11-2021

Color Layer Scissioning in See-Through Augmented Reality

Tucker Downs
Rochester Institute of Technology

Michael Murdoch
Rochester Institute of Technology

https://doi.org/10.2352/issn.2169-2629.2021.29.60

Color Layer Scissioning in See-Through Augmented Reality

Tucker Downs, Michael Murdoch
Rochester Institute of Technology, Rochester, NY, USA

Abstract

Color appearance of transparent objects is not adequately described by colorimetry or color appearance models. Despite the fact that the retinal projection of a transparent object is a combination of its color and the background, measurements of this physical combination fail to predict the saliency with which we perceive the object's color. When the perceive color forms in the mind, awareness of their physical relationship separates the physical combination into two unique perceptions. This is known as color scissioning.

In this paper a psychophysical experiment utilizing a see-through augmented reality display to compare virtual transparent color samples to real color samples is described and confirms the scissioning effect for lightness and chroma attributes. A previous model of color scissioning for AR viewing conditions is tested against this new data and does not satisfactorily predict the observers' perceptions. However, the model is still found to be a useful tool for analyzing the color scissioning and provides valuable insight on future research directions.

Introduction

In see through displays where observers can directly perceive the background behind the display as well as the emissive stimuli on the display an interesting visual phenomenon arises called "color scissioning" whereby the observers can separate and saliently perceive the color of the display separately from the color of the background, despite the two layers overlapping on the retina. This is true of all transparent objects, and in some cases, can even be faked by means of the transparency illusion on a printed figure. In the case of transparent objects, there is a perception of "the body color" and observers may be asked, "what is the color of this piece of stained glass?" In essence, for augmented reality (AR) viewing conditions with an emissive transparent display, the intent is to understand "what is the (body) color of the AR stimulus?"

Some of the early work on the phenomenon of color scissioning comes from Fabio Metelli, who asked the question, "How is it that two shades of gray give rise to the same shade of gray in the transparent layer that is perceived?" and used colored slips of paper to demonstrate this. By arranging at least three slips of paper of different colors he could create the illusion that, rather than three colors on a single layer there were two colors in separate layers, where one layer was partly transparent (Figure 1).

Metelli found that the luminance relationships between the layers were the primary contributor the illusion. If the luminance ratio between the middle color and either the left or right was too great then the illusion would be broken [1]. It had to *seem* plausible that the middle color could be ascribed to a mixture of both the left and right colors.

However, rather than the illusion of transparency, in optical see-through AR the imagery is truly physically transparent. This research addresses how the transparent AR layer is perceived against background viewing conditions. In AR displays, an emissive stimulus is displayed by reflecting from a partly reflective and partly transparent glass. Other display geometries / technologies are possible but not discussed in this paper. For a simple and high-quality research display for stationary viewing, a thin teleprompter mirror can be used between the observer and the desired background and overlay an image displayed on a common LCD display.

Previous work has attempted to explain the perception of AR displays by simply adding together the tristimulus values of the background and the AR layer. This is known as the "proximal colorimetry" and is representative of the light that reaches the first surface of the eye [2]. This simple sum describes the physics, but does not include any psychological effect of color scissioning which recent work has shown to be important for the perception of lightness in AR display contexts[3]–[5].

Figure 1. Transparency illusion as studied by Metelli. Three colors, which would have been paper cut outs, are arranged to give the illusion of a moon shape and a transparent circle in separate layers.

The failure of this physical match has been studied by Hassani and Murdoch. Hassani proposes a model of additivity utilizing two independent coefficients, α and β for the contributions of the foreground and background respectively (Equation 1). In a physical combination these coefficients are each strictly equal to 1, producing the proximal colorimetry. Hassani hypothesizes that by allowing each coefficient to vary independently as a positive number, the color appearance may be predicted [6]. Murdoch later found the ability of this model to predict lightness matches under various conditions very promising [4]. This study expands the investigation to include lightness and chroma simultaneously.

$$XYZ_{effective} = \alpha * XYZ_{arDisplay} + \beta * XYZ_{background}$$
$$\alpha, \beta \geq 0$$

Equation 1. Hassani's αβ model for color appearance in AR. The XYZeffective value can be combined with typical color appearance models such as CIECAM02, CAM16 or others to predict the color appearance of the AR stimulus.

