rendering and spatial arrangement need to be considered in color management. This includes properties like gloss, shadow, and occlusion, that need to be re-rendered according to the relative position of the virtual content, background, and viewpoint.

1.1 Motivation

The appearance prediction and color management in AR OST displays requires different approaches from conventional display monitors and surface colors. It raises demands for color researches on this innovative technology. Color appearance models and material appearance models specific to the OST AR environment are required to accurately predict appearance and improve color management. As the first step towards the appearance model, we focus on the three attributes modulated by the same physical dimension: luminance. The three inseparable attributes: lightness, brightness, and transparency, all related to the interaction between the virtual content and the background, demand identification and quantification in OST AR.

Lightness and brightness are two fundamental perceptual attributes in color appearance. They have been studied and modeled since the 1860s, starting from the Weber-Fechner law. The lightness and brightness theories have evolved over time considering not only the stimulus itself, but the background, adaptation, context, media, etc. Appearance models on surface colors have both lightness and brightness attributes and are applicable to physical and digital colors. But these models are not designed for the context that physical colors and digital colors are presented simultaneously as AR colors do. This context requires more comprehensive models on lightness and brightness, considering effects the real world background might bring in.

The transparency of AR content is different from subtractive filter transparency. In the subtractive filters, the transparency is modulated by the physical properties like reflectance, transmittance, absorbance, and scattering. The transparency of subtractive filters can be assessed in two dimensions: perceptual transmittance and clarity. Perceptual transmittance is defined as overall transmittance of the filter, associated with the amount of light subtracted from the background. And clarity is defined as the loss of background contrast through the filter, associated with both the light subtracted from the background and the additional light reflected by the filter. The two transparency dimensions cannot be applied to the AR transparency directly because it is not induced by the physical properties. The measured transmittance will be over 100% if we calculate the transmittance as in subtractive filters, since more light is added to the overlaid area compare to the background. The clarity is always high because there is no scattering in the optical combiner to introduce the cloudiness to the AR content.

Perceived transparency in AR is not incorporated in existing color appearance models such as CIECAM02 and CAM16, which predict and manage displayed colors to improve image quality. Effort has been made to improve AR image visibility problem due to the high transparency. But there

has been little research on quantifying perceived transparency directly in OST AR, though it is the nature of OST displays and requires thorough understandings to improve AR image quality.

The missing components in lightness, brightness, and transparency encouraged this research focusing on quantifying the perceptual dimensions targeting at the interaction between the AR rendering and the background. This research, as the start of an AR total appearance model, meant to provide fundamental knowledge from subjective evaluation data on AR appearance, help determine the luminance dependent perceptual attributes, supplement important components to the AR color appearance model. The extended mathematical modeling on these perceptual dimensions could provide psychophysical support for future improvement in accurate imaging and rendering in OST AR.

1.2 Approach

This research will mainly focus on the luminance induced perceptual effects while setting chromatic properties aside. It is essential to develop accurate lightness and brightness scales on neutral colors before extending them to chromatic colors for an integrated color appearance model. We separated the research into two parts, with one targeting at lightness and brightness, and the other at perceived transparency. In each part, we primarily focused on building perceptual scales. In order to test different background conditions in luminance and spatial variation, we adopted a display monitor as the background. For the AR overlay, we utilized an LCD screen paired with slanted beam splitter to reflect and overlay the rendering to the background.

In the first part, we measured the perceived brightness scales on AR rendering against various backgrounds with two experiments. The first experiment focused on building the brightness scale

from partition scaling, which is a psychometric method constructing perceptual scales from equal-appearing intervals. The bisection task was performed by observers selecting the mid-point in a subrange, where the two end points were the pair of references. The bisecting gave two equal intervals at the first subdivision level, and generated two new pairs of references from the mid-point along with the two original references. The bisection progressed to the next subdivision level on the new reference pairs. Multiple subranges were included to reduce the difficulty in bisecting references with large perceptual difference. We used square patches in 2°x2° FoV as stimuli. We used three unbalanced overlapping subranges with three subdivision levels inside each subrange. Observers were asked to adjust the center patch in a triplet to be halfway between the other two references. The subranges were pooled and modeled to build brightness scales. The impact of background conditions were examined on the brightness scales. The modeled scales were also compared to existing predictors to determine the applicability of existing models in AR context. The second experiment was designed to determine the absolute luminance level required for diffuse white appearance on AR 2D and 3D stimulus with method of adjustment. We adopted a simple shading model with Lambertian surfaces on a cube for the 3D stimulus in addition to the 2D square patches in the first experiment. We offered the freedom of rotating the cube during the experiment for observers to examine the appearance. The impact of background luminance and contrast, as well as the virtual rendering dimensionality, were examined on the diffuse white luminance. The combined result from the two experiments were discussed with hypothesis anchoring the relative brightness scales with the white levels.

In the second part of the research, we targeted the perceived transparency through three psychophysical experiments with patterned backgrounds emphasizing the transparency. The transparency modulation was approached from the background Michelson contrast reduction. Two types of patches were included to verify if contrast is the fundamental mechanism modulating the transparency. AR overlay patches reduced the contrast by increasing the average luminance, and LCD

patches reducing background contrast from luminance differences at constant average luminance. The first experiment measured the transparency detection threshold to confirm the possibility of a patch being perceived as transparency and non-transparent over a range of contrast. The perceived transparency was then extended beyond a single threshold through a direct anchored scaling experiment. Observers were asked to rate the test stimulus transparency on a continuous scale between ”completely transparent” to ”completely opaque”. The result was used to build perceived transparency scales from the patch contrast in relation to the background's. A contrast based transparency model was proposed based on the direct scaling experiment result from the contrast ratio between the stimulus and the background and the stimulus type. A transparency equivalency experiment were carried out between the two types of patches to further inspect the mechanism on transparency perception. AR patches as test target were compared to reference LCD patches using staircase method by varying AR patch luminance. The data was also used to validate the proposed contrast based model. The underlying mechanism of transparency perception was discussed with the alternative modeling method. Existing transparency and translucency models were compared to our proposed model and experiment data to demonstrate the uniqueness of the perceptual dimension in AR environment.

Future work should address chromatic virtual and real background colors for their luminance dependent perceptual attributes and chromatic properties like chroma, hue, saturation, and colorfulness. Furthermore, the AR color appearance model can be combined with studies on texture, depth, environment adaptation towards an AR total appearance model, improving AR color management, graphic rendering, and image quality.

1.3 Dissertation Structure

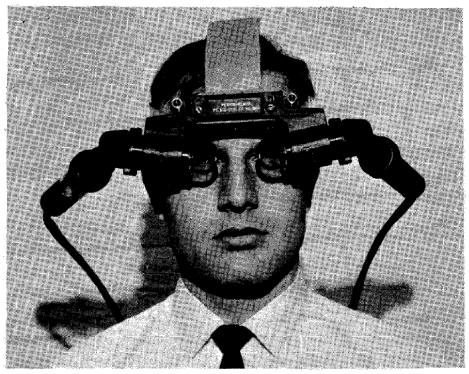
The dissertation comprises 6 chapters. The next chapter emphasizes the related historical research from three aspects: AR system history and color appearance research, lightness and brightness studies, transparency perception and modeling. Chapter 3 described the experiment environment and test bed for our studies. Chapter 4 and 5 each focuses on one of the two major topics of our studies: brightness and transparency. Within each chapter, we will walk through the individual experiments one by one from the experimental setup, procedure, result, and modeling. The last chapter summarizes the studies, discusses the application and limitation, proposes further work extending the current research.

Chapter 2

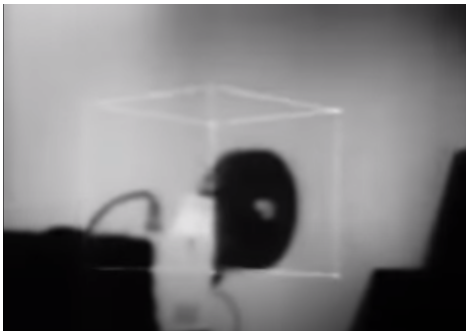
Background

2.1 Augmented Reality Technology

The idea of merging reality with virtual content can be traced back to the 1860s, when John Henry Pepper and Henry Dircks introduced the Pepper's ghost to the theaters by projecting an actor's image with a sheet of tilted glass [4]. Although the projected images were live performance, but the boundary between "real" and "virtual" was not clear in the Pepper's ghost. The predecessor of the modern AR head-mounted displays (HMDs) was from more than 50 years ago. Sutherland had the vision of an ultimate display with perspective images according to viewing position on a 3D display [1]. The prototype of it was published in 1968 by him and his colleagues [5]. It was composed with miniature monochrome CRTs, half-silvered mirrors, and the tracking system known as the "Sword of Damocles" as in Figure 2.1. The research was not followed up until 1985. The colored version appeared with LCD displays and half-silvered mirrors [6]. But the term "augmented reality" was not introduced until 1992, emphasizing the information augmentation in manufacturing [7]. Furthermore, AR was defined in the "virtuality continuum" [8]. Virtuality continuum described the mixture of classes of objects presented in any particular display situation, where one end consists solely real objects and the other end solely virtual. AR was defined as real environment being



(a)



(b)

Figure 2.1: (a) The AR head-mounted 3D display in 1968 by Sutherland and his team.[5] (b) A virtual cube with real world background through the display. Image from ACM Fall Joint Computing Conference, San Francisco 1968.

augmented by means of virtual objects. AR has been prospering and maturing since the 1990s, with numerous technologies supporting its development. In the last two decades, both commercial and personal AR devices had leaped over the research labs to the market.

Currently, optical see-through (OST) displays and video see-through (VST) displays are both wildly adopted. OST displays, also known as transparent displays, utilize an optical combiner to superimpose virtual rendering on the real world. It often appears in two forms: near-eye displays (NEDs) or head-up-displays (HUDs) depending on the relative position of the displays. NEDs usually appear in smart glasses and head-mounted devices, while HUDs are used in automotive industry, museums, and education [9, 10]. VST displays are comprised of scene cameras and conventional displays. The real-word background is captured and displayed in real-time while the computer generated images are added through the imaging pipeline. The displayed world correspond with the immediate outside real world, creating a see-through effect analogous to the OST displays [8]. VST AR are wildly adopted on mobile platforms like smartphones and laptops.

2.1.1 Optical See-Through Display Architecture

OST displays at minimum require two components: a light engine and an optical combiner. A light engine is usually a micro-display with relatively small size but high resolution[11]. These requirements come from the small form factor of NEDs that often incorporate magnifying optics. The brightness level and contrast requirement are also demanding due to the visible background and high dynamic range use cases. Common light engines used in OST displays include liquid crystal on silicon (LCoS), micro organic light emitting diode (μOLED), digital light processor (DLP), and laser beam scanner (LBS). Figure 2.2 are six example main stream OST AR devices using various light engines.

The other essential component is the optical combiner, which directs the light coming from the light engine to the designated pupil location while allowing the light from the real world to pass through. In general, there are two types of optical combiner: reflective optics, and waveguide-based optics. Reflective optics have at least one optical surface that is semi-reflective, which combines the light from the light engine and the real world as shown in Figure 2.3(a)-(d). The reflective optics can appear in different shapes and thickness, for example, like flat or curve half-mirror, beam splitter cube[12], or free-form prism [13, 14]. The waveguide-based optical combiner is semi-transparent, and utilizes the total internal reflection (TIR) to guide the light from the micro-display towards the eyes as shown in Figure 2.3(e). Diffraction gratings are used as couplers for the input and output light before and after transiting through the waveguide. Diffraction grating options include surface relief gratings (SRG)[15], polarization volume gratings (PVG) [16], holographic volume gratings (HVG) [17], and holographic planar gratings [18]. The choice of couplers along with the optical design affect the diffraction efficiency and the display performance in channel luminance and spatial uniformity.

(a) **(b)**

(c) **(d)**

(e) **(f)**

Figure 2.2: OST head-mounted devices and light engine types: (a) Magic Leap One (LCoS). Image from Magic Leap. (b) Google Glass Enterprise Edition 2 (LCoS). Image from Google. (c) Nreal Light (μOLED). Image from Nreal. (d) EPSON Moverio BT-40S (μOLED). Image from EPSON. (e) Microsoft HoloLens 2 (laser beam scanner). Image from Microsoft. (f) Vuzix Blade (digital light processor). Image from Vuzix. All images accessed in 2021.

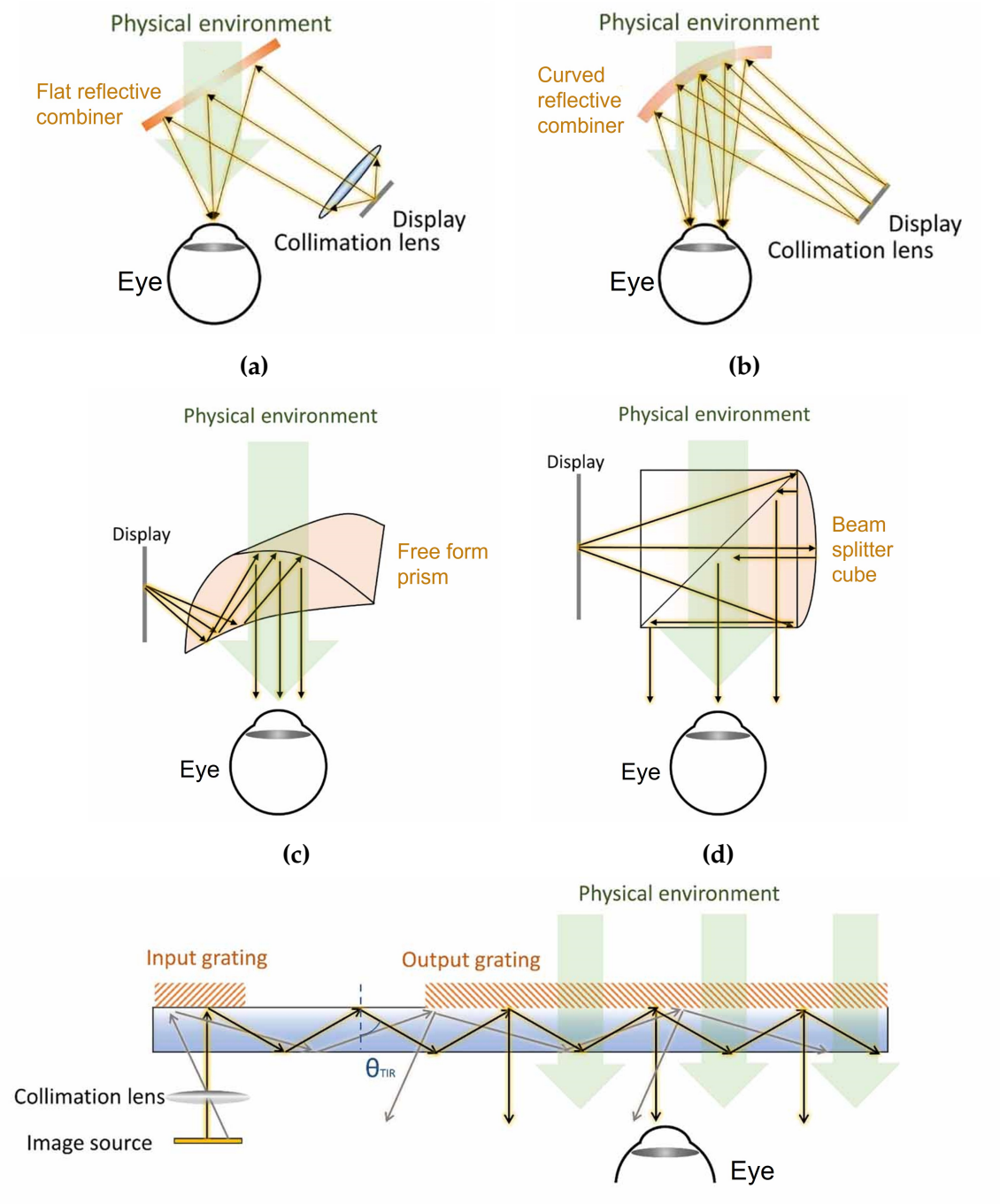


Figure 2.3: Schematic diagram of OST AR architectures with example combiners. Images redrawn from Yin et al. Figure 5, 6 and 9 [12]. (a) Flat reflective combiner. (b) Curved reflective combiner. (c) Free form prism. (d) Beam splitter cube. (e) Grating-based waveguide. θ_{TIR} is the total internal reflection angle inside the waveguide.

2.1.2 Objective Appearance Correction in OST AR

OST AR, as a not yet mature technology, faces lots of challenges. Problems exist in both experimental and off-the-shelf devices. Some issues are often approached from hardware including expanding FoV, improving display spatially uniformity, limiting channel cross-talk, achieving rendering occlusion [19], etc. Other issues like the color distortion from the real world background are usually approached from software. One research group creatively utilized this distortion and chromatic adaptation to create colors out of the physical gamut of the display [20].

The distortion from the background on AR colors appearance, which arises from background elements visible through the transparent optical combiner, is well recognized. Several researches have quantified and attempted to correct the background blending with existing colorimetric models. The general approach is to compensate for the background in colorimetry with the assumption of additivity. First, the target color C_{mix} and the background color C_{bg} are determined in one color space. The modification is then applied to adjust the virtual color C_{AR} so that the additive combination with the background result in the target color.

$$C_{mix} = I * C_{bg} + R * C_{AR} \tag{2.1}$$

The transmittance I and the reflectance R of the optical combiner are also included in the compensation. The background color or image can be measured directly [21, 22] or inferred from a scene camera [23, 24]. Different color spaces were explored to achieve better compensation including CIE XYZ [23–26], CIELUV [21], CIELAB [22]. The modification methods include direct subtraction of the transmitted background color from the intended color mixture [22, 23, 26, 27]:

$$C_{AR} = \frac{(C_{mix} - I * C_{bg})}{R}. \tag{2.2}$$

Figure 2.4: Image captured by a webcam through the OST display rendering a bust with color checker on the background by Weiland et al. [27] Left is the uncompensated rendering with color distortion. The right is compensated rendering with background subtraction.

Adaptive methods were developed to limit gamut clipping from the subtraction by setting a maximum cap of pixel alternation from the original image [24], or by performing local gamut mapping with linear regression [25]. Weiland et al. proposed a two-stage method to modulate the virtual image to compensate for not only the background, but also the luminance adaptation [27]. In the colorimetric stage, the photo receptor response on the background was sparsely sampled and used to compute the adaptation state. The pixel value of the virtual image was adjusted according to the time-dependent adaptation. In the photometric stage, the background image was subtracted from the colorimetric compensated virtual image to minimize the background color distortion. The comparison between the uncompensated and compensated rendering is shown in Figure 2.4.

Color management algorithms were developed to compensate for the contrast and gamut change due to the existence of the background. Real-time tools help select color with the best visual contrast according to the background was developed based on the colorimetric compensation[26]. More sophisticated color optimization algorithm was developed, enhancing contrast with power consumption constrain [28]. Most researches defined the background color using luminance and chromatic-

ities, without specifying if the change is from ambient lighting or background reflectance. Lee focused on the distortion from ambient light and proposed a two-step gamut mapping method for ambient light supporting both gamut compression and expansion [29].

Projected color cannot be categorized as OST, but the appearance problem caused by the ambient light and projection surface share common ground with OST AR. Generally, the projection screens are white flat diffusive surfaces viewed in dim rooms. But some studies focused on the color compensation for non-traditional screens like textured screen [30], chromatic screen [31, 32], dynamically and spatially changing screen [33]. Menk and Koch had a series of publications on truthful color reproduction in AR with projectors [34–36]. They calculated the projected RGB values based on the measured the spectral properties of the ambient light, the material, background surface normal, and the projector. The projected color was compensated in tristimulus values using a physically-based rendering, resulting in compelling visualization (Figure 2.5).

Projection mapping is a special form of projection, creating the mixed reality using real objects as the projection screen. The immersive sensation is especially strong with large scale architectures. With the natural texture and color on architecture exterior surfaces, projection mapping faces the same problems as AR background blending. The workflow for accurate color reproduction is to photograph the projected surface and compute a corrected image that prepares the projection surface to be achromatic [37].

2.1.3 Subjective Appearance Evaluation in OST AR

Before applying colorimetric compensation to OST AR displays, it is questionable that if the conventional color theories are applicable to this novel additive media. Subjective evaluation focused

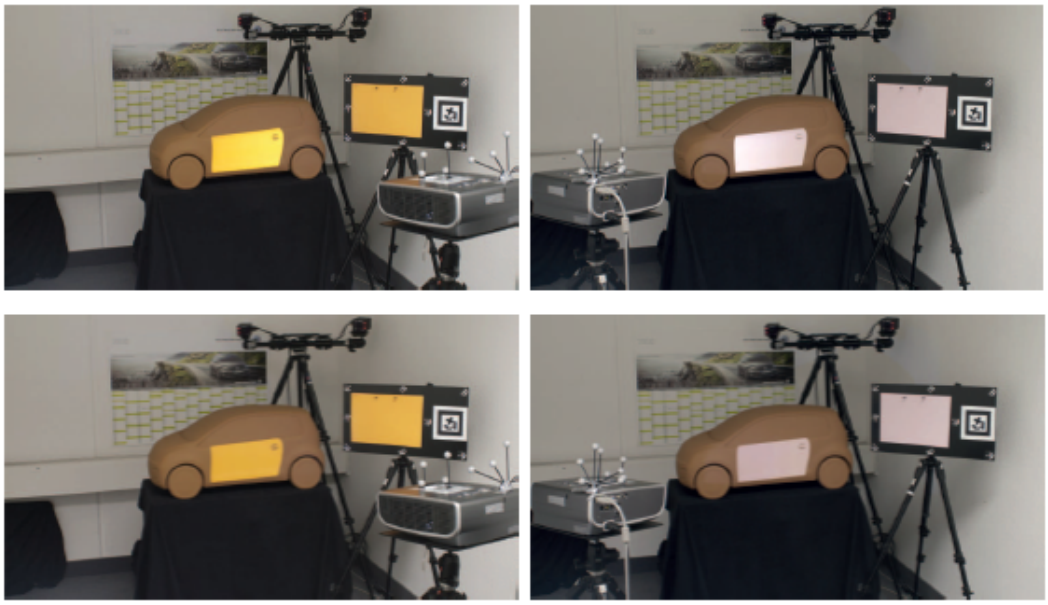


Figure 2.5: Color projection correction from ambient light by Menk and Koch [36]. The top row shows the color projection on the model car before correction, and the bottom row after. The color on the black boards were the reference color to be matched to.

on measuring the human visual perception in OST AR directly to offer support for appearance compensation and enhancement.

For the appearance of achromatic virtual colors, Kim et al. measured the equivalent brightness on a transparent OLED display between the see-through area and non-see-through area through a matching experiment [38]. The non-see-though area had a non-transmissive sticker behind the display and was used as the reference. The result showed when the reference luminance was 2.5 times higher than the background luminance, the summation of background luminance and observer adjusted the transparent patch luminance matched to the reference, indicating physical luminance additivity. But in region near background luminance, observers adjusted the transparent patch so that the luminance summation of the transparent patch and the background were higher than the reference (Figure 2.6). The region was defined as the "transparent effective range" where brightness was affected by the perceived transparency, occupying the range when reference luminance is lower than 2.5 times of the background luminance. A 2-gamma correction model was proposed to compensate for the deviation with a reduced gamma value of 1.7 in the effective region compared to the rest of the range with gamma value of 2.2.

Murdoch focused on how the AR overlay affected the real background cube brightness [39]. Different alignment between the overlay and the background cube was tested, from tightly aligned hexagons to overly large rectangles covering both the cube and the adjacent area. Observers were asked to adjust the overlay on darker real cubes to match the area brightness to the reference. The result showed that the large rectangles were adjusted to be higher than reference. With the reduced overlay area and improved alignment to the background cube, the adjusted area luminance reduced and were close to the reference luminance. The background cube brightness also affected the matched luminance.

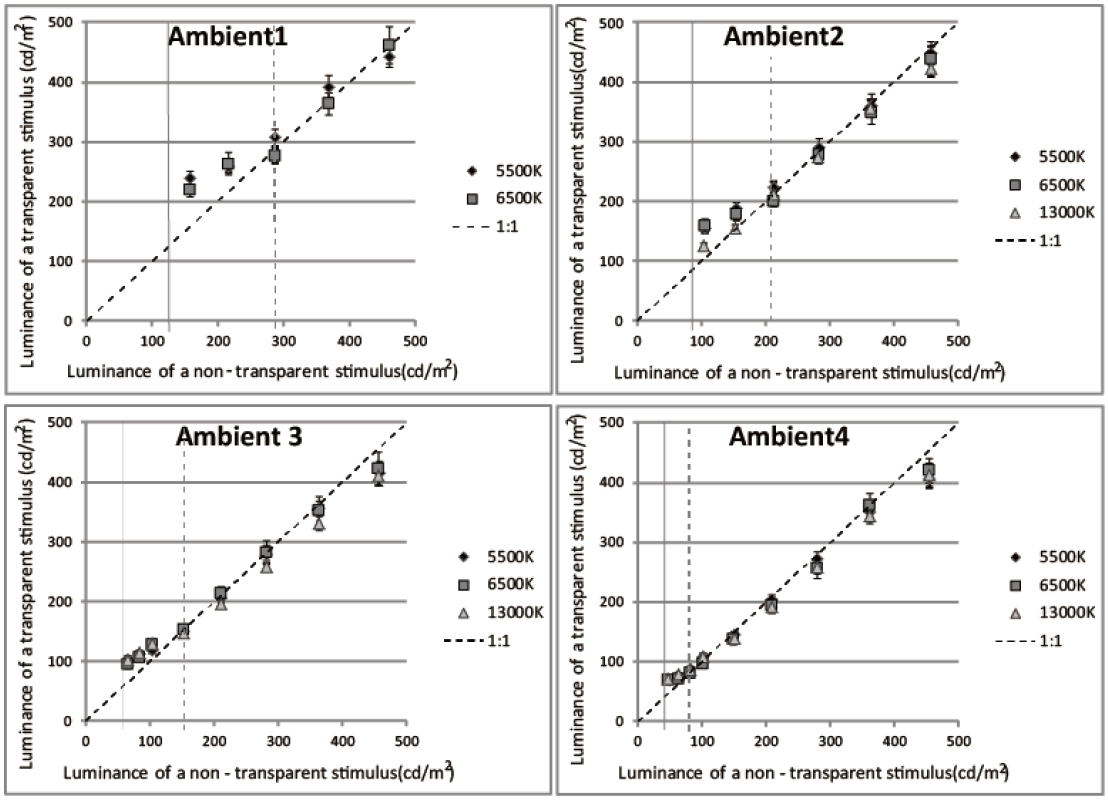


Figure 2.6: Equivalent brightness between transparent and non-transparent patches under different ambient light conditions from Kim et al. [38]. The background luminance (gray solid vertical lines) under ambient light 1-4 were at approximately 124 cd/m^2, 84 cd/m^2, 65 cd/m^2, 40 cd/m^2. Higher luminance was adjusted on the transparent patches when the non-transparent reference patches luminance were close to the background. The black dashed vertical lines represented the luminance ceiling of the transparent effective range, which is about 200%-250% of the background luminance.

The achromatic appearance with chromatic background was assessed by Huang et al. The chromatic adaptation level when viewing the OST virtual images was demonstrated to be dependent on the background and adapting luminance level. The perceived achromatic color shift towards higher CCT compare to the adapting CCT, demonstrating a lower level of chromatic adaptation than viewing the real background [40].

For chromatic virtual colors, Hassani and Murdoch demonstrated chromatic simultaneous contrast needs to be considered when predicting AR color appearance with chromatic backgrounds using CAM16 model[41] (Figure 2.7). In the following study, an AR appearance model was proposed with perceptual weighting factor α on virtual content and real background tristimulus values.

$$C_{mix} = \alpha(1) * I * C_{bg} + \alpha(2) * R * C_{AR} \tag{2.3}$$

When $\alpha(1) = \alpha(2) = 1$, it's equivalent to Equation 2.1 [42]. Their result showed that observers discounted the background color with $\alpha(1) < \alpha(2)$. The values of α varied according to the virtual content complexity in shape and texture.

Zhang et al. demonstrated through a color matching experiment that the perceived mixed color appearance shifted towards the background color, implying the transparent display might not be perceptually additive despite simultaneous contrast and chromatic adaptation compensation [43]. Moffitt and Browne examined the visibility of chromatic symbols against bright achromatic backgrounds with numerical rating [44]. Difference metrics were evaluated against subjective data: luminance contrast ratio, ΔL^* in CIELAB, ΔE^*_{uv} in CIELUV with a size-correcting factor k, and CIEDE2000 with factor k between the symbol and the background. ΔL^* was recommended for predicting symbol visibility. In simulated OST with bright achromatic background, an image quality improvement was demonstrated by decreasing the transparent display gamma [45].

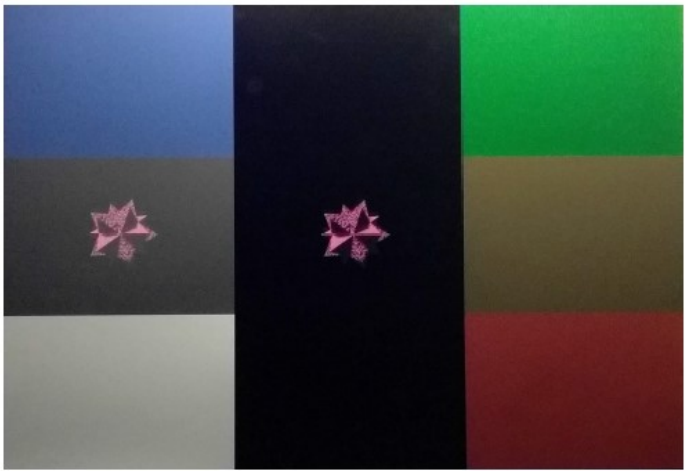


Figure 2.7: Appearance matching on 3D rendered complex object among chromatic backgrounds from Hassani [42]. The result challenged the transparency display additivity with perceptual weighting factors.

Virtual text visibility was also of interest of researchers in additional to images. Kim et al. measured the appropriate monochromatic virtual text brightness against background luminance using categorical rating [46]. The background luminance was modulated from both ambient light illuminance and the achromatic background reflectance, examining the effect of the two components (Figure 2.8). They demonstrated that the measured illuminance at eye position from the combined ambient illuminant and the background reflectance can be used to predict the comfortable virtual text luminance. The ideal virtual text luminance was modeled with a power function of the eye illuminance with exponent of 0.49.

Apart from modifying text luminance to improve legibility, high contrast text box were also studied in OST AR. Huang et al. evaluated the perceived text clarity under difference surround lighting and virtual text box luminance to model the readability of high frequency information [47]. A polynomial regression model was proposed predicting text clarity based on the background, text, and text box luminance.

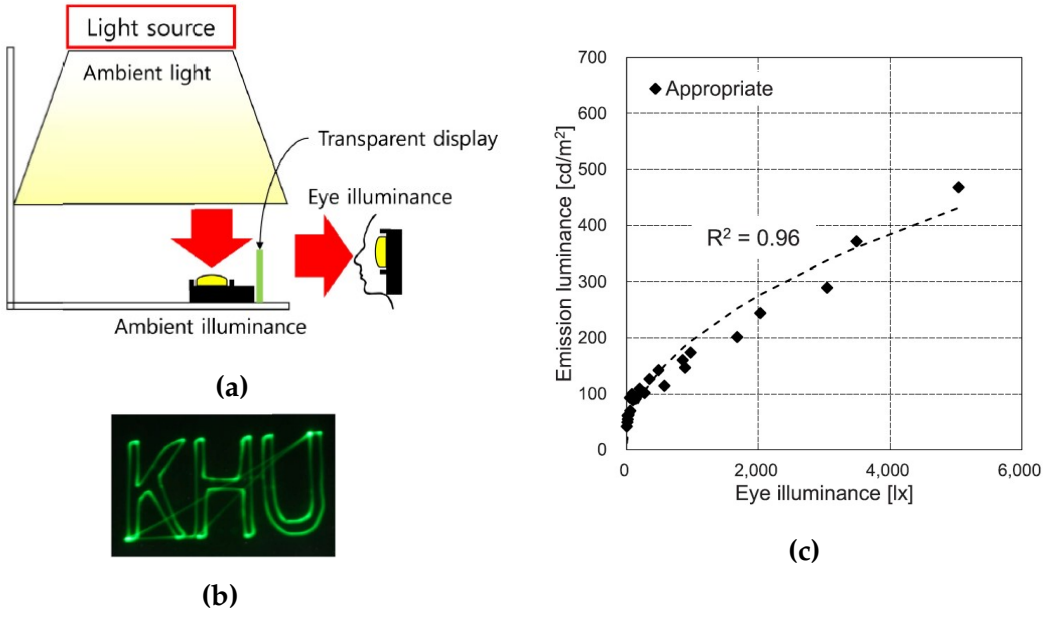


Figure 2.8: Appropriate virtual text luminance based on the illuminance measured at eye position from Kim and Lee [46]. (a) Experiment Setup showing illuminance measurement positions. (b) Monochromatic virtual text stimulus from laser beam scanner. (c) Modeled appropriate brightness from eye illuminance with power function with exponent of 0.49.

2.2 Lightness and Brightness

In CIE International Lighting Vocabulary, brightness is defined as "attribute of a visual perception according to which an area appears to emit, transmit or reflect, more or less light", while lightness as "brightness of an area judged relative to the brightness of a similarly illuminated area that appears to be white or highly transmitting" [3]. But there isn't always a consensus among researches on the definition. Particularly, lightness is sometimes described as the "perceived reflectance" [48, 49]. And in some models, it's not computed as the relative brightness as defined above [50, 51]. The two attributes are generally considered independent of the media. They are studied on conventional media like surface colors, emissive digital colors on display monitors, and far beyond reflective white luminance levels.

2.2.1 Surface and Digital Color Models

The relation between brightness perception and luminance can be traced back to the 1800s, when scientists just started exploring the relationship between physical intensity and perceptual sensation. Fechner proposed that sensation ψ can be represented with a weighted logarithmic scale of stimulus intensity Φ [52]:

$$\psi = klog(\Phi) \tag{2.4}$$

Steven's power law describes the relation of sensation ψ and stimulus intensity Φ as exponential relation [53]:

$$\psi = k\Phi^a \tag{2.5}$$

The brightness scale can be converted to the lightness scale when normalized to the brightness of the adopted white. This is used to describe the surface color lightness and then extended to emissive colors. In 1976 the Commission Internationale de l'Eclairage (CIE) recommended the CIE 1976 (L^*, a^*, b^*) color space known as CIELAB [54]. In CIELAB, the lightness L^* is written as cube-root relation to the luminance L normalized by the luminance of the reference white L_n:

$$L^* = 116\left(\frac{L}{L_n}\right)^{1/3} - 16 \tag{2.6}$$

Another color space, *IPT* by Ebner and Fairchild, was designed for digital image processing purpose [55] and includes the nonlinearity on cone responses *LMS* with a power function. Cone responses are calculated from tristimulus values *XYZ* with a linear transform. I, describing the light-dark scale, is computed from the nonlinealy corrected cone responses $L'M'S'$.

$$\begin{aligned}
\begin{bmatrix} L \\ M \\ S \end{bmatrix} &= \begin{bmatrix} 0.4002 & 0.7075 & -0.0807 \\ -0.2280 & 1.1500 & 0.0612 \\ 0 & 0 & 0.9184 \end{bmatrix} \begin{bmatrix} X \\ Y \\ X \end{bmatrix} \\
L'M'S' &= \begin{cases} LMS^{0.43} & for\ L, M, S \geq 0 \\ -(-LMS)^{0.43} & for\ L, M, S < 0 \end{cases} \\
I &= 0.4L' + 0.4M' + 0.2S'
\end{aligned} \tag{2.7}$$

Various color appearance models (CAMs) were proposed over the years, as predictors of relative color appearance attributes from tristimulus values. Those models were able to account for the appearance change due to chromatic adaptation and viewing conditions on simple stimuli. Generally, in CAMs, an achromatic response A is calculated first from rods and cones sensitivity functions. Lightness J and brightness Q are then computed from A with additional steps accounting for other visual effects like background induction [56](contrast change with surround) and Helmholtz-

Kohlrausch effect [57](chromaticity dependent brightness).

Hunt proposed his model for practical applications in 1995 [58], in which achromatic response A had embedded correction for luminance adaptation, chromatic adaptation, and cone bleach. Brightness Q is calculated from A, colorfulness M that account for the H-K effect, and surround induction factor N_b that account for the different surround conditions.

$$\begin{aligned} Q &= N_1 \left[7 \left(A + \frac{M}{100} \right) \right]^{0.6} - N_2 \\ N_1 = \frac{(7A_w)^{0.5}}{5.33N - b^{0.13}},\ N_2 &= \frac{7A_w N_b^{0.362}}{200} \end{aligned} \tag{2.8}$$

A_w is the achromatic response of the reference white. Lightness J is written as exponential relation to normalized brightness with an index of background luminance factor z.

$$\begin{aligned} J &= 100 \left(\frac{Q}{Q_w} \right)^z \\ z &= 1 + \left(\frac{Y_b}{Y_w} \right)^{0.5} \end{aligned} \tag{2.9}$$

Q_w is the brightness of the reference white. Y_b and Y_w are the luminance of the background and reference white respectively. In particular, he proposed a different formula to describe the lightness of projected transparencies $J_p$from brightness Q, due to "the top-most part of the lightness scale from 85-100 is much steeper than the rest of it":

$$J_p = 100 \left(\frac{Q}{Q_w} \right)^{1.2} \left\{ 1.14 * \left[1 - \left(\frac{Q}{Q_w} \right)^{3.6} \right] + \left(\frac{Q}{Q_w} \right)^6 \right\} \tag{2.10}$$

Two CIE recommended color appearance models, CAM02 [50] and CAM16 [51] were also proposed aiming to provide simplified appearance models for practical use. CAM02 parameterized the surround factors c, N_c, F based on the viewing conditions to determine if discounting the illuminant.

