

Colour in AR & VR

Michael J. Murdoch

Abstract

Head-mounted virtual reality (VR) and augmented reality (AR) systems deliver colour imagery directly to a user's eyes, presenting position-aware, real-time computer graphics to create the illusion of interacting with a virtual world. In some respects, colour in AR and VR can be modelled and controlled much like colour in other display technologies. However, it is complicated by the optics required for near-eye display, and in the case of AR, by the merging of real-world and virtual visual stimuli. Methods have been developed to provide predictable colour in VR, and ongoing research has exposed details of the visual perception of real and virtual in AR. Yet, more work is required to make colour appearance predictable and AR and VR display systems more robust.

Introduction

Virtual reality (VR) and augmented reality (AR) systems offer tantalizing potential to extend our visual experience with immersive imagery and rendered graphics. Sci-fi-like visions of immersing oneself in an alternate world, with other people and/or with AI agents, as well as concepts of seamlessly blending virtual objects and remote people with the real world, have captivated scientists, science fiction authors, and content makers for decades. Presently, the pace of technological improvement in AR and VR display and system technologies is as fast as ever, and AR and VR are being quickly adopted in applications including entertainment, education, medicine, and industry. As in any visual technology, colour is a crucial characteristic, and many aspects of AR and VR display systems impose challenges to creating robust colour.

VR, AR, and MR (mixed reality) are all variations of extended reality (XR), meaning that they allow a user to experience something beyond real reality. VR refers to a display environment in which all content is virtual, which typically means that the real world is blocked out, as in a typical VR head-mounted display (HMD). The terms AR and MR are not universally defined or distinguished, but for the sake of this chapter, they both refer to a display environment where the real world remains visible, and the display system can render virtual content in front of, or mixed in with, the real environment. The chapter will proceed only with the term AR, for efficiency. AR can be further subdivided into optical see through (OST) AR and video see through (VST) AR. OST-AR uses an optical combiner such as a beamsplitter to create a transparent display capable of overlaying the virtual content on the real world, while VST-AR uses a non-transparent display to present the viewer with a live video of the real world into which virtual content is composited using image processing. Some HMD hardware may be designed to provide both VR and VST-AR.

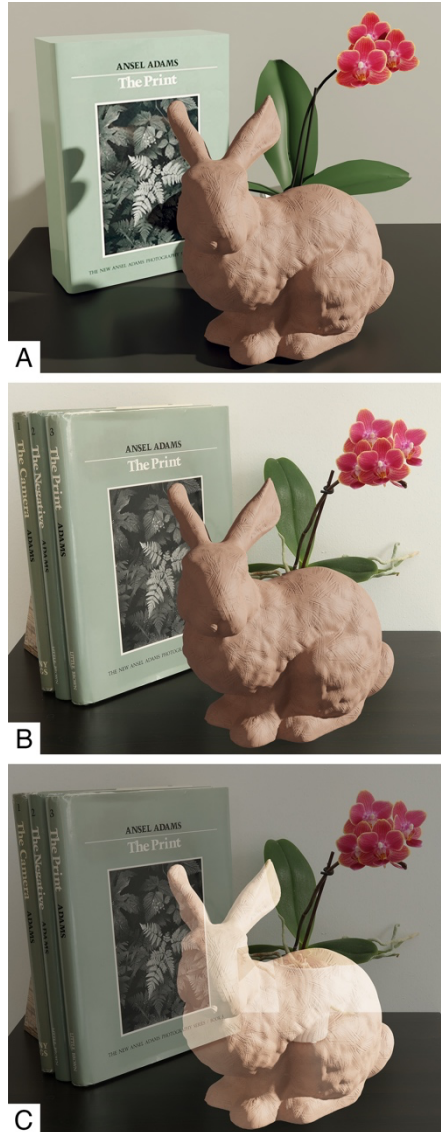


Figure 1: AR & VR Examples: image A simulates a VR view in which all scene elements are virtual. Image B simulates VST-AR, in which the virtual bunny is alpha-matted into an image of the scene with real books and plant. Image C simulates OST-AR, in which the virtual bunny is presented on a transparent display, through which the real-world background books and plant remain visible. [Source: Murdoch]

The distinction between OST and VST is important because OST-AR displays remain transparent, so even where virtual objects are displayed, the real world bleeds through from behind. In VST-AR, true occlusion is provided in image processing, but only in trade for moderating the real world through a live video camera view, which may have issues like temporal lag and dynamic range limitations. Examples of these are shown in Figure 1. Image A depicts a VR display, in which all scene elements are virtual, rendered with computer graphics. Image C simulates an OST-AR environment, in which the virtual bunny, rendered with graphics, is displayed on a transparent display, allowing the real-world background books, plant, and table to remain visible, though visibly attenuated by the low transparency of the display. The AR bunny appears more transparent in regions where it is dark and the background is bright. Image B simulates VST-AR, which composites the virtual bunny onto a live image of the real-world scene,

rendering the bunny opaque through the use of alpha matting. The VST-AR image in the centre may look the best in this comparison, but in real systems the live video view of the real world often comes with artifacts including lag and perspective distortion which may be worse than the challenges of transparency in OST-AR. Whether the future trends to VST- or OST-AR remains to be seen.

Most AR and VR hardware takes the form of an HMD. VR HMDs are typically boxy devices that seal around the user's eyes to block out ambient light that would otherwise cause flare on their optics and display(s). Orientation sensors in the HMD allow the system to respond to the user's head motions to make the virtual world appear fixed, so the user can look and move around naturally. Some VR systems include forward-looking cameras which make them capable of VST-AR, though this is still relatively uncommon. VST-AR is more commonly implemented with a mobile device such as a phone or tablet, using its forward-looking camera and orientation sensors. In this case, mobile VST-AR is much like holding a "window" through which the real and virtual components are visible, aligned with the real world so that objects appear fixed. OST-AR HMDs vary significantly, from minimal glasses-like systems with a small, even monocular, display, to systems as large as VR HMDs with binocular transparent displays and transparent cases through which the real world remains visible. The latter OST-AR HMDs typically include the orientation sensors necessary to fix the virtual objects in real-world positions, while the simpler systems may just use the transparent display as an information overlay, sometimes called "data glasses."

Because of the range of variants and the fast pace of development in the AR and VR industry, generalizing colour in AR and VR display systems is difficult. However, the basic components of such a system include a display, similar to either a mobile device or a projector, optics to bring it to the user's eyes, and a graphics engine to render virtual content [1]. These are the most important components for colour performance, but are not the whole story. In HMDs, displays and optics are integrated in a wearable device along with cameras, sensors, processors, and batteries, making for a complex set of human factors trade-offs to be sure that weight, heat generation, cost, and power usage are minimized while visual experience is maximized [2][3][4]. As progress on all of these system components continues, colour quality emerges as a driver of image quality and user experience.

This chapter will focus on colour performance in AR and VR systems, first addressing colour synthesis in AR and VR display systems, then colour appearance characteristics as influenced by the systems' viewing environments. Finally, the graphics and imaging pipeline will be described, followed by a forward-looking assessment of necessary progress. Distinctions between VR, OST-AR, and VST-AR will be made where appropriate.

Colour Synthesis in AR & VR Displays

Despite the wide range of possible AR and VR system architectures, at the core of any system is a display: typically, either a flat-panel display much like that used in a mobile phone, or a micro-display in the form of a small projector. The behaviour and colour performance of this display is not unique to AR and VR, which means that the same kinds of measurement, modelling, and characterization methods used for other displays can be applied with minor modifications for

the form factor and the optics that are between the display and the user. Characterizing and modelling a display's colour output is essential to using the display to create colour predictably.

The best general model of colour synthesis in displays in general is the gamma-offset-gain (GOG) model (also discussed in other chapters in this book, including 8 and 22), initially developed for cathode-ray tube (CRT) displays and successfully applied to many types of displays, often generalizing the “gamma” with an empirical one-dimensional look-up table (LUT) to describe the system nonlinearity [5][6]. A GOG display model makes several assumptions that should be confirmed, and which may not hold for some OLED displays and some types of HDR LCDs: constant primary chromaticity (*aka* scalability, meaning that the chromaticity of each primary channel does not depend on luminous intensity), channel independence (meaning that the output intensity of each primary channel is not affected by the intensity of the others), and the separability of system non-linearity (or gamma, meaning that additive colour can be modelled with linear algebra after separately handling nonlinear input signals).

A model of colour synthesis can be thought of as describing a pixel, or set of single-color sub-pixels, in a display, in a given condition of operation at a given moment in time. However, there may be spatial and temporal variations, intentional or not, in the performance of a display that must be added to a display colour model in order to adequately predict its performance. AR and VR display systems may be especially prone to variations such as spatial nonuniformity.