The present study was conducted using a color selection task between see-through AR stimuli and real physical Munsell color

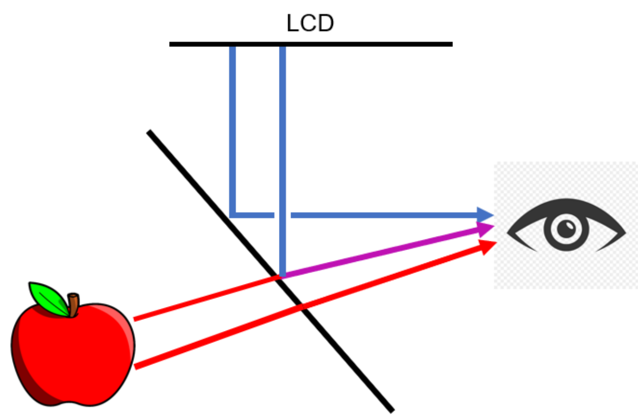

Figure 2. Description of proximal colorimetry. Blue lines represent light from the display, forming the display color. Red represent light from the background forming the background color. When the two images are combined on the beam splitter (shown in black, angled), they form the purple arrow, forming the proximal color. Calibrations were made by measuring the red and blue components in isolation.

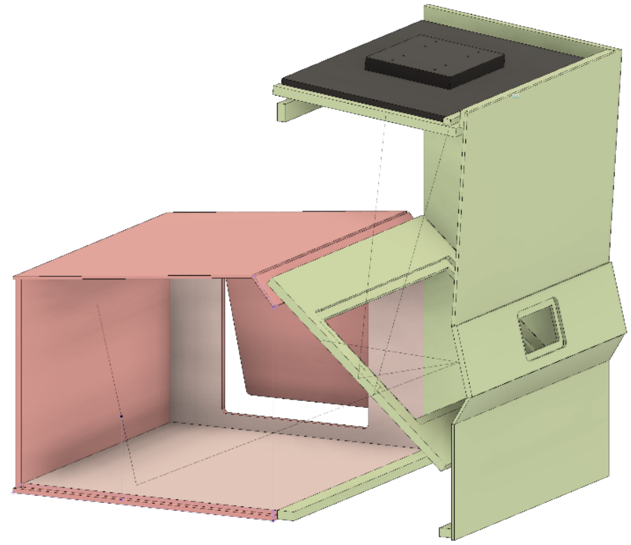

Figure 3. Cutaway view of AR Display. Viewed through the hole on the right, a 5-LED light booth optimized for D65 (pink) is visible through the AR display system employing a teleprompter mirror and 27" LCD display (green).

samples. Observers were asked to look at a color patch displayed in AR and select "the most similar color" from a one-hue Munsell page while focusing on lightness and chroma, which are the forefront attributes studied. Four Munsell pages were used, corresponding to unique hues. The "most similar color" Munsell sample selections made by observers were used to calculate lightness and chroma values in CAM16, and scissioning effects were analyzed.

The αβ as used in previous research did not satisfactorily predict the results of this experiment. However, after further analysis of the model, it is shown that it may still be promising to pursue and does provide insights into the problem of predicting color appearance in AR.

This study was approved by the Institutional Review Board for human subject's research at Rochester Institute of Technology, Rochester, NY, USA.

AR Display Design

To support this experiment, a new teleprompter-mirror style AR display was designed and built, illustrated in Figure 3. It employs a 27" LCD display and a dedicated light booth with a 5-primary LED system optimized for color fidelity at 6500K (IES TM30 Rf = 91.3 [7]). The light booth has a side access door which allows for objects to be changed in and out without disturbing the display alignment. The beam-splitter mirror employed is a thin 2mm sheet of glass that provides a proximal viewing condition of 60/40 background / AR display. The interior of the display frame is covered in black felt to reduce reflections and flare, whereas the interior of the light booth is painted grey, approximately L* = 50.