Partial adaptation factor D was introduced. CAM16, as the succession of CAM02, further simplified the model, completing luminance and chromatic adaptation in the same space. In both CAM02 and CAM16, the lightness calculation is performed before brightness, unlike in Hunt's model. But the method of calculating the two attributes are similar in the two CIE CAMs. Achromatic response A is calculated based on the adapted cone responses R, G, B with background induction factor N_{bb}:

$$\begin{aligned} A &= (2R + G + 0.05B - 0.305)\, N_{bb} \\ N_{bb} &= 0.725 \left(\frac{Y_w}{Y_b}\right)^{0.2} \end{aligned} \tag{2.11}$$

where Y_b is the background relative luminance, and Y_w is the relative luminance of the reference white. Lightness J is calculated with the normalized achromatic response A with surround factors. And then brightness Q was calculated with lightness J, achromatic response A, and luminance adaptation factors F_L for achromatic stimuli when colorfulness M is zero:

$$\begin{aligned} J &= 100 \left(\frac{A}{A_w}\right)^{1.48c + c\sqrt{Y_b/Y_w}} \\ Q &= \frac{4}{c} \left(\frac{J}{100}\right)^{0.5} (A_w + 4) F_L^{0.25} \\ F_L &= k^4 L_A + 0.1(1 - k^4)^2 (5L_A)^{1/3} \\ k &= \frac{1}{5L_A + 1} \end{aligned} \tag{2.12}$$

A_w is the achromatic response of the reference white, and L_A is the adapting luminance. CAM16-UCS uniform color space was also proposed for color difference evaluation. These two color appearance models are the widely accepted, state-of-art models for predicting color appearance considering the adaptation and viewing conditions.

2.2.2 Considering the Context

It was noticed that the background context has effect on the perceived lightness and brightness on simplest stimulus, for example: simultaneous contrast, crispening, and background induction. Some of the color appearance models above have considered context, including background and surround effect. In addition to those color appearance models, numerous studies have been exploring the effect of the context on lightness and brightness, focusing only on neutral colors.

In the early 70s, Land and McCann developed the retinex theory, proposing the lightness judgment in Mondrian patterns is done by computing a chain of reflectance ratio from the sharp edges and anchored to the maximum lightness in the scene [59]. This local contrast approach is supported by the lateral inhibition mechanism. Rudd et al. extended this edge integration model to be a weighted sum of steps in logarithmic luminance across spaces [60].

There are also ideas that the perceived images are parsed into layers to infer scene properties. Lightness here is considered as the perceived reflectance of an object in the scene. Bergstrom suggested that retinal images are separated into three components: reflectance, illumination, and 3D form [61]. Gilchrist proposed that edges need to classified according to illumination or surface reflectance [62]. The classified edges are integrated into surface lightness layer and illumination pattern layer. The intrinsic image model proposed in 1978 describes that the visual system infers lightness by decomposing the scene into layers of reflectance, illumination, and depth [63].

Another branch of research describing lightness in complex scenes is the anchoring theory. Anchoring is necessary whenever mapping relative perceptual attributes to absolute. Wallach suggested the highest luminance in a scene is judged to be white as the anchor [64]. This was supported by Cataliotti and Gilchrist [65]. The theory was expanded by Gilchrist, combining the local and global

anchoring [66]. Areas with common properties can be classified into one framework. The perceived lightness J can be derived from the summation of weighted lightness of all n frameworks by specifying the highest luminance anchor within the frameworks:

$$J = \sum_{i=1}^{n} W_i \left(\frac{L}{L_{max,i}} \right) \tag{2.13}$$

where W_i is the weighting of the framework i, L is the luminance of the interested area, L_{max} is the highest luminance of the framework. But it was pointed out that in complex scenes, it is likely that highest lightness instead of luminance serves as anchor, indicating layer decomposition before anchoring [48].

Crispening, on the other hand, is the effect that the perceived colors are biased away from the background. The effect was named by Takasaki from an experiment measuring perceived lightness [67]. A typical illustration of lightness crispening is the lightness difference of a pair expands when the background brightness is between the pair, as in Figure 2.9. Crispening can also happen on chroma scale besides brightness [68]. It was pointed out that the lightness crispening effect can be reduced by minimizing the stimulus spacing [69], introducing a coarse noise background [69], introducing a hue difference on the background from the stimulus [70], or introducing a black annulus [71]. Semmelroth studies the brightness crispening and modeled the effect [73]. He proposed that there are two psychohysical responses to a stimulus in crispening: Steven's brightness power function with exponent m, and luminance difference to its background power function with exponent n. When the stimulus is brighter than the background as in AR, the perceived lightness J is written as:

$$J = L^m + K|L - L_B|^n \tag{2.14}$$

where L is the stimulus luminance, L_B is the background luminance, m is the brightness exponent, n is the exponent affect response to background-stimulus difference, K is the proportion parameter

Figure 2.9: Crispening effect illustration from Hajdek and Hajdek [72]: the perceived lightness difference between the two gray patches increases when the background being very similar to the two patches.

of the two psychophysical responses.

Kane and Bertalmio proposed crispening can be described with a gain model [74]. The Gain factor $G_i = 1 + k|L_i \check{} L_b|$, where L_i and L_b are the luminance of the stimulus and the background. A detection threshold function τ_i is modulated by G_i to calculate the local magnitude threshold ΔI_i. The stimulus magnitude M_i is obtained by integrating magnitude threshold ΔI_i over the luminance range. M_i is then normalized with the maximum stimulus magnitude M_{max} to obtain the lightness estimation:

$$\begin{aligned} \Delta I_i &= \tau_i G_i \\ M_i &= \int_0^i \frac{1}{\Delta I_i} \\ Q_i &= \frac{M_i}{M_{max}} \end{aligned} \tag{2.15}$$

2.2.3 Wide-range Lightness

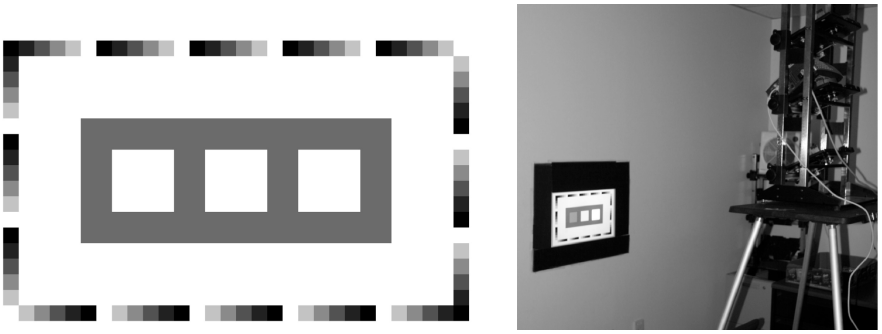
Chen and Fairchild ran a study in 2011 measuring the wide-range lightness above and below diffuse white [75] [76]. They used a projector to modulate the stimulus luminance on paper white to

enable high dynamic range. The projection surface contains a repeating gray scale forming a frame surrounding the test area. Inside the test area, there were three squares as stimulus with medium gray background (Figure 2.10a). The base illumination shines on the background, gray scales, and one of the white square offering information of reference white. With the second square projected to be a different luminance than the reference white, observers were asked to adjust projected luminance on the third square to make equal lightness difference between the two pairs. The experiment was done with one set below diffuse white and the other above. The result showed the exponential lightness compression relation to the luminance still hold for lightness exceeding diffuse white, but crispening effect showed at both gray background ($L* = 50$) and diffuse white ($L* = 100$). The result of this experiment was used to improve the lightness scale in high-dynamic-range (HDR) imaging by replacing the cube-root relation in CIELAB with the optimized Michaelis-Menten equation to model the nonlinear decay:

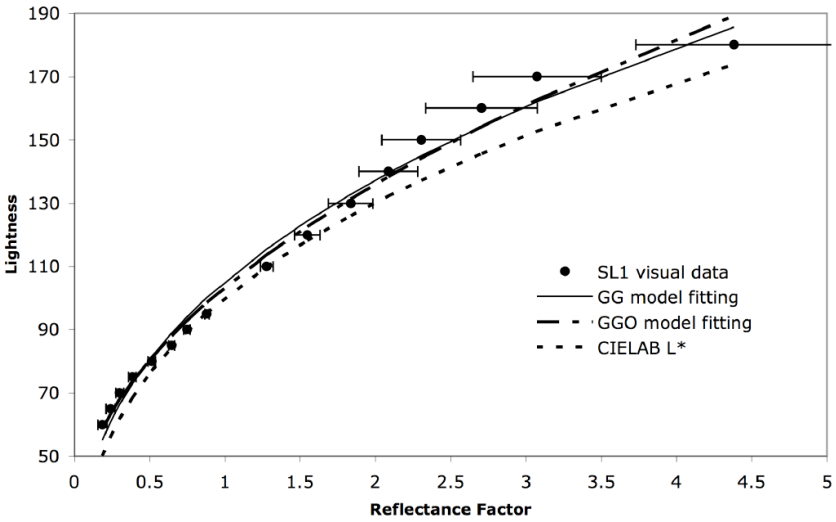
$$\begin{aligned} f(Y) &= 247 * \frac{Y^{\varepsilon}}{Y^{\varepsilon} + 2^{\varepsilon}} + 0.02 \\ \varepsilon &= 0.58 \end{aligned} \tag{2.16}$$

where Y is the stimulus relative luminance, and ε can be adjusted to compensate for surround relative luminance or absolute luminance level.

Abebe et al. demonstrated the effect of reference white luminance on lightness scaling in 2017 using similar configuration as Chen and Fairchild's [77]. Their result showed that when changing reference white luminance from 100 to 1000 cd/m^2, the normalized lightness-luminance relation would have a higher exponential index as an effect of high dynamic range. They illustrated that both Steven's power law and Michaelis-Menten equation are adequate to describe the lightness with optimized parameters according to reference white luminance range.



(a) a **(b)**



(c) b

Figure 2.10: Wide range lightness scaling from Chen [75]. (a) The projection surface used in the experiment with gray scale reference and medium gray background. (b) The experiment setup with the projectors on the right side. (c) The result on wide range lightness is denoted as "SL1 visual data". The gray background and diffuse white have the reflectance factor of 0.5 and 1 respectively.

2.2.4 Brightness Contrast

Weber's law describes the contrast with just-noticeable-difference (JND) being proportional to the base stimulus intensity [78]:

$$\frac{\Delta I}{I} = k \tag{2.17}$$

where I is the intensity of the stimulus, ΔI is the JND of I. k is a parameter depends on the types of sensation. Michelson contrast C is another method commonly used in vision research, also known as peak-to-peak contrast [79]. It's defined with the maximum luminance I_{max} and minimum luminance I_{min} and should be used on patterns that bright and dark areas take similar fractions like sinusoidal gratings:

$$C = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} \tag{2.18}$$

Some color spaces are designed to describe the color difference, for example CIELAB [54]. The lightness difference in CIELAB is calculated with the difference in L^*. New color difference formulas were proposed to be more uniform in color difference ΔE. CIE DE94 [80] is one of them, with the lightness difference is calculated the same manner as in CIELAB. CIE DE2000 is another color difference formula [81], in which lightness difference is computed with the weighted L^* in CIELAB (see chapter 2.1.1):

$$\begin{aligned} \Delta L_{2000} &= \frac{\Delta L^*}{k_L S_L} \\ S_L &= 1 + \frac{0.015(\overline{L}^* - 50)^2}{\sqrt{20 + (\overline{L}^* - 50)^2}} \end{aligned} \tag{2.19}$$

where ΔL^* and $\overline{L}^*$ are the lightness difference and lightness average of the stimulus pair, respectively. k_L is a scalar usually set to 1, and S_L is the lightness difference function to equalize the lightness difference over the lightness range.

Whittle's contrast [70][82] was based on a series of experiments measuring brightness matching, scaling, and discrimination of a patch on a uniform background. It was demonstrated that brightness JNDs can be scaled to the brightness difference scale. The brightness contrast is captured with $\log W$:

$$W = \frac{\Delta L}{k + L_{min}} \tag{2.20}$$

where ΔL is the luminance difference between the patch and the background, and L_{min} is the luminance of the dimmer one. k is regarded as internal noise level. This relation is proven to be working on grating patterns too [83], and extended to measure equal perceptible differences (nEPD) on self-luminous devices with the noise level factor set to 0.39 [84][85].

$$nEPD = a * log_{10}[1 + b(1 - k)W] \tag{2.21}$$

where a, b, and k are parameters determined by the stimulus context.

2.3 Perceived Transparency

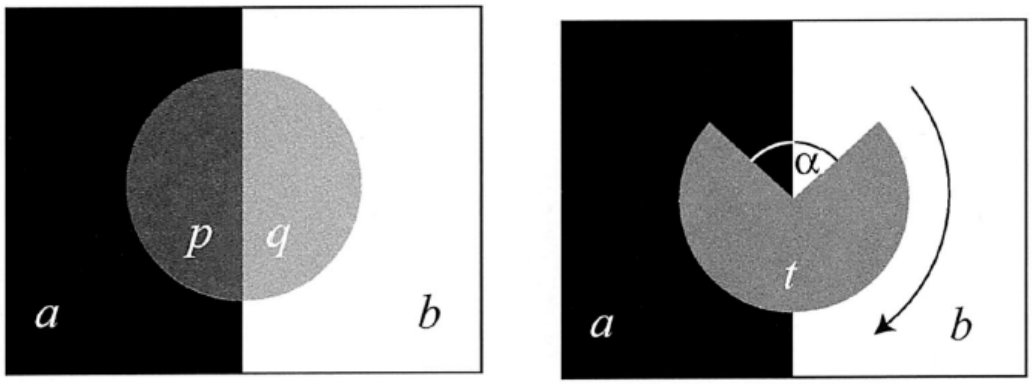
When discussing physical interaction between light and matter, transparency-related concepts are usually well-defined. Physically, transparency and translucency are both material properties governed by the amount of diffuse transmission and reflection. In radiometry, the material is characterized as transparent when transmission and reflection are mainly regular, while translucent when diffuse [86]. The difference between the regular and diffuse depends on if the material has surface and internal boundaries as scattering centers [87]. The percentage of transmitted light that is scattered, resulting in translucency, can be described as haze. Haze dictates the apparent contrast reduction of an object viewed through a material. Clarity, the opposite concept of haze, describes the ability to perceive the fine details of an object through a transparent material [88–90]. Thus,

translucency is a descriptor of the combined effects of clarity and haze.

But physical properties like transmittance and scattering are not enough to define the perceived transparency [91]. Studies have shown that highlight, contrast, caustics (imperfectly focused rays pattern formed by translucent objects) could impact the perceived transparency [92] [93]. Furthermore, the transparency sensation in OST AR is not even evoked by physical diffuse transmittance and scattering, but through additive light. The increased virtual image luminance decreases the background contrast, and reduces the perceived transparency. The increased AR luminance increases haziness without decreasing clarity, as the fine detail in the background stay retained. In this section, we cover both perceptually evaluated transparency and translucency of see-through media, since the perceived transparency in OST AR cannot be defined via physical properties of transmittance and scattering.

2.3.1 Episcotister Model and Scission

Metelli proposed a theory in the 1970s that the human visual system could decompose images into distinct layers with carefully arranged reflectance and boundaries [94]. The decomposition into transparent and underlying layer is refereed to as "scission" [95]. An example is shown in Figure 2.11 (a): four regions with respective reflectance of a, b, p, q, along with appropriate boundaries, could result in transparency perception of a disk on the bi-part background. It illustrated that transparency perception, which distinct from physical transmittance, could be evoked from non-transparent materials. To describe the perceived transparency, Metelli proposed the episcotister model. It reproduces the four-region transparency from a fast rotating disk with an open section on a bipart background (Figure 2.11 (b)). α, the relative area of the open section on the disk, is also the transparency descriptor of the disk. t is the reflectance of the open-section disk. The episcotister



(a) **(b)**

Figure 2.11: Metelli's episcotister model on transparency illusion [94]. (a) Four regions with respective reflectance of a, b, p, q could result in transparency perception of a disk on the bi-part background. (b) a fast rotating disc with an open section of relative area α. t, a, b are the reflectance of the disk and backgrounds.

transparency model describe the four-region reflectance and the open-section disk as:

$$\begin{aligned} p &= \alpha a + (1-\alpha)t \\ q &= \alpha b + (1-\alpha)t \end{aligned} \tag{2.22}$$

Thus the transparency α could be solved from the four-region reflectance a, b, p, q:

$$\alpha = \frac{p-q}{a-b} \tag{2.23}$$

IIn this model, transparency is computed from reflectance. Gebino proposed that it is luminance of the four areas in instead of reflectance should be used in those equations [96]. Also, the latter term $(1-\alpha)t$ should be replaced with F representing the transparent layer effective luminance, which could be independent of the α.

Though, there are limitations to the episcotister model. The reflectance values of the four region are restrained to produce meaningful α values. Furthermore, the reflectance values on edges and

junctions, as well as their sequence also decide if evoking transparency perception [97].

2.3.2 Contrast-based Model

Anderson proposed that the lightness and transmittance can be estimated from contrast for simulated patches with depth difference to the background using stereo disparity [98]. A generative model derived by Singh and Anderson [99] can be described as:

$$\begin{aligned} \alpha &= \frac{L_{range}}{A_{range}} \\ t &= \frac{A_{range}L_{mean} - L_{range}A_{mean}}{I(A_{range} - L_{range})} \end{aligned} \tag{2.24}$$

where α and t are the transmittance and reflectance of the transparent layer, respectively. I is the illuminant intensity. L represents the luminance of the region where the transparent layer overlaps with the background, while A represents the luminance of the background in plain view without occlusion. $range$ denotes the peak-to-peak luminance difference of the area, and $mean$ the average luminance. The α calculation is essentially the Michelson contrast ratio of the transparent region and the background. Metelli's episcotister model can be derived from this general model with assumptions of homogeneous reflectance and transmittance of the transparent layer. They then conducted experiments matching the transparent layer brightness or transparency on a sinusoidal background. The result showed transparency of the layer appears to decrease as the average luminance of the transparent area increases (with its luminance range kept fixed). The transparency could be described with only the Michelson contrast of the transparent region, which align with their general model. In the followup study, they proposed that the perceived contrast should be used to estimate the slope instead of Michelson contrast [100].

In addition to the lightness and transparency perception, Anderson and Winawer also demonstrated in 2005 that there is strong effect on layered image representations on lightness with textured

background [101]. They showed that the lightness of a moving transparent layer on textured backgrounds is perceived higher than physical signal, indicating partial background discounting [101]. They concluded in a later paper that "When viewing a transparency display, two dimensions of perceptual experience can be readily distinguished: transparent surfaces appear to have a particular opacity (or hiding power), and they appear to have a particular lightness (or, more generally, color)" [102]. This decomposition of transparency and lightness dimension is also supported by Petrini and Logvinenko [103].

The transparency matching by Robilotto and Zaidi concluded that observers may match brightness contrast when asked to match transparency using neutral density filters [104]. For colored transparent media, the interested region being brighter or darker than the surround resulted in different perception [105]. With bright region and dark surround giving the spotlight-like look, the matched bias towards the overlay color can be explained by the gray world model with higher weight on the bright region. But with the target region darker than the surround creating the filter-like perception, minimal chromatic presented in the illuminant color matching.

2.3.3 Subtractive Filter Models

Beck proposed that Metalli's transparency can happen in subtractive colors [106]. A filter based model with only first order reflectance was proposed [107]:

$$\begin{aligned} d &= f + (t^2 a)/(1 - fa) \\ c &= f + (t^2 b)/(1 - fb) \end{aligned} \tag{2.25}$$

where t, f is the transmittance and reflectance of the filter. a, b are the reflectance of the background. d, c are the apparent reflectance of the transparent regions p, q respectively in Figure 2.11a. The

squared transmittance t^2 captured the fact that the light is transmitted twice through the filter: once before reflected by the background and once after. He suggested from the experiment that when reflectance of the overlapping area are higher then the non-overlapping area, transparency maybe not be perceived. Minimum 4 sections are required to determine filter parameters.

D'Zmura et al. developed a convergence model describing the chromatic changes of the color behind the filter as translation and convergence in DKL color space without requiring the systematic luminance change [108]. The model is verified with psychophysical experiments [109] and applied to predict color behind the filter [110].

Another scaling filter model was proposed by Faul and Ekroll based on the physical internal and external reflectance [111]. :

$$\begin{aligned} P_i &= \tau_i(A_i + \delta I_i) + \mu\delta I_i \\ Q_i &= \tau_i(B_i + \delta I_i) + \mu\delta I_i \end{aligned} \tag{2.26}$$

where A, B, P, Q are the cone excitation of the four regions in Figure 2.11a. i denotes L, M, S cone. τ is the transmittance factor, roughly represents the squared total transmittance. δ represent the amount of direct reflection and μ the indirect reflection. I is the color of the illumination. For the simplified four luminance case, μ can be set to 0 and I is estimated from the backgrounds $I = (A + B)/2$. It was also illustrated that this filter model works on chromatic transparency. The model was then associated with visual attributes in their later work [112]. The triplet transmittance factor τ is associated with hue (H), saturation (S), and transmittance (V), while direct reflection factor δ is associated with clarity (C). The model was demonstrated to be able to predict the transparency layer constancy in change of background reflectance and illumination through asymmetric matching experiment [113]. In Faul's newest research, the four attributes are transformed into a more uniform,

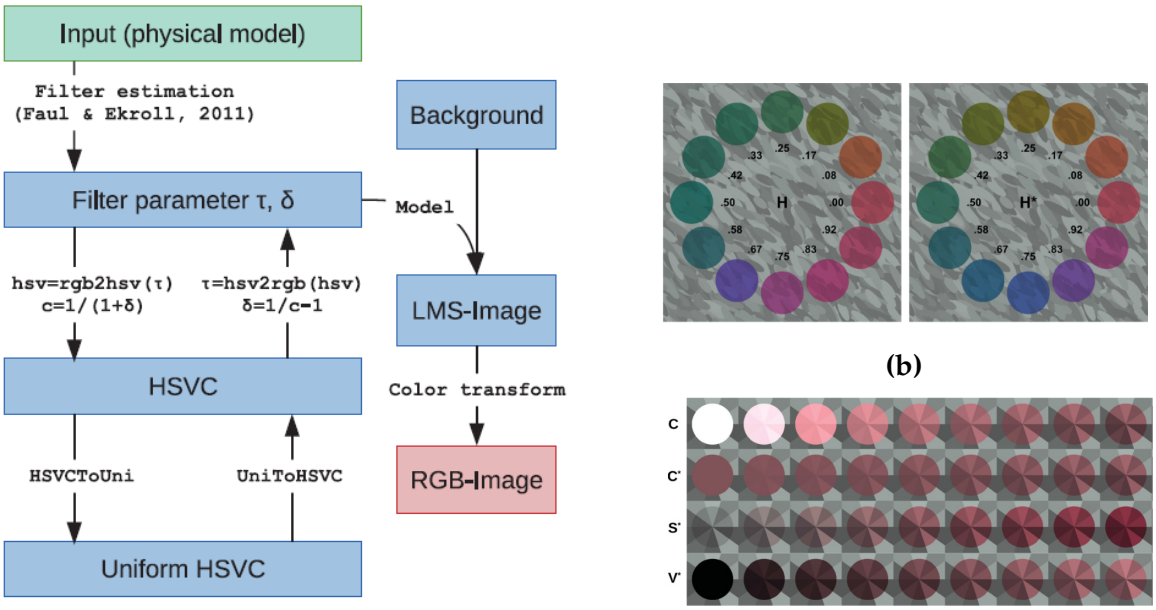
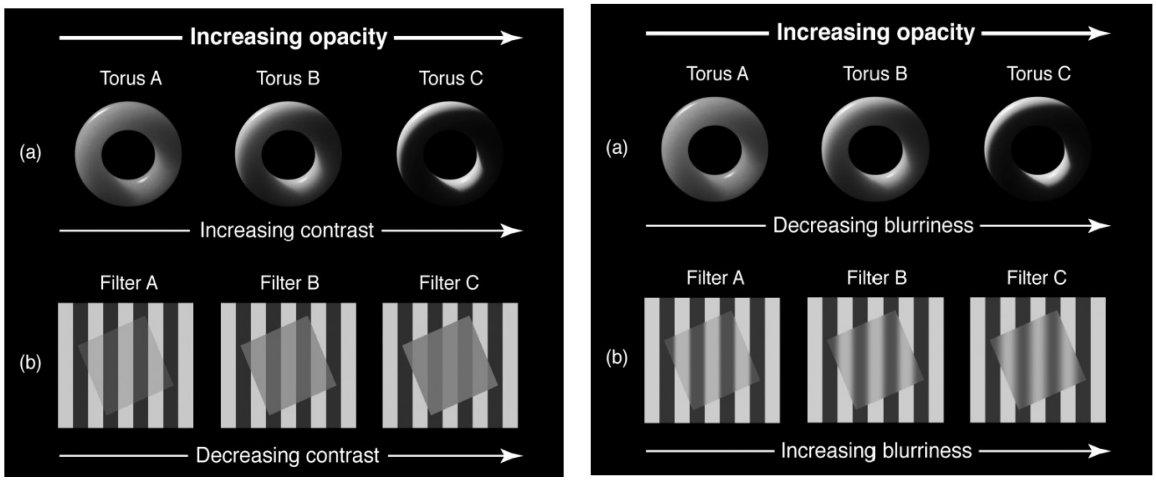


Figure 2.12: Subtractive filter appearance model from Faul and Ekroll [112, 114]. (a) Model parameters τ and δ were transformed to a uniform and independent HSVC space associating with perceptual attributes. (b) Hue uniformity before (left) and after (right) correction. (c) Native clarity parameter C compared to independent and perceptually more uniform scales: clarity $C*$, saturation $S*$, and transparency $V*$

independent, invertible and perceptually intuitive space with H*, S*, V*, C*. (Figure 2.12) [114]

2.3.4 3D object transparency

3D object transparency is more complex than thin-film media. In computer graphics, alpha composition is commonly used to achieve perceived transparency by blending the object color with the background colors. More sophisticated methods use ray tracing to simulate physical light paths including transmitted, scattered, and reflected rays. Both internal and external factors have been demonstrated affecting the perceived transparency or translucency. The same image property like



(a) **(b)**

Figure 2.13: The same image property could change the perceived transparency towards opposite direction on 3D and thin-film media. Images from Fleming [92]. (a) The increased opacity required increased contrast on the 3D torus but decreased contrast on the filter. (b) The increased opacity required decreased blurriness on the 3D torus but increased blurriness on the filter.

contrast or background blur can result in the opposite transparency perception on thin-film media and 3D object depending on the other factors (Figure 2.13[92]). When a non-uniform background is presented, factors impacting perceived transparency include background blur, contrast, color shift, relative motion [91]. Without the presence of the background, transparency, and translucency sensation can still be evoked through object shape, surface roughness, specular, caustic [93], spatial shading pattern, lighting direction and geometry. [92, 93]. The color matching experiment by Ennis and Doerschner demonstrated the curved 3D transparency object color is mainly driven by the ratio of the mean cone excitation on the filtered and unfiltered background region [115].

2.4 Summary

Lightness, brightness, and transparency have been studied through the past two centuries. They were studied through traditional media like physical samples or display monitors. There are well-developed and tested models on all three aspects. Numerous systematic appearance models, now serve as standards in related fields, correlates lightness and brightness with the perception. Supplemental models were also developed for stimuli under certain context approached parametrically, psychologically, or physiologically. For the transparency perception, the interest has moved from the traditional thin-film media to real 3D object, from subtractive filters to new media like AR, from rigorous lab controlled environment to more sophisticated natural environment. Some recent studies have tried applying existing models and predictions on colors in AR, but few studies focused on measuring the perception directly before applying the models. The ideal of layers appeared in both lightness and transparency and matches well with the background-virtual color layers in OST AR. The separation of lightness and transparency in simulated transparent target are also very relative to our research. These studies offered great inspiration and foundation for this research.

Chapter 3

Experimental Environment

For the purpose of our experiment, we required an OST AR system with accurate control on presented color, size, and resolution. Simple optical structure were preferred to minimized artifact like glare, ghost images, channel cross-talk. The digital background from display monitors enabled easy test on the background blending. The system had the flexibility for swapping out the optical combiners and display monitors of difference properties for different tasks.

3.1 System

The desktop AR system is custom-built adopting the flat reflective optical combiner as in Figure 2.3((a)) instead of diffraction grating to reduce the spatial non-uniformity and channel breakups. It uses a flat beam splitter at 50° as the optical combiner in front of a viewing booth, as shown in Figure 3.1. The beam splitter reflect about 30% of the light and transmit about 70%. The light engine is an upward facing 27-inch display monitor (TRU-VU SRM-27R-12), namely the AR display, with resolution of 1920×1080 at the bottom of the box. The focal distance is decided by the distance between the optical combiner and the light engine display. The 27-inch background display (Dell

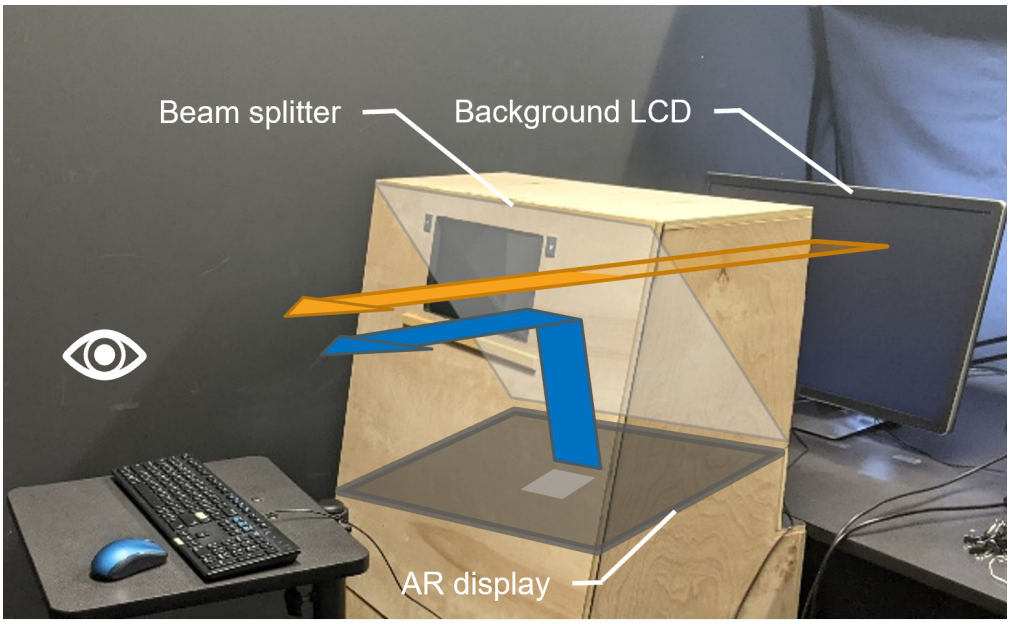


Figure 3.1: Experiment setup showing a cutaway of the optical blending of the image from the background LCD through the beam splitter (orange arrow), and the image from the AR display reflected by the beam splitter (blue arrow).

P2715Q) with the same resolution, namely the background LCD act as the real-world background. The two displays were set to be at the same focal distance at 110 cm from the designated observer's eye position, so they appeared approximately co-planar at the same depth. This was to minimize the vergence-accommodation conflict and create a conformable viewing experiment when switching between the AR and the background. The AR rendering generated from the bottom display monitor was reflected by the beam splitter (Figure 3.1 blue arrow). The background LCD was also visible through the beam splitter (Figure 3.1 orange arrow). Both the AR rendering and the background are combined and visible through viewing port, occupying roughly 25°x13° field of view (FoV).

3.2 Depth and Alignment

Since the rendering from the AR display and the background LCD were viewed simultaneously, it is important to ensure they are at the same depth to minimize the vergence-accommodation conflict and create conformable viewing experiment when switching between the two displays. In this setup, the viewing distance was set to be 110 cm from the designated observer's eye position. We visually minimized parallax to ensure the AR display and the background LCD presented images at the same depth. A white circle with diameter of 60 pixels (px) was positioned at the center of the background LCD. Meanwhile, a red circle of the same size (60 px diameter) was presented at the center of the AR LCD. We adjusted the background LCD distance so that there is no relative movement between the two circles with left-right head movement (parallax free). Thus, the two displays appeared approximately co-planar at the same depth.

The depth calibration only locked the center of the display aligned at the same depth. Since the background LCD could swivel, aligning the center did not grantee the alignment for the rest of the screen. To ensure the alignment between the two displays, a 4×2 alignment grid with unit size of 300×350 px were created on the background LCD positioned at the center. On the AR display, haircross cursor was moved to select the aligned pixel to the grid corners for a list of corresponding pixel locations. The selection was repeated 5 times and the average values were recorded. The average x and y values of the 4×2 corresponding pixels are used as the centers of the screen for the background LCD and the AR display, respectively. An average pixel size of a grid unit on the corresponding pixels were calculated on the AR display. The background LCD angular resolution was calculated with the display size (598 mm horizontal, 338 mm vertical) and viewing distance (1100 mm), resulting angular resolution of 0.0158°/px horizontally and 0.0162°/px vertically. The angular size of the alignment grid unit (300×350 px) calculate 4.7521°x5.6612°. The grid unit angular size combined with the corresponding grid unit pixel size was used to calculate the AR

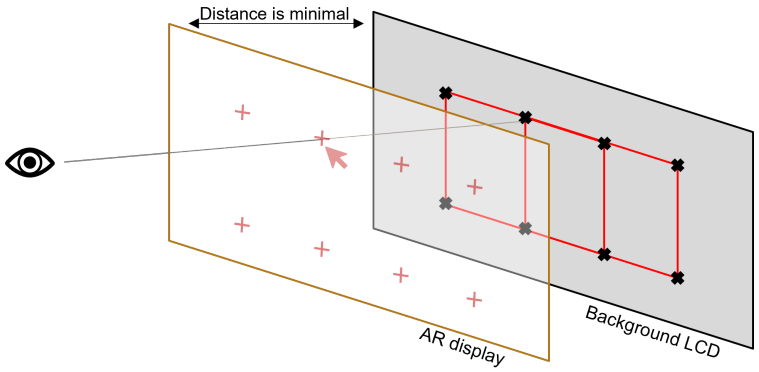


Figure 3.2: Display alignment with 4×2 grid. Cursor movement was constrained to the AR display only. It was moved to select pixels aligned with the grid corners on the background LCD. The distance between the two displays are enlarged for visualization only.

display angular resolution. The AR display resolution was 0.0157°/px horizontally and 0.0162°/px vertically. The center points and the angular resolution on the AR display and background LCD were used in the experiment to align stimulus and backgrounds.

3.3 Color Characterization

The system characterization was performed before the experiments for color management. The system was treated as a 6-primary display where 3 were from the AR display, and the other 3 from the background LCD. A rectangle target occupying 2° FoV surrounded by black was displayed at the center of the AR system. The target spectral radiance was measured at center by a spectroradiometer (Konica Minolta CS-2000) with 1° aperture. A ramp with input from 0 to 255 was measured for each of the 6 primaries. The dark measurement combined from the AR display, background LCD, and room ambient light are subtracted from the measured primary SPD (Figure 3.3). A set of 50 random combination of the 6 primaries were generated for verification and measured with the same setup.

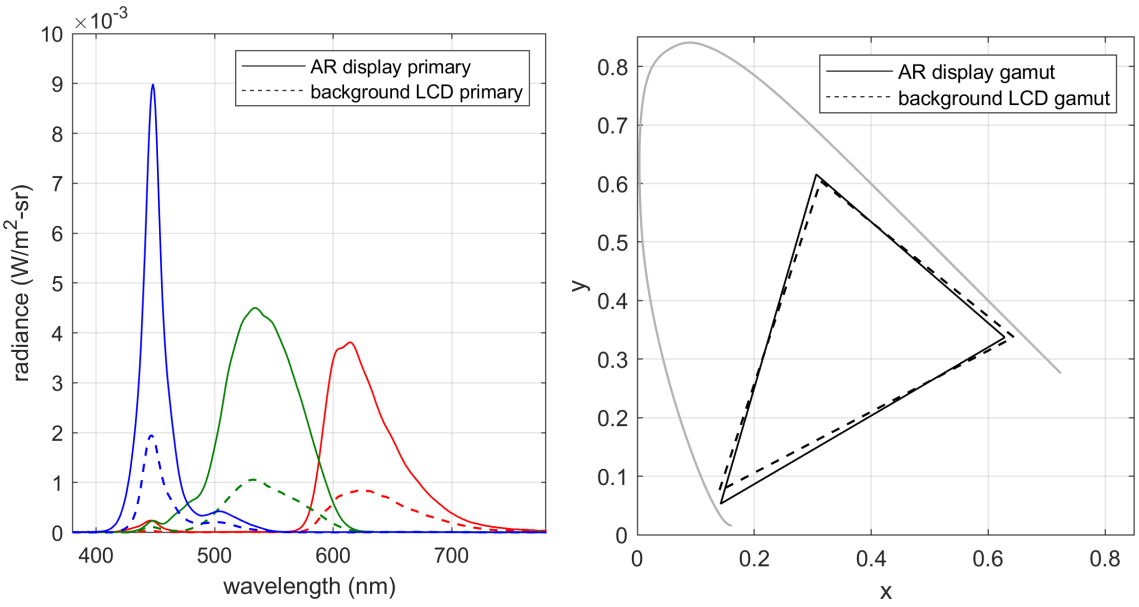
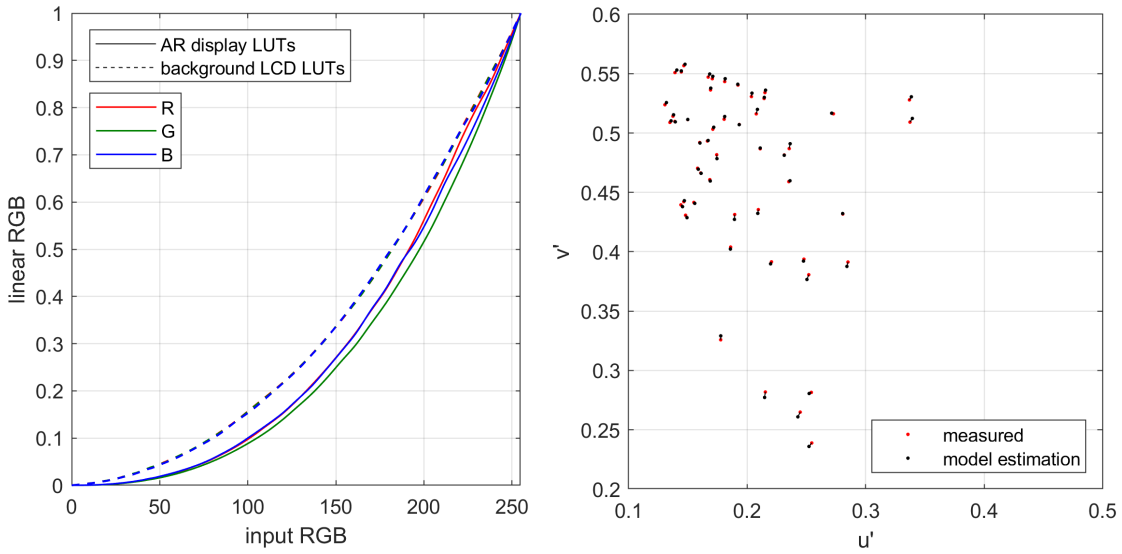


Figure 3.3: AR system characteristics. (a) Spectral power distribution of the 6 primaries. (b) Chromaticity gamut of the AR display and the background LCD calculated with CIE 2015 2° cone fundamental-based color matching functions. The gray curve is the spectral locus.