GOG Display Model

A GOG forward model predicts the colorimetric output (CIE XYZ) of a display with a given input signal, typically a vector of red, green, and blue values, RGB . The system nonlinearity is handled first on a channel-by-channel basis, resulting in linear intensity values $R'G'B'$, where the functions f_i could take the form of a power-law gamma function, arbitrary piecewise function, or a 1D LUT based on measurements or idealized system behaviour:

$$\begin{aligned} R' &= f_R(R) \\ G' &= f_G(G) \\ B' &= f_B(B) \end{aligned} \tag{1}$$

The $R'G'B'$ linear intensity values are then multiplied by the display's primary matrix, which contains the CIE XYZ values of the displays red, green, and blue primaries. This means that the resulting XYZ colorimetric output is a linear combination of the three primaries, each scaled by the corresponding linear intensity value. An additive black offset XYZ_k is used to account for any constant “flare” light produced by display system even when set to zero or black:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} X_R & X_G & X_B \\ Y_R & Y_G & Y_B \\ Z_R & Z_G & Z_B \end{bmatrix} \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} + \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_k \tag{2}$$

A reverse model can be made by inverting the matrix and gamma functions, allowing the computation of the required RGB values for a desired XYZ colorimetric output. In case of more than three display primaries, the forward model may be easily extended, adding nonlinear functions and matrix columns for each additional channel. However, the resulting non-square

matrix cannot be inverted, instead requiring further constraints, which may be heuristics for power savings or artifact reduction [7][8][9].

It is important to note that a colorimetric model is a shortcut for a spectral model, which may be valuable to predict radiometric output or useful for iterative optimization to reach primary intensity solutions with goals such as spectral matching or the reduction of inter-observer differences. Methods for multiprimary intensity solutions, which are essentially the same for both display and lighting systems, may be constrained to also be colorimetric matches or handle multiple objectives [10][11]. Following the GOG structure, the linear portion of a spectral forward model may be written similarly to Eq. (2), but with the matrix comprised of m -band spectral vectors $S(\lambda)_i$ in place of colorimetric XYZ_i for each of the n primary channels, the RGB vector replaced with a vector of n intensity values p_i , and a spectral vector for the black offset. Equivalently, it may be expressed as a weighted sum of n primary channels with intensity values p_i as the weights:

$$S(\lambda) = \sum_{i=1}^n p_i S(\lambda)_i + S(\lambda)_k \quad (3)$$

Idealized Display Models

Standard colour spaces, or colour encodings, such as sRGB and Adobe RGB, are essentially models of an idealized display, which take the same form as a GOG display model [12]. For example, the sRGB standard defines a relationship between input RGB vectors and normalized display colorimetry XYZ using a gamma-like nonlinearity and a primary matrix [13]. The nonlinearity of sRGB is a piecewise function with a linear portion near zero and a power-law for higher values. The result is very similar to a power law with exponent 2.2, which is commonly referred to as “gamma 2.2,” but the linear portion is included to avoid approaching zero slope. sRGB defines the CIE 1931 chromaticity coordinates of the red, green, and blue primaries and the display white point (CIE D65), resulting in the following linear matrix equation:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9505 \end{bmatrix} \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} \quad (4)$$

Wherein the vector $R'G'B'$ indicates the linear RGB intensity values, in the range $[0, 1]$, obtained through the nonlinear gamma transformation of the nonlinearly-encoded RGB values typically used in image file formats (which in sRGB are defined for 8-bit unsigned integers in the range $[0, 255]$). The sRGB-predicted XYZ output vector is normalized so Y represents CIE 1931 luminance factor in the range $[0, 1]$, meaning the XYZ of the D65 white point is $[0.9505 \ 1.000 \ 1.0888]$.

An idealized display model such as sRGB provides an invertible transformation between RGB and XYZ , giving a colorimetric definition to RGB vectors. This allows both the prediction of a display’s output for a given RGB input and a way to compute the required RGB for a desired XYZ output. Note that sRGB and similar colour encodings provide only colorimetric definitions, not spectral; therefore, there is no spectral version (like Eq. (3)) of Eq. (4). This spectral ambiguity gives implementation freedom to a device maker, but it means that there is no standard spectral power distribution associated with the colour encoding.

Spatial and Temporal Independence

A GOG model describes a single colour synthesis unit, perhaps a single pixel, in a single operating condition. If all of the pixels of a display behave in the same way, then the model may be applied to all pixels. However, if there is any spatial dependence to the spectral or colorimetric output of the display, additional measures may be required to adapt the GOG model output for different spatial regions. For example, if an LCD backlight is temporally static but not spatially uniform, a simple luminance map may be multiplied by the GOG model colorimetric output to describe the display's behaviour (or used to compensate an input image to make it appear uniform).

Similarly, the model remains valid only while the display is temporally stable, while its operating conditions remain fixed. If the display changes over time, such as with automatic brightness or white point changes, then the model must account for this as well [14]. Simple dimming may be accounted for with a luminance scale factor, but other temporal changes may require deeper understanding of the system behaviour. More complicated systems may have temporally-varying spatial differences, including 2D-dimming HDR LCDs and OLED displays with current limits; these are more difficult to model and compensate, and are beyond the scope of this chapter [14][15].

HMD Optics

While the above description of spatial independence is general for any display, HMDs incorporate optical elements that may cause more extreme spatial variations. A typical VR system uses one or two flat-panel displays, either LCD or OLED, with simple lenses to make the displayed image appear further from the user's eyes, for example at a focal distance of 1-2m. These lenses are necessarily thin, lightweight, and inexpensive, often plastic Fresnel lenses, with spatial and chromatic distortions. AR displays incorporate various types of optical combiners which must be transparent to allow the real environment to remain visible while bringing the displayed virtual content to the user's eyes [15]. As in all HMDs, these must be lightweight and inexpensive, and typically result in spatial and chromatic distortions.

Optical distortions in AR and VR displays have been characterized in terms of visual performance [16][17], and display driver software commonly includes some compensation for geometric distortion and chromatic aberrations. In measuring and characterizing AR and VR displays, special attention should be paid to spatial nonuniformity, with measurements made in the same operating condition as the intended use. For example, if the image rendering pipeline includes compensation for chromatic aberration, the measurements should be made with the exact same compensation in operation.

Measuring AR & VR Displays

AR & VR systems typically involve near-eye displays (NEDs), which require careful measurement because of their size and form factor. A good practice when measuring displays is to place the colour measurement instrument, such as a spectroradiometer, where the user's eyes will be. For handheld, desktop, and larger displays, a normal viewing distance is 40 cm or more, and the angular sensitivity of the display itself is relatively small, so the placement of the instrument

does not require high precision. A tele-spectroradiometer at approximately the user's head position can simply be focused on the surface of the display screen.

For NEDs such as those used in HMDs, system optics present a virtual image that must be focused by a user's own eye, only possible when the eye is positioned in a zone known as the eye box [1][15]. For a measurement to avoid errors, a colour instrument must have an aperture or entrance pupil similar in size to that of the eye (1-5mm) and be positioned accurately in the eye box [18]. Recent developments in NED-specific colour measurement emphasize the need to locate and measure the size of the eye box by scanning the relevant region using a small-aperture luminance probe. A conventional tele-spectroradiometer may be used with a 1-3mm diameter aperture aligned and attached in front of its lens, and the luminance and spectral radiance of the display may be accurately measured while the instrument aperture is in the eye box [19]. In some situations, the image projected by the NED can be seen on the instrument aperture plate, which may aid in centring the instrument in the eye box [18].

Most importantly, the light projected from a NED must completely and uniformly fill the aperture or entrance pupil of the colour measurement instrument. Either a reduced-aperture conventional instrument or a NED-specific specialized measurement instrument may be used to accurately measure spectral radiance or colorimetry. Note that placing an aperture in front of the instrument reduces the light intensity, requiring longer integration times [18].

Measurements of different regions of the displayed image may be made by rotating the instrument about its aperture, if care is taken to keep the aperture in the eye box. Spatial characteristics including resolution and optical distortion may be measured using an imaging colorimeter with a NED-specific lens, which includes an aperture of about 3mm in diameter at the front of the lens.

With any display, colour characterization based on measurements of a single spatial location at a moment in time may not be sufficient if there are spatial or temporal variations, whether intended or not. Because of the display optics and the interactive nature of AR and VR displays, they may be more variable in uniformity and with use than other displays. Spatial uniformity should be verified at the very least, and may require spatial masks (for example, implemented as a uniformity correction image multiplied by the model output) or region-specific characterization models. An example of the latter is provided later in this chapter.

The result of a carefully-made colour measurement of an AR or VR system is a physical measurement of the light reaching a user's eyes, which is also known as the proximal stimulus. Measurements characterize the colour performance of an AR/VR system, and are essential to build or verify a model, as shown in the following example. Later, the complications of colour appearance will be discussed.

Example: Measuring and Characterizing an AR Display

Measurements can be used to characterize a display's colour behaviour, allowing accurate prediction of its output. Note this not the same as calibrating a display, which would involve modifying its behaviour to match a standard; the goal is to model its behaviour for predictable colour. To illustrate the characterization of an AR display, consider the following example of a transparent display, measured in a dark room with black fabric behind it to prevent real-world

background distorting the measurements. First, a list of 8-bit *RGB* input vectors was created to test the GOG model assumptions, build a GOG characterization, and verify its performance. These input vectors comprise “ramps” of 52 steps from 0 to 255 in 5-value increments for each primary colour alone, a neutral ramp in which $R=G=B$ at each step, and a $6 \times 6 \times 6$ “grid” of points in *RGB* space. A sparse selection of these *RGB* values and their data are listed in Table 1.

