

Perception in HDR and Displays

Although the world has been high dynamic range (HDR) all along, commercial HDR displays' burgeoning availability provides new and interesting challenges for colorimetry and perceptual modeling. The biggest challenge stems from the way that large luminance changes spatial and temporal scales often compressed relative to the natural world. This results in veridical luminance/chromaticity reproduction not representing appearance reproduction. Here, the authors review some of the perceptual and colorimetric issues that create challenges in HDR image reproduction.

by Mark D. Fairchild, Fu Jiang, and Hao Xie

IT IS WELL KNOWN THAT THE HUMAN VISUAL SYSTEM can function effectively over roughly 14 orders of magnitude in luminance level, ranging from starlit nights to bright sunny days. However, the visual system is not capable of detecting useful information about the world at all of those luminance levels simultaneously. Instead, mechanisms of light and dark adaptation function as a sort of automatic exposure control to place the mean response of the visual system near the mean luminance level for any particular scene or environment. These mechanisms include changes in pupil diameter, transition from rod (night) to cone (day) photoreceptors, physiological gain control in the photoreceptors, and other physiological and psychological processing in the neural circuits of the retina and brain.

Thus, when we think about the challenges of perception in high-dynamic-range (HDR) displays, we do well to remember that every natural day-to-day perception in our world is of HDR. It is really historical standard dynamic range (SDR) displays that were often the aberration. There are several aspects of visual perception that become important when considering HDR displays. These range from the initial visual responses of the photoreceptors, to

mechanisms of adaptation, to spatiotemporal variations in the stimuli and how they are presented.

The first step is to define the cone responses, or color matching functions (CMFs), for average and real observers, and analyze the variance that can be produced by various spectral stimuli (or sets of display primaries). This is known as observer metamerism (OM)—differences in color perception across observers because of differences in their visual sensitivities. The next step in the process concerns how observers adapt to the illumination in scenes and displays and the differences between them. A complexity in adaptation is that it can be localized in both space and time, which increases the overall dynamic range of the visual system beyond what can be observed in a single, static scene or image. Lastly, assumptions about whether that state of adaptation is fixed or dynamic can have a significant impact on the best choices for color metrics for describing and/or encoding color information.

We discuss these topics in a bit more detail later in the text. The history of their study and application of perceptual concepts to the reproduction of HDR perceptions can be understood from several texts in the field.¹⁻⁴ Despite this history often predating

Fig. 1.

XYZ-like color matching functions of the 10 two-degree categorical observers from Asano & Fairchild.⁶

HDR display technology, the fundamental concepts of perception have not changed.

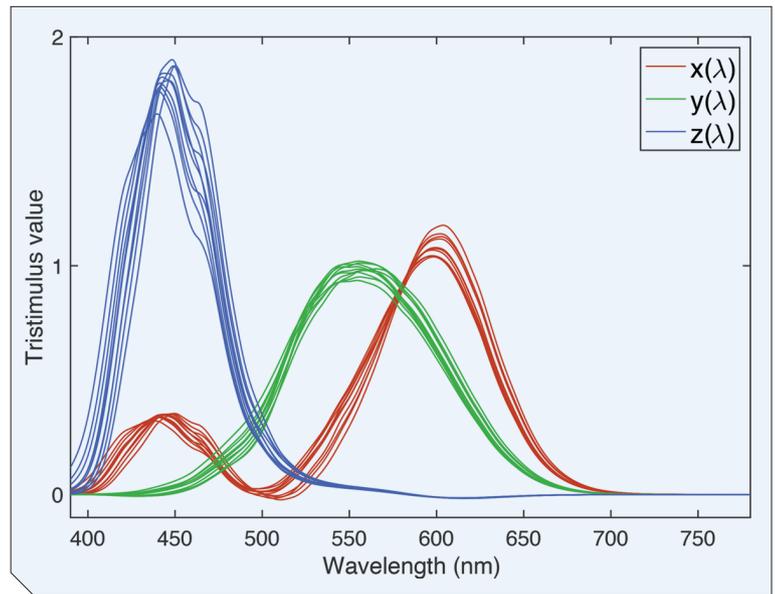
Observer Metamerism

As with all human attributes, color vision varies from person to person, even among those with what is considered normal color vision. For colorimetry, these responses are quantified with CMFs (Fig. 1). Typically average functions, meant to represent an average human observer, such as the CIE 1931 Standard Colorimetric Observer functions, are used in practical colorimetry. Recently, physiological models of the components that define human CMFs have been developed that allow prediction of the range of CMFs that can be found in individuals.⁵ Such models can then be used with information about population demographics to create large collections of CMFs for statistical analysis. One such analysis has led to the collection of 10 categorical observers (Fig. 1).⁶ These categorical observers represent most of the variation found in the population and represent a practical method of evaluating individual differences in color matches, which can become critical in HDR and wide-color-gamut (WCG) displays.