The CIE 2015 2° cone fundamental-based color matching functions [116] were used to calculate the tristimulus values from the measured SPDs. The primary peak luminance of the AR display are 66.4 cd/m^2, 187.1 cd/m^2, 15.4 cd/m^2 for the red, green, blue channel respectively. The peak luminance of the background LCD are 14.5 cd/m^2, 42.1 cd/m^2, 5.2 cd/m^2 for the red, green, blue channel respectively. A model with two 3×3 transformation matrixes and 6 one-dimentional lookup tables (LUTs) was used to predict the combined AR system color [117]. The 6 input RGB values are first converted to linear RGBs using the 6 LUTs as in Figure 3.4((a)). The two matrices, correspond to the AR display and the background LCD normalized to their white point respectively, then



(a) **(b)**

Figure 3.4: Characterization model and verification. (a) 6 one-dimentional LUT. The three curves on the background LCD are almost identical and overlap with each other. (b) Model prediction verification on 50 random samples with average CIE DE2000 of 0.84 and luminance difference of 1.4 cd/m^2

transform the linear RGB values to tristimulus values:

$$M_{AR} = \begin{bmatrix} 0.6285 & 0.3060 & 0.1425 \\ 0.3370 & 0.6155 & 0.0528 \\ 0.0345 & 0.0785 & 0.8046 \end{bmatrix}, M_{BG} = \begin{bmatrix} 0.6439 & 0.3136 & 0.1400 \\ 0.3374 & 0.6040 & 0.0744 \\ 0.0187 & 0.0824 & 0.7856 \end{bmatrix} \tag{3.1}$$

The model estimation on the 50 verification samples showed good prediction 3.4((b)) with DE2000 of 0.84±0.49 and luminance difference of 1.4±1.2 cd/m^2 between the measured and predicted samples. Due to the background blending, the verification samples did not cover the full gamut.

The spatial variance of the system were examined by measuring additional targets in the system surrounding of the center target, described above. Two spots were measured at the top and the bottom of the center, deviating by 3° (Figure 3.5). Another two spots were measured at the left and

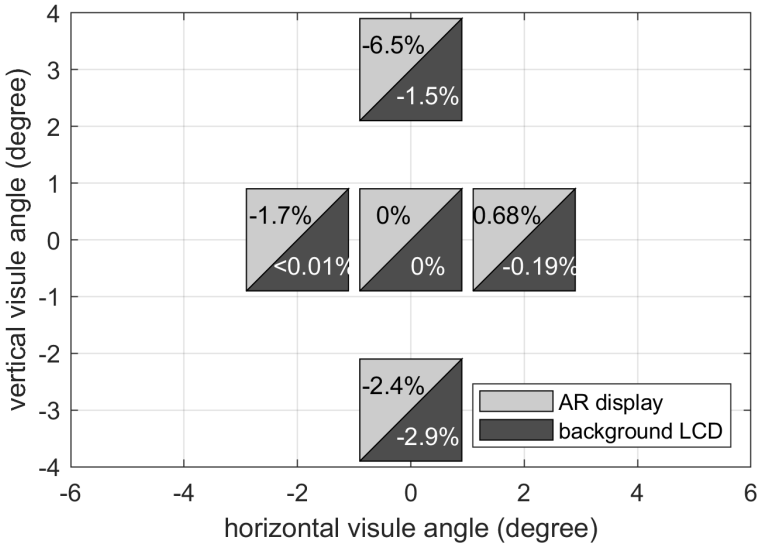


Figure 3.5: Spatial variation measured in peak luminance deviation percentage from the center target. Each square represented one measurement location, with the two respective numbers on the AR display and background LCD.

the right of the center, deviating by 2°. The peak luminance is mostly impacted on the AR display by the vertical relative position, with the top and the bottom peak white luminance dimmer than the center by 17.2 $\mathrm{cd/m^2}$ (6.5%) and 6.4 $\mathrm{cd/m^2}$ (2.4%) respectively. The left is dimmer than center by 4.4 $\mathrm{cd/m^2}$ (1.7%) and the right brighter by 1.8 $\mathrm{cd/m^2}$ (0.68%) on the AR display. The background LCD had better spatial uniformity with the peak luminance deviation less than 1 $\mathrm{cd/m^2}$ for the left, right and top location, and dimmer at bottom by 1.7 $\mathrm{cd/m^2}$ (2.9%). Due to the larger vertical luminance variation, we placed our stimulus mostly at the center area along horizontal directions.

The experiments were conducted in a dark room. Both the background LCD and the AR display white point were set to be D65 for the CIE 2015 2° observer. Behind the background LCD were two lamps illuminating the black drop cloth or white board, depend on the experiment to create a more natural viewing condition than completely dark [118]. The lamps used the Philips Hue bulbs with tunable color profiles. We tuned the HUE bulbs so that the measured the chromaticities of the illuminated board matched to the background LCD white point (D65). The brightest area visible to

the observers also matched the background LCD white point luminance to extend the digital background to the real world. On the black drop cloth, the brightest area visible to the observers were at $5.3\,\mathrm{cd/m^2}$. The ambient light level was low for our setup, comparing to the real user cases. But this is limited by the background display peak luminance.

Overall, the system showed capability of overlaying AR rendering over a large luminance range on the digital background, with peak luminance over 4 times of the background peak luminance. The colorimetric performance was reasonably good, especially considering our stimuli were achromatic. There was some spatial variation, but we reduced its effect by placing stimuli strategically.

Chapter 4

Lightness and Brightness

4.1 Introduction

The objective of the research in this chapter was to establish brightness scales for rendered stimuli in OST AR. Two experiments were included: brightness partition scaling, and diffuse white estimation. In both experiments, backgrounds varied in luminance, spatial pattern, and contrast levels. Brightness scaling experiments aim to build relative brightness scales from partition scaling with bisection tasks on various background conditions. The diffuse white estimation experiment aims to find the appropriate luminance for AR rendering to appear as a reflective white object. Both 2D patches and 3D cubes were included in this experiment to examine the rendering dimensionality effect. The result also help anchor the relative brightness scale from the first experiment, building lightness scales. The combined result from the two experiments offers insight on the perceived lightness, brightness, and 3D object appearance in OST AR.

4.2 Brightness Partition Scaling

4.2.1 Methodology

Partition scaling

Partition scaling is a psychometric method for constructing interval scales from equal-appearing perceptual intervals [119]. One approach to partition scaling is through bisection tasks. A pair of references are given, and a third stimulus is adjusted to be the mid-point between them, giving two equal intervals from the triplet stimuli. This is called a subdivision level. In the first subdivision level, the two endpoints of the set range presented as the references. The observer-adjusted bisection point was recorded to generate the second subdivision level with two set of references: the lower endpoint – bisection point, and the bisection point – the higher end point. The median of the repeated trials was used as the bisection result (new reference) and the bisection continued to the next subdivision level. Three subdivision levels were used in the experiment (Figure 4.1). The bisection was repeated 5 time on the first level, 3 times on the second, and 2 times on the third. At the end of the third subdivision level, nine points of equal brightness steps were generated, including the initial reference points for each sub-range.

Background and stimulus specification

The experiment stimulus and background were set to be monochromatic for the CIE 2015 2° observer [116]. The chromaticities of the achromatic stimulus were set to match the standard illuminant D65 at (0.3134, 0.3308). The white board backdrop illuminated with two lamps bidirectionally. The radiance reflected by the board was calibrated to match the background LCD white chromaticities with luminance of $58\,\mathrm{cd/m^2}$ by tuning the HUE bulb profiles. We used solid and sinusoidal gratings as the background on the LCD, as shown in Table 4.1. Three solid gray backgrounds at

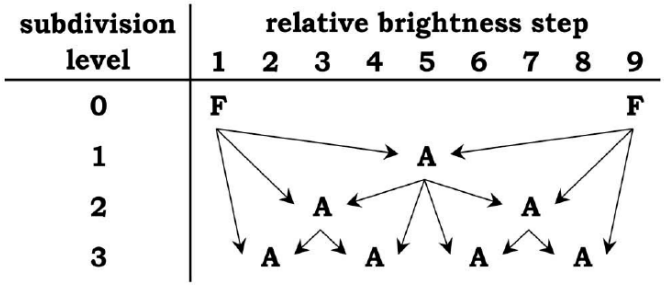


Figure 4.1: Partition scaling method: 3 levels of subdivision result in 9-equal-interval steps from a pair of reference from Perre et al. [119] **F** is the initial reference in the first subdivision level, and **A** notes all bisected points used as the new references for the next subdivision level.

Table 4.1: Partition scaling background specification

Background	LL	ML	HL	SLL
Type	solid			sinusoidal
Contrast	-	-	-	0.5
Avg luminance (cd/m^2)	2	9	50	2

low luminance (LL), medium luminance (ML), and high luminance (HL) at $2\,cd/m^2$, $9\,cd/m^2$, and $50\,cd/m^2$ respectively. An additional patterned background (SLL) had the same average luminance as LL at $2\,cd/m^2$, but in sine wave at 3 cpd and Michelson contrast of 0.5. The backgrounds covered 6°x15° field of view, surrounded with 20% medium gray and repeated 6-step neutral scales for reference (Table 4.2).

Table 4.2: Surround and neutral reference specification

	Surround	Neutral reference scale					
		1	2	3	4	5	6
Luminance (cd/m^2)	11.7	1.3	5.2	12.4	23.3	38.5	58.3

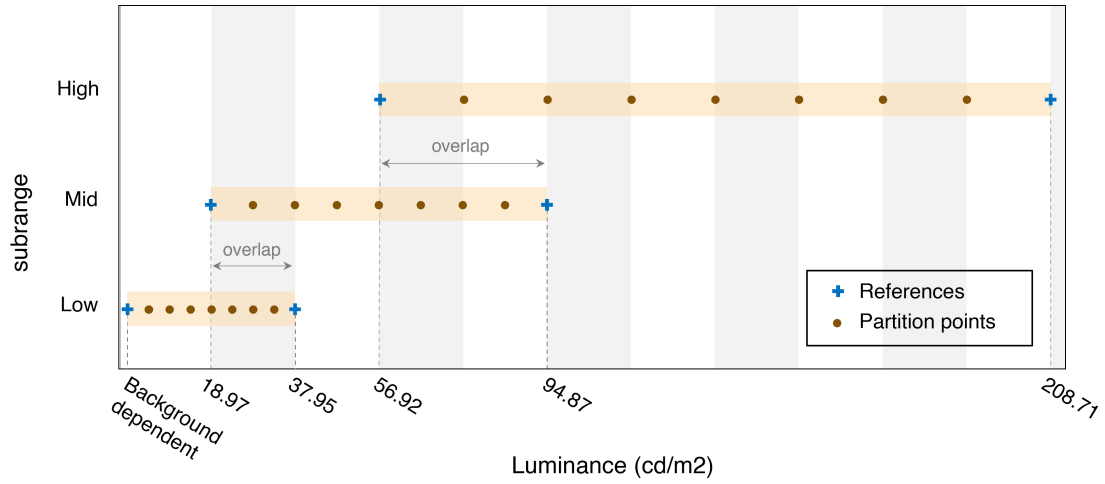


Figure 4.2: Unbalanced subranges used in the experiment compensating for the increased JND with higher luminance. Each subrange started with the two end points as the reference and produced 9 equal-brightness-interval points after 3 subdivisions.

Three unbalanced luminance subranges were used in the experiment to improve the bisection accuracy, as in Figure 4.2. The luminance specification was from the AR display only. Bisection tasks are retained within each subrange. The initial reference pair for a bisection task was the two ends of each subrange (Figure 4.1 F in subdivision level 0). The luminance range of each subrange covers 200% of the lower sub-range luminance range. Sub-ranges also overlap, with 50% of luminance on the lower subrange. The lowest reference luminance level was set to be at the incremental brightness JND from the background using Barten contrast sensitivity function [120] to ensure the visibility of the reference patches. The minimum AR luminance were $0.14\,\mathrm{cd/m^2}$, $0.60\,\mathrm{cd/m^2}$, $3.32\,\mathrm{cd/m^2}$, $0.14\,\mathrm{cd/m^2}$ on the background LL, ML, HL, SLL respectively. The highest luminance level was set to $208.7\,\mathrm{cd/m^2}$).

Procedure

The experiment was conducted in a dark room. In the experiment, the observer adapted to the surround for 1 minute first. A triplet of patches each in 2°x2° squares generated from the AR overlay were presented in the center of a background with 0.5° separation (Figure 4.3). The left and the right patches were the reference pair, and the center patch was the test patch for brightness adjustment. The location of the reference pair on the left and the right was randomized. The observers were asked to adjust the center patch brightness using the mouse scroll wheel, so that its brightness was halfway between the left and the right reference patches. The initial luminance of the center patches was a random value between the two references. The luminance adjustment range was also bounded by the references. All trials on the first subdivision level on all background were randomized and assessed before moving to the next subdivision level. Each observer assessed 228 trials in total. Observers went through a training session before the experiment comprised of only background LL in low subrange to familiarize with partition scaling. Seventeen observers participated in the experiment, of whom 9 are males and 8 are females.

4.2.2 Result

Trial duration and difficulty

We examined the time spent on adjusting the test patch for the bisection task and used it as an indicator of trial difficulty (Figure 4.4(a)). Trial duration outside three standard deviations from the mean was excluded, as they could be breaks observers took. Twelve trial duration were removed, which takes 1.3% of all trials. The mean duration spent on adjusting one trial is 11.7s and median 10.5s. We ran a 3-way analysis of variance (ANOVA) on the time spend on a trial against subrange levels, subdivision levels, and backgrounds conditions. The result showed that the higher subdivision level resulted significantly shorter trial duration ($p < 0.01$) as shown in Figure 4.4(b). As

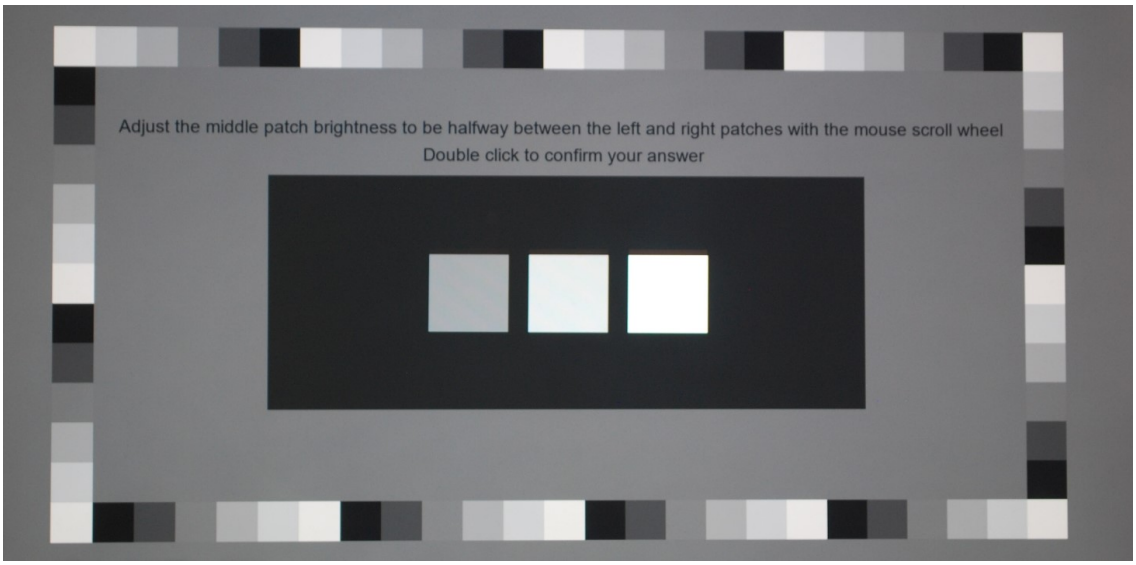


Figure 4.3: Photograph of the partition scaling experiment user interface. Observers adjusted the center patch brightness to be halfway between the left and the right reference patches.

a difficulty indicator, it showed that as the partition division progress to smaller brightness interval between references, the bisection task become easier and less time was required to complete a trial. Although some observers also reported after the experiment on indistinguishable brightness difference among the three patches in the higher subdivision levels, suggesting that at the 3rd levels of bisection may have reached the brightness JND. Longer trial duration was also found in the low luminance subrange compared to the mid and high subranges ($p < 0.01$, Figure 4.4(c)). The mid and high subrange average duration are not significantly different from each other. Although we set unbalanced subranges to compensate for the difficulty difference among subranges (see Figure 4.2), it persists in some trials. There is no significant effect from either background luminance or pattern (Figure 4.4(d)).

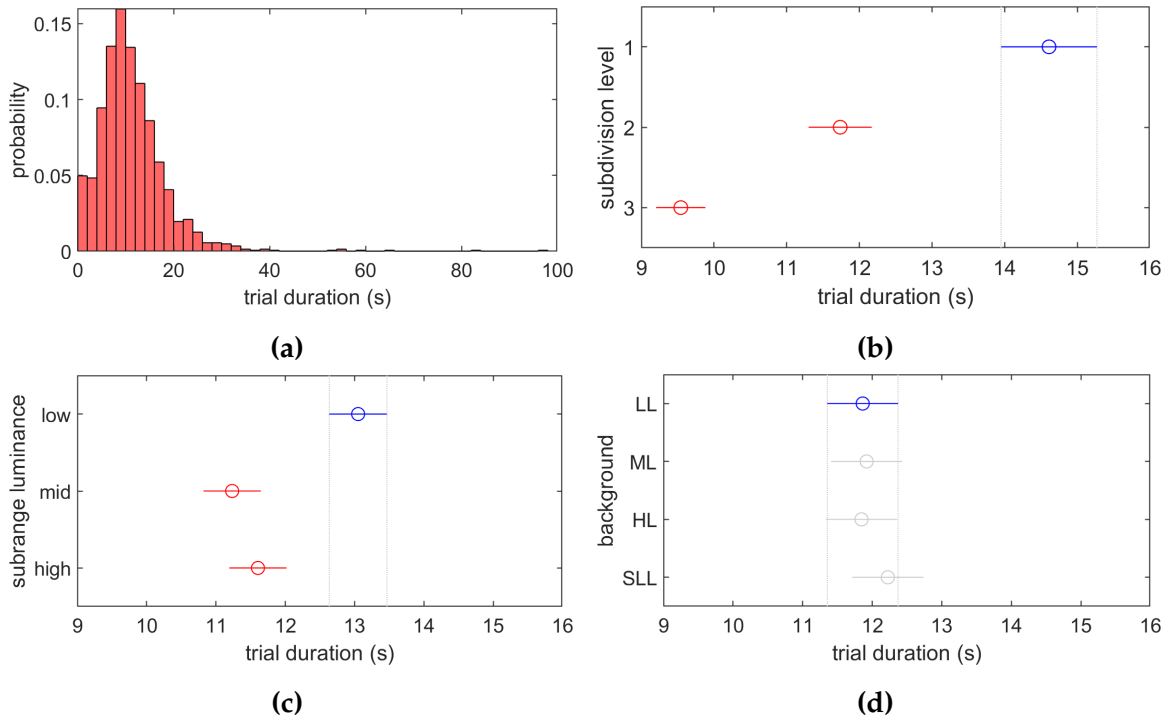


Figure 4.4: Trial duration in the partition scaling experiment. (a) Histogram on all observers' trial duration. (b) Anova on trial duration against subdivision level. (c) Anova on trial duration against subrange luminance level. (d) Anova on trial duration against Background condition

Table 4.3: Individual power law fitting parameters b value.

Background	SSL	LL	ML	HL
b	0.3375	0.3556	0.4418	0.7063

Individual power law pooling and fitting

We averaged the two repeated trials in the last subdivision levels. The four bisected points along with five references generates the final nine points of equal brightness steps in each subrange and background. The nine equal brightness points are marked as relative brightness from 0 to 1 with interval of 0.125. Data from all observers were pooled together for fitting and stitching subranges. A general model of brightness bases on Stevens's power law [53] is used for fitting the subrange curves:

$$Q_{bg} = a{L_{AR}}^{b} + c \tag{4.1}$$

, where Q is the relative brightness for a background condition, and L_{AR} is the observer adjusted AR overlay luminance. Since the brightness scale we are building are interval scales, each subrange can be considered as a section of the brightness scale, and rescaled to 0-1 as in the original data. The model parameter reflected it with a and c subject to each subranges, stitching the subrange data together for a continuous smooth curve. Parameter b is subject to each background, reflecting the overall brightness scale for the background condition. The three parameters were optimized for each background minimizing two objective functions at the same time: 1) the brightness prediction error in the two overlapping region across the three subranges as in Figure 4.2. 2) the brightness prediction error for all observers' rating scores. The fitted exponent b can be found in the Table 4.3. The fitting performance was evaluated in coefficient of determination R^2 to the original data and root-mean-square error (RMSE) (see Table 4.4)

The pooled brightness scales are normalized to 0 to 1 as shown in the Figure 4.5. The low,

mid, and high subranges are marked in red, green, and blue respectively. Individual observer's data are marked as large transparent circles. We calculated the standard error of the mean (SEM) of observers for each trial condition for the inter-observer variability. The mean SEM is 2.03 cd/m2, median 1.54 cd/m2, and maximum 6.07 cd/m2 (in background HL with high luminance subrange). ANOVA showed no significant impact on SEM from subdivision level, background condition, or subrange, with $p > 0.05$ for all factors. An average observer's data is obtained from the pooled subranges by averaging all observer's luminance data and marked as the small opaque circle in Figure 4.5 with SEM as the error bars.

The curvature difference among the four background conditions indicates that the background had impact on the relative brightness scale. When the background luminance was low, as in background LL ans SLL, the brightness resolution was higher on the low luminance area. When the background luminance was relatively high as in background HL, the luminance-brightness relation was closer to linear relation. In the pooling and fitting step, the exponent value b captures this feature. We modeled the relation between the exponent b and background luminance L_{bg} on solid backgrounds with a linear relation:

$$b = m\left(\frac{L_{bg}}{L_w}\right) + n. \tag{4.2}$$

The background luminance factor is used relative to the background reference white L_w at $58.3\,\mathrm{cd/m^2}$. m and n are fitting parameters with values of 0.4087 and 0.3587 respectively. The fitted line is shown in Figure 4.6. It indicates that on black solid background, the AR brightness curve has the exponent b of 0.4087. On white background, the b value increased to 0.7674. The background type being patterned resulted in a slightly smaller curve exponent b compared to the solid background at the same luminance.

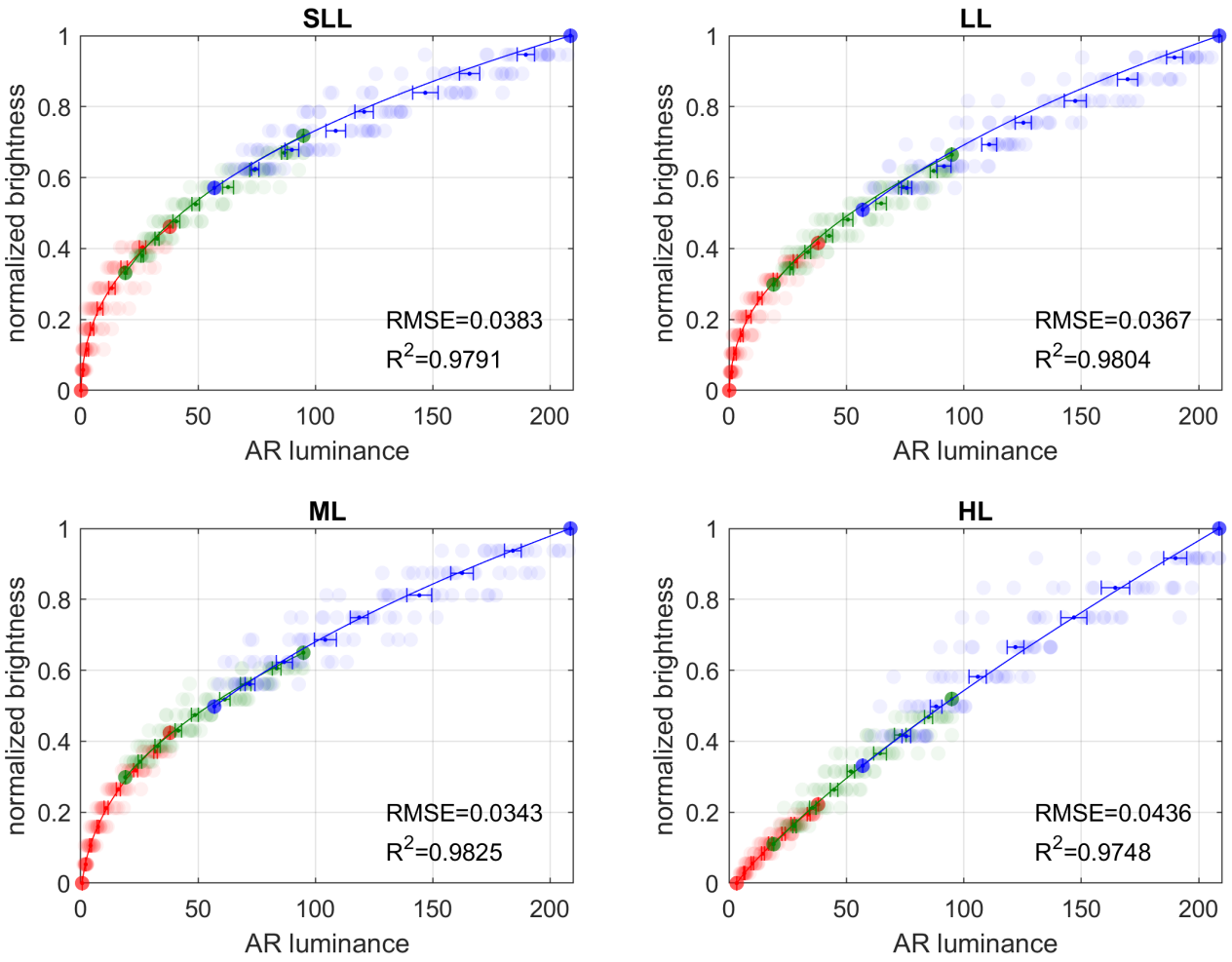


Figure 4.5: Relative brightness scales from pooled and fitted subranges in four backgrounds normalized to 0 to 1. Red, green and blue colors denote the data in low, mid and high subrange respectively. The transparent circles are individual observer results. The small dots are average observer result with 95% CI error bars. The solid line is the fitted model.

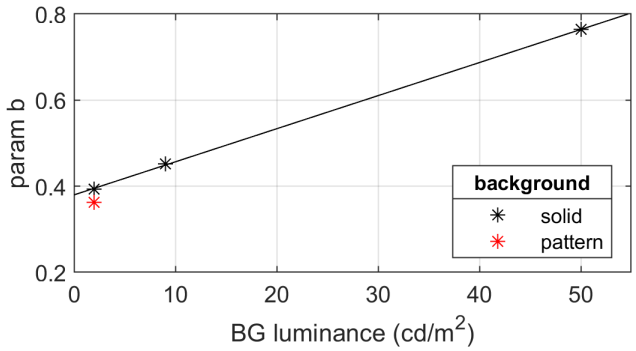


Figure 4.6: Background condition affects the power law fitting exponent value b. b increases linearly with the background luminance. b on the solid background (black star) is slightly higher than it on pattern background with the same average luminance (red star).

Universal power law fitting

Considering the changes on exponent b in Equation 4.1 was mainly modulated by the background luminance, an alternative model bearing the background luminance as the independent variable could result in one universal model on all background luminance instead of individually. So, the proximal luminance of the patch were used as the independent variable instead of the AR luminance. Proximal luminance L_{prox} was the additive luminance summation from the background and the AR overlay. The universal model was described as

$$Q = a' {L_{prox}}^{b_{uni}} + c' \tag{4.3}$$

, where Q is the relative brightness for regardless of background conditions normalized to 0-1. Parameter a' and c' are subject to each background, and parameter b_{uni} is universal to all backgrounds. The three parameters were optimized minimizing the error on overlapping proximal luminance regions for backgrounds, and error to observers' data. Because SLL and LL had the same background average luminance input to the proximal luminance, only LL was included in parameter optimization to avoid over-weighting the low luminance background conditions. The optimization was only run

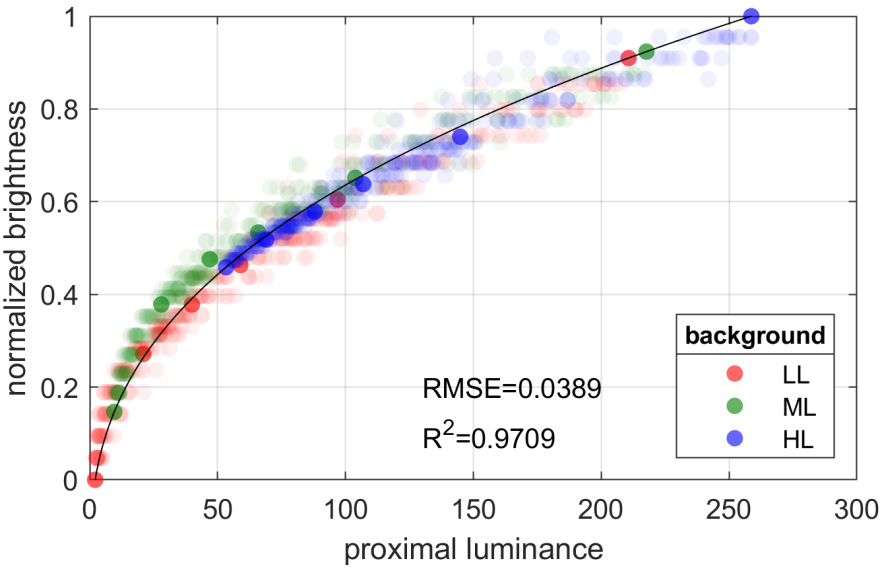


Figure 4.7: Universal power law fitting on proximal luminance. Relative brightness scales were pooled from the different backgrounds and normalized to 0 to 1. Red, green and blue colors denote the data from background LL, ML, and HL, respectively. The transparent circles are individual observer results. The solid line is the fitted model.

on background LL, ML, and HL, but evaluated also on SLL. The fitted exponent b_{uni} was at 0.3856, with overall R^2 at 0.9709 and RMSE at 0.0389. The universal fitting resulted comparable RMSE and slightly worse R^2 comparing to individual fit. The fitting evaluated for each background can be found in the Table 4.3. The increased background luminance decreased the individual model performance (bigger RMSE) but increased the universal model performance (smaller RMSE), which is the major contributor to the overall good fitting. The universal fit failed to capture the curvature, especially in background ML, with RMSE of 0.0439 and R^2 of 0.9526. Also, it could not differentiate the brightness curves in background SLL and LL since the proximal luminance did not incorporate the background spatial information. The universal model still gave good overall estimation and did not need to infer the exponent from the background conditions.

Table 4.4: Model performance comparison on the individual and universal models in coefficient of determination R^2 and RMSE.

Background	R^2			RMSE (e-2)		
	individual model	universal model	universal overall	individual model	universal model	universal overall
SSL	0.9791	0.9685	0.9709	3.83	4.24	3.89
LL	0.9804	0.9716		3.67	4.03	
ML	0.9825	0.9526		3.43	4.39	
HL	0.9748	0.9675		4.36	2.68	

Comparison to CAMs

We compared the experiment brightness scale with popular color appearance model (CAM) lightness and brightness attributes. Four CAMs with brightness or lightness attributes were included: CAM16 brightness Q, CAM16 lightness J, CIELAB L^*, and IPT I. The proximal luminance was used as the target luminance with chromaticities at (0.3134, 0.3308). The background average luminance was used as the background and surround luminance. The display white was used as the reference white point. The viewing condition is set to "Dim". Here, the calculated scales were normalized on the luminance range for comparison. The scale comparison between the experiment data and the CAM prediction are shown in Figure 4.8. The attributes difference between the model prediction and the experiment result in average RMSE are shown in the Figure 4.9. The result of CAM02 were almost identical to CAM16 thus not included in the comparison.

All four CAMs had overall higher RMSE compared to the individually fitted brightness scaled in the previous section, and cannot reflect the scale curvature change from the background pattern. The result showed the CAM16 lightness J performed better than brightness Q on low luminance backgrounds (SLL and LL), but worse on brighter backgrounds (ML and LL). CIELAB L* showed similar trend as CAM16 Q but slightly better on dark backgrounds (LL and SLL). IPT I performed similarly as CIELAB L^*, with prediction on the background LL compare to other backgrounds.

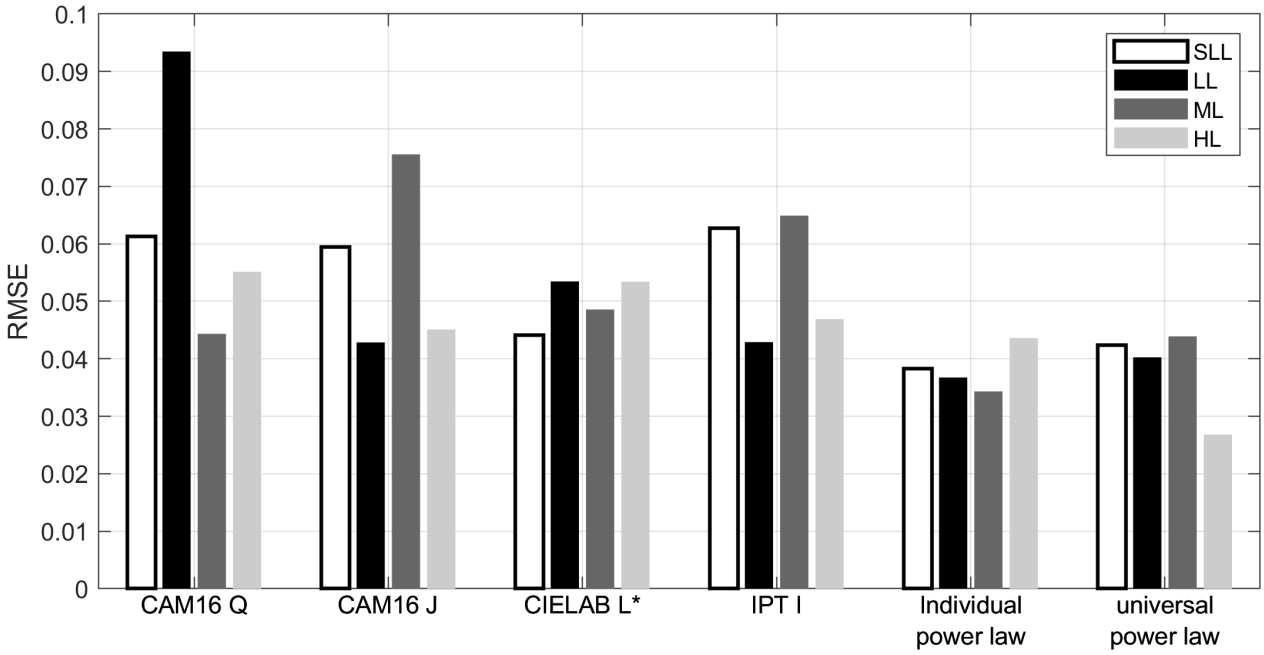


Figure 4.8: CAMs predictions on the observer adjusted proximal luminance compared to the experiment data.