The display provides accurate color reproduction in the sRGB color space with a white point of D65 @ 175 cd/m^2, measured after reflection from the beam splitter. Against 222 tested colors, the mean CIE ΔE2000 score was 0.64 ΔE and the 90th percentile was 1.05 ΔE. The light booth provided an effective illumination of approximately 555lux with a CCT of 6514K + 0.004 Duv, measured after transmitting through the beam splitter. Lighting was provided by diffuse lighting strips from the top and sides of the light booth. The side lighting helps reduce the illumination gradient to approximately 50lux across the Munsell page.

The display was characterized while the light booth was powered off, and a black felt sample was placed inside, making the display and lighting system calibrations independent. An explanation of the measurement geometries and proximal colorimetry can be found in Figure 2.

Experiment Design

A color selection experiment was designed where emissive transparent stimuli (AR stimuli) were presented on the AR display. Observers were tasked with selecting "the most similar color" from an available Munsell page displayed in the light booth. Stimuli were presented on the left side of the light booth, against either a grey or black construction paper background. An AR user interface was developed to overlay the Munsell page and users would use the arrow keys and space bar of a wireless keyboard to indicate their selection. Care was taken to ensure the AR user interface had minimal effect on the Munsell page, active elements of the GUI were only displayed for 100ms and persistent elements were displays away from any color samples.

The user interface provided arrows near the edge of the page as well as a rectangle selector around selected patch. The rectangle persisted for around 1/10s to prevent or reduce the effect of a surround on the appearance of the Munsell samples. Observers used the arrow keys to move the selector to the patch that they felt had the most similar color as the AR stimuli. Once they were satisfied with their selection, it was confirmed by the space bar or enter key (Figure 4). Observers did not report any excessive difficulty using the experiment interface after some training.

Selections using the Munsell system restrict observers to only selecting value and chroma attributes directly, meaning hue cannot be independently varied during the experiment. A single Munsell page was placed into the light booth on an easel. Once the observer was finished making comparisons with that page, another hue page would be placed inside. The Munsell system was selected because of the availably of physical samples, and their roughly perceptually uniform spacing in lightness and chroma. Some observers were not experienced at color selection and received additional instruction about the definitions of value and chroma.

Red (h = 20.9, 5R)

J	40	20	40	60	20	40	60	20	40
C	90	20	20	20	40	40	40	60	60

Yellow (h = 95, 5Y)

J	80	20	40	60	40	60	80
C	60	20	20	20	40	40	40

Green (h = 154.7, 2.5G)

J	80	20	40	60	80	40	60	80	60
C	70	20	20	20	20	40	40	40	60

Blue (h = 237, 10B)

J	60	20	40	60	80	40	60
C	50	20	20	20	20	40	40

Table 1: Stimuli list for the four hues. The hue angle, J and C values are to generate the AR layer stimuli using CAM16. White point for stimuli generation is D65@175 cd/m^2

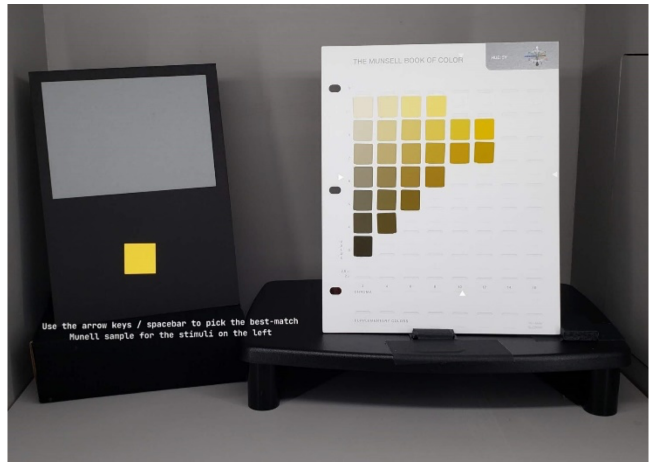

Figure 4. Example of the observing conditions and user interface. Left: A placard has black and grey construction paper to form the two backgrounds used in this experiment. Left below: A yellow square is displayed on the black background. Beneath this square are the instructions to pick the best matching color from the Munsell page on the right. Right: the 5Y page from the Munsell Book of Colors. Users used a keyboard to move selectors, displayed in AR, to their perception of the most similar colors. Although it is difficult to see in this image, the markers were plainly visible to observers and are currently showing the selection 5Y 8/10 (value/chroma).