To quantify OM and thus optimize display primary design, color conversion, and other display parameters that may affect spectral outputs,⁷ different metrics have been proposed. CIE recommended an index based on the CIE Standard Deviate Observer,⁸ whose deficiency of underestimated observer variations is addressed by the Asano model.⁵ Long and Fairchild suggested two types of OM indices, OM_{max} and OM_{var} .⁹ For a test population of CMFs, the two metrics quantify the worst color mismatch and the variance volume of color differences, respectively, yielding different implications. More recently, Xie et al. conducted experiments with commercial sRGB displays and suggested

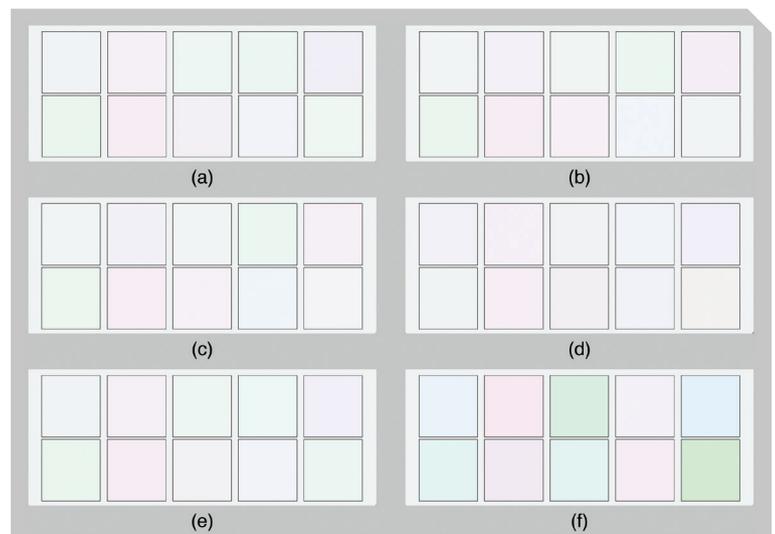
Fig. 2.

Simulation of observer metamerism between CIE D65 spectrum versus its metamers. (a)–(f) correspond to six displays with different primaries, whereas (a)–(d) are four sRGB-gamut displays chosen from Xie et al., 2020,¹⁰ (e) is a DCI-P3 display, and (f) is a simulated Recommendation 2020 display with monochromatic primaries. Each square represents a D65 metamer matched for one of the 10 categorical observers in Fig. 1, the order left-to-right then top-to-bottom following the observer's importance ranks. The square's color and its proximal background are rendered to visualize how the categorical observer's metamer and CIE D65, respectively, appear to the 1931 standard observer.



a better psychophysically correlated OM index, POM_i , which was revised from OM_{max} and counts the percentage of observers that perceive a mismatch larger than a threshold, i .¹⁰ They also pointed out the importance of OM distribution across the gamut, so an image quality metric, which is a function of observer CMFs and image contents, is envisioned. Fig. 2 illustrates the importance of such metrics in displays. The range of colors presented illustrate stimuli that would be color matches to a neutral daylight illuminant (D65) for each of the 10 categorical observers presented in Fig. 1. The key point is that the range of mismatch is significant and depends on the display primaries. Recommendation 2020 primaries (monochromatic lights) have the most sparsely populated spectral power distributions and thus the largest potential for observer differences. This point should be considered as a trade-off with color gamut volume. WCG displays might introduce more observer variability for a meager gain in the range of reproducible color appearance.

One final point that can be considered with individual, or categorical, CMFs, is the development of personalized colorimetry or color management systems. Such systems, based on prototype observer calibrators that have been developed, would allow key stakeholders in color-imaging processes or production



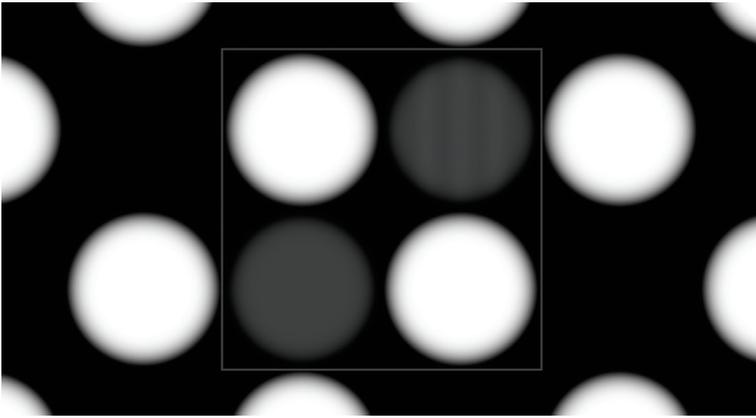


Fig. 3.

▲ Example stimulus used in a dynamic-range perception experiment. The luminance of the darker patches is lowered until the pattern on one of them can be detected. The luminance ratio of the bright patches to the determined dark patch luminance defines the simultaneously perceptible dynamic range for a given spatial scale.



to have displays and other imaging devices calibrated to their personal colorimetry. This can become more important with HDR and WCG displays that have the potential for large observer differences. While personalized colorimetry will not solve problems with mass distribution or multiple simultaneous observers, it provides a potential improvement to current workflows and end-users' experiences.