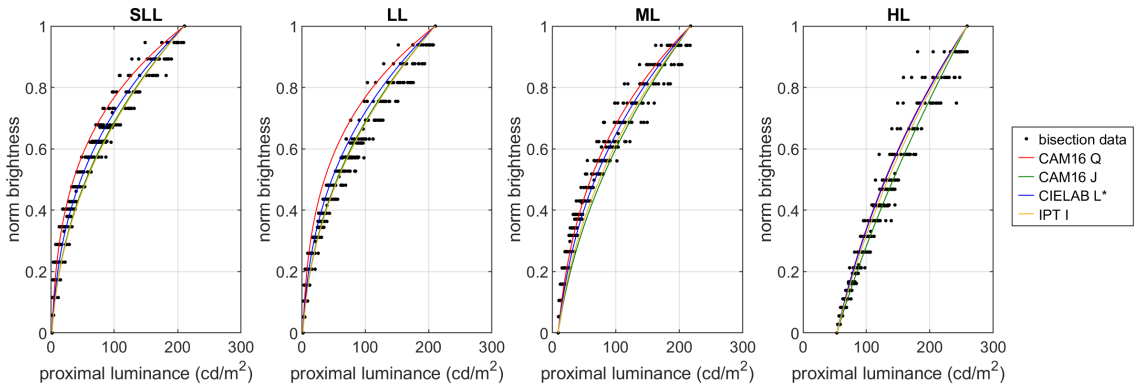


Figure 4.9: CAMs performance compared to the experiment fitted models in RMSE.

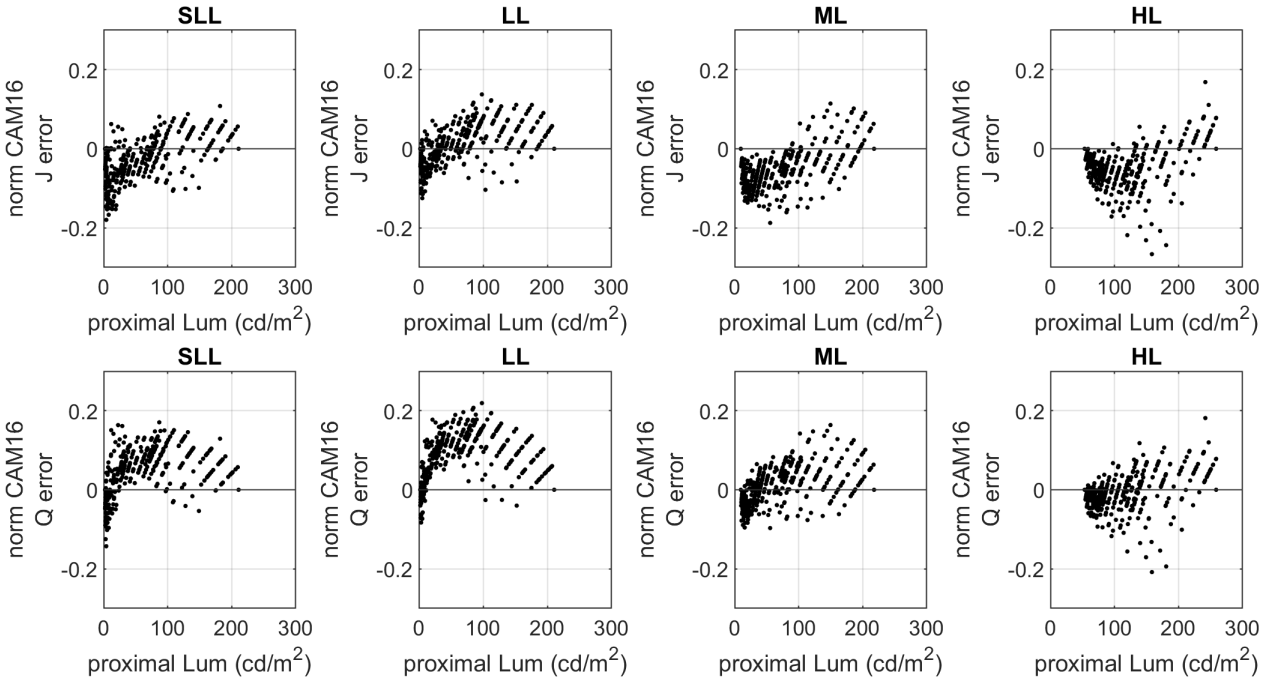


Figure 4.10: CAM16 lightness and brightness prediction error. Neither attribute could estimate the brightness scale accurately over the proximal luminance range and backgrounds.

The common problem in these models, including the universal power law fit, was over- or underestimating the scale curvature (Figure4.10). This was most evident in either lowest or highest luminance backgrounds. In CAM16, the ideal background and surround condition was 20% of display white at 11.6 $\mathrm{cd/m^2}$, which was between the background LL and ML. In backgrounds darker than medium gray, CAM16 J overestimated the scale curvature, and in brighter backgrounds underestimated. More discussion between J and Q can be found in the discussion section. Although the patterned background (SLL) might introduce the transparency appearance to the test patches, the model prediction on background SLL were not systematically worse than on the solid background LL.

4.2.3 Discussion

Error source and propagation

One drawback of the method with bisection tasks is the error propagation. Since the adjusted results serve as the references in the new subdivision level, the variance in the adjustment is also carried to the next level. Trials were repeated with more repeats on early stage of partition scaling to reduce the variance being amplified to the result. The median of the repeated trials was used instead of the mean to prevent deteriorating the bisection data quality from one biased or mistake trial in the repeats. Figure 4.11 illustrated the adjustment SEM largely dependent on both the subdivision level and the step position in the subrange. In this figure, the y-axis was the average observer's SEM from repeated adjustment, with the error bar representing the SEM of observers. Each color represented one subrange. The average SEM increased almost linearly with the adjusted average luminance for each subrange. Inside each subrange, it is obvious that the middle step from subdivision 1 had higher SEM, with subdivision 2 lower, and subdivision 3 the lowest, resulting the zigzag curves. Another source was from the subrange pooling. The pooling was base on minimizing the overlapping region between subranges, which largely affects the resulting scale curvature.

Another model that's commonly used to describe the brightness is based on the generalized Michaelis-Menten equation which is used in CAM02, CAM16, and HDR-CIELAB:

$$Q(bg) = \frac{a(kL_{AR})^b}{(kL_{AR})^b + c} + d \tag{4.4}$$

Since the relative brightness scale was arbitrary in the experiment context and can be scaled, parameter a and d will be omitted in the optimization. We made attempt for both individual and universal fitting to optimize. For individual optimization, the function was very sensitive to the initial points, and could not converge on the overlapping regions error. The universal optimization on M-M equa-

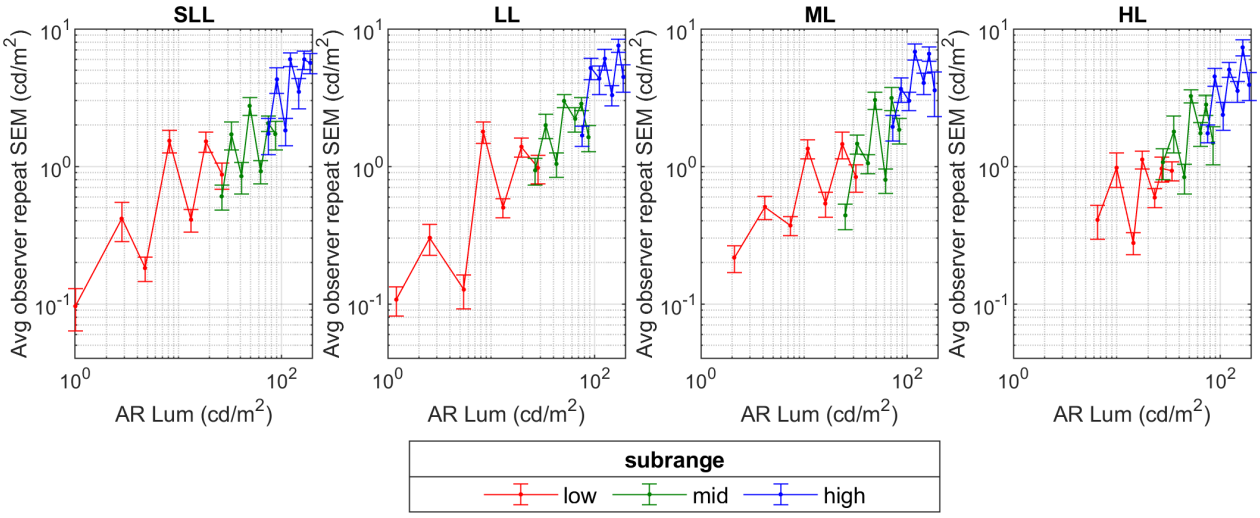


Figure 4.11: Average observer SEM on repeated trials. SEM increased linearly with log AR luminance. The zigzag shape indicates increased SEM with increased subdivision levels.

tion, k and c were replaced by the additional normalization step, resulting the M-M equation being equivalent to the universal power law model.

Below just-noticeable difference reference pairs

Some observers reported difficulty assessing the luminance difference between the reference pair, resulting in the same brightness on the adjusted test patch as the reference pair. This was more frequent on higher subdivision levels as the luminance difference reduced. We examined the average observer on the third subdivision level. The luminance difference between the reference pair were compared to brightness just-noticeable difference (JND) using Barten contrast sensitivity function (CSF) model [120]. The spatial frequency was set to 0.25 cpd, since the stimulus size are 2° visual angle as half a cycle. The background luminance was used as the surround luminance, and the dimmer one in the reference pair as the base luminance. The result was shown in Figure 4.12. For background SLL, LL, and ML, the reference pair luminance differences were well above the JND.

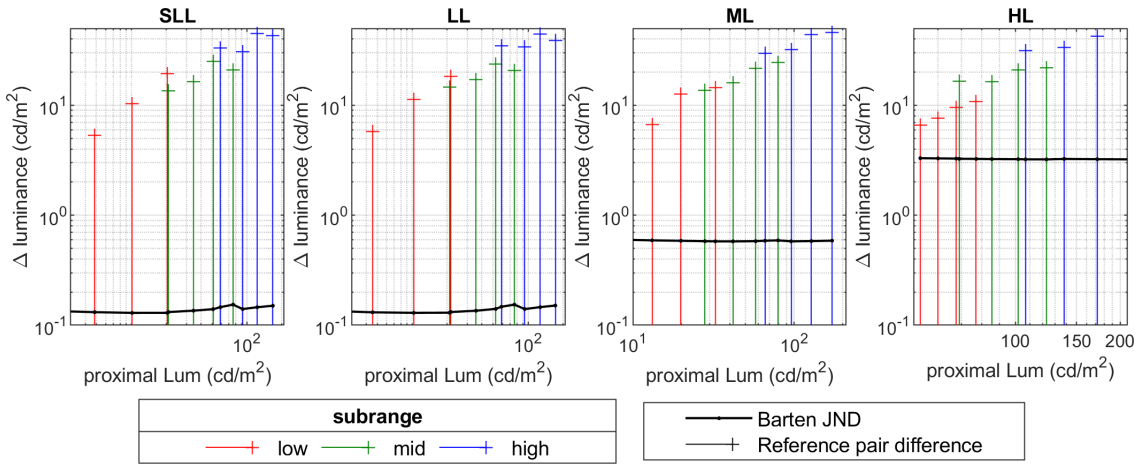


Figure 4.12: Luminance difference between reference pairs on an average observer compared to JND computed from Barten CSF. In some trials on background HL, the luminance difference between the reference patches might be lower than twice of the brightness JND, making the trials difficult to bisect.

But in background HL low luminance region, the reference differences were close to JNDs, thus the intended two equal-interval bisection were likely to be under JNDs. This result was based on the average observer. Considering the observer variability and the subdivision error propagation, some of the third division bisection tasks on the low- and mid-subranges could also have been under JNDs.

Lightness vs brightness in CAMs

In the experiment, we asked observers to bisect brightness. But both brightness and lightness are examined in CAMs. In CAM02 and CAM16, lightness J and brightness Q give opposite performance on the dark background (LL and SLL) and bright backgrounds (ML and HL). The difference between the two attributes came from the additional square root computing Q from J. Lightness J was first calculated from M-M curve based achromatic response A with nonlinearity applied with

exponent cz to compensate for the background and viewing condition effect.

$$J = 100(A/A_w)^{cz} \tag{4.5}$$

With the viewing condition in our experiment (dim surround, various background luminance), the exponent cz were 0.9837, 1.1077, and 1.4258 in the dark to bright backgrounds (LL, ML, HL respectively). But when calculating the brightness Q, an additional square root was applied to the normalized achromatic response, reducing the exponent to 0.4919, 0.5538, and 0.7129.

$$Q = \frac{4}{c}(A_w + 4){F_L}^{0.25}\left(\frac{A}{A_w}\right)^{0.5cz} \tag{4.6}$$

Depending on the prediction from J relative to the experiment data, this reduction improved the performance of Q on bright backgrounds in which scale curvatures were underestimated. But worsened the performance on dark backgrounds, in which scale curvatures were already overestimated. The square root was inherited from the Hunt appearance model conversion from achromatic response to brightness, with the exponent at 0.6:

$$Q_{Hunt95} = N_1\left[7\left(A + \frac{M}{100}\right)\right]^{0.6} - N_2 \tag{4.7}$$

The CAM16 increased z and decreased the exponent from 0.6 to 0.5. But it was not explained in Hunt model why the achromatic response was compressed and how was the parameter chosen.

Summary

We measured the brightness scale of AR stimulus with partition scaling experiment. The result showed that the background average luminance largely affected the brightness curve, while background spatial variance did not. The brightness curve can be modeled with power law functions with the exponent linearly increasing with the background average luminance. Alternatively, a universal

power law model on proximal luminance as the summation of the AR luminance and the background luminance could also predict the brightness scale. The commonly used Michaelis-Menten equation in color appearance models produced slightly worse result with noisier parameters. Existing color appearance model brightness or lightness did not predict the result well.

4.3 Diffuse white level in AR

This experiment aimed to determine the appropriate luminance level for a virtual object on various background conditions to appear as diffuse white or made of diffuse white material. The method of adjustment was used on simple plat patches and 3D rendered cubes.

4.3.1 Methodology

Background and stimuli specification

The similar experiment environment was used as in the brightness partition scaling experiment. Ten backgrounds were examined in this experiment, including the four backgrounds in the brightness partition scaling (Table 4.5). For the solid background, 5 different luminance levels were used at $1.3\,\mathrm{cd/m^2}$, $2\,\mathrm{cd/m^2}$, $9\,\mathrm{cd/m^2}$, $25\,\mathrm{cd/m^2}$, and $50\,\mathrm{cd/m^2}$, noted as L1-L5. For the patterned background, three backgrounds at low luminance (SLL), mid luminance (SM), and high luminance (SHL) at the same contrast of 0.5 were used. The spatial average luminance were at $2\,\mathrm{cd/m^2}$, $9\,\mathrm{cd/m^2}$, $25\,\mathrm{cd/m^2}$ for SLL, SM, and SHL respectively. Their average luminance levels were the same as solid background L2, L3, and L4. Additionally, two patterned background had the same average luminance at $25\,\mathrm{cd/m^2}$ but with lower contrast at 0.1 (SLC) and higher contrast at 0.8 (SHC) respectively. All patterned backgrounds are sinusoidal gratings at 3 cpd. The same surround image and neutral reference scales were used as in the brightness partition scaling (see Table 4.2).

Table 4.5: Background used in diffuse white adjustment experiment.

Background	L1	L2	L3	L4	L5	SLL	SM	SHL	SLC	SHC
Type	solid					sinusoidal				
Contrast	-	-	-	-	-	0.5	0.5	0.5	0.1	0.8
Avg luminance (cd/m^2)	1.3	2	9	25	50	2	9	25	25	25

Two types of stimuli rendered from the AR overlay were used in this experiment: 2-dimensional patches, and 3-dimensional cubes. The 2D patches were 2°x2° visual angle squares, same as in the brightness bisection experiment. The cubes were rendered with OpenGL in Psychtoolbox 4.13. The camera in OpenGL was set to look at the cube horizontally along the -z direction. The cube material was set to diffuse (Lambertian surface), with the material albedo adjustable by the observers. The size of the cube was set to 2°x2° visual angle for each surface. The lighting in OpenGL used a combination of directional light and ambient light with intensity ratio of 3:1. The directional light illuminated the cube from 45° above the scene camera to simulate a light booth viewing geometry. All light albedo were set to (1,1,1), so that the color of the cube is only modulated by the material color. The cube took input from both the mouse scroll wheel and the mouse movement. The scrolling data was taken as the rendering luminance, and converted to tristimulus values with chromaticities at (0.3134, 0.3308) for CIE 2015 2° observer. The tristimulus values were then converted to RGB values from the AR system characterization model and fed to the cube material albedo to adjust the brightness. The two-dimensional mouse movement data were translated into rotation along x-axis and y-axis on the cube in right-handed coordinate. Flat shading was used to render the cube. The luminance of each side of the cube depended on both the rotation angles and the material albedo.

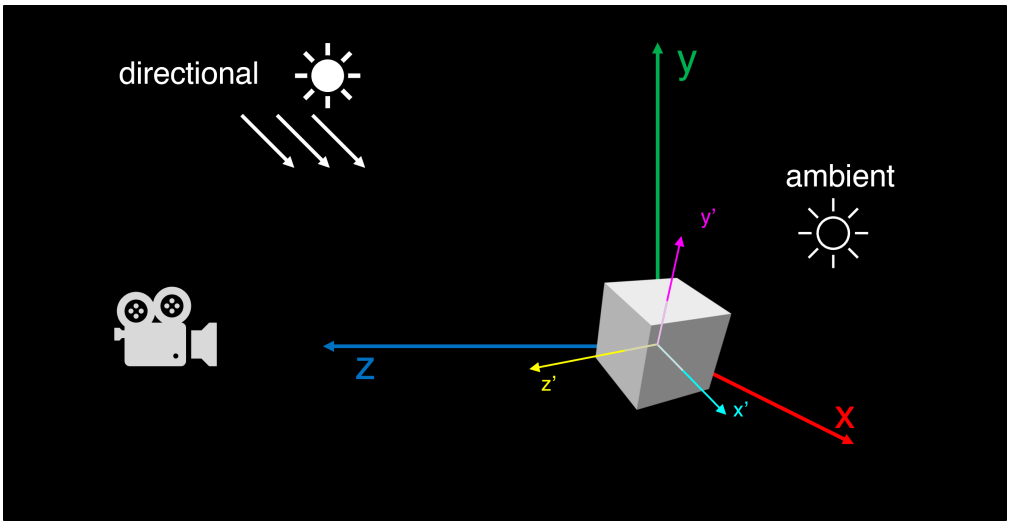


Figure 4.13: OpenGL rendering geometry for the cube. The lighting intensity on the directional and the ambient light was set to 3:1. The cube was at the origin and rotate around x and y axes in world coordinate.

Procedure

Before the experiment, observers were provided with the definition of diffuse white as "a matt reflective white color that is not gray nor emit light". Examples of diffuse white were presented including a sheet of copy paper, a perfect reflecting diffuser (PRD), a semi-transparent film filter, and a foam cube. Observers were also suggested that they could use the small white patch on the surround neutral scale ($58.3\,\mathrm{cd/m^2}$) for reference during the experiment. The experiment is composed with two sessions with a break between them: the 2D patch session, and the 3D cube session 4.14. Before each session, the observer adapted to the surround for 1 minute.

In the first session, a patch at random luminance appear at the center of a background and the observer was asked to adjust the patch luminance using the mouse scroll wheel so that it appears as diffuse white, then double click to confirm. The adjustment was repeated five times on each background. The backgrounds were presented in either a luminance incremental or decremental sequence. The sequence of backgrounds at the same average luminance were randomized, regardless

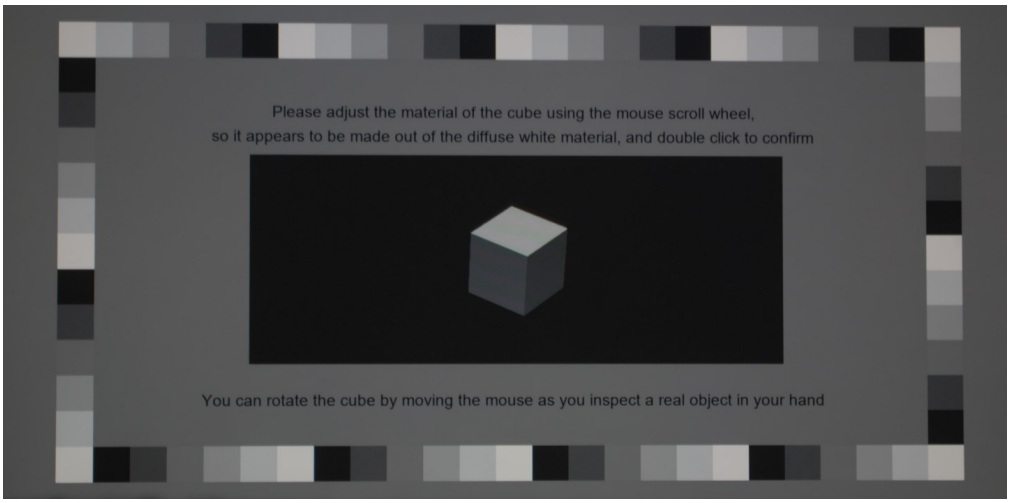


Figure 4.14: Photograph of the diffuse white experiment user interface on the cube stimulus. Observers were asked to adjust the cube brightness to appear diffuse white with the freedom to rotating it.

of the background type (solid or pattern) or contrast. In the second session, the cube appeared at the center of the background with a random luminance and initial rotation. The observer was again asked to adjust the cube luminance using the mouse scroll wheel so that it appeared as diffuse white, then double click to confirm. In addition to the luminance adjustment, observers were also encouraged to move the mouse to rotate the cube and feel the luminance change. The background sequence, and trial repeats, were the same as in the first patch session. In total, 100 trials were assessed for each observer, with half of them on the 2D patch and the other half on the 3D cube. Fifteen observers with normal or corrected to normal vision participated in the experiment, of which 9 were male and 6 female.

4.3.2 Result

The observer-adjusted patch luminance and cube material albedo were recorded. The average of the 5 repeated trials were recorded as the result of the stimulus-background-observer combination. The

average result of a combination were taken from the mean of all observers. For the 2D patch stimulus on the solid background, the average adjusted AR luminance were accurate match to the reference white on background L1-L4, and slightly below reference on the brightest background (L5) (Figure 4.15 (a)). It demonstrated that on the simplest stimulus-background condition, observers were able to estimate diffuse white luminance accurately by perceptually excluding the background from the stimulus. The estimating failed when the background was too bright, and the proximal luminance from the patch and the background resulted in the perception that the patch was emissive. The comparison between the solid and the pattern backgrounds of same average luminance were shown in Figure 4.15 (b). The adjusted patch luminance were lightly lower on the pattern backgrounds than the solid by 3.48, 0.87, 5.88 for background SLL, SM, and SHL respectively, but were still close to the reference white luminance. The 95% CI were slightly larger on the pattern backgrounds. The effect of the pattern background contrast are shown in Figure 4.15 (c). The adjusted patch luminance decreased with the increased pattern contrast, and all were below the reference luminance value (black dash line).

On the 3D cube stimulus, the visible surfaces and their luminance varied with the rotation angle. We calculated the cube luminance from the albedo and lighting, which was the luminance of the brightest surface facing the directional light. On all backgrounds, the average adjusted AR luminance were well above the reference white luminance (shown as the black dash line in Figure 4.15) and significantly higher than the patch stimuli. The 95% CI shown as error bars (avg=23.3) were significantly larger than the patch stimulus (avg=11.8). But the same three trends shown in the patch stimulus were also shown on the cube: 1. Lower AR luminance were adjusted on the brightest solid background. 2. Lower AR luminance were adjusted on the pattern backgrounds of the same luminance compared to the solid. 3. The increased pattern contrast decreased the adjusted AR luminance.

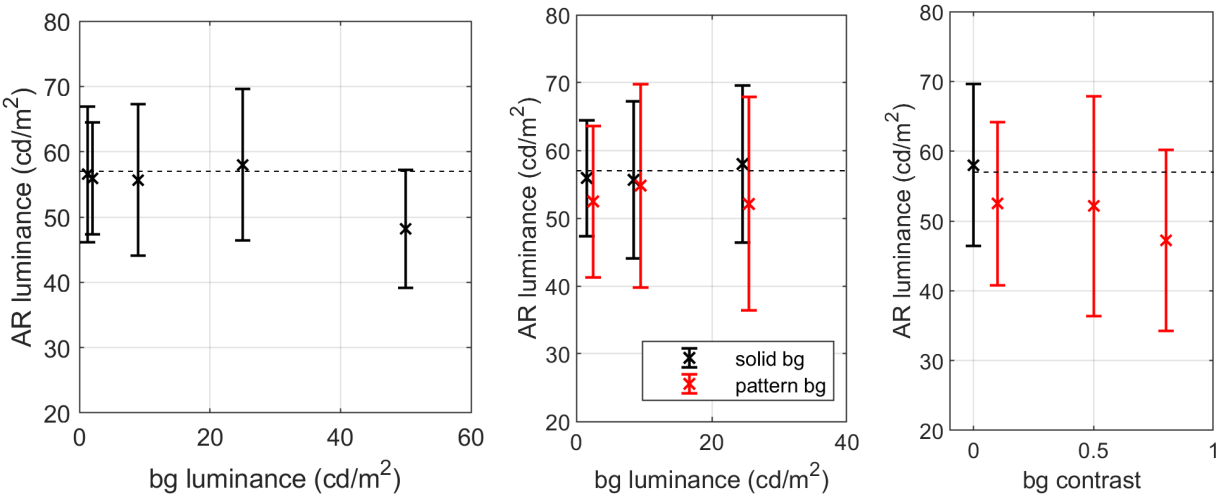


Figure 4.15: Observer adjusted AR luminance of the patch stimulus. The error bars are 95% CI of the average observer. The reference luminance is marked in black dash lines. (a) On solid backgrounds (L1-L5). (b) On pattern backgrounds (SLL, SM, SHL) compare to solids backgrounds of same luminance (L2-L4). The two types of backgrounds are offset for illustration only. (c) On backgrounds of same luminance but different contrast (L4, SLC, SM, SHC). The solid background L4 is included with contrast of 0.

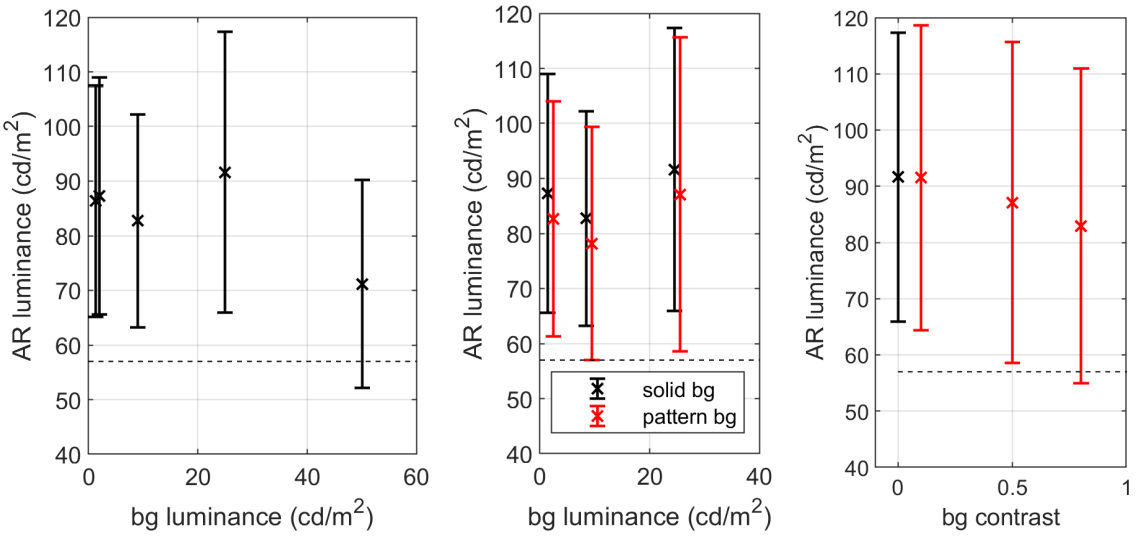


Figure 4.16: Observer adjusted AR luminance of the cube stimulus. The error bars are 95% CI of the average observer. The reference luminance is marked in black dash lines. (a) On solid backgrounds (L1-L5). (b) On pattern backgrounds (SLL, SM, SHL) compare to solids backgrounds of same luminance (L2-L4). The two types of backgrounds are offset for illustration only. (c) On backgrounds of same luminance but different contrast (L4, SLC, SM, SHC). The solid background L4 is included with contrast of 0.

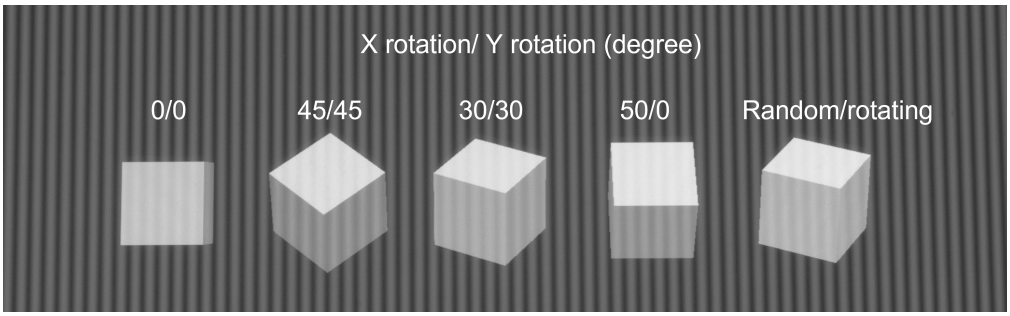


Figure 4.17: Photograph of observer reported common angles used to assess the cube brightness.

Luminance variation from cube rotation

At the end of the experiment, observers were asked if there was a certain rotation angle often at which they assessed the cube diffuse white level. Several rotation angles were reported, often used to assess the cube, as shown in the figure below. Almost all reported angles include one surface facing up, with high luminance and contrast to other surfaces. The experiment result suggested that the brightest cube surface adjusted by observers were much brighter than the reference. When the brightest surface were visible to observers, other shaded surface(s) were also visible, resulting in luminance contrast among surfaces. Since the rotation angle changes the overall cube luminance and contrast, it is possible that the cube material albedo is not the best metric representing the cube appearance. We calculated the luminance ratio of the patch stimulus to the cube for each background condition. The average luminance ratio for all backgrounds is 0.635 with SD of 0.044.

We simulated the projected polygon from rotation angle around x-axis and y-axis in world space. In this simulation, we focused on the shading of the cube, and no background is considered here. The rotation angles range from 0° to 90° on both axes covered all possible projected polygon luminance conditions. To reduce the repeated polygons, the rotation calculation are ceased at 89° on both axes. The lighting condition and surface material were the same as in the experiment. The luminance factor was the average pixel luminance of the projected polygon at a given angle, normalized to the

highest possible luminance, which was the surface faced towards the directional light. The calculated the Michelson contrast of the projected surfaces (Figure 4.18 (b)) ranged from 0 to 0.6. The minimum contrast appeared when there is no luminance differences among only visible surfaces. It appeared at two rotations: at only one surface visible to the observer at (0,0), and at two surfaces visible with the same luminance at (0,45). The maximum contrast appears at (45,0) with two surfaces visible and one of them facing towards the directional light reflecting the maximum amount of light. We also calculated the spatial average luminance factor of the projected surfaces. The projection area and luminance factor of each surface was calculated first. The average luminance factor was then taken as the sum of surface luminance factor weighted by the surface area fraction of the total projection area. The average luminance from the rotation angles were shown as a heatmap in Figure 4.18 (a) ranging from 0.56 to 0.78 with average of 0.647. The maximum average luminance appeared at when only one surface was visible to the camera (observer), and minimum when at rotation angle (40,33) and three other symmetrical rotation location. The average luminance (0.647) in the simulation from all rotation angles agreed well with the luminance ratio (0.635) calculated above between the patch and the cube stimuli. It indicated that there is no fundamental mechanism difference when judging the 3D cube compare to the 2D patch. Observer used the spatial average luminance on the 3D rendering as the metric matching to the reference white. And observers are capable of separating the background from the rendering almost completely for this task. Figure 4.19 illustrates the cube luminance correction matched well to the patch result after multiplying the luminance factor of 0.647.

Trial duration

We recorded the trial duration for an observer, adjusting the diffuse white level. On average the trial duration on the cube were 5.2s longer than on the patch (15.4s) mainly due to the additional time spent on rotating it. The background being solid or pattern did not affect the average trial duration

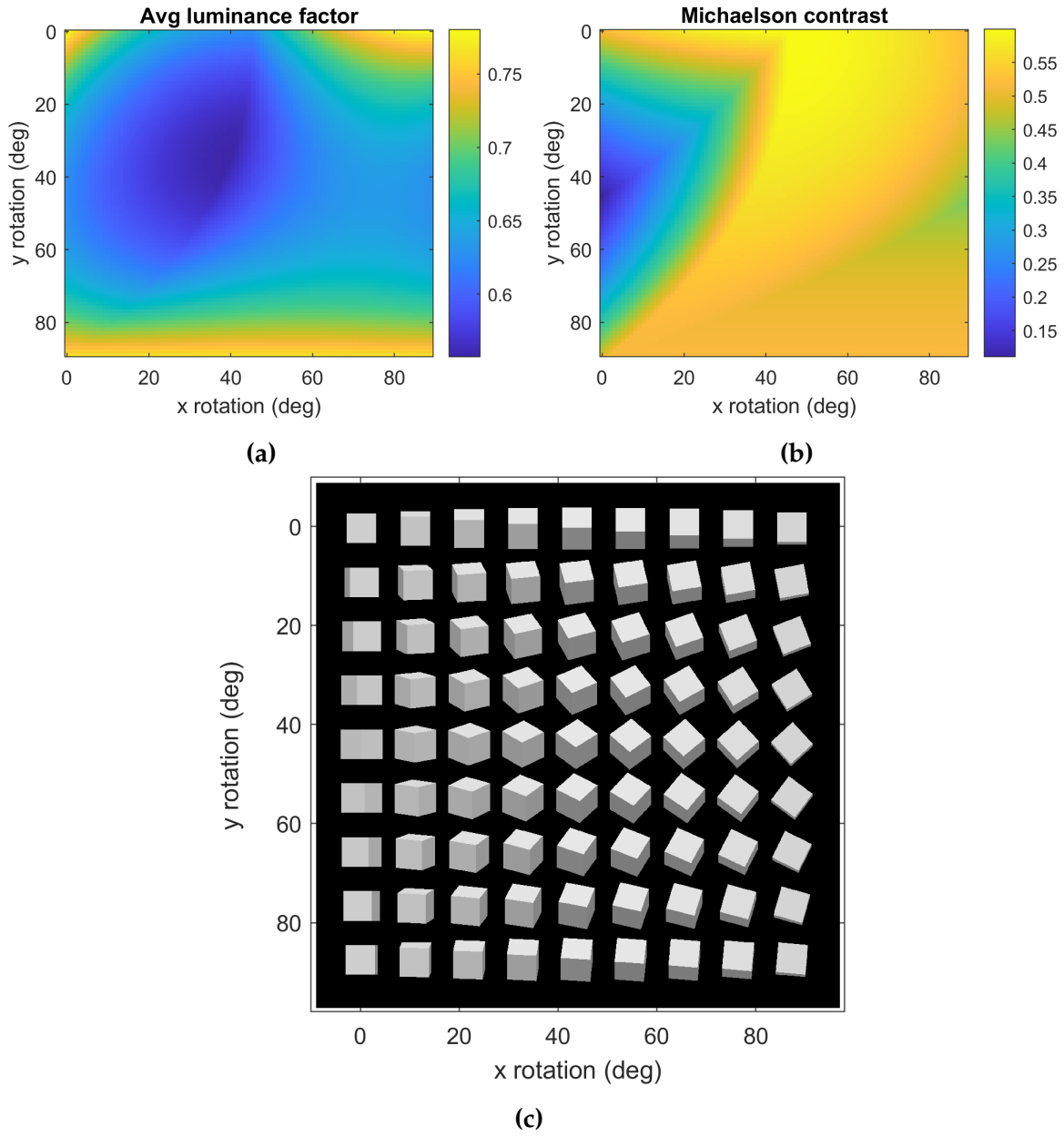


Figure 4.18: Simulated cube appearance change from rotating around x and y axes. (a) Spatial average luminance factor normalized to the maximum surface luminance. (b) Michelson contrast from rotation angles. (c) Rotation illustration of the same cube.

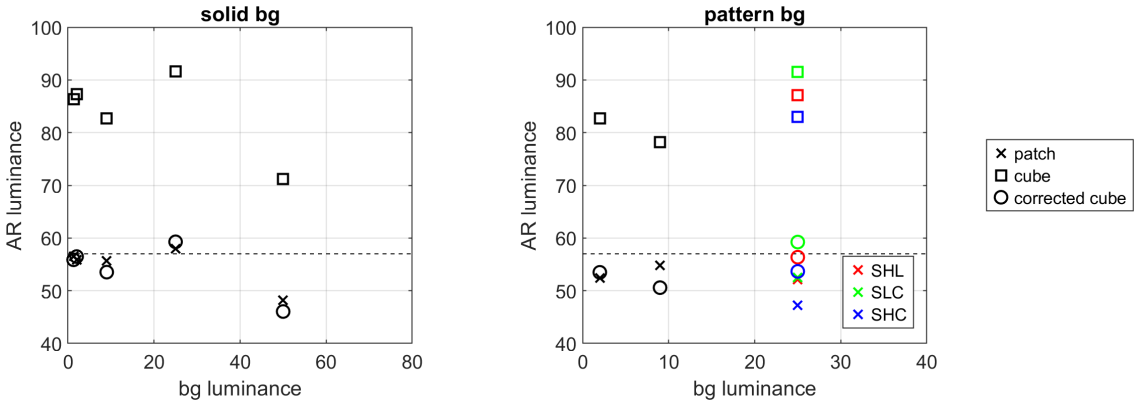


Figure 4.19: The patch result compared to the cube and the luminance factor corrected cube. The average luminance factor is a constant of 0.647 calculated from the cube rotations projection polygons.