Because this particular display was being prepared for psychophysical experiments that would be implemented using MATLAB, the characterization was also done using MATLAB: this avoids any differences in colour management that the computer operating system might employ between programs. The display and beamsplitter geometry were fixed, and the display was dimmed slightly using its on-screen controls to approximately match its luminance level to the background illumination that will be used. A tele-spectroradiometer was placed on a tripod with its lens at the viewer’s eye position. Note this was a large, desktop-sized AR system, which avoids some of the complications of NED measurement. A short MATLAB program was used to loop through the list of *RGB* colours, display each one as a full-screen figure, pause for a moment to ensure stabilization, then make the measurement using the tethered spectroradiometer. Spectral measurements were converted to CIE 1931 *XYZ*; a selection of these measurements is given in Table 1.

The first step involves subtracting the black, or flare, measurement from all other measurements, based on the assumption that flare light within the display provides a constant offset regardless of the input *RGB*. Display black (*RGB* [0, 0, 0]) was measured multiple times, so the *XYZ* of all black measurements were averaged to produce a low-noise estimate: [0.0425, 0.0435, 0.0968]. Next, the GOG model assumptions were tested [5][6].

To test scalability, which means the chromaticity of each primary is constant regardless of intensity, the CIE $u'v'$ chromaticity coordinates of each of the black-subtracted primary ramp steps were computed, and the average chromaticity of each primary, over all steps. As expected, the difference between the individual steps and the average was somewhat larger for darker steps; at lower light levels the display may be less stable, any problem with the assumption of flare offset would manifest itself, and measurements are noisier. Regardless, the average Euclidean distance from each of the 51 red ramp steps to the average chromaticity of the red was found to be $0.00028 \Delta u'v'$. For green, the value was $0.00024 \Delta u'v'$, and for blue $0.00085 \Delta u'v'$; these values are extremely low, implying satisfactory scalability, or constant chromaticity, of the display.

Channel independence was tested by comparing the black-subtracted *XYZ* of the neutral ramp with the sum of black-subtracted *XYZ* values of the separately-measured *RGB* ramps. At each step, the sum of primary ramps was consistently slightly lower than the measured neutral ramp, indicating a slight non-additivity. However, this was determined satisfactory because the average deviation over all ramp steps was 0.46% in luminance *Y*, with a maximum deviation of 0.89%.

The primary matrix was constructed using the black-subtracted *XYZ* values of the three primaries, which coupled with the flare offset results in the following equation for the linear portion of the GOG model:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 13.31 & 11.44 & 6.554 \\ 7.185 & 24.47 & 2.454 \\ 0.6741 & 5.049 & 45.91 \end{bmatrix} \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} + \begin{bmatrix} 0.0425 \\ 0.0435 \\ 0.0968 \end{bmatrix} \quad (5)$$

Three 1D look-up tables (LUTs), each describing the nonlinearity of one of the primary channels, were computed from the neutral ramp using the inverse of Eq. (5). It is possible to compute LUTs from the separately-measured primary ramps, and in a perfect system these would be equivalent; however, a slight advantage in model error for near-neutral colours is typical using the neutral-LUT method.

Just three components -- LUTs, matrix, and flare offset -- comprise the GOG forward model, allowing the prediction of XYZ output for arbitrary RGB input vectors. As in any model evaluation, the prediction should be evaluated with independent data, which is the purpose of the 6x6x6 grid of verification colours in RGB space. Because of the perceptual uniformity of RGB and XYZ colour spaces, it is valuable to transform the measured and modelled colorimetry to CIELAB, which is commonly used for quantifying colour differences and setting tolerances. In Table 1, both measured and modelled XYZ are provided, along with the perceptually-scaled colour difference between them – the error of the model – in CIEDE2000 (ΔE_{00}). The table includes only a few of the 216 verification colours, with errors ranging from 0.1 to 0.9 ΔE_{00} , but statistics from the whole set were computed. Overall, the mean error was found to be 0.30 ΔE_{00} and the maximum error 1.18 ΔE_{00} , both low values for display applications, indicating an excellent fit and providing confidence that the display will produce colour reliably when the model is used.

Table 1: Example Data: A selection of data from the characterization of an AR display are listed, with approximate color indicated at left. 8-bit RGB values were the input to the display; linear R'G'B' values were computed using the LUT; measured XYZ values were measured from the perspective of a viewer; modeled XYZ values were computed using the forward model; and error is the difference in ΔE_{00} between measured and modeled XYZ.

	8-bit RGB			Linear RGB			Measured XYZ			Modeled XYZ			Error ΔE_{00}
	R	G	B	R'	G'	B'	X	Y	Z	X	Y	Z	
	255	0	0	1.00	0.00	0.00	13.35	7.228	0.771	13.41	7.263	0.774	0.114
	0	255	0	0.00	1.00	0.00	11.48	24.52	5.146	11.54	24.64	5.171	0.113
	0	0	255	0.00	0.00	1.00	6.597	2.497	46.00	6.629	2.509	46.23	0.070
	255	255	255	1.00	1.00	1.00	31.46	34.28	51.93	31.49	34.32	51.98	0.053
	205	205	205	0.62	0.62	0.62	19.38	21.18	31.75	19.38	21.12	32.00	0.854
	155	155	155	0.33	0.33	0.33	10.50	11.49	17.25	10.48	11.42	17.31	0.851
	105	105	105	0.14	0.14	0.14	4.481	4.892	7.412	4.467	4.866	7.396	0.333
	55	55	55	0.03	0.03	0.03	1.122	1.220	1.860	1.128	1.227	1.888	0.203
	0	0	0	0.00	0.00	0.00	0.043	0.044	0.097	0.043	0.043	0.097	0.004
	191	31	31	0.53	0.01	0.01	7.158	4.074	0.984	7.243	4.119	1.007	0.230
	231	191	31	0.81	0.53	0.01	16.85	18.73	3.716	16.88	18.76	3.799	0.126
	31	151	71	0.01	0.31	0.06	4.155	7.905	4.354	4.158	7.923	4.439	0.220
	71	71	111	0.06	0.06	0.16	2.574	2.349	7.673	2.559	2.330	7.662	0.140
	111	231	111	0.16	0.81	0.16	12.44	21.34	11.45	12.41	21.33	11.50	0.113
	151	151	151	0.31	0.31	0.31	9.846	10.77	16.11	9.831	10.71	16.24	0.905
	191	71	191	0.53	0.06	0.53	11.11	6.599	24.52	11.18	6.586	24.92	0.320
	231	231	191	0.81	0.81	0.53	23.50	26.99	28.60	23.48	26.91	28.88	0.410
	31	191	231	0.01	0.53	0.81	11.44	14.98	39.16	11.50	14.99	39.85	0.601

Colour Appearance in AR & VR

Colour appearance is a perception that can depend on context, surroundings, illumination, scene understanding, and recent visual stimulation. The human visual system is extremely good at adapting to different lighting conditions and interpreting complex scenes, generally managing to disentangle what is object and what is lighting. Humans see reflective objects in a surface viewing mode, while emissive displays are perceived as a source of light in illuminant mode [20]. Modern colour appearance models (CAMs) such as CIECAM02 and CIECAM16 are designed to compute values that correlate with perceived lightness and brightness (J and Q), chroma, saturation, and colourfulness (C , s , and M), and hue (h), depending not only on the colour stimulus, but also on its surroundings and illumination [21][22][23]. These CAMs include accounting for adaptation, illumination level and colour, and the luminance of the region surrounding the stimulus.

Colour appearance models compute appearance correlates that may be used to predict correspondences between colours in different viewing environments. For example, consider a pair of corresponding colours comprising one reflective colour sample under D65 and another reflective colour sample under Illuminant A. These illuminated, reflective samples are presumably different in both spectral reflectance and in spectral radiance, and are unlikely to be a colorimetric match; however, they match in appearance when viewed separately. Many experiments have been performed to measure corresponding colours between different illuminant colours, between different illumination levels, with different surroundings, and between reflective and emissive displayed colours, and recent CAMs have been evaluated with respect to how well they predict measured appearance attributes [22][23][24][25]. Despite CIECAM02 and CIECAM16 being the most deeply studied and widely used CAMs, their prediction errors remain appreciable, with colour differences well above a just-noticeable difference (JND) and other measures of residual error far from zero; these discrepancies are presumably due in large part to the measurement noise inherent in all of the visual corresponding colour datasets.

Limitations of CAMs for AR & VR

CAMs remain imperfect, perhaps in part due to visual measurement noise, and in part due to formulations chosen to balance models of visual processes and computational simplicity. Research continues on all aspects of CAMs, including data generation, modifications and extensions to existing models, and new formulations based on the human visual system. For example, recent appearance studies have addressed incomplete adaptation [26][27], high dynamic range stimuli [28], and post-CAM computations to model chromatic simultaneous contrast [29]. Other research has borne cone-based formulations, one focused on HDR data and uniform colour differences [30], and one oriented toward the long-term contrast adaptation expected between individual observers [31]. An image-based CAM has been proposed, but has so far not been widely applied [32].

CIECAM02, CIECAM16, and related models have proven useful in predicting appearance for reflective objects and displays, but it is not obvious that they may be directly applied to AR and VR viewing conditions without further study. VR is relatively similar to traditional displays,

except that the field of view is much larger and the content is typically rendered interactively in response to the user's head position. Initial research has addressed colour constancy in VR [33] and colour management in VR [34], essential steps toward approaching colour appearance. AR, however, is much more complex due to the mix of real and virtual stimuli, which can create unusual viewing experiences that were not anticipated in the development of current CAMs.