AR stimuli were generated using CAM16 based on a white point of D65 @ 175 cd/m^2 [8]. AR stimuli corresponding to each Munsell page were matched in hue by a pilot experiment with 8 expert observers. The set of stimuli for each hue always included the highest chroma patch that the display could produce, as well as all samples at multiples of 20 J (lightness) and C (chroma) that were in gamut for the display. Stimuli also included “no surround” or a white surround.

Each stimulus was displayed as a small square, approximately 1.5 degrees of visual angle. When a surround was present, this was a further 1 degree of visual angle on each side. Exact sizing depended on observer position. Observers viewed the display through a small 1in. by 5in. opening which restricted large differences in viewing position. Mainly, observers could be a few centimeters closer or further from the display. As seen in Figure 4, the stimuli were only slightly larger than the Munsell samples.

In total the stimuli list included all the J and C values from Table 1 in 4 groups: grey construction paper background, black construction paper background, no AR surround, or a white AR surround. The black and grey backgrounds had a luminance of 8 cd/m^2 (J = 18) and 55 cd/m^2 respectively. With the above AR stimuli list, two backgrounds, and two surround conditions the total number of stimuli was 128. Each observer repeated a different hue based on their assigned permutation.

Observers would view all the stimuli for a particular hue in a random order before switching to their next hue. One observer might observe red, blue, yellow, green, red. Each observer was assigned a unique permutation of red, yellow, green, and blue, to which a repetition of the first hue was added. Some care was taken to avoid having any observers repeat the same permutation as another, however due to an error in the data collection protocol two permutation repetitions were made. The remaining observers each had a unique permutation.

For each observation, the Munsell value and Munsell chroma were recorded. Observers were allowed to choose values or chromas that were beyond the gamut of physical samples on the page, and they were instructed to use the samples as their reference scale for extrapolating value and chroma. For example, they might see a yellow AR patch that looks like it has a Munsell chroma of 12 but is brighter than the available patches on the Munsell page, which had a maximum value of 8 for that column. In that case they would indicate that the observed AR stimulus was value 9 and chroma 12.

Munsell matches were then converted to CAM16 lightness and chroma values for comparison with the AR stimuli by fitting a two-dimensional second order polynomial function to measurements of the available Munsell samples (Equation 2). A polynomial surface fit was selected because of the requirements to provide extrapolation for selections that were outside of the available gamut in the Munsell page. Each hue had a unique set of functions fit using MATLAB v. 2021a “poly22” fittype based on measurements of the available patches for each Munsell page. Each function had an R^2 value of greater than 0.99.

$$(J, C) = (\, f(v, c), g(v, c)\,)$$

Equation 2. Model for converting Munsell value and chroma (v, and c respectively) to CAM16 J and C. MATLAB was used to fit functions f and g separately for each equation and hue page. The fit data came from measurements of the in-gamut Munsell samples.

Finally, the J and C values for each stimulus and observation were averaged across observations to give the average J and C values for each AR stimulus. These average values are expected to correspond to the appearance of AR stimuli and notably does NOT correspond to the appearance computed from proximal tristimulus values due to the color scissioning effect as explained in the results section.

Results

18 color normal observers participated in a color selection experiment where stimuli were presented on the AR display and they were tasked with selecting “the most similar color” from an