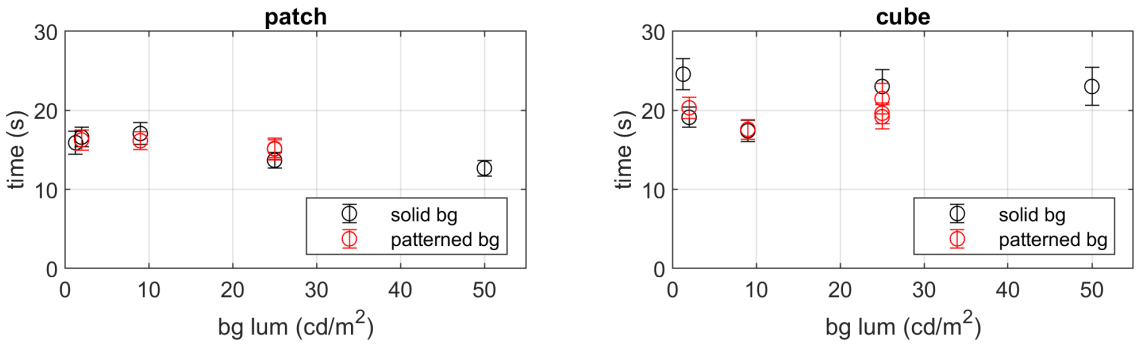


Figure 4.20: Trial duration for an average observer on the patch and cube stimulus with the error bars as the SEM.

on either patch or the cube stimulus. On the patches, trial duration on lower luminance backgrounds were longer than higher luminance backgrounds. But on the cube with the increased background luminance, the trial duration first decreased then increased (Figure 4.20).

Observer repeatability

The inter-observer variability were shown in Figure 4.15-4.16 error bar, representing the 95% CI of the average observer result. We also calculated the intra-observer variability from the repeated trials. The 95% CI of the repeated 5 trials were calculated for each observer in each condition. An average

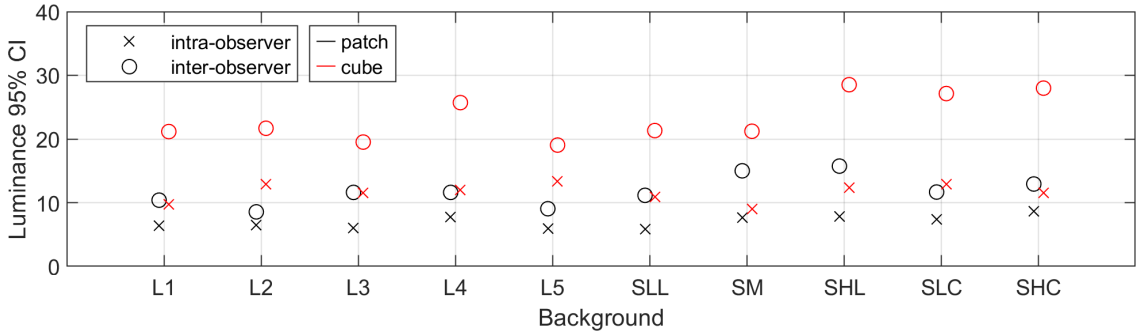


Figure 4.21: Inter- and intra-observer variability in 95% CI. Along x- axis are the 10 backgrounds in the experiment, with L1-L5 as solid, and SLL-SHC sinusoidal gratings. The background specification is in Table 4.5. The patch and cube results are offset for visualization only

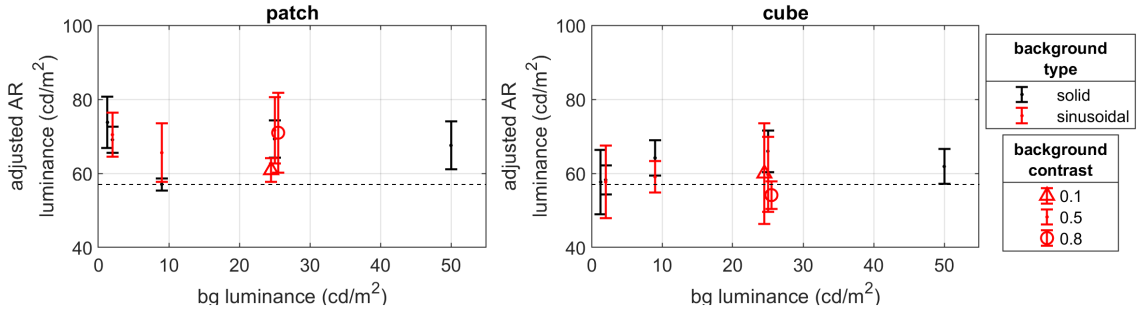


Figure 4.22: An example obsercer's result on the patch and cube stimuli. The observer had the best consistency on repeated trials among all observers.

CI was then taken from all observers as the intra-observer variability, shown in Figure 4.21. On average, the inter-observer CI were about twice the value of intra-, indicating large variation among observers. Observers were much more consistent on the patch stimulus compare to the cube. There was no significant impact from the background on the variability CI. An example observer's data is shown in Figure 4.22 who had the smallest intra-observer variance.

4.3.3 Discussion

Background separation or blending

In OST AR, the biggest challenge is the background light blending. Previous researches on compensating it were almost all based on additive light from the rendering and the background [21–27]. Though some researchers have proposed that observers could partially discount the background [38, 42, 43]. And the result on our diffuse white experiment also suggested observers could exclude the background almost completely, except the brightest background close to the reference white. Previous researches along with our result indicated that the amount of physically combined light from the AR overlay and the background may not be sufficient to predict the comparable AR brightness to non-AR references. The perceptual discounting can be explained by the layering theory. Observer's recognition of the coexistence of the background layer and AR layer in the scene helps them retain their brightness inside their media, minimizing the interruption with each other. It can be interpreted as a form of adaptation to the AR medium. Its own medium brightness scale is not necessarily associated with the real world black and white. There are other cues in the experiment helped enhance the layering perception. In the brightness matching experiment by Murdoch [39], the precisely aligned overlay showed reduced the amount of discounting at 5% compared to 24% on the rectangle overlay larger than the real cube, due to the lack of cues. The dark AR overlay on the patterned background resulted in the strong transparent appearance, which could also be cues for separated layers. Transparency and scissioning have been linked to perceptual layering [94, 121]. The relative motion between the AR content and the background also enhanced the layering perception. In our experiment, we had two extreme cases, maximizing (pattern background with rotating cube) and minimizing (solid background with patch) the layering cues. But both indicated complete background separation in achromatic attributes. Depth is another strong cue that contribute to the separation of layers, which is not included in our experiments, but did present in real user cases for OST AR. The problem introduced by the background blending is cut down by the perceptual dis-

counting. But further investigated, focusing on the amount of separation cues provided to observers is required to predict the perceptual discounting reliably.

On the other side, previous research found only partial background discounting or discounting within certain luminance ranges. The brightness matching result from Kim et al. [38] showed maximum of 60% of background luminance discounting within the transparent effective region. The amount of discounting decreased with increased reference luminance, and remained at zero outside the transparent effective region. For a single reference in our experiment, background L4-L5 were within the transparent effective region and L1-L3 outside according to their definition. But our experiment result showed 100% background discounting on background L1-L4 and about 80% of discounting on background L5. The percentage of discounting were much higher than their result. The only fundamental difference between the two experiments is the presentation media of the reference. Their reference and test patches were on the same display, while ours different. The medium-retained brightness might explain the conflicted result from the two experiment. The amount of background discounting within the same medium could be much lower than comparing to a different medium.

Simultaneous contrast

In the experiment, the test stimulus (patch or cube) were surrounded by the test background, while the small reference white patch was surrounded with other neutral scales and medium gray (20% of reference white). The background luminance difference between the stimuli and the reference could result in simultaneous contrast, pulling the perception away from the background. We considered this possibility resulting in the background luminance effect on the adjusted stimuli luminance shown in Figure 4.15(a) and 4.16(a). If simultaneous contrast presented, the bright background would make the stimulus appear darker compared to the reference. Thus, the observers should ad-

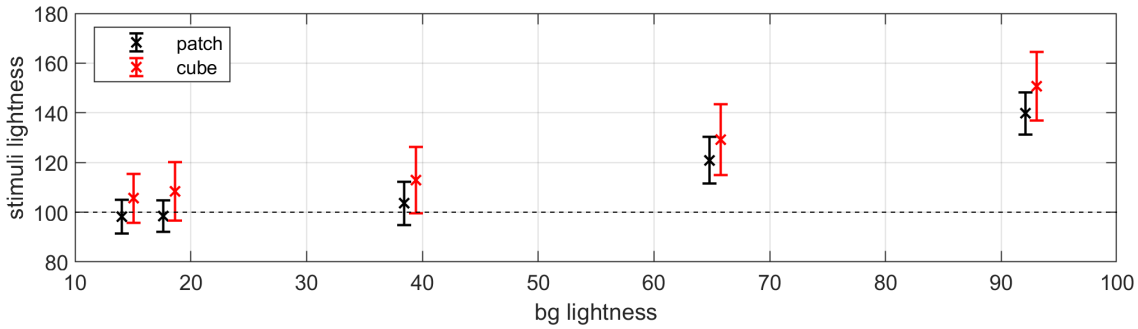


Figure 4.23: CAM16 lightness prediction on observer adjusted stimuli on solid backgrounds (L1-L5) compare to the reference (black dash line). The patch and cube results on the same background are offset on the lightness axis for visualization only.

just to higher luminance. But the experiment result showed lower stimulus luminance, contradicting the hypothesis. If the background was not completely discounted by the observers, the combined proximal luminance being much higher than the reference luminance can be explained partially with simultaneous contrast. We examined this with CAM16 lightness J with "Dim" viewing condition and the test background as Y_b. Only solid backgrounds L1-L5 are included here, since CAM16 does not consider the background spatial variance. The resulted showed inaccurate prediction over predicting the stimulus lightness, especially on brighter backgrounds. The complexity of the setup and stimuli were indeed beyond the scope of CAM16. Other factors could also contribute to the lightness perception, including simultaneous contrast and partial background discounting.

Assumptions and scalability

We proposed an explanation in the result section that observers were able to exclude the additional light from the background when assessing the AR overlay stimuli. On the 3D object with shading and rotation, the spatial average luminance was used to assess the brightness in the same manner of simple flat stimuli. These conclusions were drawn from the comparison between the two types of stimuli from the experiment.

Several assumptions were prerequisite here: the lighting condition (including the directional light angle and directional to ambient light ratio), The shading model (flat fragment shading), projection method (orthographic projection), simple symmetrical rendering object (cube), diffuse object material (Lambertian surfaces), fixed viewer position relative to the directional light, and dim viewing environment. These prerequisite were not representative of assessment in higher degree of freedom or the most realistic rendering. The lighting condition affects the rendered object appearance dramatically. When the directional to ambient light ratio varies, both the cube average luminance and contrast will be altered. The simulated spatial average luminance in Section 4.3.2 may not hold anymore. For 3D object without edges, the graduate changing luminance on the curved surface may result in difference perception. If the material includes specular highlight and shadow, the overall brightness may be different. Before extending the conclusion, more assessment is required to test these assumptions.

Summary

We measured perceived diffuse white levels on AR rendered 2D patch and 3D cube stimuli from adjustment experiment. Both solid and sinusoidal grating backgrounds were included in the experiment. The result showed the adjusted AR luminance almost always matched to the reference white patch regardless of the background luminance when matching the 2D stimulus. This means that observers were able to discount the background completely when assessing the overlaid AR content on the background. There was no significant difference between solid and patterned backgrounds of the same average luminance. The increased background contrast further decreased the adjusted AR patch luminance. 3D cube were adjusted to higher luminance compare to the 2D patch due to the shading. The average luminance over cube surfaces and rotation angle showed good match to the 2D patch adjustment result, indicating average spatial luminance on 3D object were used to assess

the overall brightness instead of the brightness surface.

4.4 Brightness Scale and Diffuse White Anchor

In color appearance models, the lightness J is usually considered as a relative scale while the brightness Q as the absolute. For surface color and emissive colors, brightness is considered as how much light is perceived from an area. Lightness is the brightness relative to the brightness of white. The two experiments in this chapter provided a brightness scale and a diffuse white anchor. It provided a possibility of anchoring the relative brightness to lightness scale, since we modeled a range of background-AR luminance combination in both experiments. To maintain the brightness scale shape, only linear scaling and offset were allowed for such interval scale.

The individual power law model was used for brightness scales, as it performed the best among tested models. The exponent values were calculated from Equ.4.2. The diffuse white AR luminance was set to be the same as the reference white, which is consistent with the experiment result suggesting complete background discounting. The calculations in this section are simulated values based on our models. The reference white luminance was set to be $50\,\mathrm{cd/m^2}$. A series of backgrounds with luminance from 0 to the reference white were tested, resulting background luminance factor of 0-1. AR luminance range was set to have a large range with maximum luminance factor of 6 to the reference white. To property scale and offset brightness, we need two points on the scale with known lightness. One is the lightness value of diffuse white at 100, below which are reflective colors and above emissive. The other one conveniently is the lower end of the brightness scale, where AR luminance is at zero. The question is, what should be the lightness value for the lower end of the scale.

One option is considering it as a pure AR lightness scale, where no AR overlay means zero light-

ness regardless of the backgrounds. It indicates that observers completely discount the background, regardless of the AR and background luminance. In this option, the lower end of the lightness scale are anchored to zero and the result shown in Figure 4.24 (a). All curves meet at the origin and lightness of 100, but the background affects the curve slopes. Backgrounds with higher luminance factor results steeper slope. This is suitable for user cases where the brightness comparison are among AR renderings within the same background. But a problem appears when comparing the same AR luminance on different backgrounds when AR luminance factor is below 1 (lightness below 100). The same AR luminance on the black background (blue line) would have higher lightness value than on the white background (yellow line), which is unlikely to be true.

The other option is considering the background has its lightness levels and use it as the anchor for the lower end of the scale. In this case, we adopted the J in CAM16 for the background lightness anchor. The brightness scale is then offset and scaled to fit the background lightness and diffuse white lightness. The result is shown in Figure 4.24 (b). On the lower end of the scale, the estimated lightness is the background lightness. The big change from the last option discussed above is the curve slopes. The increased background luminance now decreases the lightness scale slope. In the section below lightness of 100, the lightness cannot be lower than the background, which fixed the problem mentioned in the first option but also eliminated the possibility of rendering dark color with AR. The new concern is when the AR luminance is higher than diffuse white level, the same AR overlay would appear darker on the brighter background, which is also unlikely. This option essentially suggests that observers do not discount the background at all when the AR present zero luminance, but discount the background completely when the AR present the same luminance as the reference white. For AR luminance factor between 0 and 1, we calculated the amount of background discounting needed to produce such curves (Fig. 4.25). The zoomed area plotted in the small panel is with AR luminance factor of 0-1. From this figure, we can see that only non-emissive colors (AR luminance factor < 1) on brighter backgrounds (background luminance factor > 0.5) showed

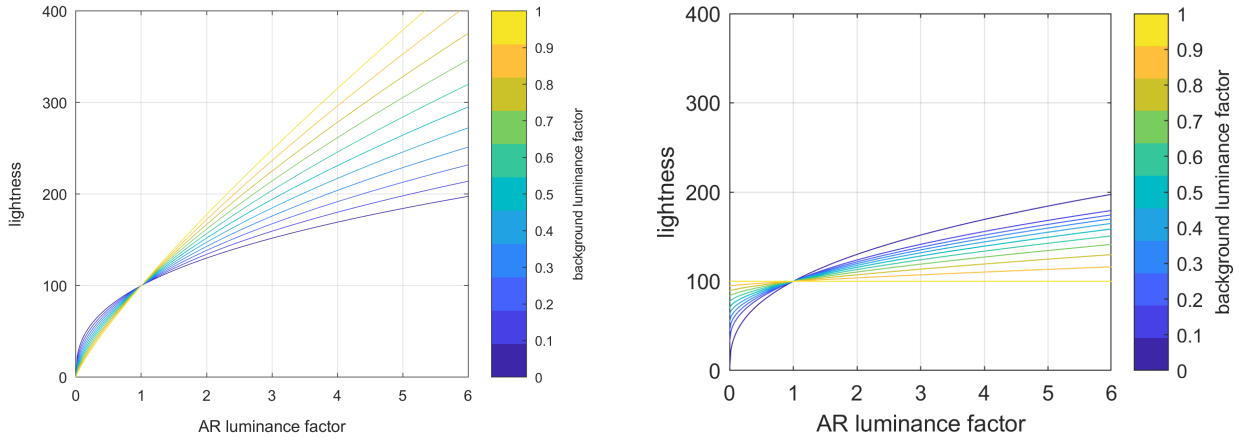


Figure 4.24: Two methods of constructing lightness scale from brightness scale and diffuse white luminance. (a) Assuming the AR lightness is at zero when AR luminance is zero, regardless of the background. (b) Assuming the AR lightness is the same as the background lightness when AR luminance is zero.

background discounting amount of 0 to 100%, which is a meaningful range in our context. Emissive colors showed more than 100% of discounting, meaning observers not only need to discount the background, but also need to increase the AR luminance perceptually. Non-emissive colors with dark backgrounds showed negative amount of discounting. These evidences theoretically reject the option. But it provided a new direction, which is the amount of background discounting could be inconsistent or vary non-linearly.

The two options mentioned above are likely to be the extreme cases, and the truth could be somewhere between them. So far, we adopted the definition of lightness being the relative brightness. But lightness is not always treated this way. A few color appearance models with both lightness and brightness attributes have non-linear relationship between them (Hunt95, CIECAM02, CAM16). An example in CAM16 describe the relation between lightness J and brightness Q:

$$Q = \left(\frac{J}{100}\right)^{0.5} \cdot \left(\frac{4}{c}\right) \cdot (A_W + 4) \cdot F_L^{0.25} \tag{4.8}$$

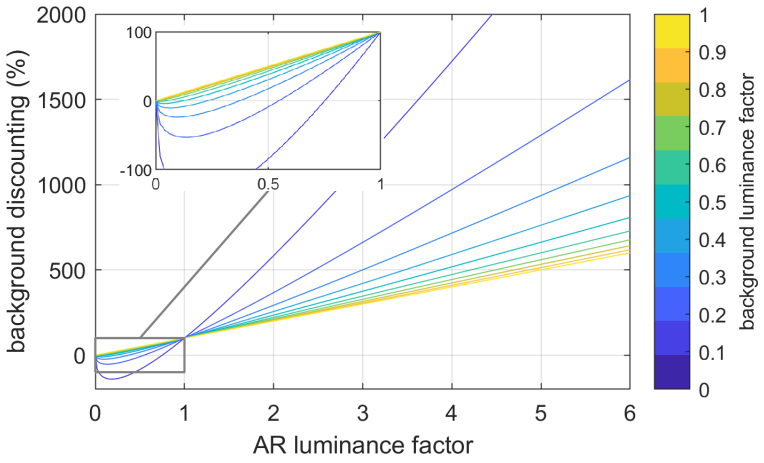


Figure 4.25: Amount of background discounting required for anchoring the lower end of the brightness curve with the background lightness. The boxed region is zoomed in the small plot.

The second to the last terms are constants for a given viewing condition, leaving the brightness being proportional to the square root of lightness. This was discussed in the brightness scaling experiment when comparing the CAM prediction to experiment data. For surface colors and emissive colors, having the nonlinearity contradict their definitions by CIE. And their calculations are challenged by new voices supporting linear relations now [122]. At this state, we do not have evidence to support or reject the nonlinear relation.

If the meaning of lightness is different in AR context compared to surface colors, the nonlinearity might be the key to the two option dilemma. The definition of lightness depends on what is the "AR white" that lightness relate to in AR environment. If AR lightness simply describe the relative brightness inside AR environment without connection to the real world, the linear lightness-brightness relation from the first option above is more reasonable. The "AR white" is just a transition point from reflective to emissive colors inside an AR environment, analogous to the reference white in HDR display monitors. Both the lightness and "AR white" are independent of the background, avoiding the background additive blending issue. This is the easier method of defining AR lightness,

but it's lack of attention on the unique background blending feature in OST AR. Alternatively, AR lightness can be defined to associate the perception of rendered content to a real world object. So the rendered and the real object with the same lightness level should be perceived as nearly the same. This rationale addresses the connection between the virtual rendering and the real world, which is also the purpose of the diffuse white experiment. "AR white" is supposed to be the same as the real world reference white perceptually. It would help AR rendering blend with the reality for immersive experiences. The second option, anchoring the lower end of AR brightness to be the background lightness, align with this rationale. But the linear lightness-brightness prevented it from a plausible solution. If we do consider a nonlinear brightness-lightness relation, especially for the AR case where the lightness scale is presented in one medium (virtual AR colors) and the reference white in a different medium (real world), this nonlinearity may also help with the unreasonable amount of background discounting in Figure 4.25. The amount of background discounting could depend on the relative luminance between the rendering and the background. For example, if the amount of background discounting increases with the increased AR luminance within the reflective range, and stop at the reference white luminance with complete discounting, the result curves would be between the two options: the curves in reflective range (AR luminance factor < 1) the same as Figure 4.24 (b), and the curves in emissive range (AR luminance factor > 1) would be the same for all backgrounds. The lightness-brightness relation is now not only nonlinear, but also are piecewise-defined. The discussion here are speculations. Accurate lightness scale will rely on more subjective evaluations on below and above diffuse white AR appearance and cross-background brightness comparisons. The scales could also be user case or viewing condition specified, for example only the reflective overlay are of interest. We did not include the long luminance adaptation in our experiments, which could also impact the brightness scale and diffuse white level.

4.5 Summary

In this chapter, we studied the perceived brightness and lightness with two experiments. The brightness partition scaling experiment was used to build relative brightness scale on different background luminance levels. The diffuse white adjustment experiment were used to determining the absolute luminance level for AR 2D and 3D stimulus. In both experiments, backgrounds varied in luminance, spatial pattern, and contrast levels. The experiment result showed dependency of the relative brightness scale on background average luminance, which can be modeled with power law functions. The absolute diffuse white level, both 2D and 3D stimuli, showed almost complete background discounting with the 3D cube using the spatial and angular average as the overall brightness. The background contrast also showed an effect on the adjusted diffuse white luminance levels. The combined result from the two experiments provided a possibility anchoring the modeled relative brightness to absolute. Two possible anchoring theories were proposed and discussed.

Chapter 5

Perceived Transparency

5.1 Introduction

In OST AR, the decreased virtual image luminance reduces the perceived brightness. But due to the background blending, the decreased luminance also increases the perceived transparency. In this chapter, we explored the contrast modulated AR transparency in related to different background conditions with three experiments. Focusing on the transparency, only patterned backgrounds are included in all experiments of this chapter. First, we measured the required contrast threshold for the AR overlay to be perceived as not transparent through a thresholding experiment. Secondly, a direct anchored scaling experiment was used to measure the perceived transparency at various contrast levels. In both experiments, in addition to the AR overlay patches, we also included patches reducing background contrast with the another method: reducing luminance amplitude while maintaining the average luminance. This is noted as the LCD patches, and was presented on the background LCD directly without any AR overlaying. This type of patch was tested to verify if contrast is the fundamental mechanism modulating the transparency. The result from the scaling experiment was used to model the perceived transparency. The third experiment was a paired comparison between the two types of patches (AR patches increasing average luminance from overlaying, and LCD patches

reducing pattern peak-to-peak luminance difference on background LCD). The experiment was to verify the equivalent transparency between the two types of patches, and the transparency model from the scaling experiment. The result of all studies were used to supplement a new perceptual dimension in transparency for better prediction of color appearance in AR.

5.2 Contrast-transparency space

In subtractive filters, the transparency and translucency perception is from physical properties including reflectance, transmittance, absorptance, and scattering. Perceived transparency in AR is different from subtractive filter transparency due to the additional light overlaid instead of the these physical properties. The AR transparency related to the luminance of the AR overlay and the background. With the increased AR luminance, the background contrast decreases, and AR patch transparency decreases. If we only consider the luminance contrast (neutral achromatic colors), on a sinusoidal pattern background, Michelson contrast C calculates as:

$$C = \frac{\Delta L}{2\bar{L}} \tag{5.1}$$

There are two methods to modulate Michelson contrast: by varying the pattern average luminance $\bar{L}$, or by varying peak-to-peak luminance difference ΔL. This contrast modulation can be described in a Cartesian space (Figure 5.1) with the base pattern as the background. At the origin, the pattern is the same as the base pattern. Along the x-axis is the change in peak-to-peak luminance difference ΔL, and along the y-axis is the change in average luminance $\bar{L}$. Simple additive AR rendering lies along the positive y-axis, reducing background contrast and AR rendering transparency by increasing average luminance $\bar{L}$. Physically, infinitely high luminance is required to reach zero contrast to achieve opaque AR rendering. But perceptual opacity could possibly be reached at certain contrast

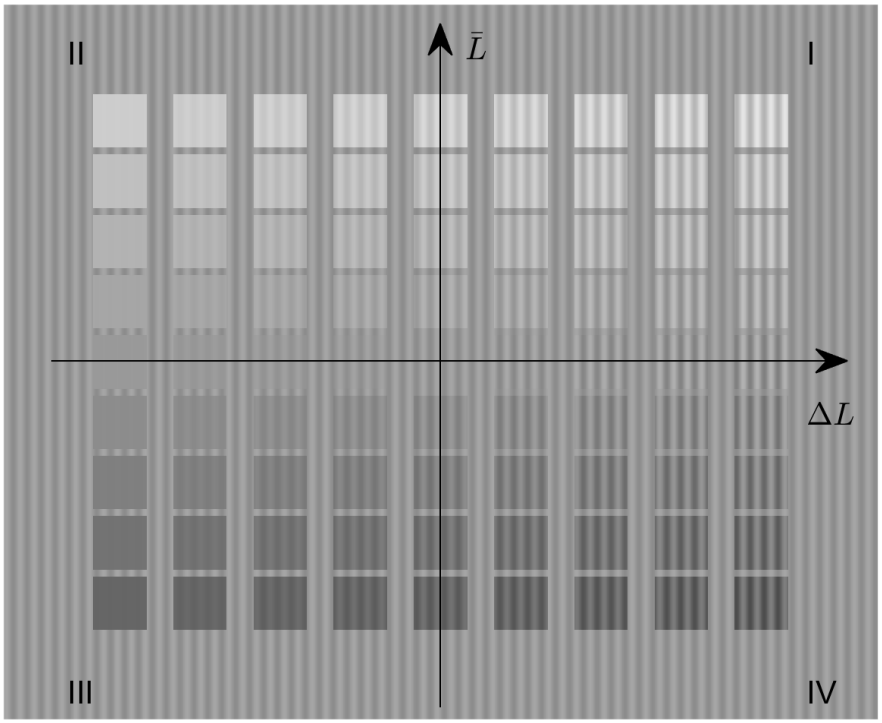


Figure 5.1: Contrast modulated transparency space. The small patches change from the base pattern in two dimensions: pattern luminance amplitude change is along the x axis, and average luminance along the y axis. The pattern at the origin is the same as the base pattern in the background. Simple AR overlays are along the positive y axis.

to reach. With very low AR luminance, the overlay appears to be vague and highly transparent. Additionally, with careful registration to the background, AR patches can be extended to quadrant I and II. Physical subtractive filters are in the quadrant II and III depending on the filter reflectance, transmittance, absorbance, and scattering.

5.3 Transparency threshold

The main purpose of the experiment is to determine the threshold of an image in OST AR being perceived as transparent or opaque when its rendering is naive to the background. The method of

Table 5.1: Background specification in transparency thresholding experiment.

Background	LL	MS	HL	LC	HC
Avg luminance (cd/m^2)	2	18	40	18	18
Contrast	0.2	0.2	0.2	0.05	0.4

limits were used to measure the threshold of perceptual transparency with Yes or No answer.

5.3.1 Methodology

Background and stimulus

We selected 5 backgrounds specified in Table 5.1. Three average luminance levels were chosen on the sine wave background: low luminance (LL), mid luminance (MS), and high luminance (HL) at $2\,cd/m^2$, $18\,cd/m^2$, and $40\,cd/m^2$ respectively, but with the same Michelson contrast at 0.2. Three contrast levels were selected on the sine wave background: low contrast (LC), mid contrast (MS), and high contrast (HC) with respective contrast at 0.05, 0.2, and 0.4, but the same luminance at $18\,cd/m^2$. The background MS served as both the mid luminance and the mid contrast condition. The specification was from the observer's eye position, meaning the reflectance and transmittance of the beam splitter were already accounted for in the values provided. The background gratings had the spatial frequency at 3 cycles per degree (cpd) which is at human visual system peak sensitivity. The background grating covered 6°x15° FoV representing the extended parafoveal field. The grating was surrounded by medium gray at 20% of the display white luminance and neutral scales for reference, which are the same as in the previous chapter.

Test patches are 2°x2° squares in the center of the backgrounds. There are two groups of test patches. One generated from the AR overlay (AR test patch) reducing the contrast by increasing the overall luminance, as along the positive y-axis in Figure 5.1. And the other from the LCD (LCD test

Table 5.2: AR and LCD test patch contrast under different backgrounds.

Patch level		1	2	3	4	5	6
AR patches							
BG	LL	0.002	0.022	0.041	0.060	0.078	0.200
	MS	0.014	0.033	0.051	0.068	0.085	0.200
	HL	0.029	0.046	0.063	0.078	0.093	0.200
	LC	0.004	0.008	0.012	0.016	0.020	0.050
	HC	0.028	0.076	0.119	0.159	0.195	0.400
LCD patches							
BG	LL, MS, HL	0.000	0.017	0.033	0.050	0.067	0.200
	LC	0.000	0.004	0.008	0.013	0.017	0.050
	HC	0.000	0.033	0.067	0.100	0.133	0.400

patch) by decreasing the peak-to-peak luminance difference as along the negative x-axis in Figure 5.1 while maintain the average luminance. When an LCD patch was presented, it was rendered on the background LCD directly and the AR display was black. Six test patch levels were selected in each background (Table 5.2). On AR patches, the lowest contrast was at patch level 1 with maximum available AR luminance. The highest contrast (patch level 6) was without AR overlay, so the test contrast was the same as the background. On LCD patches, the lowest contrast (patch level 1) was at 0 as solid patches with the same luminance as the background average luminance. The highest contrast (patch level 6) was again the background contrast. In the experiment, a test patch from either group was presented randomly in a trial.

Procedure

During the experiment, the observer first adapted to one test background for 10s. Then a test patch was presented for 500 ms preceded by a notice sound at a random location in the center 2°x8° area

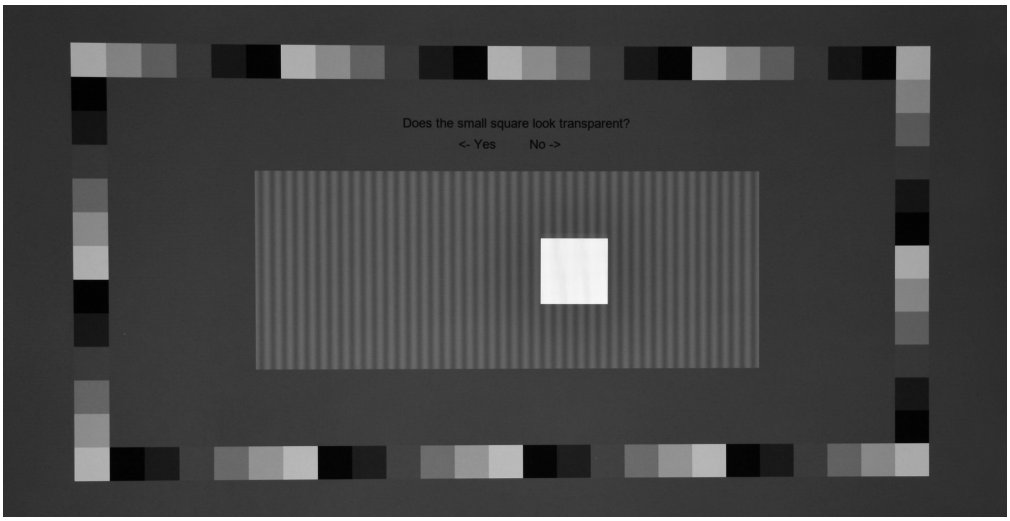


Figure 5.2: Photograph of the transparency threshold experiment UI. The small patch was presented for 500 ms on the pattern background and observers were asked to respond if it appeared transparent or not.

on the background. The observer was asked to indicate if the patch they saw appeared to be transparent with keyboard input. The limited stimulus presentation time was selected to comply with the general procedure in threshold experiment and reduce adaptation to stimuli. Considering the average saccade velocity at 2.7 ms per degree visual angle [123], the eye movement would take maximum 29 ms if the gaze was inside the background pattern area, and 45 ms if anywhere on the display. All test patches were examined randomly on this background before moving to the next background. Each patch level was repeated 5 times. All observers had training sessions to practice before the experiment, which contained all patch levels in background MS without repeat. After the experiment, observers were asked how they judged the transparency of the stimuli. In total, 2 patch types x 5 backgrounds x 6 contrast levels x 5 repeat resulted 300 trials each observer. Twenty observers with normal or correct to normal vision participated in the experiment (15 males and 5 females). Observers both with and without experience with mixed reality were included.

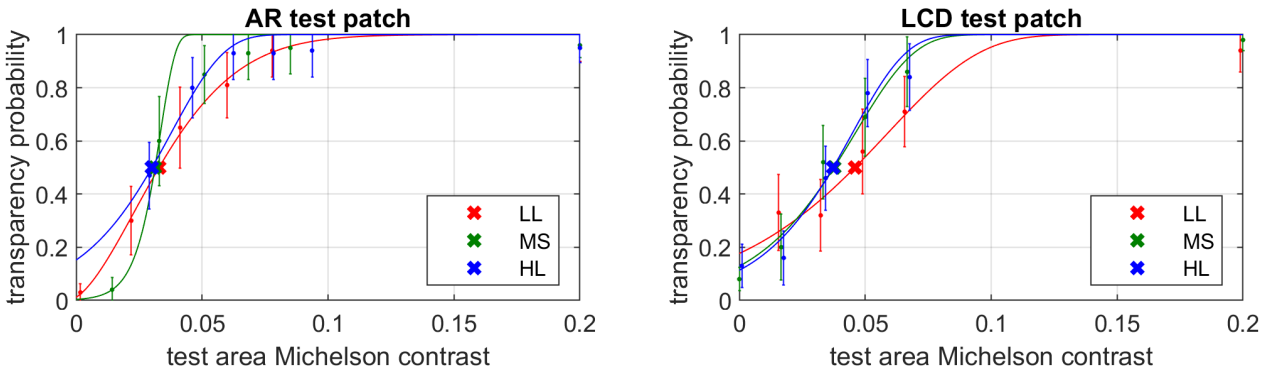


Figure 5.3: Psychometric curve of the AR and LCD test patch on the low (LL), mid (MS) and high (HL) luminance background.

5.3.2 Result and discussion

First, each observer's responses at each test contrast level were pooled from the 5 repetitions and the probability of responding "Yes" to the transparency was calculated. The mean probability from all observers at different contrast levels were fitted to a psychometric curve with accumulated 3-parameter Weibull function on the test patch contrast:

$$y = 1 - e^{-(\frac{x-c}{a})^b}. \tag{5.2}$$

The threshold is taken at 50% probability on perceived transparency from the fitted curve. The results were shown in Figure 5.3 - 5.4. The colored dot represented the average probability, with 95% CI as error bars. Each color corresponded to one background condition. The 50% threshold are marked as crosses of the same color. The threshold values are in Table 5.3.

Results showed that there was no significant difference between the two test patch types for AR and LCD (Fig. 5.5). There was no significant effect of the background average luminance on the perceived transparency threshold. The background contrast has significant impact on the transparency threshold, as the threshold elevated with the increased background contrast. This is

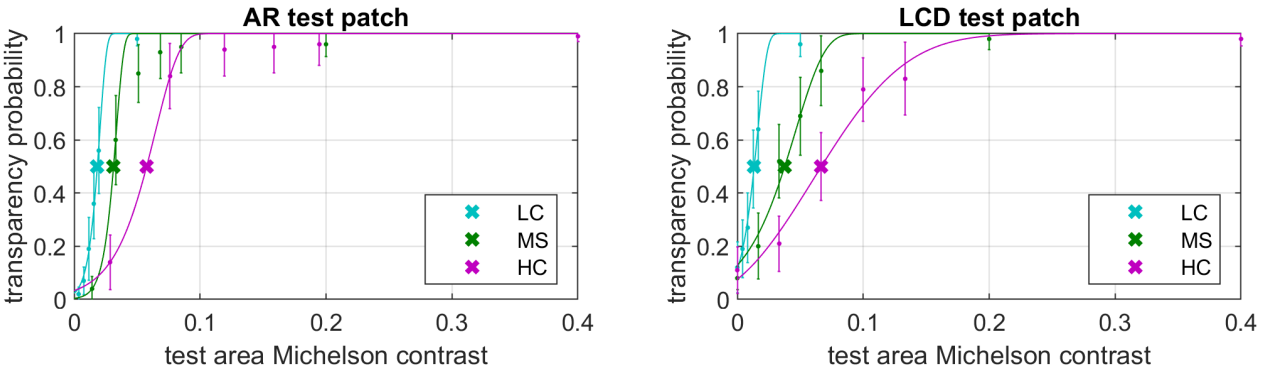


Figure 5.4: Psychometric curve of the AR and LCD test patch on the low (LC), mid (MS) and high (HC) contrast background.

Table 5.3: Threshold contrast values at 0.5 probability from the psychometric curve and 95% CI.