Chromatic Adaptation

One of the most important features of the human visual system, chromatic adaptation includes both retinal and neural processes working toward a goal of colour constancy: the concept that the perception of an object's colour can remain relatively fixed despite changes in illumination colour and intensity. In general, people can adapt fully to a wide range of whitish light sources in an interior environment, including both luminance adaptation and chromatic adaptation. Chromatic adaptation from one light source to a second can be modelled by separately scaling the long-, medium-, and short-wavelength sensitive (LMS) cone signals by the ratios of the *LMS* signals of a representative neutral (or white) illuminated by the two sources, as proposed by von Kries [35]. The transform may be implemented as in the following equation, using a diagonal matrix in which the diagonal elements are these ratios, subscripts *2,w* indicating the response to a white object under the second light source, and *1,w* indicating the response to a white object under the first light source:

$$\begin{bmatrix} L_2 \\ M_2 \\ S_2 \end{bmatrix} = \begin{bmatrix} L_{2,w}/L_{1,w} & 0 & 0 \\ 0 & M_{2,w}/M_{1,w} & 0 \\ 0 & 0 & S_{2,w}/S_{1,w} \end{bmatrix} \begin{bmatrix} L_1 \\ M_1 \\ S_1 \end{bmatrix} \quad (6)$$

Incomplete or partial adaptation to an illuminant can occur in low-illumination environments or if the illuminant colour is chromatic. It can also happen when viewing a display in a typical interior environment, because the state of adaptation depends partly on the display and partly on the ambient illumination [26][27][38][39]. Because the display is seen in the illuminant mode and not part of the rest of the scene in object mode, it can be difficult to make cross-media comparisons between the display and reflective colour patches. Modern colour appearance models implement a chromatic adaptation transform (CAT) utilizing sharpened cone signals, and also account for incomplete adaptation [20][22][36]. However, CIECAM02 and CIECAM16 implement a degree of adaptation that depends only on the light level, but not on the illuminant chromaticity, which can overestimate adaptation to chromatic lighting.

Contemporary display systems in mobile devices and laptop computers commonly implement both auto-brightness and auto-white adjustment algorithms. One goal of an auto-white algorithm is to keep the display white point similar to the colour of the ambient illumination, which would presumably eliminate a potential conflict in chromatic adaptation [37]. It also results in the display appearing more like reflective paper, which can be preferred for reading. Note that these algorithms can render a GOG characterization model useless unless the model is augmented with parameters accounting for the adjustments the algorithms make. For research uses of displays, it is best to turn off automated algorithms and use the display in a fixed condition.

Chromatic Adaptation in VR

VR systems typically block out the real world and provide complete control of the user's field of view; this may make it possible to direct their adaptation more fully or more completely than would be possible with a normal display [34]. Unusual situations, like rendering a simulated emissive display within a VR environment, have not been tested, though research on display- and VR-rendered architectural interiors and lighting environments indicates a high degree of adaptation and perceived realism [40]. VR presentation of a rendered building interiors, both artificially-lit and daylight, have been shown to convey lighting attributes comparably to the real space itself [41]. The focus of many of these studies on lighting, rather than reflective objects, may be a limitation in generalizing their positive results.

Several perspectives are possible for choosing the optimal white point for VR display: on one hand, assuming more complete adaptation in VR, the choice of white point may not matter much, therefore it would make sense to keep the display white fixed at a standard like D65. On the other hand, because a user regularly takes the VR HMD off to interact in the real world, it would perhaps be better to adaptively adjust to match ambient white just as many mobile displays do so, to make those transitions less jarring. Another perspective is that a content creator may want to make use of a wide range of display white points (or, equivalently, image white balance settings) as a creative tool to make transitions between different virtual environments more impactful.

Chromatic Adaptation in AR

AR provides a more complicated situation than VR because of the visibility of the real world behind the AR content. The user's adaptation will be partly influenced by their real-world environment and partly influenced by the stimuli presented on the AR system itself [43]. In studies of appearance in AR, the choice of reference white when using uniform colour spaces such as CIELAB and CIECAM02 has involved some trial-and-error, settling on an average between a perfect white in the illuminated real-world environment and the white point of the transparent AR display [44]. The luminance of the real and AR white points, as well as the relative size and salience of the AR stimuli, presumably affect the extent to which each of these affect the user's chromatic adaptation. In AR viewing environments, it seems likely that matching the chromaticity of the real-world illumination would be a good choice to avoid conflicts in adaptation, but this has not yet been adequately tested.

Scission & Transparency in AR

Recent research on colour appearance in AR has revived the concept of perceptual scission, or scissioning, which is the cognitive separation or decomposition of colour combinations into layers [45][46]. In a variety of situations where the spatial structure of visual stimuli can be interpreted as layers, perceptions of transparency and induction, among others, can result [47]. In a simple OST-AR situation, the proximal stimulus (meaning the radiometric flux reaching the user's eyes) is the sum of the real-world background, attenuated by the transmission of the transparent display, and the AR overlay itself [48]. The proximal stimulus is a good physical representation of the optical combination of AR overlay and background, and measurable considering the complications mentioned previously. However, due to perceptual scission,

which can vary in extent depending on the viewer's interpretation of the situation, it is generally a poor representation of the appearance of the perceived stimulus.

Existing CAMs have no accounting for transparency or scission, making them very incomplete for OST-AR stimuli. Two possibilities exist to adapt CAMs to predict empirical AR appearance data: using the CAM as usual, and transforming the resulting lightness, chroma, and hue (*JCh*) correlates; or, transforming the input colorimetry (*XYZ*) values before applying the CAM. The latter appears to fit experimental data better, despite the counter-intuitiveness of a transformations in the linear *XYZ* stimulus space to predict a cognitive effect. A flexible foreground-background blending model has been proposed to predict an "effective" colorimetric stimulus, generally different from the proximal stimulus, as a weighted sum of the colorimetry of the AR foreground and that of the real-world background [49]:

$$XYZ_{eff} = \alpha XYZ_{FG} + \beta XYZ_{BG} \quad (7)$$

Vectors XYZ_i represent measured (from the user's perspective) colorimetry, though *LMS* cone signals or even spectral radiances may be substituted instead. The subscripts refer to effective (eff), AR foreground (FG), and real background (BG). Importantly, the designation "background" refers to the real-world scene behind the AR display including the attenuation of the display's transparency; thus β is not a measure of transparency. α and β are the weights of the relative contributions of the foreground and background to the perceived colour.

The weighting scalars α and β would both equal one in order to describe the physical, proximal stimulus. However, recent visual experiments have shown that α and β can each range from about 0.5 to 6 depending on the situation [52], which seems to indicate that the visual system can "discount" either the foreground or background to some extent, which is also evidence of scission. A series of experiments have shown that the model of Eq. (7) can be helpful in describing appearance in AR, but future research may uncover a better model; in any case, more research is required to predictably determine the α and β weights for specific situations.

Experimental Evidence for Scission in OST-AR

The foreground-background blending model in Eq. (7) was developed through a series of visual experiments designed to measure colour appearance in a variety of AR situations. Visual cues can induce perceptual scission and cause the perceptual discounting of one layer or the other, depending on the situation and visual task. To simplify the accurate presentation of colour stimuli in experiments, a desktop OST-AR system, rather than an HMD, was constructed using a conventional 27" LCD and a large beamsplitter, as shown in Figure 2 and described in detail elsewhere [44].

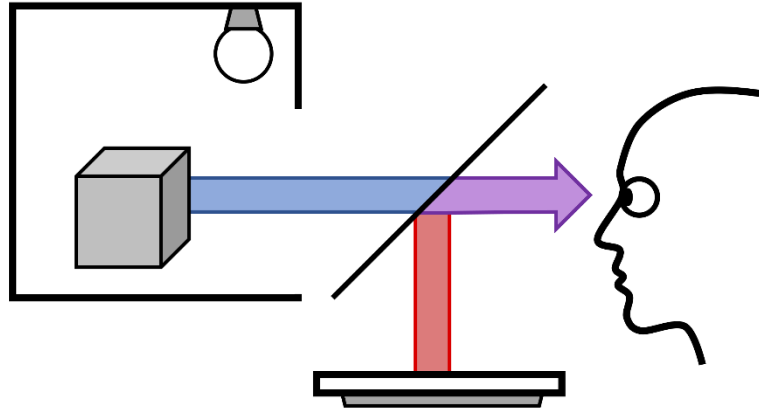


Figure 2: Desktop OST-AR Setup: The observer, at right, sees an optical blend (indicated by purple arrow) of real and virtual content in which the virtual content appears as a semi-transparent, bright AR overlay. At the centre of the diagram is a diagonal beamsplitter that transmits light from the real-world objects illuminated in the viewing booth at left (blue arrow) and reflects light from virtual content displayed on the LCD display below (red arrow). Adapted with permission from [51] © The Optical Society. [Source: Murdoch]

The earliest set of studies asked observers to match the colour appearance of an AR rectangle overlaid on a different-coloured backgrounds by adjusting the colour of an AR rectangle on a black background, as shown in Figure 3A. A variety of colours were used for both AR foreground and real background. The coloured backgrounds distorted the AR colours by blending with them; however, observers' visual matches were not the same as physical colour matches. Observers consistently "discounted" the background contribution to the blend, adjusting their matches to colours intermediate between the undistorted colours and the background-distorted colours [44]. The foreground-background blending model was later applied to these results, with best-fit α values between 1 and 1.1 and β values between 0.5 and 0.6 [49]. Thus, it appears that observers were able to ignore about half of the background's physical contribution to the blended colour when interpreting and making colour matches. In a follow-up study with more complex AR stimuli (rendered 3D objects rather than flat colour patches), even stronger discounting was observed, with α about 1.0 and β values about 3.1, presumably because the 3D objects induced stronger scission.