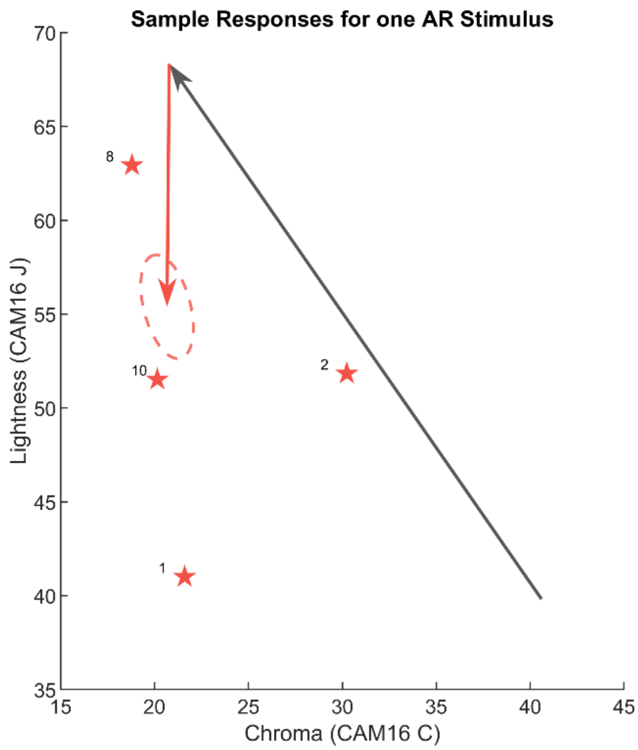

Figure 5. The first level data product from the experiment. The observer response distribution is shown with red markers and a number indicating the number of times observers selected that sample. Observer selections were constrained to the Munsell sample grid from the Munsell book of Colors and are therefore highly quantized. The grey arrow shows the difference between display colorimetry and proximal colorimetry, while the red arrow indicates the difference between proximal colorimetry and the average observer perception. The ellipse represents the 95% confidence interval on the estimate of the mean.

available Munsell page displayed in the light booth. There was a mix of expert / non-expert observers as well as a wide range of ages (avg. 36y, σ = 15y).

An example of the first level results from the analysis are shown in Figure 5, illustrating the responses for a red stimulus ($C \cong 40, J \cong 40, h \cong 20.9$) on a grey background card and a white surround in AR. The grey arrows represent the difference between the proximal color, caused by the addition of the background color, and the display color. Referring again to Figure 2, this arrow represents the difference between the display color (shown in blue Fig. 2) and the proximal color (shown in purple) which is formed by the addition of the background light (shown in red). The average observer response is plotted as an arrow to projecting from the proximal color and an ellipse is plotted around the arrowhead to show the 95% confidence interval for the estimate of the mean.

Each stimulus had a distribution of responses represented by the star symbols, each with a number indicating the number of responses that selected that Munsell value/chroma combination. Because the responses are restricted to the quantization of the Munsell Book of Colors the resulting responses shown in CAM16 C and J are also highly quantized. The fact that responses are not strictly on a rectangular grid in (C, J) coordinates is likely to be a result of differences in the appearance methods used by Munsell, now over 100 years old, and modern color appearance models.

Figure 6 shows the average responses for all stimuli grouped by hue and observation condition on lightness / chroma plots. Alternating rows indicate the background condition, either grey or black construction paper. Alternating columns indicate the surround condition, either no surround, or a bright AR surround around the AR stimulus. Again, the grey arrows indicate the difference between display and proximal color and the colored arrows indicate the average observer response. Ellipses indicate the 95% confidence intervals on the estimates of the mean responses.

For orientation, consider two hypothetical cases: a zero-length colored arrow would mean observers made a colorimetric match to the proximal stimulus, implying no scissioning and no surround effects; a colored arrow that returns to the root of the gray arrow would mean observers made a match to the display stimulus, ignoring the background and thus exhibiting perfect scissioning. A range of scissioning levels (including the direction and magnitude of scissioning) is apparent from this data and is at least depends on the background and surround conditions as discussed in the following sections.

Noteworthy observations include that each hue behaves rather similarly to each other. Most of the confidence intervals around the mean responses exclude the proximal color, indicating a significant color scissioning effect. The plots for the "white surround" condition typically have less difference between the chroma of the mean color and the chroma of the proximal color, however there is still a strong lightness effect against the brighter grey background. Finally, in the case where the background is inducing only a small change in the proximal color (mainly the black background) many of the data points still show a significant color scissioning effect.

Background effect

As expected, the stimuli observed against the grey background showed significant compression in their overall gamut shape compared to those shown on the black background. For the grey background condition there was compression in the lightness perception in the negative direction for the brightest stimuli and in the positive direction for the darkest stimuli. Whereas the gamut area used by observer responses for the black background condition is larger and mainly shows a lightness effect in the positive direction. Additionally, the chroma compression of responses was larger for the grey background than for the black background.

For the black background, it is noteworthy that the white surround helped preserve the proximal appearance the best out of all the conditions. However, it is important to recognize that the proximal color still lies outside of the 95% confidence interval for most of the stimuli, indicating a significant color scissioning effect for this experiment.