Background		LL	MS	HL	LC	HC
AR patch	threshold contrast	0.0329	0.0313	0.0301	0.0184	0.0576
	95% CI	0.0078	0.0020	0.0206	0.0079	0.0050
LCD patch	threshold contrast	0.0459	0.0376	0.0372	0.0134	0.0667
	95% CI	0.0445	0.0192	0.0262	0.0062	0.0296

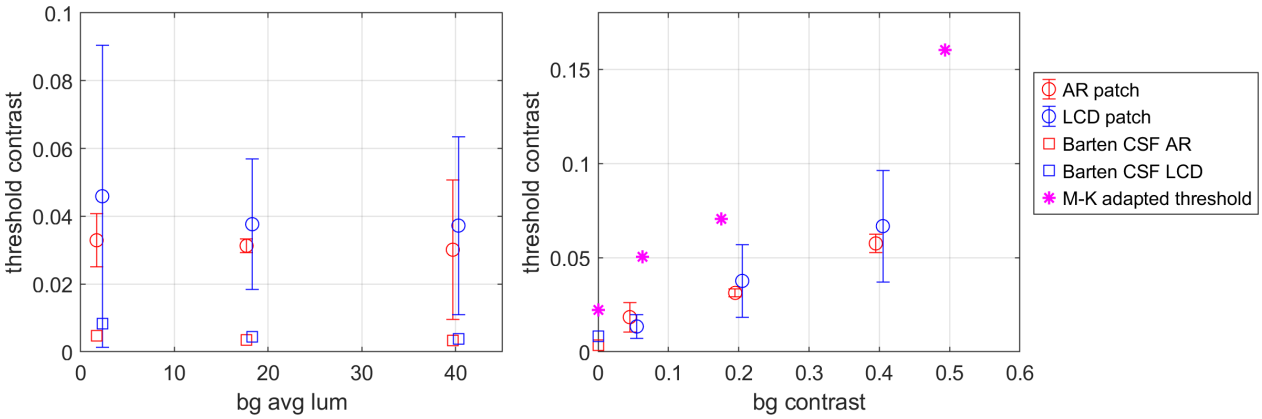


Figure 5.5: Transparency threshold and 95% CI against background luminance and contrast. The background luminance did not impact the threshold (left figure) while the background contrast did (right figure). Our result was compared to the thresholds from Barten CSF (sqaures), and CSF with contrast adaptation by Maattanen and Koenderink (stars). The threshold change on the background contrast can be explained with contrast adaptation.

likely due to the contrast adaptation mechanism in the human visual system.

Comparison to contrast sensitivity functions and contrast adaptation

Some observer reported that their threshold criteria was judging if the background pattern is still visible. Thus, we considered the hypothesis that the perceived transparency threshold is a function of contrast sensitivity function (CSF). First, we compared our result with Barten CSF model, which predicts CSF according to the target visual degree, target luminance, spatial frequency and surround luminance. The Barten CSF thresholds were plotted in the Fig. 5.5 as open squares, while our experiment data in open circles. The AR patches and LCD patches were colored in red and blue respectively. Our measured thresholds were significantly higher than the model prediction because Barton CSF is only applicable to uniform surround, while our surround is patterned. It is reasonable that the CSF threshold is lower after adapting to a plain background. For comparison, we plotted the contrast adaptation result from Maattanen and Koenderink as magenta stars. They used similar

setup as in our experiment with Gabor pattern at 5 cpd in 2.4° FoV. Their measured threshold were elevated with adapted contrast compared to plain-field-adaptation. Their result agrees with ours with similar elevation slope, supporting the change in the threshold is due to contrast adaptation. But the absolute levels of threshold levels are different from our transparency threshold. This was likely because of the slow falloff in the Gabor pattern elevated the threshold. There was also target luminance difference ($80\,\mathrm{cd/m^2}$) and spatial frequency difference between the two experiments.

Summary

AR transparency thresholds were measured through a detection experiment on sinusoidal pattern backgrounds with various average luminance and contrast levels. The result showed possibility of the AR patch to be perceived non-transparent with low enough contrast on the test area. The background luminance did not impact the threshold but background contrast did. There is no significant difference by reducing contrast from luminance amplitude or average luminance. The transparency threshold can be modeled with contrast sensitivity functions with background contrast adaptation.

5.4 Transparency scaling

The previous threshold experiment demonstrated that the AR rendering can be perceived as transparent within a range of contrast. To further study the perceived transparency, we examined it using an anchored scaling experiment with constant stimuli.

Table 5.4: Background specification in transparency scaling experiment.

Background	LL	MS	HL	LC	HC	HHC	MQ
Avg luminance (cd/m^2)	2	18	40	18	18	18	18
Contrast	0.2	0.2	0.2	0.05	0.4	0.93	0.2
Type	sinusoidal						square

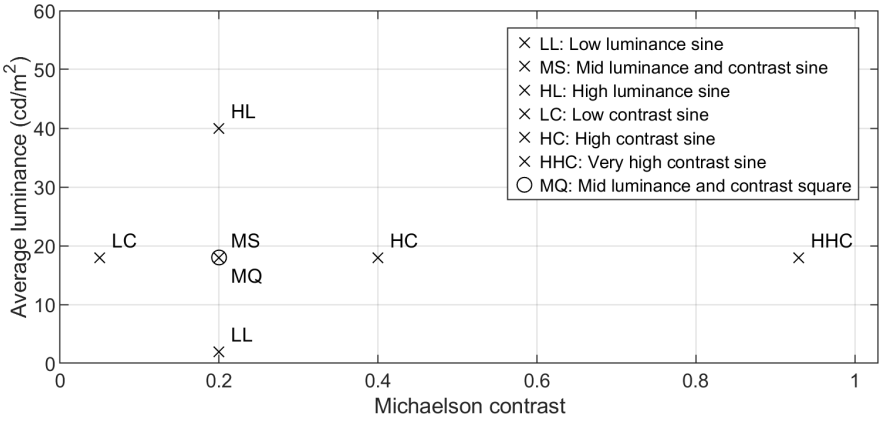


Figure 5.6: Background specification in average luminance and contrast.

5.4.1 Methodology

Background and stimuli specification

All backgrounds used in the thresholding experiments were included in the scaling experiment (Table 5.4). Extending the coverage of the contrast range, a higher contrast background (HHC) of 0.93 are included with average luminance of 18 cd/m^2. An additional square wave was included (MQ) to examine if the high frequency edge effect the perceived transparency. Background MQ had the same contrast and average luminance as sine wave MS. The average luminance and contrast of can be found in Figure 5.6. All backgrounds spatial frequency (3 cpd), FoV (6°x15°), surround (medium gray at 20% of the display white luminance with neutral scales) were the same as in the thresholding experiment.

Test patches were again 2°x2° squares in the center of the backgrounds. Both AR patches and

Table 5.5: AR and LCD test patch contrast under different backgrounds.

Patch level		1	2	3	4	5	6	7	8	9	10
AR patches											
BG	LL	0.002	0.015	0.031	0.048	0.067	0.088	0.112	0.138	0.167	0.200
	MS, MQ	0.014	0.022	0.037	0.054	0.073	0.093	0.115	0.140	0.169	0.200
	HL	0.029	0.030	0.045	0.061	0.079	0.099	0.120	0.144	0.170	0.200
	LC	0.004	0.005	0.008	0.012	0.016	0.021	0.027	0.033	0.041	0.050
	HC	0.028	0.051	0.089	0.129	0.169	0.212	0.256	0.302	0.350	0.400
	HHC	0.068	0.168	0.267	0.367	0.467	0.566	0.666	0.765	0.865	0.965
LCD patches											
BG	LL, MS, MQ, HL	0.000	0.013	0.025	0.050	0.075	0.100	0.125	0.150	0.175	0.200
	LC	0.000	0.003	0.006	0.013	0.019	0.025	0.031	0.038	0.044	0.050
	HC	0.000	0.025	0.050	0.100	0.150	0.200	0.250	0.300	0.350	0.400
	HHC	0.000	0.060	0.121	0.241	0.362	0.482	0.603	0.723	0.844	0.965

LCD patches were included. Ten test patch levels were selected in each background (Table 5.5). For AR patches, the lowest contrast was at patch level 1 where maximum AR luminance available. The highest contrast (patch level 10) was with zero AR overlay, so the test contrast was the same as the background. For LCD patches, the lowest contrast (patch level 1) was at 0 as solid patches with the same luminance as the background average luminance. The highest contrast (patch level 10) was again the background contrast. In the experiment, a test patch from either group was presented randomly in a trial.

Procedure

Here, the transparency is defined as a layer one can see through to the background. If it does not appear to have a layer between the observer and the background, it is categorized as "completely

transparent". If the background is not visible through the layer, it is categorized as "completely opaque". Physical samples with different transparency levels were presented to the observer to familiarize with the perceived transparency. The physical samples include filters with decreased transparency to both dark and white opaque samples (Figure 5.7). In the experiment, the observer first adapted to one test background for 10s. A test patch from either AR or LCD type was presented randomly in a trial, appearing in a random location within a 2°x8° area at the center of the background to stimulate eye movement and minimize afterimage of the background pattern. The observer was asked to "rate the patch transparency by moving the slider with a mouse between the two anchors (completely transparent and completely opaque)" as seen in Figure 5.8. Observers had unlimited time to assess a patch before making decisions.

Test patches of both types at 10 contrast levels were examined with each background. Each contrast was repeated for 3 times. All test patches were examined on one background before moving to the next background, and each patch was repeated three times. The observer had a training session to practice before each experiment, which contained 5 randomly selected trials from the main experiment. After the experiment, observers were asked how they judged the transparency of the stimulus. The experiment was separated into two sessions: 5 backgrounds (LL, MS, HL, LC, HC) in the first session, and 2 backgrounds (HHC, MQ) in the second session. In total, 2 patch types x 7 backgrounds x 10 contrast levels x 3 repeat resulted 420 trials each observer. Twenty observers with normal or correct to normal vision participated in the experiment (15 males and 5 females). Observers both with and without experience with mixed reality were included. All observers participated in the first session for the 5 backgrounds. Five observers returned and participated in the second session for the additional 2 backgrounds.

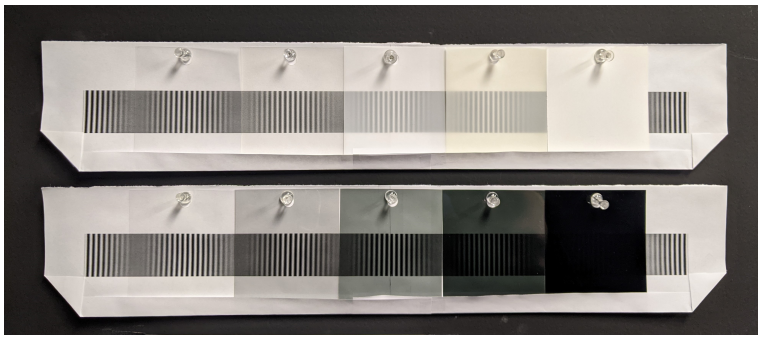


Figure 5.7: Transparent samples provided to observers.

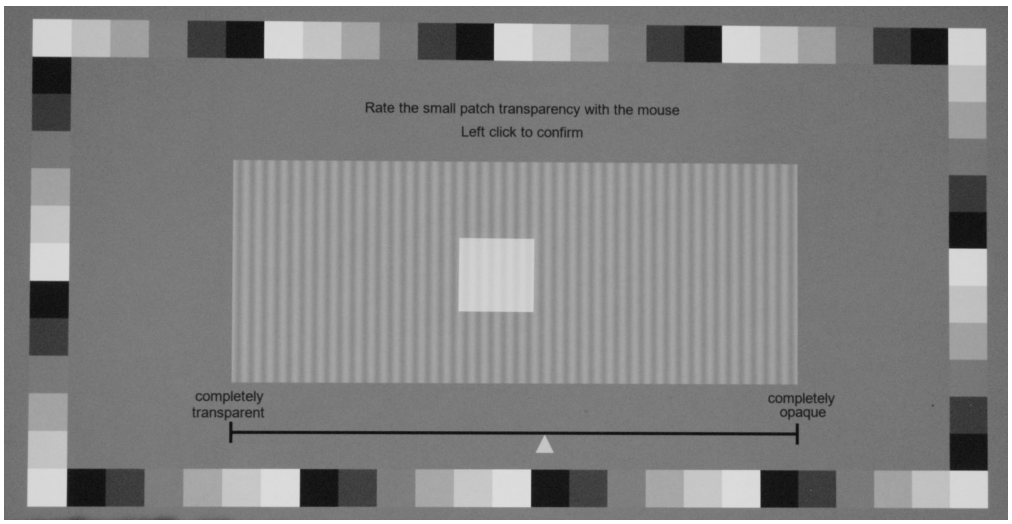


Figure 5.8: Transparency scaling experiment user interface. Observers were asked to indicate the patch transparency using the slider between the two anchors: completely transparent and completely opaque.

5.4.2 Result

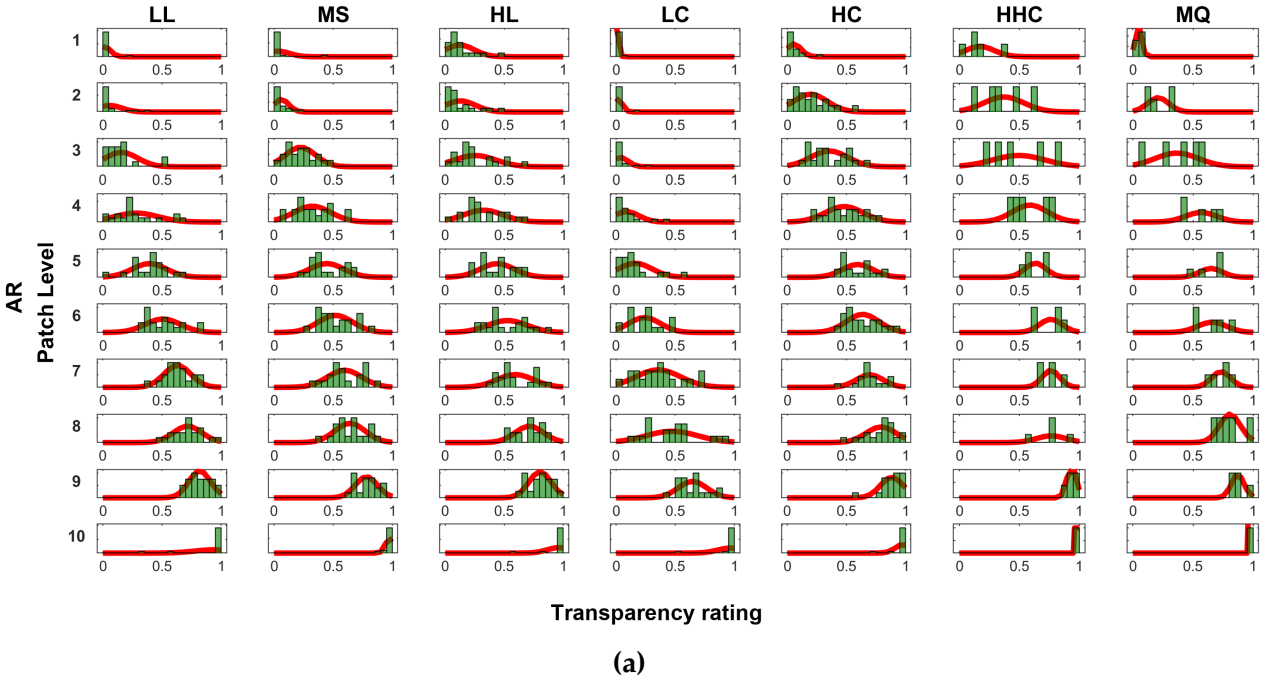
Outlier test and duration

Observers' transparency ratings were converted into numerical scales from 0 to 1, with "completely opaque" as 0, and "completely transparent" as 1. First, we examine the observer's rating consistency on the 3 repetitive trials to exclude observer's mistakes. A trial is discarded if it meets both requirements: 1. Its median absolute deviation is larger than three. 2. Its standard deviation (SD) is larger than a random distribution standard deviation (0.288). The outlier detection test removed 1.0% trials in AR test patch group, and 3.6% trials in LCD group. Observers took 3.6s on average to assess the appearance before giving a response with SD of 4.3s. On both types of patches, an increase in background contrast resulted in increased average response time (LC 3.5s, HC 3.8s, HHC 5.7s). On AR patches, increased background luminance also increased the average response time (LL 3.2s, MS 3.5s, HL 3.8s), but this trend was not clear in LCD patches.

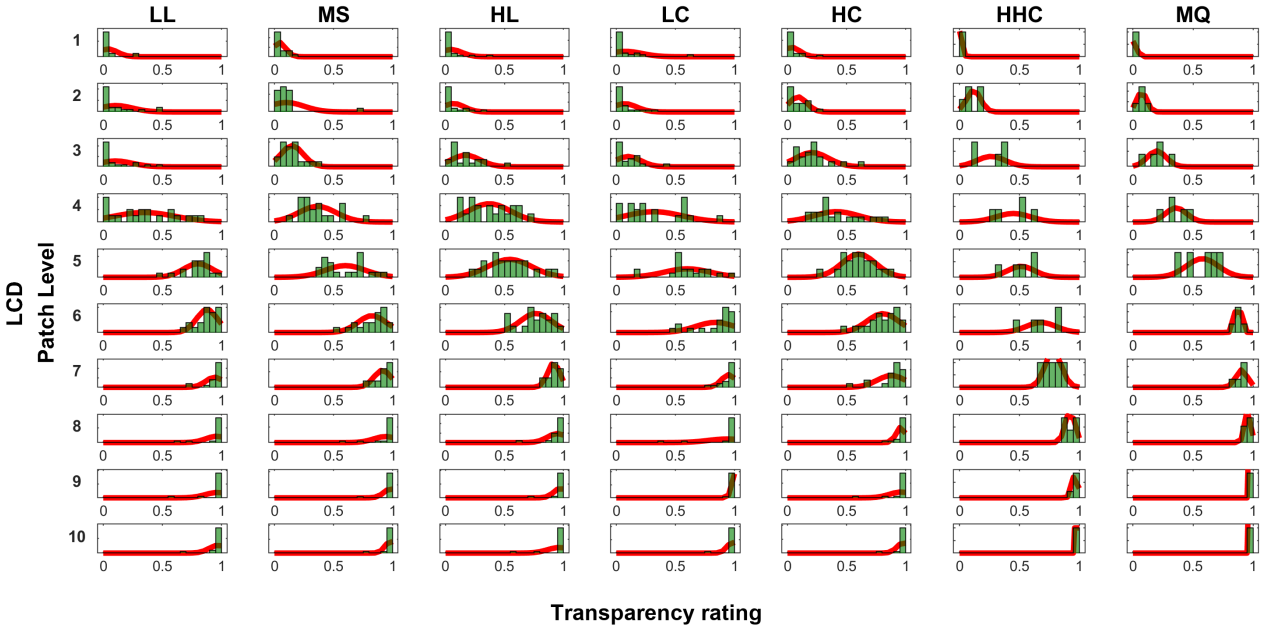
Observer rating and background factors

Observers' responses were averaged over repeated trials for each patch level and background. All observers' response were pooled and fitted with normal distributions using maximum likelihood estimation, and used μ in the fitted distribution as the average rating. The response histogram and the fitted distribution are shown in Figure 5.9. The standard error of the estimated rating was calculated from the variance-covariance matrix of the estimator. The estimated standard error on the background HHC and MQ are higher than other backgrounds due to the smaller observer population. This is most evident on AR patches with low contrast (low patch levels). The estimated SEM was smaller on both the high and low contrast ends compared to the middle.

Since the patch level contrast are different on different backgrounds, we calculated contrast ratio



(a)



(b)

Figure 5.9: Observers' transparency rating probability density histograms and normal distribution estimation. Each row corresponds to one test patch level. Each column corresponds to one test background. (a) Results on AR test patches. (b) Results on LCD test patches.

$\frac{C}{C_b}$ by normalizing the patch contrast to the background contrast. On the AR patches, the contrast ratio is equivalent to the average luminance ratio of the background $\bar{L_b}$ and the additive luminance on the AR area ($L_{AR} + \bar{L_b}$).

$$\left(\frac{C}{C_b}\right)_{AR} = \frac{\frac{\Delta L_b}{2(\bar{L_b}+L_{AR})}}{\frac{\Delta L_b}{2\bar{L_b}}} = \frac{\bar{L_b}}{L_{AR} + \bar{L_b}} \tag{5.3}$$

On the LCD patch, the contrast ratio is the ratio of the luminance amplitude between the LCD patch ΔL_{LCD} and the background ΔL_b:

$$\left(\frac{C}{C_b}\right)_{LCD} = \frac{\frac{\Delta L}{2\bar{L_b}}}{\frac{\Delta L_{LCD}}{2\bar{L_b}}} = \frac{\Delta L_{LCD}}{\Delta L_b} \tag{5.4}$$

The average estimated transparency ratings with 95% confidence interval (CI) error bars were plotted against the patch contrast ratio in Figure 5.10. The transparency rating curves for the AR patches are always different in shape from the rating curve for the LCD patches, which suggests that reducing contrast by the two methods activates different mechanisms in transparency perception. On backgrounds with the same contrast but different luminance levels (top two plots), the transparency ratings were not different from each other on either AR patch or LCD patch, as their 95% CI bars overlap. The transparency ratings were similar on sinusoidal wave and square wave at the same luminance and contrast (middle two plots). Looking at the bottom two plots, it is clear that for the AR patches, the perceived transparency on contrast ratio was still affected by the background contrast as their 95% CI bars do not overlap. But on the LCD patches, background contrast did not impact the perceived transparency significantly except a small deviation on the highest contrast background (HHC).

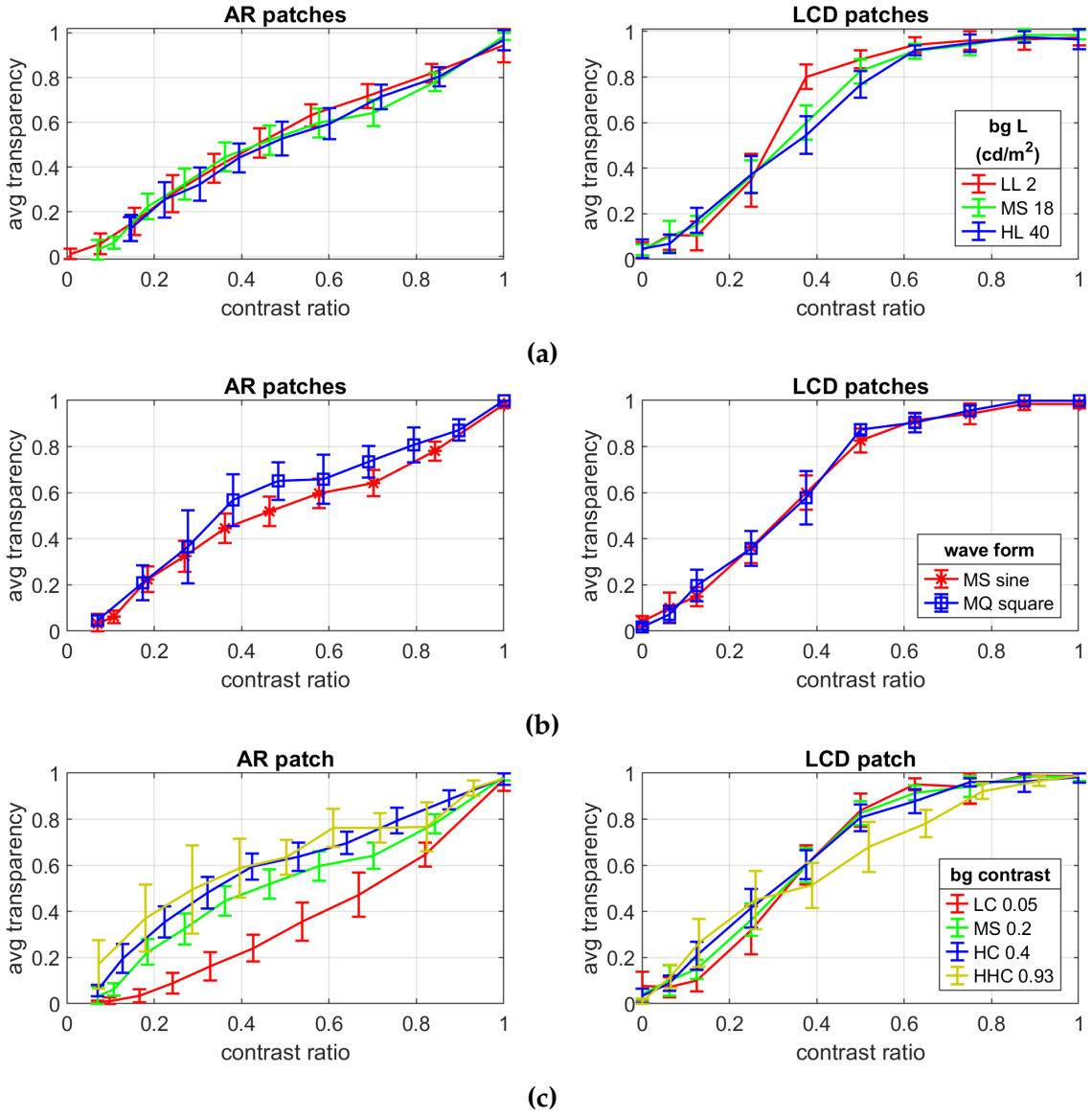


Figure 5.10: Transparency rating on relative contrast ratio to background. The left and the right subfigure in each subfigure shows AR patch and LCD patches respectively. The error bars represent the 95% CI of the mean. (a) Effect of background average luminance on the transparency ratings. (b) Effect of background wave form on the transparency ratings. (c) Effect of background contrast on the transparency ratings.

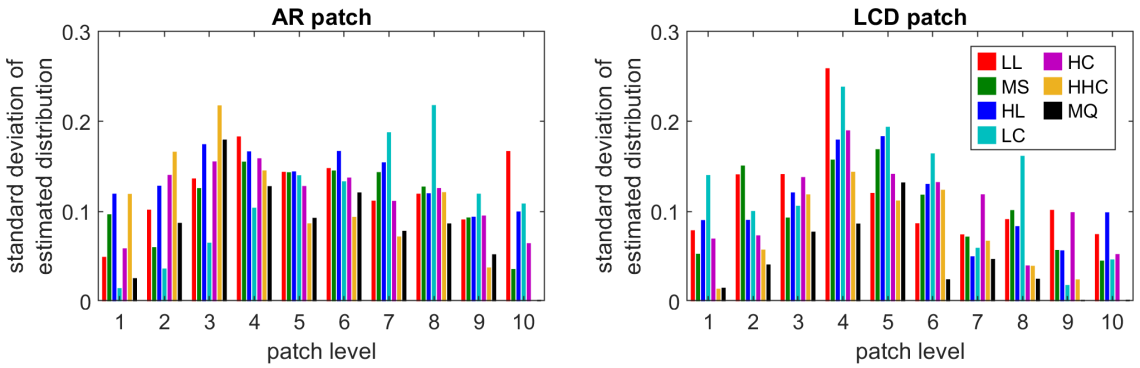


Figure 5.11: Standard deviation of the estimated mean of the normal distributions.

5.4.3 Discussion

Transparency rating distribution

In the data pooling, we had the assumption that transparency rating among observers follows a normal distribution. In some conditions in Figure 5.9, the normality of the distribution is questionable. For example, on LCD patch with background LL and patch level 4, the transparency rating distribute almost evenly across the 0-0.81, resulting a large standard deviation of the fitted normal distribution of 0.26. It indicates a large inter-observer variance on the perceived transparency of the patch. The estimated standard deviation of all conditions are shown in Figure 5.11. In some other conditions, it seems to be more than one peak in the histogram, for example on AR patch with background MS and patch level 5. It inspired the hypothesis that the transparency rating could be bi-modal instead of normally distributed. More discussion on the bi-modal transparency can be found in the contrast based transparency model section.

Scaling vs Threshold

We compared the opacity threshold from the previous detection threshold experiment with the transparency rating in the current experiment. The contrast ratio of the opacity threshold is 0.3683,

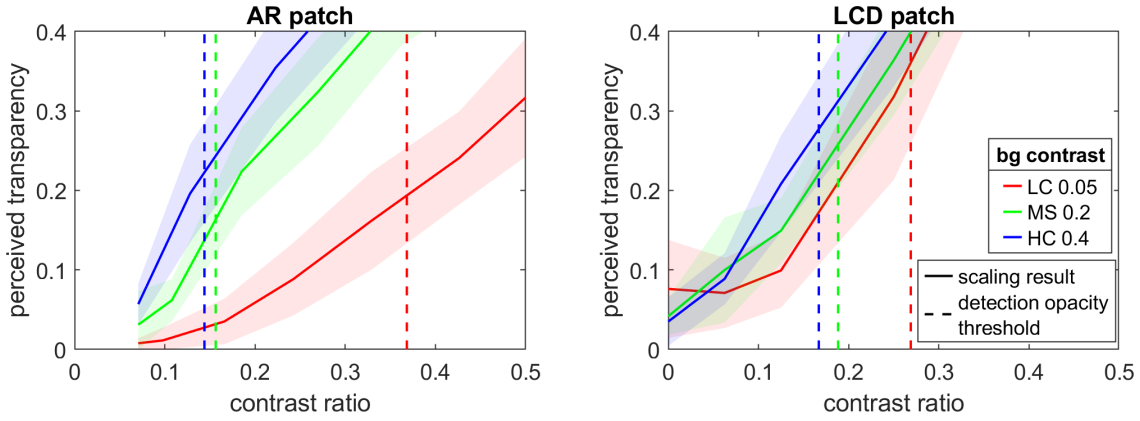


Figure 5.12: Comparison between the scaling and thresholding experiment on the contrast ratio of opacity threshold. The local zoomed image of Figure 5.10(c) is shown with the solid lines as the scaling experiment result with the shaded area as 95% CI. The dashed lines of the same color are the transparency/opacity threshold from the thresholding experiment of the same background.

0.1564, 0.1439 for background LC, MS, and HC respectively on the AR patches. While on the LCD patches, the respective contrast ratio of threshold is 0.2688, 0.1882, 0.1669 for background LC, MS, and HC. The rated transparency level of the these threshold contrast ratio in the scaling experiment are not 0 in perceived transparency scale. The comparison is shown in Figure 5.12. The opacity threshold in the thresholding experiment are rated with transparency level 0.16-0.22 on the AR patches and 0.25-0.35 on the LCD patches. This is likely because of the experimental method difference. The constant stimuli method in the scaling experiment offered more time to observe and adapt to the local contrast of the patch area, resulting in a more noticeable background pattern.

Summary

Perceived transparency scales were studied through direct anchored scaling with constant stimuli. The subjective transparency rating showed dependency on the stimulus types and the contrast ratio to the background. Reducing contrast through luminance amplitude (LCD patch) and average luminance (AR patch) presented different transparency perception mechanism. Background absolute contrast value also impact transparency on AR patches. Transparency scales can be described with

the contrast ratio between the stimulus area and the background. Background average luminance and pattern wave form did not affect perceived transparency.

5.5 Equivalent Transparency

The purpose of this experiment is to check if there is a perceived equivalency in transparency between the two types of patches (AR and LCD) through direct comparison. Though they could be rated at the same level of transparency in the scaling experiment, it was not clear if two types of patches were evaluated with the same criterion.

5.5.1 Methodology

Three backgrounds of different contrasts were selected in Experiment 2 (LC, MS, HC) for verification. The experiment used paired comparison with an interleaved multiple simple staircase method. The same experimental setup as in the transparency scaling was used. Three backgrounds (LC, MS, HC) of different contrasts were used in this experiment. LCD patches were used as the reference at 7 contrast levels relative to its background at 0.078, 0.204, 0.257, 0.311, 0.369, 0.4343, 0.636, respectively. Individual reference patch contrast was dependent on the background contrast (Table 5.6). Each reference contrast of each background was associated with one staircase. Observers adapted to a background for 10s first. In each trial, one of the seven staircases was randomly selected. Two patches were displayed on the background: the reference LCD patch and the test AR patch. Both patches were side by side with 6° separation, with their positions on the left or the right randomized (Figure 5.13). Observers were asked to select the patch that appeared more transparent to them using the arrow keys. The initial AR patch contrast was either minimum possible in our setup, or maximum as the background contrast without additional AR overlay. One staircase would termi-

Table 5.6: LCD patch contrast as references.

reference level		1	2	3	4	5	6	7
Relative contrast to background		0.078	0.204	0.257	0.311	0.369	0.4343	0.636
Background	LC	0.0039	0.0102	0.0129	0.0156	0.0184	0.0217	0.0318
	MS	0.0156	0.0408	0.0515	0.0622	0.0738	0.0869	0.1272
	HC	0.0312	0.0817	0.1029	0.1245	0.1475	0.1737	0.2545

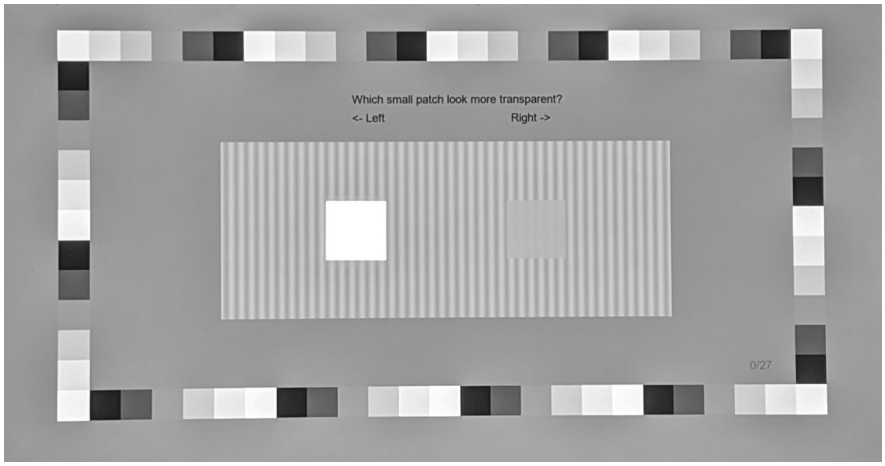


Figure 5.13: Photograph of the transparency equivalency experiment user interface. Observers were asked to indicate which of the two patches appear more transparent.

nate when seven reversal points occurred or the maximum limit of 25 trials was reached. Observers finished all seven staircases before moving to the next background. Ten observers with normal or correct to normal vision participated in the experiment (8 males and 2 females). Observers had a training session composed of 14 random comparison pairs in background MS.

5.5.2 Result and Discussion

All staircases for all observers terminated with seven reversal points before reaching the trial limit. Some observers reported difficulty making decisions on low contrast test patches and low contrast

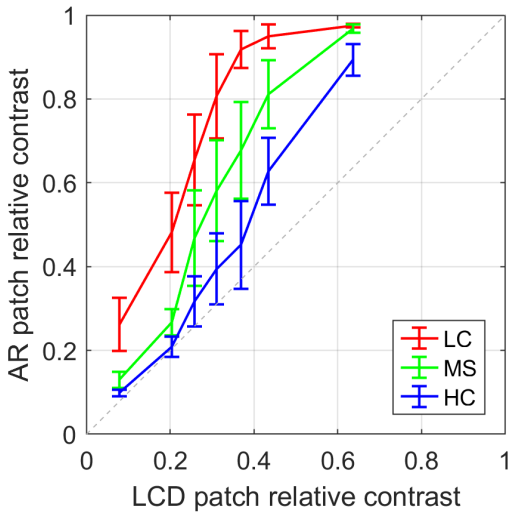


Figure 5.14: Experiment result on equivalent transparency between AR patches and LCD patches on different contrast backgrounds (LC, MS, HC). The error bars represent the 95% CI of the mean.

backgrounds. Observers spent 5.4s on average comparing the patch transparency with median of 3.5s and SD of 6.9s. The matched contrast was again pooled and estimated with maximum likelihood. We plotted the relative contrast of equivalent transparency between the reference LCD patches and the test AR patches in Figure 5.14 with error bars representing 95% CI. The deviation from the diagonal dash line on all three backgrounds supports the separated mechanisms between two types of patches in the transparency scaling experiment (Figure 5.10). The separation of the three curves confirmed the effect of the background contrast on the perceived transparency.

5.6 Contrast based transparency model

The perceived transparency models are proposed based on the experiment result in this chapter. The result of the scaling experiment was used to optimize the model parameters. The result from the equivalency experiment was used to evaluate the model performance.

5.6.1 Modeling Perceived Transparency

We proposed a model predicting the perceived transparency based on the patch type and background contrast. On the AR patches, the perceived transparency is described with exponential curves as a function of the normalized contrast to the background (contrast ratio). The perceived transparency T on the AR patches can be described as:

$$T = \left| \frac{C}{C_b} - a \right|^b \tag{5.5}$$

, where C is the Michelson contrast of the additive patch area, and C_b is the Michelson contrast of the background. Since the additive AR overlay does not affect the peak-to-peak luminance difference, the term C/C_b can also be written as $\bar{L}_b/(\bar{L}_{AR} + \bar{L}_b)$. The parameters a and b were optimized with the least square fitting to the scaling experiment measured transparency, where a is a constant and b is a reciprocal function of the background absolute contrast:

$$\begin{aligned} a &= 0.0592 \\ b &= \frac{0.05737}{C_b} + 0.4315 \end{aligned} \tag{5.6}$$

On the LCD patches, all backgrounds results from the scaling experiment were pooled to optimize the model parameters. The transparency curve is modeled with a single cumulative Weibull function, described as:

$$T = 1 - e^{-\left(\frac{2.5C}{C_b}\right)^{1.75} - 0.05} \tag{5.7}$$

In the LCD patch condition since the average luminance $\bar{L}$ on the test patch is the same as the background, the term C/C_b can also be written as peak-to-peak luminance difference ratio of the patch and the background $\Delta L/\Delta L_b$.