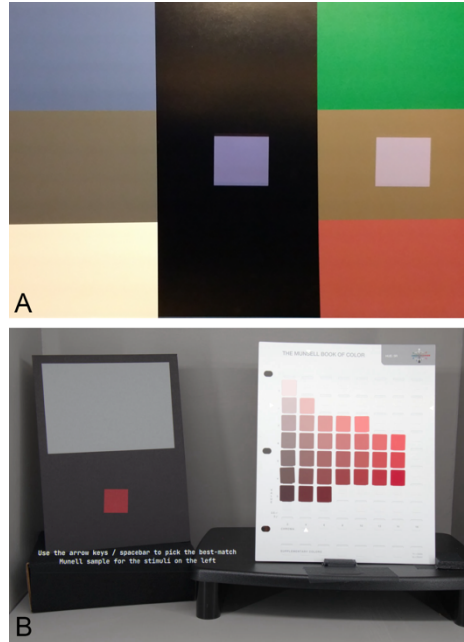


Figure 3: OST-AR Experiment Stimuli: Image A shows a stimulus presentation for AR-to-AR colour matching across different backgrounds. Observers were asked to adjust the colour of the bluish rectangular AR patch on the black background to match the appearance of the rectangular AR patch on the brown background. Image B shows a stimulus presentation for a matching experiment from AR to real-world colour patches. Observers were asked to select the best match from the arrangement of small Munsell colour patches at right to the red rectangular AR patch on the black background at left. Image B adapted with permission from [52] from IS&T: The Society for Imaging Science and Technology sole copyright owners of, "CIC29: Twenty-ninth Color and Imaging Conference 2021" [Source: Murdoch]

Background discounting was also shown in an OST-AR experiment with 2D and 3D rendered stimuli, where chromatic mismatch between foreground and background induced greater ratios of α to β [50]. When the AR foreground and illuminated background were at similar chromaticities, α to β values were both found to be close to 1, indicating no discounting, potentially because no conflict was seen. With differences in either luminance or chromaticity, α : β ratios increased, and the 3D stimulus induced slightly higher ratios yet. Thus, it seems that the visual cues of mismatch and dimensionality lead to scissoring and discounting.

Another study looked at brightness matching using AR overlays on real 3D cubes. Results showed that when the AR overlays were more obviously a transparent layer, presented oversized relative to the cubes, the AR foreground was discounted, with α values as low as 0.76 while β values remained at or slightly above one [51]. Compared with the previous matching experiments, these results show that differences in the matching task, focusing on either the AR foreground or the real-world background, result in discounting the other layer.

A later experiment using a similar AR setup tested the correspondence between AR colour patches and physical reflective colour samples, asking observers to select the best match from pages of Munsell colour patches, as shown in Figure 3B. Several sets of stimuli were used, each set covering a range of lightness and chroma at a fixed hue, corresponding to a page of Munsell patches. Black and gray backgrounds were used, and in some presentations a thick white border was included around the AR colour sample. Both background and border had large effects on the observers' colour matches, and a satisfactory global fit using the foreground-

background blending model was not found. However, looking at individual matches, patterns were clear. First, observers apparently discounted the gray background, with β values mostly between 0.5 and 1. Second, when the white border was present, α values very close to 1 resulted; with no border, α values ranged from 1 to nearly 6, with higher values for darker colours [52]. The implications seem to be that people interpret colours and transparency differently depending on the context, and that it will be difficult to populate the α and β weights for a wide variety of situations.

Interpretation of Transparency and Related Visual Effects

The complication of describing perceptual effects in AR is confounded by other perceptual effects which are not even accounted for in contemporary CAMs. The perception of transparency is thought to be related to chromatic induction, including the effects of simultaneous contrast, assimilation, and neon spreading [47]. Some of these related effects are shown in Figure 4. Figure 4A shows two physically-identical orange circles with different surround colours; the appearance of the circles differ, the left circle appearing lighter and yellower, the right one appearing darker and redder. This visible but mild appearance difference is known as simultaneous contrast; essentially, the surround colour pushes the appearance of the circle away from it. Figure 4B uses the same three colour stimuli in a striped arrangement – yes, the same red, yellow, and orange as in Figure 4A, which may be difficult to believe because the red and yellow stripes within the orange circles induce the much stronger effect of assimilation. The orange circle with red stripes appears redder, and the orange circle with yellow stripes appears yellower. This is similar to the Munker illusion, in which the alignment of stripes induces hue differences [53], and it is also similar to transparency, as the closely-spaced stripes may appear to blend together into an intermediate colour. Figure 4C shows the same red and yellow comprising a split background, but the circle is made up of two new orange colours, a lower-contrast pair in a spatial arrangement that induces the perception that the circle is a transparent filter on top of the coloured rectangles. It may be that the orange colours in Figure 4C appear similar to the illusory colours in Figure 4B. Figure 4D-F are analogous to the previous, though achromatic, illustrating that these effects may be obtained with or without chromatic content or contrast. Interestingly, all of the stimuli in Figure 4 can be attributed to transparency and the discounting of an interpreted background, consistent with the concept of scission [47][54].

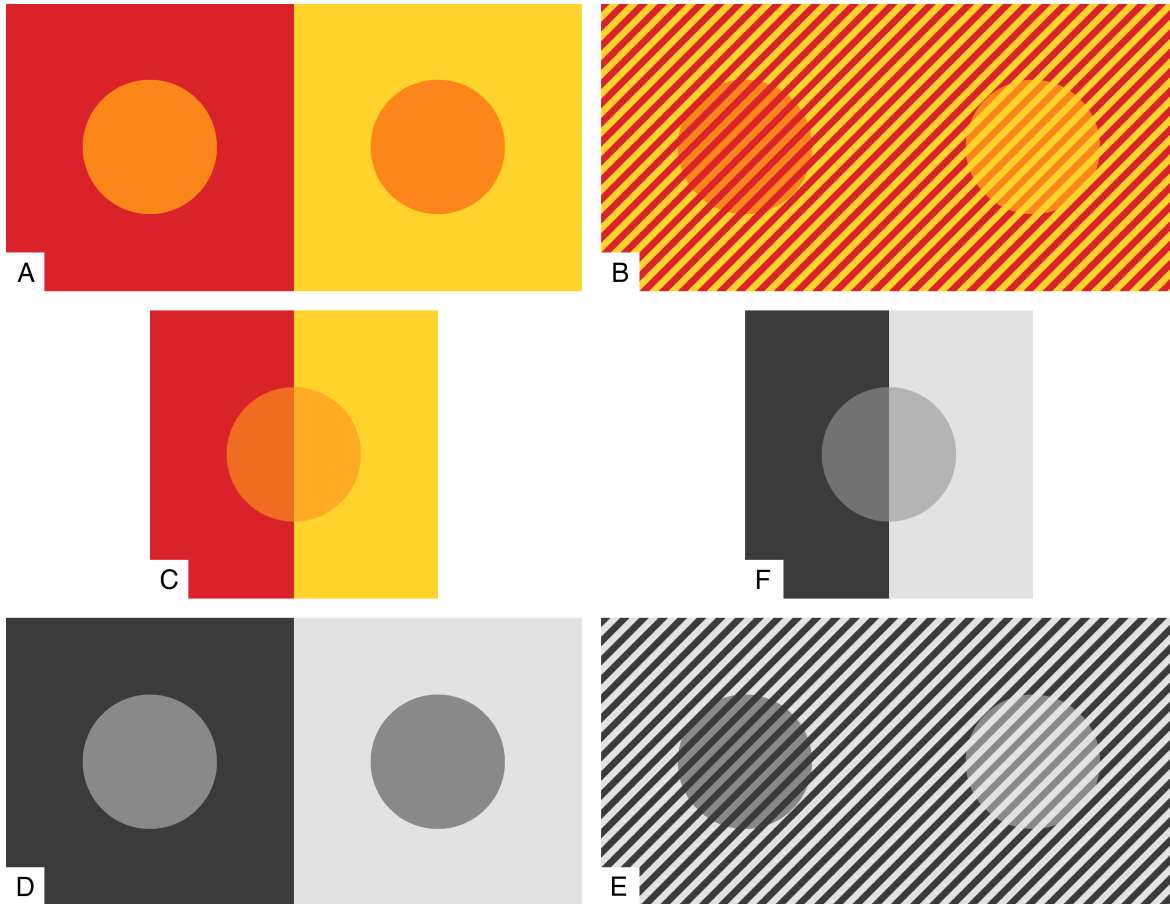


Figure 4: Induction and transparency: Images A and D illustrate simultaneous contrast, whereby the pairs of circles, identical colour stimuli, appear different because of the colours of the surrounding rectangles. Images B and E illustrate assimilation, where appearance is affected by the colour of stripes. Images C and F show a spatial arrangement of colours that may be interpreted as a transparent circle on top of the split-colour background. [Source: Murdoch]

Existing literature on transparency perception has focused on subtractive filters that reduce the visibility of background detail and scattering or reflective filters that reduce the contrast of background objects [45][47]. However, while in OST-AR the transmission of the optical system is typically non-scattering, the AR display adds light atop the background scene, resulting in a reduction of the contrast of the real-world background. This may also be compared to veiling luminance, another optical reduction in contrast, that is known to be discounted while interpreting background objects behind the veil [55]. Because OST-AR only adds light to the background, a practical implication is that AR luminance correlates with both perceived opacity and perceived brightness, which means that darker AR stimuli are also more likely to appear more transparent, a situation not seen with other media [56].