Surround Effect

The white surround vs no surround condition has a typical characteristic of shifting all the perceived colors in the negative direction along the lightness axis. This could be explained as a simultaneous contrast effect. The white surround also typically reduced the chroma differences between the surround / no surround conditions. For the no surround condition, we can see some chroma compression in most plots, but for the white surround condition the mean response has a chroma much closer to the chroma of the proximal color.

Modeling

Results from this experiment were analyzed by fitting the αβ model used in previous work and described in *Equation 1* [4]–[6]. Previously, researchers have utilized numerical or procedural optimization methods to calculate α and β for subgroups of stimuli.

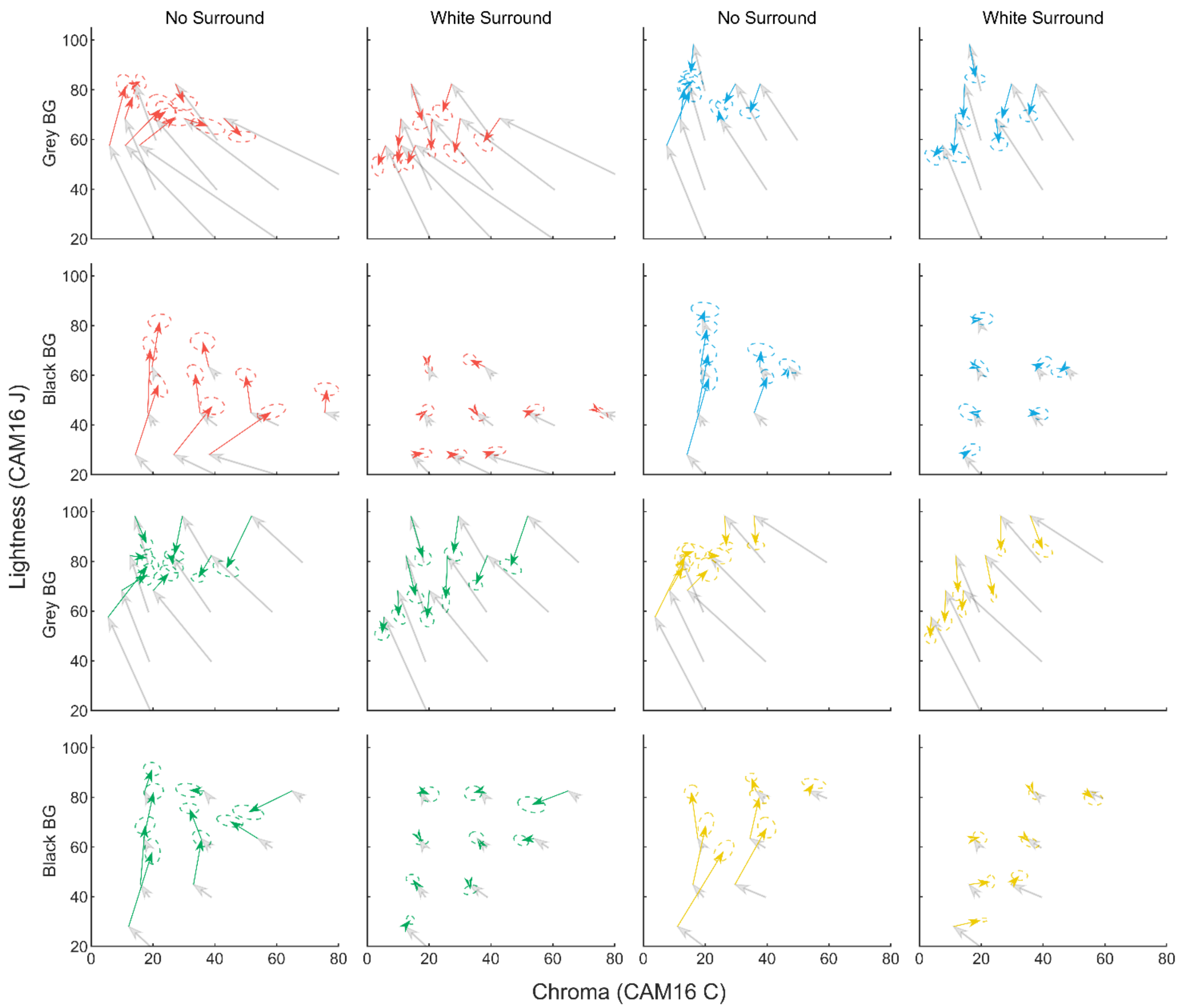

Figure 6. Lightness and Chroma plots for each stimulus grouped by hue and by viewing condition. Each color-coded arrow represents the difference from the proximal colorimetry to the mean observer response. The grey arrows indicate the difference from the display colorimetry to the proximal colorimetry. The ellipse around each arrowhead indicates the 95% confidence interval on the estimate of the mean response.

In this analysis the same attempt was made on various types of subgroups including white surround vs. no surround, black background vs grey background, and their combination. 8 subgroups in total: *surround* / *no surround*, *black* / *grey*, and each hue individually.

This did not produce satisfactory results. In most subgroups the mean color difference from the mean (MCDM), based on CAM16 distance in the J, C plane was between 10 and 15 when $\alpha = \beta = 1$. After optimizing coefficients for the 8 different sub groupings for each hue more than 50% of the fits had less than a 20% improvement on the MCDM metric.