With the modeled transparency, the discrete patch contrast and background contrast levels in the

experiment can be extended to continuous values using Equ.5.5 and 5.7. A collection of background contrast curves on AR patches are plotted in Figure 5.15 from green to purple colors. The LCD transparency is plotted in black and independent of the background contrast. All AR curves intercepted the x-axis at relative contrast at 0.05, below which the patches are perceived to be completely opaque, indicating AR patch perceptual opacity is possible even with the constantly presented backgrounds. Curves intercept with the full background contrast at different transparency levels, ranging from 0 (opaque on solid background) to 0.97 (on maximum contrast background where there is physically no addition of light from the AR overlay). With the maximum background contrast, the curve has the exponential index of 0.51. With a solid background at 0 contrast, the patch will also be solid and perceived to be completely opaque, regardless of the luminance of the additional AR overlay. On the LCD patches, the transparency ranges from 0.05 to 0.91 while the normalized contrast ranges from 0 to 1, which indicates the LCD patches do not appear to be completely transparent or completely opaque in any relative contrast levels. The transparency rating being larger than 0 when the LCD patch is indeed a solid patch (0 relative contrast) with patterned surround, representing the effect of the pattern afterimage on the test patch appearance. The curve shape difference between the two types of test patches indicates they trigger different mechanisms in transparency perception.

Perceived transparency is modulated by the AR luminance, which is conventionally evaluated in luminance or lightness in color appearance models. A comparison between linear transparency, luminance and lightness is illustrated in Figure 5.16. Since the modeled transparency is independent of background luminance as the experiment result indicates, luminance range is scaled down to illustrate the additional light from AR with sinusoidal pattern with contrast of 0.4 as the background. The first column from the top to the bottom shows equal steps of linear additional luminance from 10% to 90% maximum additional luminance. The second column shows equal steps of lightness from 10 to 90 in CIELAB style, with the maximum luminance as reference white. The third and the fourth columns show steps of the linear transparency scale from 0.9-0.1 from the proposed model

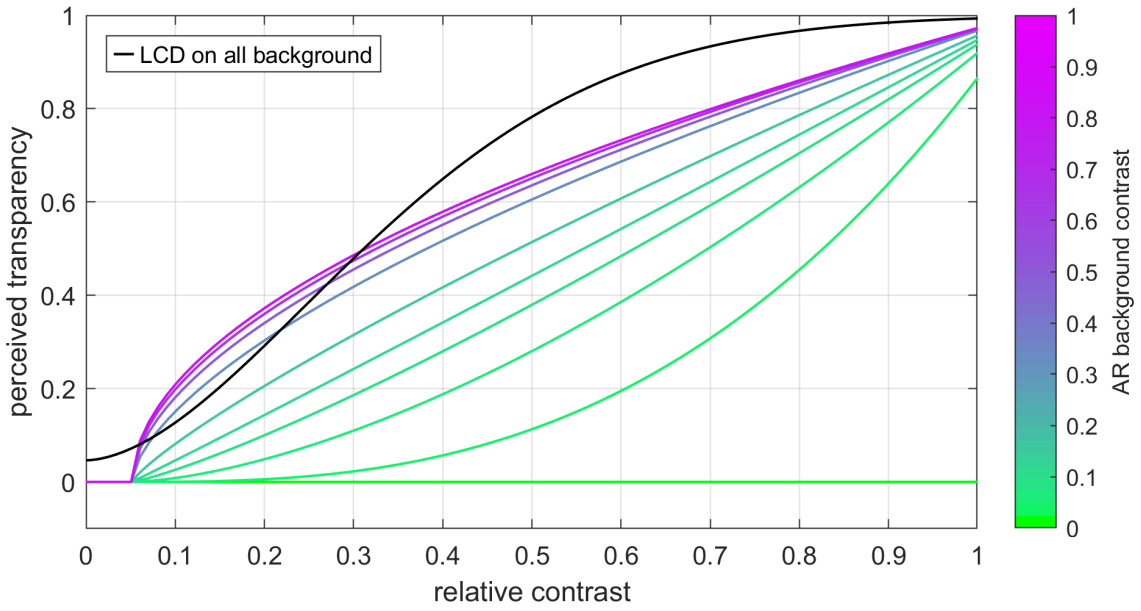


Figure 5.15: Contrast based transparency model prediction for different AR background contrast levels (see color scale) and LCD patches for any background (black).

prediction, corresponding to the AR transparency and LCD transparency respectively. This figure is for illustration purpose and valid only under specified viewing condition on calibrated displays (figure width at 7 inches, viewing distance at 15 inches, viewed in a dark room, displayed with sRGB format and peak luminance at 150 cd/m2). It demonstrates that the transparency scale in OST AR is indeed unique and cannot be accurately estimated with existing dimensions like luminance or lightness, which highlights the need of incorporating the new perceptual dimension in AR color management pipelines to improve image quality.

5.6.2 Performance Evaluation

The model prediction compared to the average experiment result on the AR patches is illustrated in Figure 5.17. Since the model is only contrast dependent and there are 4 backgrounds have the same contrast (LL, MS, HL, MQ), the results of the 4 backgrounds were pooled for the comparison in Figure 9 top right subfigure. The model performance is evaluated with the average root mean

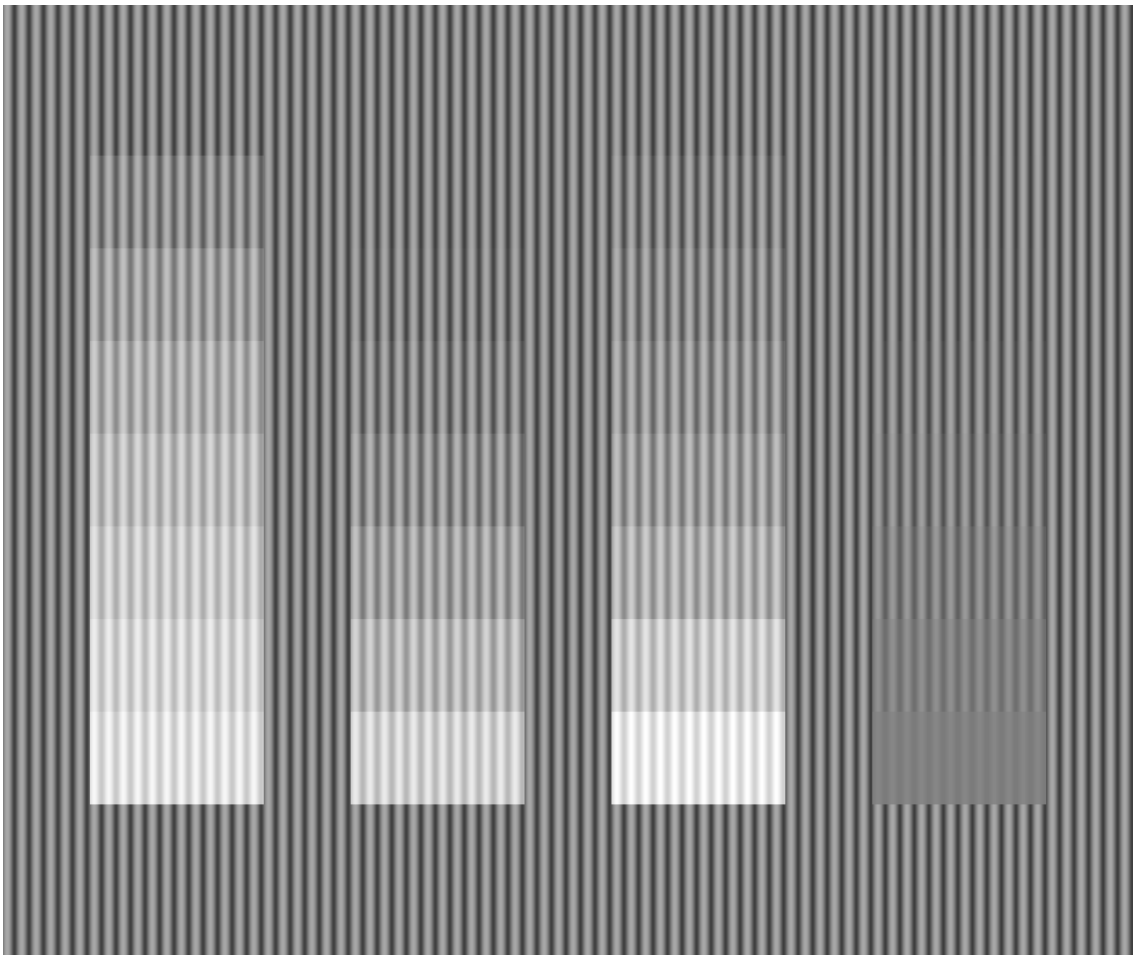


Figure 5.16: Transparency scale compared to luminance and lightness. From left to right: linearly increased luminance, linearly increased lightness, linearly decreased AR transparency, linearly decreased LCD transparency. The latter two are rendered based on our proposed transparency model. This illustration is only valid on calibrated displays and viewing conditions.

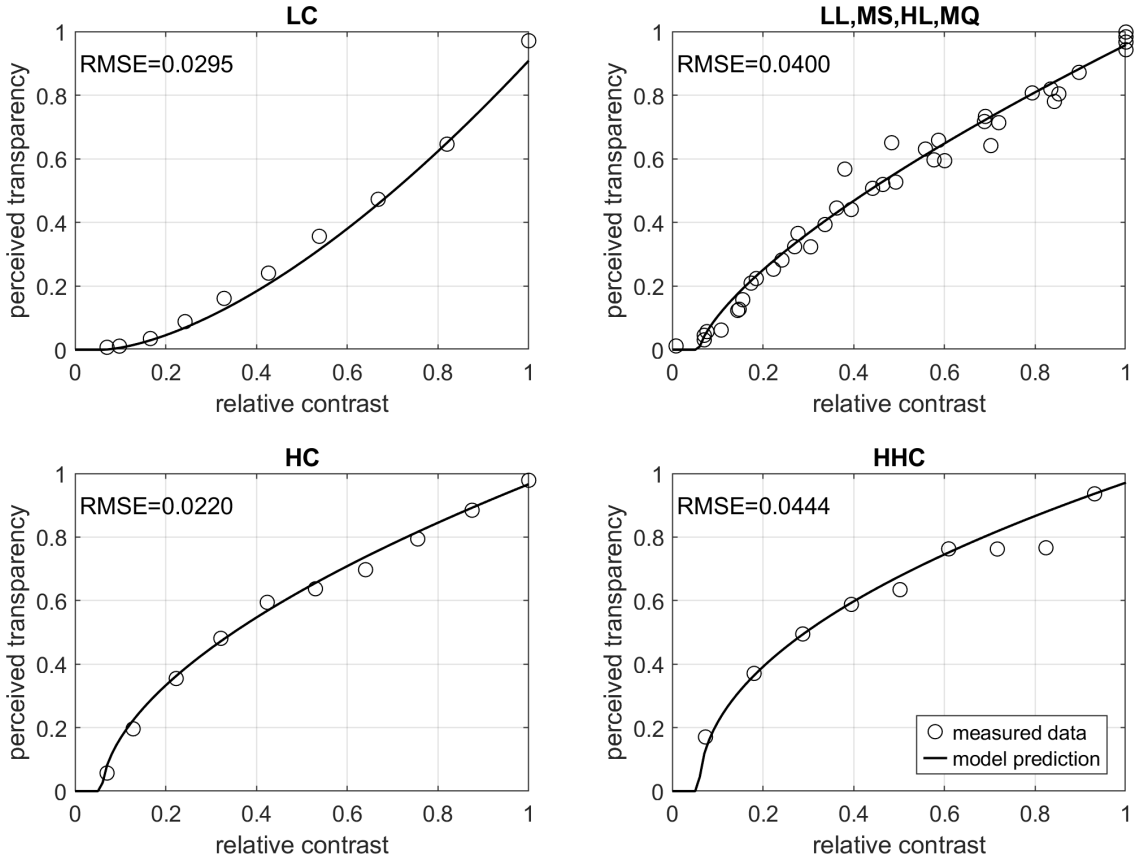


Figure 5.17: Model parameter optimization on AR patches.

squared error (RMSE) at 0.0349 among backgrounds. The largest RMSE is on HHL background at 0.0444. The RMSE of the transparency prediction are 0.0295, 0.0400, 0.0220, and 0.0444 on background contrast of 0.05, 0.2, 0.40, and 0.93 respectively. On the LCD patch, experiment results from all backgrounds are pooled for the evaluation (Figure 5.18). The overall model prediction RMSE is 0.0773.

The proposed transparency model was validated with the result of equivalent transparency experiment. The comparison between the model-estimated AR contrast against the measured is shown in the Figure 5.19. The model shows a good match to the experiment data, with slight overall underestimation of the AR patch contrast in MS background (plotted in green). The coefficients of

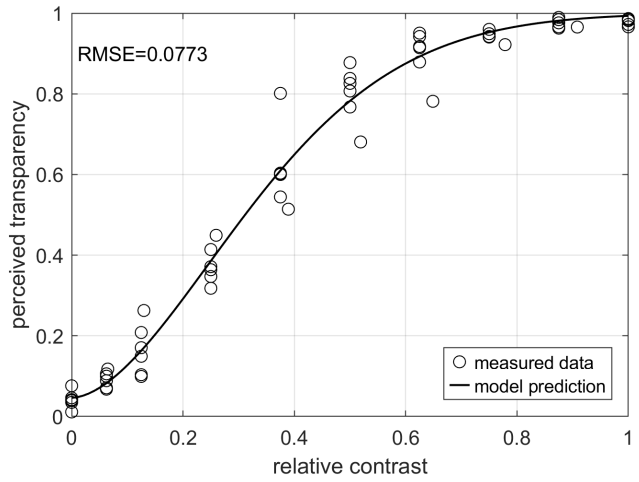


Figure 5.18: Model parameter optimization on LCD patches.

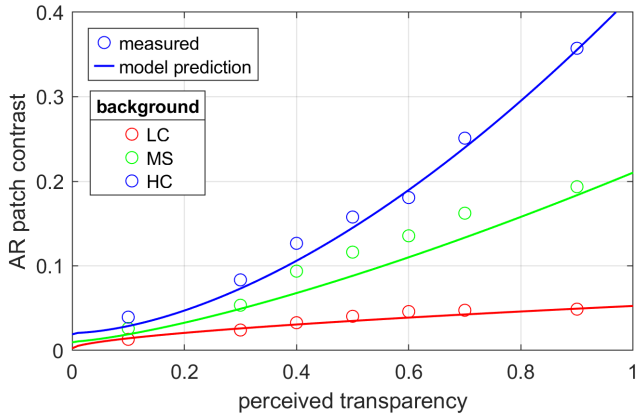


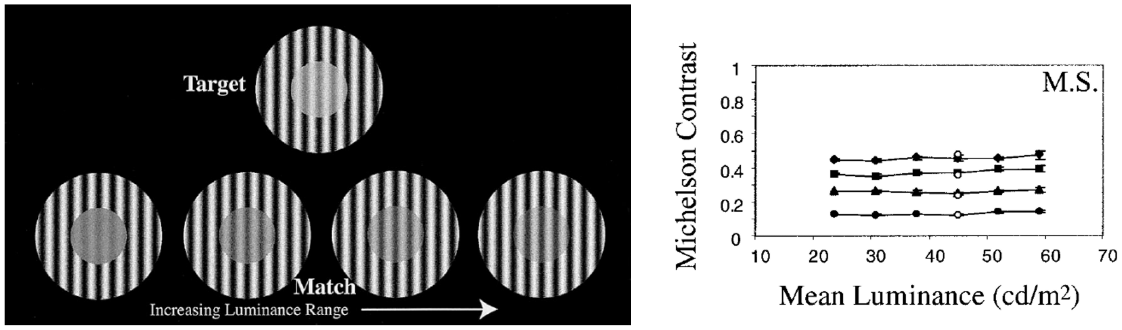
Figure 5.19: Model performance validation on equivalent transparency compared to Experiment result.

determination R^2 on the measured and predicted AR contrast are 0.9703, 0.9843, and 0.9969 on the background LC, MS, HC respectively. RMSE of the model prediction are 0.0041, 0.0210, 0.0120 on the background LC, MS, HC respectively.

5.6.3 Comparison to Existing Transparency Models

In the simulated transparency study by Singh and Anderson [99], they matched perceived transmittance at different average luminance levels by adjusting peak-to-peak luminance difference. Only high contrast (0.945) sinusoidal grating background was used with stereoscopic test patches at a nearer disparity. The result of the three observers showed the same perceived transmittance at the same patch Michelson contrast level, regardless of the average luminance. The perceived transmittance was modeled linearly to the relative contrast to the background. Our result on the AR patches showed that in certain background contrast, the perceived transparency can be linear to the test patch contrast. But this linear relation can be altered with the background contrast. The contrast modulation in their experiment was different from AR condition, where the test patch luminance is always higher than the background. Their result was mostly for transparent filters instead of additive projections. It is also debatable that if the perceived transparency in our experiment is equivalent to the perceived transmittance in their experiment.

In the subtractive filter appearance model by Faul and Ekroll [114], linear perceptual transmittance and clarity (perceived haziness) were modeled based on the physical transmittance and reflectance. The predicted transmittance stays constant at 1 regardless of the background conditions in our experiment setup, so is the clarity prediction on the LCD patch. The clarity prediction on the AR patch (Figure 5.21) does not fully predict the results in our experiment (see Figure 5.10 lower-left subfigure). The predicted clarity on backgrounds with different contrast levels (LC, MS, HC, HHC) overlap with each other and do not reflect the impact of background contrast on transparency.



(a) **(b)**

Figure 5.20: Equivalent transmittance by Singh and Anderson [99]. (a) Observers were asked to to match reference (top) transparency by adjusting the center disk contrast with fixed average luminance (bottom). (b) A typical result showing the transmittance could be predicted from the contrast only and was independent of the average luminance. Each curve represented one reference condition. The x axis is the fixed luminance of the test patch, and the y axis is the adjusted contrast that matched to the reference transmittance.

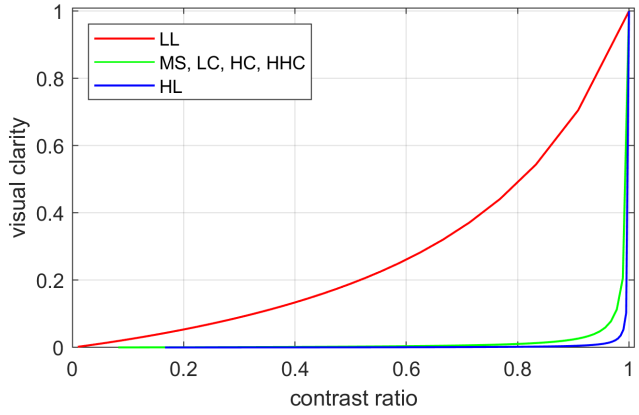


Figure 5.21: Subtractive filter model prediction on clarity by Faul and Ekroll [114] compared to our experiment result. The transmittance prediction stays at 1 regardless of the background and contrast ratio, thus not included in the figure.

5.6.4 Uses and Limitation

The contrast based transparency model was developed to represent the perceptual dimension on additive systems. There are a few assumptions involved: 1. The virtual rendering is naive to the background light intensity and spatial arrangement. It indicates that the virtual image does not have additional information for scene registration or world-locked rendering. As demonstrated in Figure 5.1, the simple additive rendering can only result in patches along the positive y-axis. But if scene registration is possible, the pattern luminance amplitude can be altered as long as the overall luminance change is positive. The experiment and model included only one dimension in the contrast space, while the rendering could be expanded to areas not covered by the experiment setup. For example, though zero physical contrast is not required for the perceptual opacity based on the experiment and model, it is possible with the additional scene information. 2. The model parameter optimization had an assumption that when the background contrast is zero (physically uniform in luminance), the rendering is perceptually opaque. This might be true with our experiment setup. But changes in rendering object dimensionality could violate the assumption. Research has shown that 3D object could be perceived transparent or translucent without the presence of the background. The perception is evoked through object shape, surface roughness, specular, caustic, spatial shading pattern, inferred lighting direction. The same physical property in 2D vs 3D object could trigger opposite transparency perception (see section 2.3.4).

5.6.5 Alternative Bi-modal Transparency

When the human visual system evaluates the transparency, we usually use the method in judging subtractive filters. Two aspects are assessed in terms of physical properties: how much light does the filter transmit, and how much light absorbed and scattered. Similarly, from the observer feedback, we noticed there were two types of criteria were used in both the scaling and the equivalency

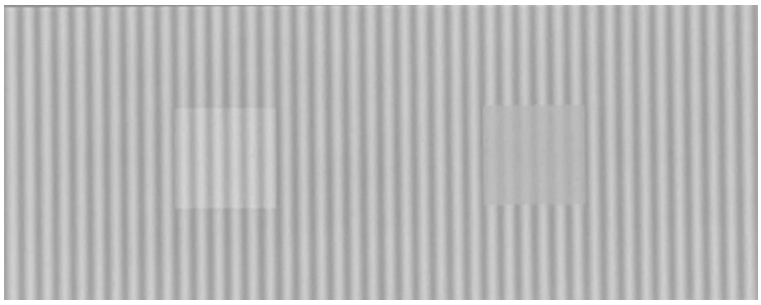


Figure 5.22: Photograph of a difficult trial in the equivalent transparency experiment. The left patch is rendered in AR, increasing overall luminance. The right patch is simulated from LCD, reducing background luminance range. There is ambiguity on which patch appear more transparent.

experiment: 1. Judge transparency according to how visible the background is. Observers used this method, described as "how well I can see the background" or "how clear the layer seems to be". 2. Judge the transparency according to how different the patch area compares to the background. The closer to the background, the more transparent the patch seems to be. In some of the comparison pairs, the two method will present conflicting results. For example, in the comparison in Figure 5.22, with the AR patch on the left and the LCD patch on the right, the left patch gives more visible background, but the right patch area is closer to the background. According to the first method, the left patch is more transparent, but the second method supports the right patch to be more transparent.

We raised a hypothesis based on this feedback to explain the model prediction error. When translating the two physical quantities into perceptual dimensions, the transmittance is evaluated in how much difference between the background and the filter area, and the absorbance is in how visible the background is. In transparency model on subtractive filters, transmittance and clarity are used to describe the two aspect of perceived transparency. The two methods were also used n evaluating AR transparency, but it is the observer's choice on which method to use. For LCD patch, the two criteria are consistent. When the background pattern luminance amplitude reduces, the background

becomes less visible and more different. Both methods are in favor of the less transparent direction. But the two methods sometimes conflict on AR patches. When the added luminance on the AR patch is still relatively low, the background actually becomes more visible and easier to see. But the pattern becomes more different from the background. The two methods give conflicted results on the AR patch in this case. Thus, there might be criteria competing and shifting in our experiment.

In the scaling experiment, the criteria competing did not show as clear as in the equivalency because one patch was presented at a time. But in the equivalency experiment, depending on the criteria of the observer, there might be a separation on the data if different criteria is used on the LCD patch and the AR patch. For the data points deviated from the model, observers needed less light on the AR patch to make it equally transparent. They used method 2 (closer to the background is more transparent). And method 1 dominant more at low contrast background because low visibility of the patch. It was hard to judge how visible the background is when there is only very small difference between the test area and the background.

In Figure 5.9 we noticed sometimes there are two peaks in the histogram. Following the hypothesis, we consider the AR patch transparency distribution being bi-modal in the scaling experiment. Instead of fitting one Gaussian distribution to the data, each distribution is fitted with the summation of two Gaussian distributions with peak contrast ratio at μ_1 and μ_2 as in the Figure 5.23. Two sets of ratings are obtained from the fitting with each background. The bi-modal fitting does not change the result from uni-modal on the effect of the background luminance and contrast on the transparency scale (Figure 5.24).

We now model the transparency based on the bi-modal hypothesis. Two functions are proposed to describe the two peaks in the experiment data. The first function follows the same trend as in the uni-modal condition $f1$, and the other is a linear function $f2$ between the perceived transparency

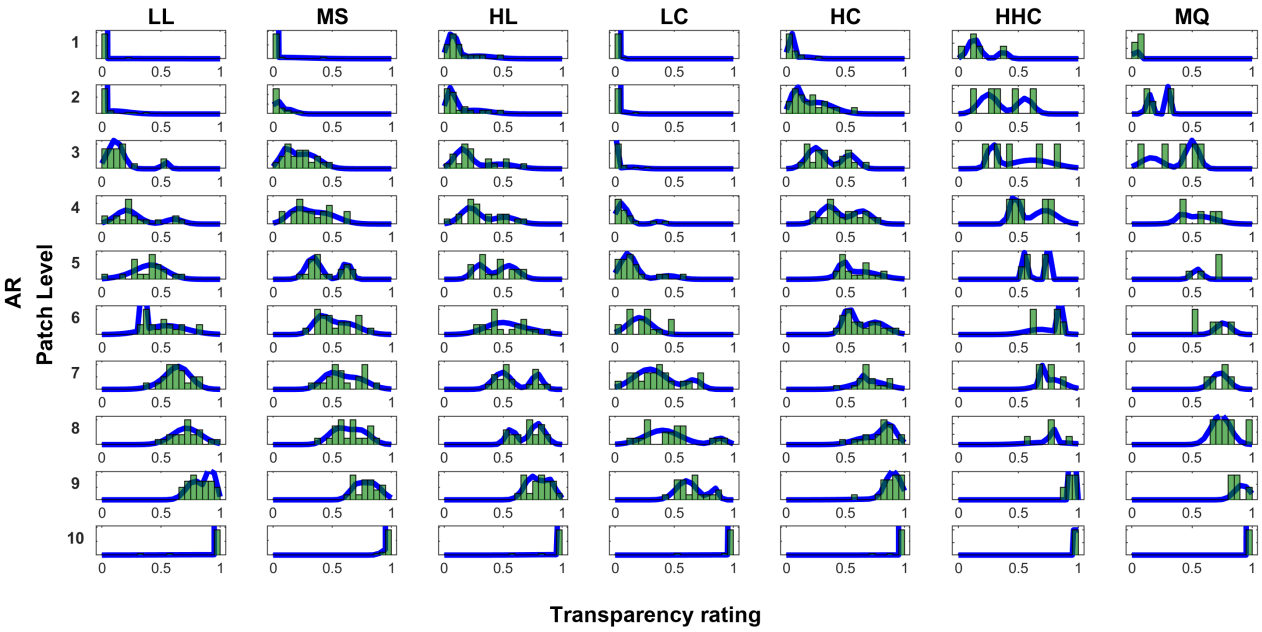


Figure 5.23: AR patch transparency with bi-modal normal distribution estimation on probability density histograms.

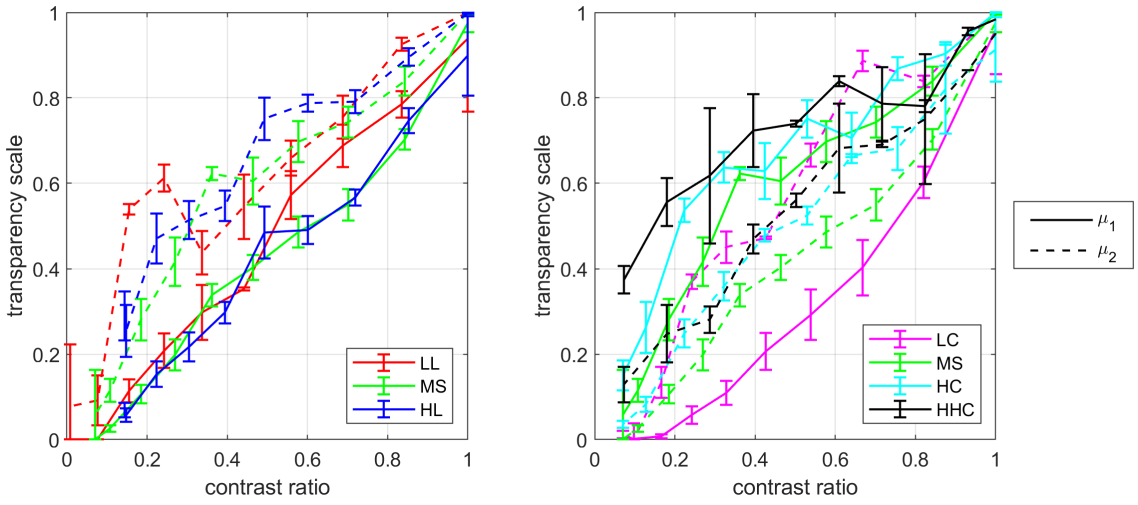


Figure 5.24: Bi-modal estimation on contrast ratio in AR patches with the effect of background luminance (left) and contrast (right). μ_1 and μ_2 are the means of the two distributions.

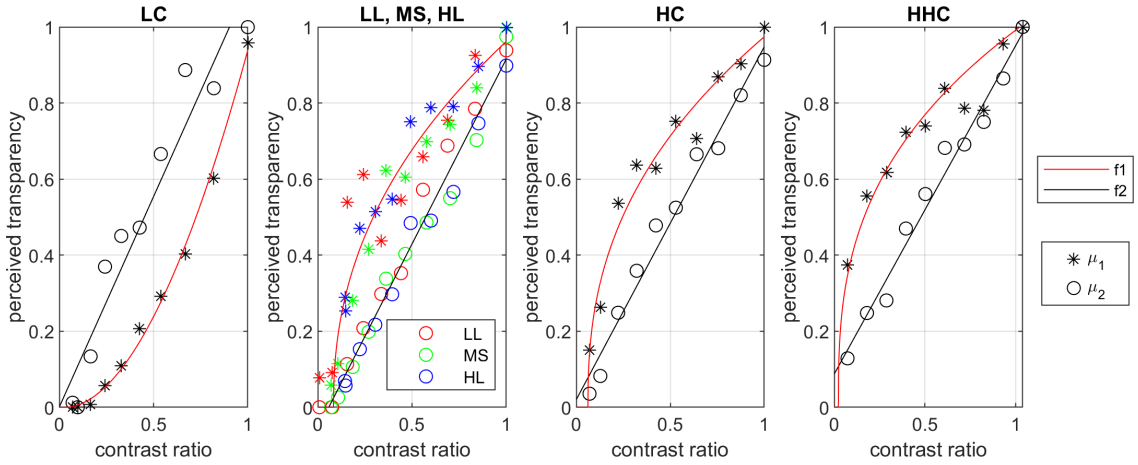


Figure 5.25: Bi-modal parameter optimization on the four background contrast levels. μ_1 was modeled with the power function $f1$, and μ_2 was modeled with a linear function $f2$.

and the contrast ratio:

$$
\begin{aligned}
f1 : T &= \left| \frac{C}{C_b} - 0.0503 \right|^{\frac{0.1292}{C_b} - 0.0188} \\
f2 : T &= 0.9677 \left(\frac{C}{C_b} + 0.0163 \right)
\end{aligned}
\tag{5.8}
$$

The parameters are optimized with the scaling experiment data and shown in Figure 5.25. The 3 backgrounds with the same medium contrast are pooled for the optimization. In each subplot, $f1$ is plotted in red and $f2$ in black. The two functions parameters are optimized for each background contrast levels, and the average values are shown in the equation.

We also plotted the predicted contrast from both uni-modal and bi-modal prediction in the measured contrast histogram to the equivalency experiment. The two predicted contrast seems expanded the matched contrast range. In some of the conditions They match the histogram peaks, but not in all.

Overall the bi-modal transparency prediction were developed based on the experiment feedback

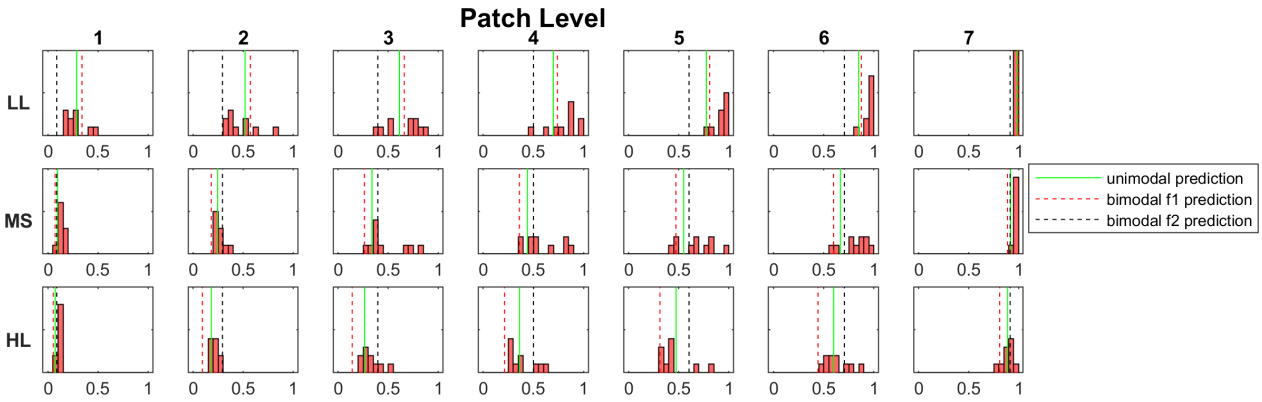


Figure 5.26: Bi-modal transparency prediction compared to the equivalency experiment result. The improvement from the bi-modal model is not significant due to the small dataset.

and subtractive filter transparency properties. The theoretical model explains the difference criteria observers reported, and the bi-modal distribution in both scaling and equivalency experiment. But the hypothesis was nether confirmed nor rejected through the two experiments. We also do not have enough data support to map which of the two functions to either mechanisms (perceiving transmittance or clarity). This is an exploration linking existing subtractive transparency model to AR transparency.

5.7 Summary

The background blending in OST AR introduced the new perceptual dimension: transparency. In this chapter we studied the perceived transparency through three psychophysical experiments. A detection threshold experiment was used to confirm the possibility of AR patch being perceived as non-transparent. The threshold can be modeled with contrast sensitivity functions with background contrast adaptation. A direct anchored scaling experiment was used to build a perceived transparency scale from the relation to background contrast. In addition to AR patches reducing

background contrast by increasing average luminance, patches reducing background contrast from luminance amplitude (LCD patch) were also included to verify if contrast is the fundamental mechanism modulating the transparency. Results showed fundamental differences between the two types of contrast reduction methods. A transparency equivalency experiment was carried out between the two types of patches and further confirms the mechanism difference.

We proposed a mathematical model predicting perceived transparency from the contrast ratio between the stimulus and the background and the stimulus type. On the AR patches the background absolute contrast were also included as a parameter to model the dependency. The model was optimized with the result from the scaling experiment and verified with the equivalency experiment. The model showed good performance on both optimization and verification data. Additionally, an alternative bi-modal transparency hypothesis was also discussed to explain the observer data distribution and underlying transparency assessment mechanism.

Chapter 6

Summary

Augmented reality is believed to be the next generation display system, replacing personal computer monitors and smartphone screens. The novel medium desires image quality improvement to deliver better experiences. Optical see-through AR displays, offering the most direct interaction between the virtual rendering and the real world, face color quality problems caused by the transparent displays with light blended from the real world background. The background blending increases the overall luminance of the virtual images, decreases the image contrast, and introduces image transparency. Lightness, brightness, and transparency are all dependent on the AR rendering luminance and the background conditions. We focused on the three attributes and their interaction with the background and each other in this research. We aimed to understand and model the perceptual dimensions of lightness, brightness, and transparency in AR OST displays, and the interaction with the backgrounds with two parts.

Part I

In the first part of the study, we focused on the perceived brightness and lightness with two experiments. In both experiments, backgrounds varied in luminance, spatial pattern, and contrast levels.

The brightness partition scaling experiment was used to build relative brightness scales on different background luminance levels. Observers were asked to adjust the center patch to be halfway between the two references. Three unbalanced overlapping subranges were used, with three subdivision levels inside each subrange. The experiment results from 17 observers were pooled from subranges to build relative brightness scales. The result showed the dependence of the relative brightness scale on background average luminance, which can be modeled with power law functions. The exponent value of the power functions increased linearly with the background luminance. The exponent value on the sinusoidal pattern background was lower compared to the pattern background, resulting in larger curvature on the brightness scale. Lightness and brightness attributes in color appearance models were demonstrated to be inadequate to predict the visual brightness scale accurately for different background conditions.

The second experiment was used to determine the absolute luminance level required for diffuse white appearance on AR 2D and 3D stimulus. The 3D cube stimuli were rendered through OpenGL with real-time shading and were freely rotated by the observers for compelling visual illustration. The adjusted 2D stimulus luminance showed almost complete background discounting with the AR luminance match to the reference patch, despite the added luminance from the background. The 3D cube luminance was adjusted to be higher than the reference on the brightest surface, but spatial and angular average AR luminance as the overall brightness matched the reference, indicating complete background discounting from integrated perception. The background contrast also showed effect on the adjusted diffuse white luminance levels, with lower luminance on the stimuli.

The result from the two experiments in the first part of the research offered not only confirmation of the conventional brightness-luminance relation, but also the uniqueness of brightness and lightness in AR context. The brightness of AR stimuli blended with patterned backgrounds also followed power law functions but with a different exponent value. The diffuse white level in AR can

be independent of the background with perceptual discounting. The brightness of 3D renderings in AR depended on the shading and the overall spatial presentation. The combined result of the two experiments provides the possibility of anchoring the modeled relative brightness to absolute. Possible anchoring theories were proposed and discussed.

Part II

In the second part of the research, we targeted the inherent appearance property in AR environment: perceived transparency. We studied this novel embedded appearance dimension through three psychophysical experiments. In this part, only patterned backgrounds were included, emphasizing the transparency. The transparency modulation was approached through background Michelson contrast reduction. AR overlay patches reduced the contrast by increasing the average luminance but maintaining the luminance difference. In addition to AR patches, LCD patches with reduced background contrast at the same average luminance were included to verify if contrast is the fundamental mechanism modulating the transparency.

The first experiment in this part measured the transparency detection threshold. The result was to confirm the possibility of the patch being perceived as transparent and non-transparent over a range of contrast provided through our system. The resulting threshold showed the dependence on the background contrast, which can be explained with contrast sensitivity functions and contrast adaptation. There was no significant difference between patch types from AR overlay or LCD. The background average luminance affected the psychometric curve but not the thresholds.