Example: Modelling an OST-AR Display and Colour Matching Results

Building on the previous example of characterizing an OST-AR display, this example shows how it has been used for research on colour appearance in AR. The example has two components: the modelling of the transparent OST-AR display as it interacts with the background and the modelling of colour-matching experiment data. The former may be interesting to anyone

building or characterizing OST-AR displays, while the latter is tailored more to the experimentalist interested in generating more AR appearance data. The reason for such experiments is to eventually build a robust model of colour appearance in AR that may be applied in the normal use of AR displays, though at the moment a complete model remains elusive.

In the AR colour matching experiments of Hassani and Murdoch [44][49], a transparent display like the one characterized in the previous example was used to present AR stimuli on various backgrounds. As shown in Figure 3A, observers were shown pairs of AR colour patches and asked to adjust the patch on black to match the appearance of the patch on the coloured background. Characterizing the visual stimuli in terms of both AR and real-world components, and their combination, was an essential part of this work, as explained here.

Referring to the coloured, inkjet-printed background in Figure 3A, measurements were made of all six colour blocks as well as the top, middle, and bottom portions of the black centre region with the LED lighting set to a warm white. Because the transparent display was rather dim (due to the low reflectance of the beam splitter), the lighting was set to a relatively low level of intensity to balance its contribution visual scene. The previous example discussed the characterization of a transparent display with a black cloth behind to ensure the background was not included in the display model, and it showed excellent prediction of colorimetry. This was an actual example for measurements made in the centre of the display. However, in preparation for this experiment it was found that the display, as measured at the three locations on the black background, did not produce the same colour; variations in XYZ colorimetry were greater than 5%, which is unacceptably large. Perhaps this was due to angular differences either in the output of the LCD itself or in the reflectance of the coated glass beamsplitter. Regardless of the reason, the solution to this situation was quite simple: a separate display model for each of the top, middle, and bottom locations, each built as in the previous example from measurements made at the corresponding location. Treated separately, each of these display models performed with similarly high accuracy.

The next step was to add in the effect of the background to accurately predict the proximal stimuli, meaning the combination of real-world printed background and virtual AR transparent foreground. Because each of the three display models predicted AR-only display colorimetry, the measured background XYZ values were simply added. This can be thought of as treating the background behind the transparent display as additional flare light.

Separate display models were not made for each of the six coloured background locations, because for this experiment, a small number of discrete stimuli were used as reference colours. For these locations, the middle display model was used to generate approximate foreground stimuli, and the actual foreground-background blended stimuli shown in the experiment were measured individually. Table 2 lists the XYZ tristimulus values of seven selected AR foreground colours: the XYZ of the AR display itself (AR-only) and the proximal stimuli as measured on the green background [49].

Table 2: AR matching data: CIE 1931 tristimulus values are listed for the seven named foreground colours. The first XYZ values are for the AR-only, meaning only the transparent AR display with a black background. The second set of XYZ values are the reference stimuli, the AR foreground colours on a green background of XYZ (1.01, 1.80, 0.38). Third set of XYZ are the average matches to the reference on a black background, and the fourth set of XYZ are the model prediction of the matches.

		AR-only XYZ			Reference XYZ			Actual Match XYZ			Modelled XYZ		
		X	Y	Z	X	Y	Z	X	Y	Z	X	Y	Z
	Brown	0.49	0.43	0.60	1.52	2.22	1.04	1.00	1.19	0.89	1.01	1.36	0.81
	Green	0.58	0.69	0.78	1.61	2.50	1.24	1.46	1.84	1.70	1.11	1.62	0.99
	Orange	3.76	2.80	1.30	5.20	4.87	1.80	5.14	4.56	1.91	4.35	3.77	1.52
	Blue	1.53	1.13	5.57	2.69	3.00	6.49	2.42	2.32	6.44	2.08	2.08	5.88
	Red	2.18	1.54	1.61	3.41	3.45	2.11	3.15	2.98	1.91	2.73	2.49	1.83
	Purple	0.35	0.27	0.85	1.36	2.04	1.29	1.60	1.73	2.89	0.87	1.19	1.06
	White	15.5	15.6	27.7	18.4	19.4	31.6	18.6	19.0	32.8	16.3	16.9	28.5

Turning from the characterization of AR stimuli to the visual experiment and its results, recall that observers were asked to adjust the colour of the AR overlay on the black background next to the reference AR overlay on a coloured background. Due to the appearance effects discussed previously, even though the pairs of AR colours matched in appearance, the XYZ values of corresponding reference and matched patches are not the same. In fact, the matches are mostly somewhere between the AR-only and reference XYZ values. Table 2 includes the Actual Match XYZ, which are the average XYZ over all observations, and the Modelled XYZ, the result of the model discussed in the next paragraph. These data are also equivalently shown in CIE $u'v'$ chromaticity in Figure 5.

While this example illustrates a subset of the experiment, visual colour matching results for the whole set of colour matches (seven foreground colours on each of six background colours) were fit using the parameters α and β in the foreground-background blending model in Eq. (7). Weighting scalars α and β were found to be 1.02 and 0.51, respectively, meaning that in the visual matches, the contribution of the real-world background was discounted by half. Or, equivalently, the visual weight of the foreground was twice as large as the background.

In the experiment analysis, the optimization of α and β was done by minimizing the average ΔE in CIECAM16-UCS, which is similar to CIELAB but takes advantage of CIECAM16's accounting for chromatic adaptation and absolute luminance. While this plot ignores any luminance differences, the results are most easily visualized with the subset of data shown in the chromaticity diagram in Figure 5. The green background is represented by the labelled green dot near the upper left. The AR-only foreground stimuli (meaning, without background distortion) are shown as labelled red dots, connected by dotted lines to black dots representing the resulting chromaticity when the background is present. The dotted lines show that the light added by the background distorts all of the foreground colours, which converge toward the background chromaticity. The degree of convergence, meaning the length of the line, is larger for the foreground colours that are relatively lower in luminance than the background. In the experiment, the reference colours were those distorted by the background (black dots), and observers adjusted AR colours on a black background to visually match. Observers' matches (plotted as +) consistently lie on the lines between the physical, distorted colour and the colour they would have seen if the background were not present, showing that they were able to discount or ignore the background to some extent. The modelled results (plotted as open circles) show the best fit using Eq. (7) with the two weighting scalars α and β mentioned above to describe the visual discounting. The fit is not excellent for all colours shown here; recall that it was optimized over the whole experiment, not just this subset of stimuli.

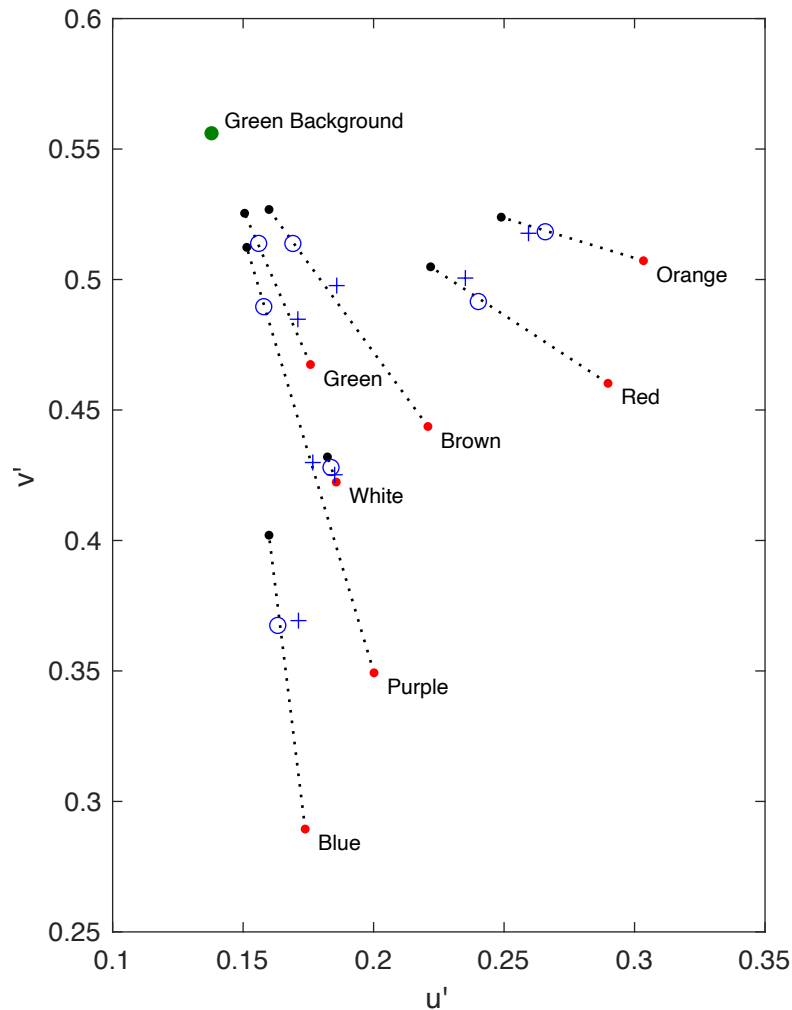


Figure 5: AR matching results: This CIE 1976 UCS diagram shows chromaticity coordinates of the real-world printed green background (labeled green circle), AR-only foreground colors (labeled red dots), and measured reference stimuli including the optical blend of AR foreground and real background (black dots). Average observer matches to the reference stimuli (+) can be compared to the modelled matches (open circles). [Source: Murdoch]

Colour Imaging & Graphics in AR & VR

AR and VR systems are not simply output devices; they are parts of an interactive ecosystem involving image capture, image synthesis, content creation, and reproduction. Colours, images, and compositions seen through AR & VR are created in real time in response to the viewer's selections and actions. The reproduction of colours both captured from the real world and synthesized using computer graphics must be considered.