The subgrouping fit of the αβ did work well for the white surround, grey background condition. For this viewing condition fitting a single αβ pair for the entire subgroup and each hue resulted in improvements of MCDM by 80%, 80%, 79% and 78% for red, yellow, green, and blue, respectively. For all 4 hues $\alpha < 1, \beta < 1$ indicating perceptual discounting of the proximal stimulus.

Next, an αβ pair was fit to each average response for each stimulus and viewing condition combination. This showed exceptional prediction power of the αβ model, with an average MCDM score of less than 1e-4 across all stimuli, a value so low that it is considered a perfect fit given the confidence intervals around the average observer responses. The resulting αβ values are shown in Figure 7.

The αβ values shown in Figure 7 are markedly different from values obtained in the previous studies. One, their range is much higher, with the largest values reaching 5 or 6 units. Though this range is much higher, it is not implausible and would explain the sometimes very deviant perceptions from Figure 6. The most extreme values of αβ are seen for the darkest stimuli, which also exhibit the largest perception differences in Figure 6.

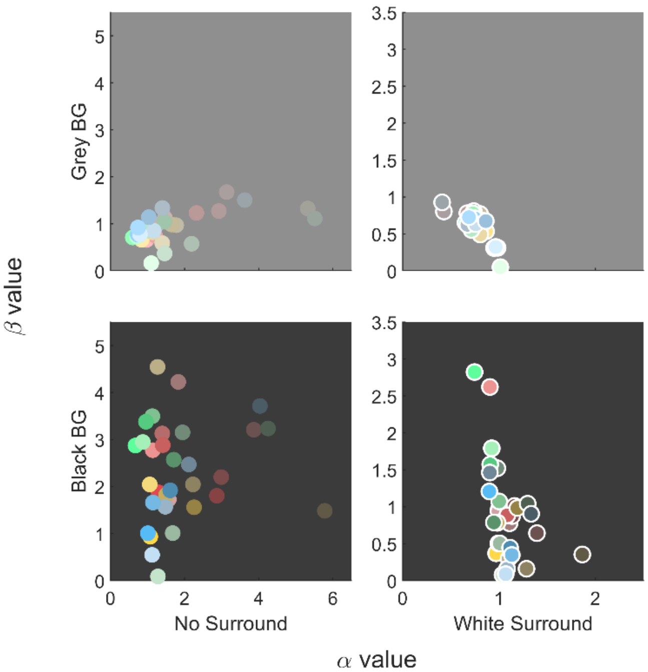

Figure 7. Each stimulus plotted in the αβ parameter space. The background color is representative of the background object placed in the light booth. The color of each circle, including the white border, is approximately representative of the proximal color shown with the AR display.

Discussion & Future Research

This experiment further examined color perception in AR viewing conditions. With a new display apparatus, a color selection experiment was performed where observers matched colors between real physical color samples drawn from the Munsell Book of Colors. The previously researched αβ model was tested against this new data and found unsatisfactory with similar application methods.