The perceived transparency was studied beyond a single threshold through a direct anchored scaling experiment. The result was used to build perceived transparency scales from the patch contrast in relation to the background's. Observers were asked to rate the test stimulus transparency

on a continuous scale between “completely transparent” to “completely opaque”. The transparency ratings were compared to the contrast ratio between the test patch area and the background to reveal the transparency modulation. Results showed a fundamental difference between the two types of contrast reduction methods with different curve shapes. On AR patches, the background average luminance, and background wave form (sinusoidal or square wave) did not affect the transparency scale, while the background absolute contrast did. On the LCD patches, none of the background conditions impacted the transparency scales. The contrast ratios rated to be “completely opaque” are significantly lower than the threshold measured in the first experiment, due to the longer stimulus presentation duration.

A transparency equivalency experiment was carried out between the two types of patches and further confirmed the mechanism difference. AR patches as test targets were compared to reference LCD patches using a staircase method by varying the AR patch luminance (and resulting contrast ratio). The result confirmed the mechanism difference between the two patch types, with equivalent transparency at different contrast ratio.

A contrast-based transparency model was proposed based on the direct scaling experiment result. It modeled the perceived transparency from the stimulus type and the contrast ratio between the stimulus and the background. On the AR patches, the background absolute contrast was also included as a parameter to model the dependency. The model parameters were optimized with the result from the scaling experiment. On AR patches, the perceived transparency was described with exponential functions of the contrast ratio, with the exponent value being a reciprocal function of the background absolute contrast. The exponent values range from 0 to 0.51, with 0 at minimum background contrast (solid background with contrast of 0), and 0.51 at maximum background contrast (contrast of 1). The model predicted opacity threshold was at contrast ratio of 0.0592, below which were perceptual opacity. With the increased contrast ratio, the patch area luminance decreased, and

the perceived transparency increased, following different paths with the according exponent values. On LCD patches, perceived transparency was described with an accumulative Weibull function on the contrast ratio. The function was independent of the background properties like spatial variance, luminance, and contrast. The model was validated with the equivalency experiment result and showed good performance with average correlated coefficient of 0.9838. Additionally, an alternative bi-modal transparency hypothesis were also discussed to explain the observer data distribution and underlying transparency assessment mechanism.

6.1 Luminance, Brightness, and Transparency

The two parts of this research (brightness and transparency) are the two perceptual sides of the single physical stimulus property (luminance) in the unique AR environment. The two perceptual dimensions are inseparable and define the AR color appearance together. Although the title includes lightness, an additional attribute, it is a relative scale based on brightness and its interpretation was discussed in Chapter 3. From the perspective of the background, its average luminance impacted the brightness perception, but its contrast did not. But the perceptual transparency was highly dependent on the background contrast and independent of the average luminance. The background and the AR combined determined the perceptual dimensions. The result from the first part confirmed existing brightness modeling with power law, with options of adjustment for backgrounds with spatial variation. The diffuse white level offered novel data suggesting complete perceptual background discounting on both 2D and 3D rendering. The diffuse white luminance, as an anchor, also related the brightness scale to lightness. The transparency part of the research supplemented AR color appearance with the new perceptual dimension and mathematical modeling.

One interesting result from this research was the complete background discounting in the diffuse

white experiment. In other researches, only partial background discounting showed on transparent displays. The amount of background discounting demonstrated the perceptual session between the virtual rendering and the real background. Though scission was only a binary concept originally for transparency illusions, we could borrow and extend it to a continuous concept. The complete background discounting indicates full scission, while partial discounting for incomplete scission. The degree of scission could depend on the number of cues indicating the virtual rendering exist in a separated setting from the background. In our setup, the cues included the cube rotation, static background, experiment information provided to observers. In our experiments, though the 3D cubes had stronger separated-layer cues compared to the 2D patch, both of them showed complete scission. This could be because observers were informed that the test stimuli were rendered virtually. It might help enhance the layering perception, hence enhanced scission. Another thing worth notice was the luminance to be matched to was for diffuse white and not highly transparent. If the virtual stimuli were highly transparent, it may give instant layer separation regardless of other cues. In addition to the cues in our experiment, other scission cues could be the depth and focus difference, lighting and tone difference, relative motion, alignment, etc.

Though the background discounting was found in brightness matching, it did not indicate the background could be ignored for OST AR content. The rendering with complete background discounting in brightness could still appear transparent, as indicated in the second part of the research. Although brightness and transparency both rely on the luminance change, they appear to be decoupled perceptually. This was also supported by Anderson's transparency theory of separated perception on lightness and opacity [102] and the multidimensional scaling result from Petrini and Logvinenko [103]. Though both researches were on rendered transparencies instead of real transparent objects or additive transparencies like in OST AR.

Overall, brightness and transparency in OST AR found their paths from centuries of visual per-

ception research. But they also peak into fresh land that color researches have not yet settled on. This research provided new data and models on brightness and transparency, answered questions on perceptual scales of basic achromatic stimuli with controlled backgrounds. The research also explored the impact of rendering dimensionality, relation to reference brightness, and perceptual theories on layered rendering interpretations. These studies are not limited to certain AR hardware but focus on the property of the rendering and background properties, offering the possibility of applying to other additive AR hardware. For vision and color researchers, the study offered new data on human visual perception in the novel OST AR media. The amount of background separation in the experiment demonstrated a distinct way the human visual system process the transparent additive overlaid images. The transparency in OST AR, being different from physical transparency and translucency, could be a unique prospect to help interpret transparency perception and material appearance in general. From the image processing and management perspective, the models in this research can be integrated into color appearance models, improving color and image quality and visual experience for future AR devices. For architects designing future OST AR devices, the numbers on the diffuse white level and opacity threshold could serve as key numbers for devices with respective abilities rendering such contents. The models also provided references on the visual effect on OST AR platform for content creators. The perceptual difference of the same rendering on conventional media compared to OST AR is especially important for cross-platform consistency. The experiment data and theories of this research could impact multiple aspects of AR, providing the perception viewpoints.

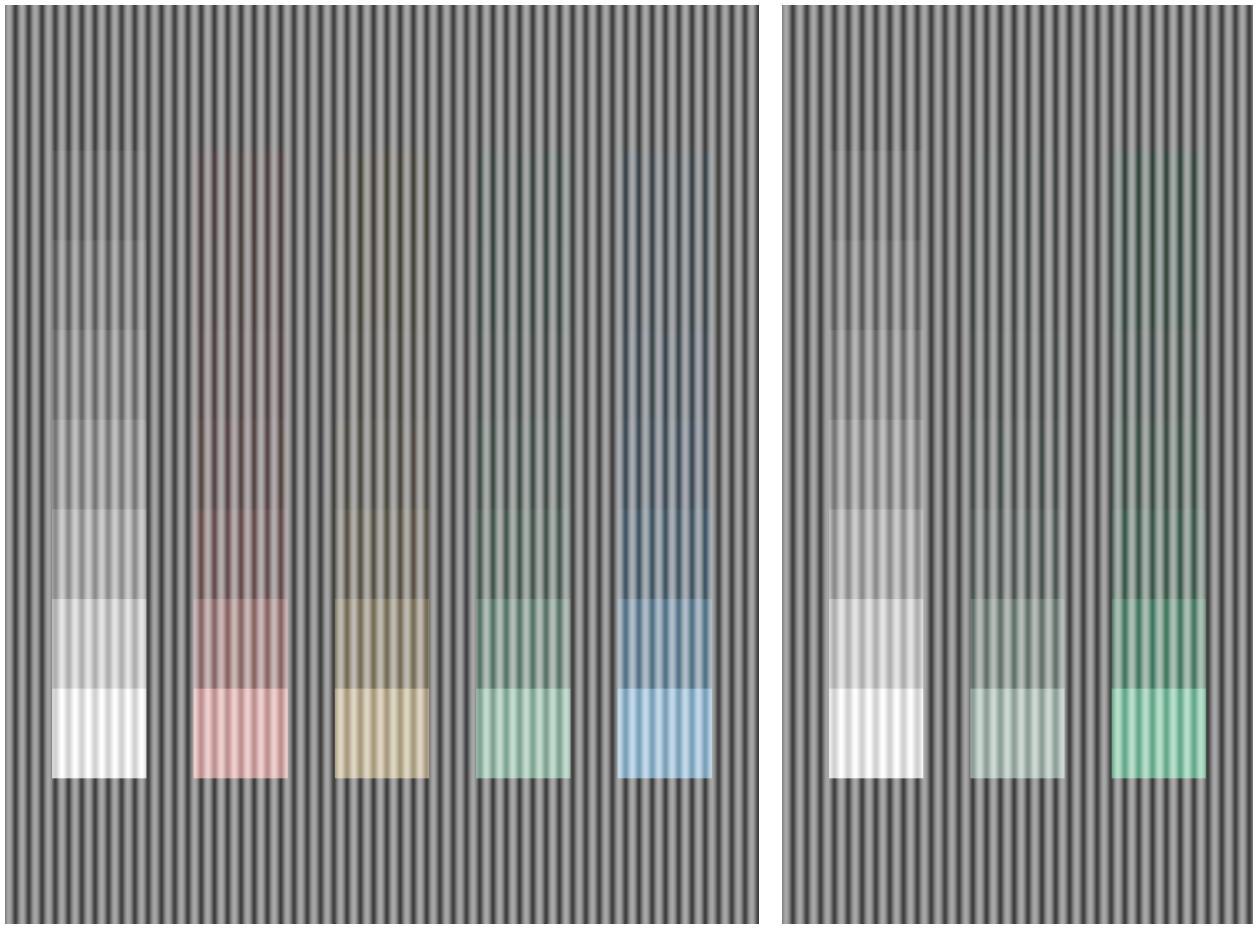
6.2 Future Work

In the scope of this study, we only focused on achromatic aspect for both the AR stimuli and the background. Naturally, this work should be extended to chromatic colors. The first step is to explore

the chromatic AR stimuli on neutral backgrounds. Figure 6.1 illustrates this chromatic transparency using the contrast based transparency model without further modification for chromaticities. Four hue angles of 24°, 90°, 162°, and 246° were selected in CIELAB in each column of Figure 6.1 (a) with the constant chroma compared to the neutral scale. In each column, the AR overlay has the same chromaticities but different luminance. The luminance of each step were set to be the same as the neutral scale. Visually, the four chromatic scales are not as linear as the neutral scale. For each hue, the achromatic background blending not only change the patch overall luminance, but also the saturation. The brightness difference were also apparent due to the Helmholtz–Kohlrausch effect. Furthermore, the Abney effect may present in certain hue angles with the wider adoption of narrowband primaries in head-mounted AR devices, leading to hue change in each column that is not visible from this illustration. Figure 6.1 (b) illustrated the other dimension in chromatic AR colors in chroma with example of green hue. Again, both brightness and transparency showed dependency on chroma from the illustration. Psychophysical examinations are required for brightness and transparency quantification on chromatic AR colors before extending conclusions to them.

Apart from the simple stimuli, the AR content in real images and complex 3D rendering are also of our interest. By introducing the spatial variation in real images and computational shading in the rendering process, paired with the background color blending, the image quality could be dramatically affected. Systematic evaluation of image quality in AR could be helpful for improving AR experience.

The other side of the problem comes from the background. For our research, all backgrounds were limited to achromatic uniform or patterned background in fixed spatial frequency produced by an LCD screen. In real use cases, backgrounds could vary from indoor black surfaces to outdoor bright complex natural scenes. The background luminance, color, spatial frequency could possibly affect the perceived AR rendering quality. The other effect from the background is the adaptation.



(a) **(b)**

Figure 6.1: Linear AR transparency scale on achromatic colors compared to chromatic colors. The chromatic scales were calculated from the same achromatic model, but show less linearity. (a) Linear transparency in achromatic colors and four hues in red, yellow, green, and blue calculated in CIELAB. (b) Linear transparency in achromatic colors and green hue at two chroma levels.

Our experimental result from transparency thresholding demonstrated the contrast adaptation effect. There is also luminance adaptation, chromatic adaptation, spatial frequency adaptation that could impact the AR image quality. For example, how should we adjust the AR display white point when the background ambient lighting is dramatically different from the default display white point? On one side, the background color blending shifts the displayed color towards the background. On the other side, the chromatic adaptation pushes the perception away from the background color. Additionally, the perceptual background discounting effect found in the diffuse white adjustment experiment could also decrease the amount of background blending. The three factors influence the AR perception from different angles. How we weigh each factor, and what the resulting AR perception is, could be key to better AR visual experience.

With the improvement of AR system hardware and the popularity of head-mounted devices, AR will eventually move to more dynamic environments. The interaction between the virtual rendering and the real background will increase to achieve a more immersive experience. Novel AR prototypes have appeared in research labs with capabilities for real world mutual occlusion and mutual shading. These features improve AR sensibility and immersivity, but may require completely different methodology for evaluating image and color appearance in these novel systems.

6.3 Publications

Journal

Lili Zhang, Michael J. Murdoch, and Romain Bachy. "Color Appearance Shift in Augmented Reality Metameric Matching." Journal of the Optical Society of America A 38, no. 5 (May 1, 2021): 701.

Lili Zhang, and Michael J. Murdoch. "Color Matching Criteria in Augmented Reality." Journal of

Perceptual Imaging 1, no. 1 (January 1, 2018): 10506-1-10506–8.

Conference

Lili Zhang, and Michael J. Murdoch. "How Bright Should It Be: Diffuse White in Optical See-Through Augmented Reality." Frameless: Vol. X : Iss. X , Article X (November 2021).

Lili Zhang, and Michael J. Murdoch. 2021. "Perceived Transparency in Optical See-Through Augmented Reality." In 2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), 115–20. Bari, Italy: IEEE.

Lili Zhang, Rachel Albert, Joohwan Kim, and David Luebke. "Developing a Peripheral Color Tolerance Model for Gaze-Contingent Rendering." Journal of Vision 19, no. 10 (September 6, 2019): 298c.

References

[1] I. E. Sutherland, "The Ultimate Display," pp. 506–508, 1965.

[2] D. Van Krevelen and R. Poelman, "A Survey of Augmented Reality Technologies, Applications and Limitations," *International Journal of Virtual Reality* **9**, pp. 1–20, Jan. 2010.

[3] C. I. de L'Eclairage, *International Lighting Vocabulary*, 2 ed., Jan. 2020.

[4] P. Burns, "The History of The Discovery of Cinematography Chapter 10," 1997.

[5] I. E. Sutherland, "A head-mounted three dimensional display," in *Fall Joint Computer Conference*, **Part I**, pp. 757–764, ACM, Dec. 1968.

[6] R. L. Holloway, "Head-Mounted Display Technical Report," tech. rep., Dept. of Computer Science, University of North Carolina at Chapel Hill, 1987.

[7] T. Caudell and D. Mizell, "Augmented reality: an application of heads-up display technology to manual manufacturing processes," in *Proceedings of the Twenty-Fifth Hawaii International Conference on System Sciences*, pp. 659–669 vol.2, IEEE, (Kauai, HI, USA), 1992.

[8] P. Milgram and F. Kishino, "A Taxonomy of mixed reality visual displays," *IEICE Transaction on Information Systems* **E77-D**(12), p. 16, 1994.

[9] P. Chen, X. Liu, W. Cheng, and R. Huang, "A review of using Augmented Reality in Education from 2011 to 2016," in *Innovations in Smart Learning*, E. Popescu, Kinshuk, M. K.

Khribi, R. Huang, M. Jemni, N.-S. Chen, and D. G. Sampson, eds., pp. 13–18, Springer Singapore, (Singapore), 2017.

[10] R. G. Boboc, F. Gîrbacia, and E. V. Butilă, "The Application of Augmented Reality in the Automotive Industry: A Systematic Literature Review," *Applied Sciences* **10**, p. 4259, June 2020.

[11] G. Haas, "Invited Paper: Microdisplays for Augmented and Virtual Reality," *SID Symposium Digest of Technical Papers* **49**(1), pp. 506–509, 2018. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/sdtp.12445.

[12] K. Yin, Z. He, J. Xiong, J. Zou, K. Li, and S.-T. Wu, "Virtual reality and augmented reality displays: advances and future perspectives," *Journal of Physics: Photonics* **3**, p. 022010, Apr. 2021.

[13] S. Yamazaki, K. Inoguchi, Y. Saito, H. Morishima, and N. Taniguchi, "Thin wide-field-of-view HMD with free-form-surface prism and applications," pp. 453–462, (San Jose, CA), May 1999.

[14] H. Hua, X. Hu, and C. Gao, "A high-resolution optical see-through head-mounted display with eyetracking capability," *Optics Express* **21**, p. 30993, Dec. 2013.

[15] B. C. Kress and W. J. Cummings, "Invited Paper: Towards the Ultimate Mixed Reality Experience: HoloLens Display Architecture Choices," *SID Symposium Digest of Technical Papers* **48**, pp. 127–131, May 2017.

[16] Y.-H. Lee, K. Yin, and S.-T. Wu, "Reflective polarization volume gratings for high efficiency waveguide-coupling augmented reality displays," *Optics Express* **25**, p. 27008, Oct. 2017.

[17] C. Yu, Y. Peng, Q. Zhao, H. Li, and X. Liu, "Highly efficient waveguide display with space-variant volume holographic gratings," *Applied Optics* **56**, p. 9390, Dec. 2017.

[18] H. Mukawa, K. Akutsu, I. Matsumura, S. Nakano, T. Yoshida, M. Kuwahara, K. Aiki, and M. Ogawa, "Distinguished Paper: A Full Color Eyewear Display Using Holographic Planar Waveguides," *SID Symposium Digest of Technical Papers* **39**(1), p. 89, 2008.

[19] Y. Itoh, T. Langlotz, D. Iwai, K. Kiyokawa, and T. Amano, "Light Attenuation Display: Subtractive See-Through Near-Eye Display via Spatial Color Filtering," *IEEE Transactions on Visualization and Computer Graphics* **25**, pp. 1951–1960, May 2019.

[20] R. Akiyama, G. Yamamoto, T. Amano, T. Taketomi, A. Plopski, C. Sandor, and H. Kato, "Light Projection-Induced Illusion for Controlling Object Color," in *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 499–500, IEEE, (Reutlingen), Mar. 2018.

[21] J. L. Gabbard, J. E. Swan, J. Zedlitz, and W. W. Winchester, "More than meets the eye: An engineering study to empirically examine the blending of real and virtual color spaces," in *2010 IEEE Virtual Reality Conference (VR)*, pp. 79–86, IEEE, (Boston, MA, USA), Mar. 2010.

[22] S. K. Sridharan, J. D. Hincapié-Ramos, D. R. Flatla, and P. Irani, "Color correction for optical see-through displays using display color profiles," in *Proceedings of the 19th ACM Symposium on Virtual Reality Software and Technology - VRST '13*, p. 231, ACM Press, (Singapore), 2013.

[23] J.-H. Ryu, J.-W. Kim, K.-K. Lee, and J.-O. Kim, "Colorimetric background estimation for color blending reduction of OST-HMD," in *2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA)*, pp. 1–4, IEEE, (Jeju, South Korea), Dec. 2016.

[24] T. Langlotz, M. Cook, and H. Regenbrecht, "Real-Time Radiometric Compensation for Optical See-Through Head-Mounted Displays," *IEEE Transactions on Visualization and Computer Graphics* **22**, pp. 2385–2394, Nov. 2016.

[25] J.-W. Kim, K.-K. Lee, J.-H. Ryu, and J.-O. Kim, "Localized color correction for optical see-through displays via weighted linear regression," in *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology - VRST '16*, pp. 11–14, ACM Press, (Munich, Germany), 2016.

[26] J. David Hincapie-Ramos, L. Ivanchuk, S. K. Sridharan, and P. Irani, "SmartColor: Real-time color correction and contrast for optical see-through head-mounted displays," in *2014 IEEE ISMAR*, pp. 187–194, IEEE, (Munich, Germany), Sept. 2014.

[27] C. Weiland, A.-K. Braun, and W. Heiden, "Colorimetric and Photometric Compensation for Optical See-Through Displays," in *Universal Access in Human-Computer Interaction. Intelligent and Ubiquitous Interaction Environments*, C. Stephanidis, ed., **5615**, pp. 603–612, Springer Berlin Heidelberg, Berlin, Heidelberg, 2009. Series Title: Lecture Notes in Computer Science.

[28] Y. Zhang, R. Wang, W. Hua, and H. Li, "Color Contrast Enhanced Rendering for Optical See-Through Head-Mounted Displays with Optimized Display Power," in *SID Symposium Digest of Technical Papers*, pp. 1258–1261, 2021.

[29] K.-K. Lee, J.-W. Kim, J.-H. Ryu, and J.-O. Kim, "Ambient light robust gamut mapping for optical see-through displays," *Optics Express* **28**, p. 15392, May 2020.

[30] S. K. Nayar, H. Peri, M. D. Grossberg, and P. N. Belhumeur, "A projection system with radiometric compensation for screen imperfections," 2003.

[31] M. Tsukada and J. Tajima, "Projector Color Reproduction Adapted to the Colored Wall Projection," *Conference on Colour in Graphics, Imaging, and Vision* **2004**, pp. 449–453, Jan. 2004.

[32] D.-C. Kim, T.-H. Lee, H.-G. Ha, and Y.-H. Ha, "Color Correction Using a Still Camera for

Images Projected onto a Light Colored Screen," *Journal of Imaging Science and Technology* **55**(2), p. 020507, 2011.

[33] C.-H. Son and Y.-H. Ha, "Color Correction of Images Projected on a Colored Screen for Mobile Beam Projector," *Journal of Imaging Science and Technology* **52**(3), p. 030505, 2008.

[34] C. Menk and R. Koch, "Physically-based augmentation of real objects with virtual content under the influence of ambient light," in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops*, pp. 25–32, IEEE, (San Francisco, CA, USA), June 2010.

[35] C. Menk and R. Koch, "Interactive Visualization Technique for Truthful Color Reproduction in Spatial Augmented Reality Applications," p. 8, IEEE, (Basel, Switzerland), 2011.

[36] C. Menk and R. Koch, "Truthful Color Reproduction in Spatial Augmented Reality Applications," *IEEE Transactions on Visualization and Computer Graphics* **19**, pp. 236–248, Feb. 2013.

[37] P. Ferschin, M. D. Angelo, and S. Niedermair, "Architectural Heritage Projection in Situ," in *International Conference on Cultural Heritage and New Technologies*, p. 9, Museen der Stadt Wien – Stadtarchäologie, (Vienna), 2012.

[38] H. Kim, Y.-J. Seo, and Y. Kwak, "Transparent effect on the gray scale perception of a transparent OLED display," *Optics Express* **26**, p. 4075, Feb. 2018.

[39] M. J. Murdoch, "Brightness matching in optical see-through augmented reality," *Journal of the Optical Society of America A* **37**, p. 1927, Dec. 2020.

[40] H. Huang, M. Wei, and S. Chen, "White appearance of virtual stimuli produced by augmented reality," *Color Research & Application* **46**, pp. 294–302, Apr. 2021.

[41] N. Hassani and M. J. Murdoch, "Investigating color appearance in optical see-through augmented reality," *Color Research & Application* **44**, pp. 492–507, Aug. 2019.

[42] N. Hassani, *Modeling Color Appearance in Augmented Reality*. PhD thesis, RIT, 2019.

[43] L. Zhang, M. J. Murdoch, and R. Bachy, "Color appearance shift in augmented reality metameric matching," *Journal of the Optical Society of America A* **38**, p. 701, May 2021.

[44] K. Moffitt and M. P. Browne, "Visibility of color symbology in head-up and head-mounted displays in daylight environments," *Optical Engineering* **58**, p. 1, Mar. 2019.

[45] S. Lee, H. Ha, Y. Kwak, H. Kim, Y.-j. Seo, and B. Yang, "Preferred tone curve characteristics of transparent display under various viewing conditions," p. 93950Y, (San Francisco, California, USA), Feb. 2015.

[46] J.-S. Kim and S.-W. Lee, "Study on how to improve visibility of transparent display for augmented reality under various environment conditions," *Optics Express* **28**, p. 2060, Jan. 2020.

[47] H.-P. Huang, M. Wei, H.-C. Li, and L.-C. Ou, "Optimal Text-background Lightness Combination for Enhancing Visual Clarity Using a Head-up Display under Different Surround Conditions," *Color and Imaging Conference* **2020**, pp. 210–214, Nov. 2020.

[48] F. A. Kingdom, "Lightness, brightness and transparency: A quarter century of new ideas, captivating demonstrations and unrelenting controversy," *Vision Research* **51**(7), pp. 652–673, 2010.

[49] A. L. Gilchrist, *Lightness, Brightness and Transparency*, Psychology Press, Sept. 2013. Google-Books-ID: GpPXNhf1bmkC.

[50] N. Moroney, M. Fairchild, R. Hunt, and C. Li, "The CIECAM02 color appearance model," in *10th Color and Imaging Conference Final Program and Proceedings*, pp. 23–27(5), IS&T, 2002.

[51] C. Li, Z. Li, Z. Wang, Y. Xu, M. R. Luo, G. Cui, M. Melgosa, M. H. Brill, and M. Pointer, "Comprehensive color solutions: CAM16, CAT16, and CAM16-UCS," *Color Research & Application* **42**, pp. 703–718, Dec. 2017.

[52] G. T. Fechner, *Elemente der Psychophysik*, Breitkopf u. Härtel, 1860. Google-Books-ID: bcA6AAAAcAAJ.

[53] S. S. Stevens, "To Honor Fechner and Repeal His Law," *Science* **133**, 1961.

[54] CIE, "Colorimetry, 3rd edition | CIE," Tech. Rep. 015, CIE, 2004.

[55] F. Ebner and M. D. Fairchild, "Development and Testing of a Color Space (IPT) with Improved Hue Uniformity," *Color and Imaging Conference* **1998**, pp. 8–13, Jan. 1998.

[56] C. J. Bartleson and E. J. Breneman, "Brightness Perception in Complex Fields," *Journal of the Optical Society of America* **57**, p. 953, July 1967.

[57] M. D. Fairchild, *Color Appearance Models*, John Wiley & Sons, June 2013. Google-Books-ID: 1BT9R6FjVhIC.

[58] R. Hunt, "A model of colour vision for practical applications," *Color Research & Application* **7**(2), pp. 95–112, 1982.

[59] E. H. Land and J. J. McCann, "Lightness and Retinex Theory," *Journal of the Optical Society of America* **61**, p. 1, Jan. 1971.

[60] M. E. Rudd, "Edge integration in achromatic color perception and the lightness-darkness asymmetry," *Journal of Vision* **13**, pp. 18–18, Dec. 2013.

[61] S. S. Bergstrom, "Common and relative components of reflected light as information about the illumination, colour, and three-dimensional form of ob-

jects," *Scandinavian Journal of Psychology* **18**(1), pp. 180–186, 1977. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-9450.1977.tb00275.x.

[62] A. L. Gilchrist and A. Jacobsen, "Lightness Constancy Through a Veiling Luminance," *Journal of Experimental Psychology: Human Perception and Performance* **9**(6), pp. 936–944, 1983.

[63] H. G. Barrow and J. M. Tenebaum, "Recovering Intrinsic Scene Characteristics fromImages," in *Computer Vision Systems*, pp. 455–471, Springer Berlin Heidelberg, (Berlin, Heidelberg), 1978. Series Title: Lecture Notes in Computer Science.

[64] H. Wallach, *On perception*, On perception, Quadrangle, Oxford, England, 1976. Pages: xviii, 490.

[65] J. Cataliotti and A. Gilchrist, "Local and global processes in surface lightness perception," *Perception & Psychophysics* **57**, pp. 125–135, Jan. 1995.

[66] A. Gilchrist, C. Kossyfidis, F. Bonato, T. Agostini, J. Cataliotti, X. Li, B. Spehar, V. Annan, and E. Economou, "An Anchoring Theory of Lightness Perception," *Psychol Rev* **106**(4), pp. 795–834, 1999.

[67] H. Takasaki, "Lightness Change of Grays Induced by Change in Reflectance of Gray Background*," *Journal of the Optical Society of America* **56**, p. 504, Apr. 1966.

[68] N. Moroney, "Chroma Scaling and Crispening," *JOURNAL OF IMAGING SCIENCE AND TECHNOLOGY* **46**, p. 8, 2002.

[69] N. Moroney, "Factors effecting lightness partitioning," pp. 35–42, (San Jose, CA), Dec. 2001.

[70] P. Whittle, "Brightness, discriminability and the "Crispening Effect"," *Vision Research* **32**, pp. 1493–1507, Aug. 1992.

[71] G. Cui, M. R. Luo, and B. Rigg, "Crispening effect on lightness differences," p. 634, (Rochester, NY), June 2002.

[72] M. Hajdek and K. Hajdek, "Color Perceptionof the Observer with the Manifestation of the Chromatic Effect of Crispening," *Tehnički glasnik* **12**, pp. 159–165, Sept. 2018.

[73] C. C. Semmelroth, "Prediction of Lightness and Brightness on Different Backgrounds," *Journal of the Optical Society of America* **60**, p. 1685, Dec. 1970.

[74] D. Kane, "Can 'crispening' be explained by contrast gain?," *Electronic Imaging* **2017**, pp. 182–187, Jan. 2017.

[75] P.-h. Chen, *Scaling lightness perception and differences above and below diffuse white and modifying color spaces for high-dynamic-range scenes and images*. PhD thesis, RIT, 2011.

[76] M. D. Fairchild and P.-H. Chen, "Brightness, lightness, and specifying color in high-dynamic-range scenes and images," p. 78670O, (San Francisco Airport, California, USA), Jan. 2011.

[77] M. A. Abebe, T. Pouli, M.-C. Larabi, and E. Reinhard, "Perceptual Lightness Modeling for High-Dynamic-Range Imaging," *ACM Transactions on Applied Perception* **15**, pp. 1–19, Nov. 2017.

[78] E. H. Weber, *E.H. Weber on the Tactile Senses*, Psychology Press, 1996. Google-Books-ID: xEd8JglYzFwC.

[79] A. A. Michelson, *Studies in Optics*, University of Chicago Press, 1927. Google-Books-ID: FXFTnQEACAAJ.

[80] R. S. Berns, "The Mathematical Development of CIE TC 1-29 Proposed Color Difference Equation: CIELCH," *C19-1* **B**, AIC COLOR, 1993.

[81] M. R. Luo, G. Cui, and B. Rigg, "The Development of the CIE 2000 Colour-Difference Formula: CIEDE2000," *Color Research & Application* **26**(5), pp. 340–350, 2001.

[82] P. Whittle, "The psychophysics of contrast brightness," in *Lightness, brightness, and transparency*, pp. 35–110, Lawrence Erlbaum Associates, Inc, Hillsdale, NJ, US, 1994.

[83] F. A. Kingdom and P. Whittle, "Contrast discrimination at high contrasts reveals the influence of local light adaptation on contrast processing," *Vision Research* **36**, pp. 817–829, Mar. 1996.

[84] R. Carter, M. Ayama, P. Green, E. Krupinski, C. Oleari, and K. Smet, "CIE 228:2018 GREY-SCALE CALCULATION FOR SELF-LUMINOUS DEVICES," tech. rep., International Commission on Illumination (CIE), Mar. 2018.

[85] R. Carter, "CIE self-luminous gray-scale calculation: inflections, parameters, and high-contrast limiting behavior," *Journal of the Optical Society of America A* **37**, p. 115, Jan. 2020.

[86] F. Grum and R. Becherer, *Optical Radiation Measurements - Radiometry*, vol. 1, Academic Press, NY, 1979.

[87] R. J. Tilley, *Colour and the Optical Properties of Materials*, John Wiley & Sons, second ed., 2011.

[88] R. S. Berns, *Billmeyer and Saltzman's Principles of Color Technology*, John Wiley & Sons, Mar. 2019.

[89] ASTM E12 Committee, "Terminology of Appearance," 2017.

[90] R. Hunt and M. Pointer, *Measuring Colour*, The Wiley-IS&T Series in Imaging Science and Technology, John Wiley & Sons, fourth ed., 2011.

[91] D. Gigilashvili, J.-B. Thomas, J. Y. Hardeberg, and M. Pedersen, "Translucency perception: A review," *Journal of Vision* **21**, p. 4, Aug. 2021.

[92] R. W. Fleming and H. H. Bülthoff, "Low-Level Image Cues in the Perception of Translucent Materials," *ACM Transactions on Applied Perception* **2**, pp. 346–382, July 2005.

[93] D. Gigilashvili, L. Dubouchet, J. Y. Hardeberg, and M. Pedersen, "Caustics and translucency perception," in *material appearance*, pp. 033–1–6, IS&T, 2020.

[94] F. Metelli, "The Perception of Transparency," *Scientific American* **230**, pp. 90–98, Apr. 1974.

[95] B. L. Anderson and J. Winawer, "Layered image representations and the computation of surface lightness," *Journal of Vision* **8**, p. 18, July 2008.

[96] W. Gerbino, "Transparent Layer Constancy," *Journal of Experimental Psychology: Human Perception and Performance* **16**, p. 18, 1990.

[97] B. L. Anderson, "A Theory of Illusory Lightness and Transparency in Monocular and Binocular Images: The Role of Contour Junctions," *Perception* **26**, pp. 419–453, Apr. 1997.

[98] B. L. Anderson, "Stereoscopic Surface Perception," *Neuron* **24**, pp. 919–928, Dec. 1999.

[99] M. Singh and B. L. Anderson, "Toward a perceptual theory of transparency.," *Psychological Review* **109**(3), pp. 492–519, 2002.

[100] B. L. Anderson, M. Singh, and J. Meng, "The perceived transmittance of inhomogeneous surfaces and media," *Vision Research* **46**, pp. 1982–1995, June 2006.

[101] B. L. Anderson and J. Winawer, "Image segmentation and lightness perception," *Nature* **434**, pp. 79–83, Mar. 2005.

[102] B. L. Anderson, M. Singh, and J. O'Vari, "Natural decompositions of perceived transparency: Reply to Albert (2008).," *Psychological Review* **115**, pp. 1144–1151, Oct. 2008.

[103] K. Petrini and A. Logvinenko, "Multidimensional scaling (MDS) analysis of achromatic transparency," *Journal of Vision* **6**, pp. 394–394, June 2006. Publisher: The Association for Research in Vision and Ophthalmology.

[104] R. Robilotto, B. Khang, and Q. Zaidi, "Perceived transparency: Tradeoffs between reflectance and transmittance," *Journal of Vision* **1**, pp. 431–431, Dec. 2001. Publisher: The Association for Research in Vision and Ophthalmology.

[105] B.-G. Khang and Q. Zaidi, "Illuminant color perception of spectrally filtered spotlights," *Journal of Vision* **4**, p. 2, Aug. 2004.

[106] J. Beck, "Additive and subtractive color mixture in color transparency," *Perception & Psychophysics* **23**, pp. 265–267, May 1978.

[107] J. Beck, K. Prazdny, and R. Ivry, "The perception of transparency with achromatic colors," *Perception & Psychophysics* **35**, pp. 407–422, Sept. 1984.

[108] M. D'Zmura, P. Colantoni, K. Knoblauch, and B. Laget, "Color Transparency," *Perception* **26**, pp. 471–492, Apr. 1997. Publisher: SAGE Publications Ltd STM.

[109] V. J. Chen and M. D'Zmura, "Test of a Convergence Model for Color Transparency Perception," *Perception* **27**, pp. 595–608, May 1998. Publisher: SAGE Publications Ltd STM.

[110] M. D'Zmura, O. Rinner, and K. R. Gegenfurtner, "The Colors Seen behind Transparent Filters," *Perception* **29**, pp. 911–926, Aug. 2000.

[111] F. Faul and V. Ekroll, "Psychophysical model of chromatic perceptual transparency based on substractive color mixture," *Journal of the Optical Society of America A* **19**, p. 1084, June 2002.

[112] F. Faul and V. Ekroll, "On the filter approach to perceptual transparency," *Journal of Vision* **11**, pp. 7–7, June 2011.

[113] F. Faul and V. Ekroll, "Transparent layer constancy," *Journal of Vision* **12**, pp. 7–7, Nov. 2012.

[114] F. Faul, "Toward a Perceptually Uniform Parameter Space for Filter Transparency," *ACM Transactions on Applied Perception* **14**, pp. 1–21, Feb. 2017.

[115] R. Ennis and K. Doerschner, "The color appearance of curved transparent objects," *Journal of Vision* **21**, p. 20, May 2021.

[116] CIE, "Fundamental Chromaticity Diagram with Physiological Axes – Part 2: Spectral Luminous Efficiency Functions and Chromaticity Diagrams," 2015.

[117] M. Fairchild and D. Wyble, "Colorimetric characterization of the Apple studio display (Flat panel LCD)," technical report, Munsell Color Science Laboratory, 1998.

[118] E. W. Jin, B. W. Keelan, J. Chen, J. B. Phillips, and Y. Chen, "Softcopy quality ruler method: implementation and validation," p. 724206, (San Jose, CA), Jan. 2009.

[119] L. Van de Perre, W. R. Ryckaert, M. Dujardin, P. Hanselaer, and K. A. Smet, "Derivation of Brightness Scales Using Partition Scaling," *LEUKOS* **17**, pp. 1–15, Oct. 2019.

[120] P. G. J. Barten, "Formula for the contrast sensitivity of the human eye," pp. 231–238, (San Jose, CA), Dec. 2003.

[121] T. Downs and M. J. Murdoch, "Color Layer Scissioning in See-Through Augmented Reality," pp. 60–65, Society for Imaging Sciience and Technology, Nov. 2021.

[122] L. Hellwig and M. D. Fairchild, "Brightness, lightness, colorfulness, and chroma in CIECAM02 and CAM16," *Color Research & Application* **(submitted)**, p. 11, 2021.

[123] R. W. Baloh, A. W. Sills, W. E. Kumley, and V. Honrubia, "Quantitative measurement of saccade amplitude, duration, and velocity," *Neurology* **25**, pp. 1065–1065, Nov. 1975. Publisher:

Wolters Kluwer Health, Inc. on behalf of the American Academy of Neurology Section: Articles.