Colour Reproduction

Some VR content is primarily or entirely derived from camera footage, such as 360-degree videos and immersive environments. The starting point for colour reproduction in VR is the same as in any imaging system, wherein the system tone reproduction should be designed to amplify image contrast and colour saturation to match the appearance of the original scene

[57]. As in cinema, artistic manipulations on top of this starting point may be used for dramatic effects to advance a story or engage a viewer.

For some applications, the use of encoding standards like sRGB, typically used for computers and web applications, or Rec 2100, which specifies image encoding and display for high dynamic range content, may be appropriate. AR and VR displays have so far not implemented high-luminance HDR, but that may be expected in time.

Virtual Colour Reproduction

In the creation of virtual content, there is no real-world scene to reproduce, but there are virtual scenes and scene elements that may be reproduced similarly to a real scene. Many different computer graphics rendering methods exist, some more physically-based than others, and addressing their differences is out of scope for this chapter. However, treating virtual cameras and imaging pipelines similarly to real imaging systems will generally benefit realism, and it will also make compositing real and virtual images easier and more natural.

In AR applications, making virtual content appear to “fit in” to a real-world scene is often the objective, which implies matching virtual lighting to an estimate of the lighting in the real scene, varying levels of accuracy depending on time, tools, and needs [58]. One specific problem to avoid is the rendering of object colours out of context of a scene, which may make them appear impossibly bright and/or impossibly chromatic. While this may be a valuable choice for dramatic effect in some cases, the result is generally very unrealistic and can easily upset a seamless composition of real and virtual.

Some recent research has addressed realism, colour constancy, and colour reproduction in VR systems and the real-time graphics engines driving them. Many details in the settings of rendering engines affect the colour output of shaded objects as well as the reproduction of images used as texture maps in rendering [59]. Such dependencies would seem to make it difficult to measure and reliably predict colour reproduction in interactive graphics environments, yet some evidence shows that scene renderings using graphics engines is sufficient to induce natural visual responses such as chromatic adaptation [34]. As in any application with complex images, it may be that focusing on the pixels themselves is less valuable than the higher-order characteristics. A general suggestion is to stay as close to physics as possible, with physically-based rendering techniques and tone reproduction characteristics similar to traditional imaging systems.

Summary

This chapter has addressed colour in AR and VR systems. To a certain extent, AR and VR displays can be treated similarly to other displays, but with more complex optics that introduce novel distortions and can make measurement more difficult. VR applications, because they fill the viewer’s field of view, may make predictions of visual adaptation simpler than in the mixed viewing environments used for many other displays. AR applications, on the other hand, are much more complicated due to the mixture of virtual content into real-world scenes, which opens questions about adaptation and interpretation. Ongoing research has resulted in a promising but incomplete model of colour appearance in OST-AR based on the concept of

perceptual scission, whereby a viewer interprets transparent content as a layered presentation and discounts one layer's contribution to the blend. Thinking of AR and VR as imaging systems, colour reproduction is difficult to isolate due to the real-time rendering engines required to drive them interactively, leaving physically-based rendering with natural tone reproduction an appropriate aim for comfortable, realistic presentation.

Open Questions in AR & VR

This chapter presents the most up-to-date understanding of colour in AR and VR systems. However, it must be acknowledged that the technology behind these systems is progressing very quickly, and research on colour and perception is ongoing. It may be that within a few years, looking back on this text will expose the folly of partial understanding. In order for such clarity to be attained, research should address open questions in at least a few categories.

Adaptation, especially chromatic adaptation in AR and VR must be given significant attention. The time course of adaptation from real-world to VR and back, as well as adaptation to changes within the virtual VR environment, should be studied, along with the extent of partial adaptation. In AR, chromatic adaptation to the mix of real-world and virtual stimuli needs to be measured and modelled, and compared to the limited data on adaptation to mixed illumination in fully real-world environments. As always in colour science, the question of “what appears white?” is deceptively simple, hard to measure, and essential to modelling colour appearance.

Further study should address the colour appearance of different kinds of objects and materials, for example reflective, emissive, and transparent, both in the real world as well as in virtual renderings. It is interesting to compare the appearance of an actual reflective object to a colorimetric rendering of it on an emissive display, then to consider the same comparison in a VR environment. Humans have a learned intuition about object properties, and it seems that graphics systems are most realistic when their output is consistent with this intuition. However, it remains unclear how all of the physical characteristics of different objects and materials contribute to colour appearance.

Finally, in order for colour appearance in AR to be robustly modelled, much more experimental data on the correspondences between virtual and real stimuli and between virtual stimuli with different backgrounds and context are needed. Some of the experiments summarized in this chapter have built a foundation for this work, but more experiments with more diversity in the kinds of stimuli presented need to be done, and models like the foreground-background blending model explained earlier need to be exercised and either verified, improved, or replaced.

The future of AR and VR systems is full of promise for many applications, and scientists, engineers, and developers are racing to improve the technology and create killer apps. The more is understood about colour perception, adaptation, and appearance in the use of these systems, the better the hardware and software designers will be able to create predictable, robust, comfortable, and natural visual experiences.

Acknowledgements

The author thanks RIT Munsell Color Science Laboratory graduate researchers Nargess Hassani, Lili Zhang, Tucker Downs, and Zilong Li, as well as RIT scholars Sara Leary and Josephine Bensa, for their thoughtful and impactful contributions to ongoing research on colour appearance in AR.

This material is based in part upon work supported by the National Science Foundation under Award No. 1942755.

References

- [1] Melzer, J., & Spitzer, C. R. (2017). Head-mounted displays. *Digital Avionics Handbook*, 256-279.
- [2] Huang, W., Alem, L., & Livingston, M. A. (Eds.). (2012). *Human factors in augmented reality environments*. Springer.
- [3] Livingston, M. A., Ai, Z., & Decker, J. W. (2018, July). Human factors for military applications of head-worn augmented reality displays. In *International Conference on Applied Human Factors and Ergonomics* (pp. 56-65). Springer, Cham.
- [4] Chen, Y., Wang, X., & Xu, H. (2021). Human factors/ergonomics evaluation for virtual reality headsets: a review. *CCF Transactions on Pervasive Computing and Interaction*, 1-13.
- [5] Berns, R. S. (1996). Methods for characterizing CRT displays. *Displays*, 16(4), 173-182.
- [6] Fairchild, M., & Wyble, D. (1998). Colorimetric characterization of the apple studio display (flat panel LCD).
- [7] Murakami, Y., Hatano, N., Takiue, J., Yamaguchi, M., & Ohyama, N. (2004, May). Evaluation of smooth tonal change reproduction on multiprimary display: comparison of color conversion algorithms. In *Liquid Crystal Materials, Devices, and Applications X and Projection Displays X* (Vol. 5289, pp. 275-283). International Society for Optics and Photonics.
- [8] Miller, M. E., & Murdoch, M. J. (2009). RGB-to-RGBW conversion with current limiting for OLED displays. *Journal of the Society for Information Display*, 17(3), 195-202.
- [9] Sun, P. L., & Lu, S. Y. (2008, December). Optimal xvYCC to RGBW Conversion for Field Sequential Color LCD. In *2008 International Conference on Computer Science and Software Engineering* (Vol. 6, pp. 278-281). IEEE.
- [10] Murdoch, M. J. (2019). Dynamic color control in multiprimary tunable LED lighting systems. *Journal of the Society for Information Display*, 27(9), 570-580.
- [11] Durmus, D. (2021, March). Multi-objective optimization trade-offs for color rendition, energy efficiency, and circadian metrics. In *Light-Emitting Devices, Materials, and Applications XXV* (Vol. 11706, p. 117061J). International Society for Optics and Photonics.
- [12] Süssstrunk, S., Buckley, R., & Swen, S. (1999, January). Standard RGB color spaces. In *Color and Imaging Conference* (Vol. 1999, No. 1, pp. 127-134). Society for Imaging Science and Technology.