Firstly, it should be acknowledged that while this is the first direct color selection experiment comparing real reality physical stimuli to augmented reality stimuli, the level of quantization makes the analysis much more difficult and naturally limits the level of precision possible for a reasonable number of observers.

Additionally, observers found this task much more difficult than the researchers anticipated. Expert observers seemed to be able to make comparisons, but non expert observers gave much more feedback about the task being difficult. The average observation time was around 15 seconds. We take this observation time as an indication of how difficult a task is for observers to respond to and found 15 seconds to be acceptable during early data collection. Future experiments should carefully examine qualitative observer feedback like this and make a concentrated effort to make the observations as simple and easy as possible. It may be the case that more training time and familiarity is required for observers, which could make data collection more expensive. Additionally, user interface decisions could be made to improve data collection.

Previous work seemed to indicate that αβ could be fit by subgroups based around figural conditions such as outline overlay shape [4] or by viewing conditions [5]. However, this type of fitting was not satisfactory for this data, possibly because the gamut of colors observed was quite large and included colors which had a much lower display luminance than previously studied. This would indicate that a predictor for αβ likely includes parameters not only for the viewing conditions and figural form of the stimulus but also the display luminance or display luminance / background luminance ratio. Figure 7 does not suggest that there is any hue dependency.

To improve on the quantization problem, an adjustment experiment would be preferable, but small adjustable reflectance stimuli are exceedingly difficult to produce. Instead adjusting an AR stimulus to match a physical sample seems more plausible. The downside to adjustment is the increased time per observation.

It is also interesting to spend some time thinking about the possible solutions to a color matching problem that the αβ model provides. It restricts solutions to the perception prediction to plane formed by two vectors in the XYZ color space. When the background color is given by a vector that is a scalar of the white point (that is to say, the background is neutral grey) this plane ends up being a plane of a single dominant wavelength determined by the stimulus color.

For grey background case, lines of constant dominant wavelength are well studied, for example we know they exhibit the Abney effect. There may be untapped previous research that can help analyze this viewing condition. When the background color is *not* grey the solution plane for the αβ model now crosses arbitrarily through XYZ space and strong hue effects should be found. This should be verified with an adjustment experiment including chromatic backgrounds.

Overall, this experiment calls into question the predictor of the αβ model, indicating that it does not simply rely on figurative (shape, lines, size, texture, etc…) cues but may also rely on the background / display luminance ratios or other factors and the range of αβ values may be larger than previous research suggested.

This material is based upon work supported by the National Science Foundation (USA) under Grant No. 1942755. The authors declare no conflicts of interest.

References

[1] F. Metelli, "The Perception of Transparency," *Sci. Am.*, vol. 230, no. 4, pp. 90–99, 1974.

[2] J. L. Gabbard, J. Zedlitz, J. E. Swan, and W. W. Winchester, "More than meets the eye: An engineering study to empirically examine the blending of real and virtual color spaces," *Proc. - IEEE Virtual Real.*, pp. 79–86, 2010, doi: 10.1109/VR.2010.5444808.

[3] N. Hassani and M. J. Murdoch, "Investigating color appearance in optical see-through augmented reality," *Color Res. Appl.*, vol. 44, no. 4, pp. 492–507, 2019, doi: 10.1002/col.22380.

[4] M. Murdoch, "Brightness matching in optical see-through augmented reality," *J. Opt. Soc. Am. A*, vol. 37, no. 11, 2020, doi: 10.1364/josaa.398931.

[5] L. Zhang, M. J. Murdoch, and R. Bachy, "Color appearance shift in augmented reality metameric matching," *J. Opt. Soc. Am. A*, vol. 38, no. 5, p. 701, 2021, doi: 10.1364/josaa.420395.

[6] N. Hassani, "Color appearance modeling in augmented reality," Rochester Institute of Technology, 2019.

[7] IES, *Technical Memorandum 30 : IES Method for Evaluating Light Source Color Rendition*. 2020.

[8] C. Li *et al.*, "Comprehensive color solutions: CAM16, CAT16, and CAM16-UCS," *Color Res. Appl.*, vol. 42, no. 6, pp. 703–718, 2017, doi: 10.1002/col.22131.