- [13] IEC 61966-2-1 Ed. 1.0 b:1999. Multimedia Systems And Equipment - Colour Measurement And Management - Part 2-1: Colour Management - Default RGB Colour Space - SRGB
- [14] Chen, J., Cranton, W., & Fihn, M. (Eds.). (2016). *Handbook of visual display technology*. Berlin, Germany: Springer.
- [15] Hainich, R. R., & Bimber, O. (2016). *Displays: fundamentals & applications*. AK Peters/CRC Press.
- [16] Beams, R., Kim, A. S., & Badano, A. (2019). Transverse chromatic aberration in virtual reality head-mounted displays. *Optics express*, 27(18), 24877-24884.
- [17] Zhang, L., & Murdoch, M. J. (2018). Color Matching Criteria in Augmented Reality. *Journal of Perceptual Imaging*, 1(1), 10506-1.
- [18] Penczek, J., Boynton, P. A., Meyer, F. M., Heft, E. L., Austin, R. L., Lianza, T. A., ... & Gacy, L. W. (2017). Absolute radiometric and photometric measurements of near-eye displays. *Journal of the Society for Information Display*, 25(4), 215-221.
- [19] Oshima, K., Naruse, K., Tsurutani, K., Iwai, J., Totani, T., Uehara, S., ... & Wakemoto, H. (2016, May). 79-3: Eyewear Display Measurement Method: Entrance Pupil Size Dependence in Measurement Equipment. In *SID Symposium Digest of Technical Papers* (Vol. 47, No. 1, pp. 1064-1067).
- [20] Fairchild, M. D. (2013). *Color Appearance Models*. John Wiley & Sons.
- [21] CIE 159:2004, *A Colour Appearance Model For Colour Management Systems: CIECAM02*, Vienna: Commission Internationale de L'Eclairage (2004).
- [22] Li, C., Li, Z., Wang, Z., Xu, Y., Luo, M. R., Cui, G., ... & Pointer, M. (2017). Comprehensive color solutions: CAM16, CAT16, and CAM16-UCS. *Color Research & Application*, 42(6), 703-718.
- [23] CIE 248:2022, *THE CIE 2016 COLOUR APPEARANCE MODEL FOR COLOUR MANAGEMENT SYSTEMS: CIECAM16*
- [24] Luo, M. R., & Hunt, R. W. G. (1998). Testing colour appearance models using corresponding-colour and magnitude-estimation data sets. *Color Research & Application*, 23(3), 147-153.
- [25] Li, C., Luo, M. R., Rigg, B., & Hunt, R. W. (2002). CMC 2000 chromatic adaptation transform: CMCCAT2000. *Color Research & Application*, 27(1), 49-58.
- [26] Smet, K. A., Zhai, Q., Luo, M. R., & Hanselaer, P. (2017). Study of chromatic adaptation using memory color matches, Part I: neutral illuminants. *Optics express*, 25(7), 7732-7748.
- [27] Smet, K. A., Zhai, Q., Luo, M. R., & Hanselaer, P. (2017). Study of chromatic adaptation using memory color matches, Part II: colored illuminants. *Optics express*, 25(7), 8350-8365.
- [28] Xinye, S., Yuechen, Z., & Ronnier Luo, M. (2021, November). A new corresponding color dataset covering a wide luminance range under high dynamic range viewing condition. In *Color and Imaging Conference* (Vol. 2021, No. 29, pp. 184-187). Society for Imaging Science and Technology.
- [29] Zhu, Y., Tian, D., & Luo, M. R. (2020, November). A new independent dataset to verify the performance of colour appearance models for predicting simultaneous contrast

- effect. In *Color and Imaging Conference* (Vol. 2020, No. 28, pp. 361-365). Society for Imaging Science and Technology.
- [30] Safdar, M., Hardeberg, J. Y., & Luo, M. R. (2021). ZCAM, a colour appearance model based on a high dynamic range uniform colour space. *Optics Express*, 29(4), 6036-6052.
 - [31] Smet, K. A., Webster, M. A., & Whitehead, L. A. (2021). Color appearance model incorporating contrast adaptation—Implications for individual differences in color vision. *Color Research & Application*, 46(4), 759-773.
 - [32] Kuang, J., Johnson, G. M., & Fairchild, M. D. (2007). iCAM06: A refined image appearance model for HDR image rendering. *Journal of Visual Communication and Image Representation*, 18(5), 406-414.
 - [33] Pardo, P. J., Diaz-Barrancas, F., & Cwierz, H. (2021). Color Constancy in virtual reality scenes. A first step toward a color appearance model in virtual reality, In *AIC 14th Congress Milano 2021*.
 - [34] Gil Rodríguez, R., Bayer, F., Toscani, M., Guarnera, D. Y., Guarnera, G. C., & Gegenfurtner, K. R. (2022). Colour Calibration of a Head Mounted Display for Colour Vision Research Using Virtual Reality. *SN Computer Science*, 3(1), 1-10.
 - [35] J. von Kries, Chromatic adaptation, Festschrift der Albrecht-Ludwig-Universität, (Fribourg) (1902) [Translation: D.L. MacAdam, *Sources of Color Science*, (MIT Press, Cambridge, 1970)]
 - [36] Fairchild, M. D. (2001). A revision of CIECAM97s for practical applications. *Color Research & Application*, 26(6), 418-427.
 - [37] Wu, J., Zhang, L., Isikman, S., & Chen, C. (2019). Enhanced Viewing Experience Considering Chromatic Adaptation. *SID*.
 - [38] Zhai, Q., & Luo, M. R. (2018). Study of chromatic adaptation via neutral white matches on different viewing media. *Optics express*, 26(6), 7724-7739.
 - [39] Fairchild, M. D. (1992). Chromatic adaptation to image displays. *TAGA Proc. Rochester*, 2, 803-823.
 - [40] Murdoch, M. J., Stokkermans, M. G., & Lambooij, M. (2015). Towards perceptual accuracy in 3D visualizations of illuminated indoor environments. *Journal of Solid State Lighting*, 2(1), 1-19.
 - [41] Chen, Y., Cui, Z., & Hao, L. (2019). Virtual reality in lighting research: Comparing physical and virtual lighting environments. *Lighting Research & Technology*, 51(6), 820-837.
 - [42] Chamilothon, K., Wienold, J., & Andersen, M. (2019). Adequacy of immersive virtual reality for the perception of daylight spaces: comparison of real and virtual environments. *Leukos*, 15(2-3), 203-226.
 - [43] Huang, H. P., Wei, M., & Chen, S. (2021). White appearance of virtual stimuli produced by augmented reality. *Color Research & Application*, 46(2), 294-302.
 - [44] Hassani, N., & Murdoch, M. J. (2019). Investigating color appearance in optical see-through augmented reality. *Color Research & Application*, 44(4), 492-507.
 - [45] Metelli, F. (1974). The perception of transparency. *Scientific American*, 230(4), 90-99.
 - [46] Heider, G. M., & Koffka, K. (1933). New studies in transparency, form and color. *Psychologische Forschung*, 17(1), 13-55.

- [47] Anderson, B. L., & Khang, B. G. (2010). The role of scission in the perception of color and opacity. *Journal of Vision*, 10(5), 26-26.
- [48] Gabbard, J. L., Swan, J. E., Zedlitz, J., & Winchester, W. W. (2010, March). 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.
- [49] Hassani, N. (2019). *Modeling Color Appearance in Augmented Reality*, PhD Dissertation, Rochester Institute of Technology.
- [50] Zhang, L., Murdoch, M. J., & Bachy, R. (2021). Color appearance shift in augmented reality metameric matching. *JOSA A*, 38(5), 701-710.
- [51] Murdoch, M. J. (2020). Brightness matching in optical see-through augmented reality. *JOSA A*, 37(12), 1927-1936.
- [52] Downs, T., & Murdoch, M. (2021, November). Color Layer Scissioning in See-Through Augmented Reality. In *Color and Imaging Conference* (Vol. 2021, No. 29, pp. 60-65). Society for Imaging Science and Technology.
- [53] Munker, H. (1970) *Chromatic grids, projection to the retina, and translation theory-based description of the color perception*. Habilitation thesis, Ludwig-Maximilians-University, Munich.
- [54] Walraven, J. (1976). Discounting the background—the missing link in the explanation of chromatic induction. *Vision Research*, 16(3), 289-295.
- [55] Gilchrist, A. L., & Jacobsen, A. (1983). Lightness constancy through a veiling luminance. *Journal of Experimental Psychology: Human Perception and Performance*, 9(6), 936.
- [56] Zhang, L., & Murdoch, M. J. (2021, October). Perceived Transparency in Optical See-Through Augmented Reality. In *2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR)* (pp. 115-120). IEEE.
- [57] Giorgianni, E. J., & Madden, T. E. (2008). *Digital color management: Encoding solutions* (2ed). John Wiley & Sons.
- [58] Kronander, J., Banterle, F., Gardner, A., Miandji, E., & Unger, J. (2015, May). Photorealistic rendering of mixed reality scenes. In *Computer Graphics Forum* (Vol. 34, No. 2, pp. 643-665).
- [59] Kim, A. S., Cheng, W. C., Beams, R., & Badano, A. (2021). Color rendering in medical extended-reality applications. *Journal of Digital Imaging*, 34(1), 16-26.