Threshold Models

Once the initial visual sensitivities are examined, one must consider the perception of changes in luminance on a display. Two types of models are relevant for this discussion: threshold models and appearance scales. Threshold models indicate the smallest changes in luminance that are perceptible in a given situation and can be important in color encoding and ensuring that artifacts are not seen across dynamic changes in adaptation. Appearance scales, discussed in the next section, indicate what the stimuli actually look like. For example, an appearance scale can be used to predict what stimulus will appear medium gray in a given viewing situation, something a threshold model cannot accomplish.

Common to both model types is the need to understand the range of luminance that can be perceived in any given instant. The human visual system can adapt to approximately 14 log units of luminance range. For low luminance level, the rods, activated photoreceptors during low light, are active from about 10^{-6} to 10 cd/m^2 . The cones, photoreceptors working for high luminance levels, are mainly active from 0.01 to 10^8 cd/m^2 . This large range is only valid when the stimulus has enough visual angle separation to avoid glare and enough time is allotted for adaptation while viewing each region. Essential to perceiving this full dynamic range are large spatial extents or temporal transitions (on the order of 20 minutes to adapt from very bright to very dark areas). Thus, calling this full range of luminance the dynamic range of the visual system is a bit of a misnomer. When the bright and dark stimuli are presented simultaneously, the dynamic range of the human visual system is quite a bit narrower because of

Fig. 4.

◀ A simplified comparison of an HDR image rendered to an RGB display (where the maximum luminance of the R, G, and B primaries sums to the white point luminance) in comparison with an RGBW display, where the maximum luminance is produced by adding a white channel to the display. This simulation assumes the white channel is producing 50 percent of the display luminance, and no adjustments are made to the image data to compensate for the reduction in colorfulness. This is a worst-case scenario for illustration purposes, not the results obtained with typical commercial displays.

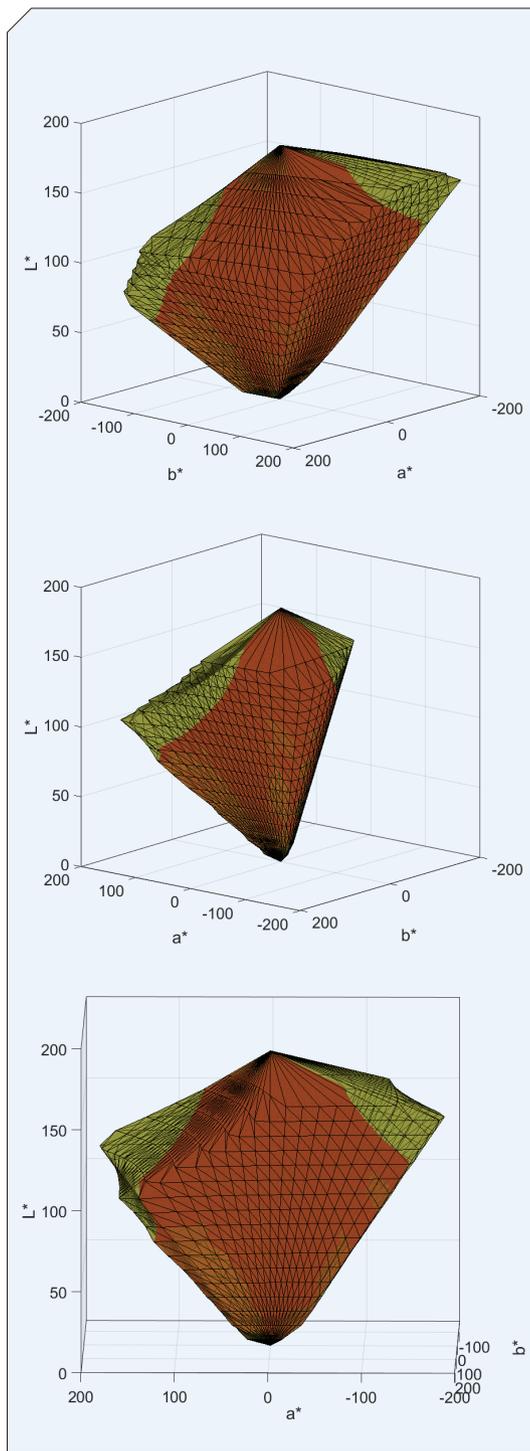


Fig. 5.

Three views of color gamuts for a typical RGB display and a RGBW display, where a white channel produces the brightest highlights. The green areas in the plots show color stimuli that can be produced on the RGB display but not the RGBW display. These plots assume the diffuse white in the image is set to $L^* = 100$, and highlights can reach $L^* = 150$.

glare caused by the bright stimulus and limited simultaneous adaptation ability. Recent experiments by the authors (yet to be published) have examined the simultaneous dynamic range of the human visual system using the pattern in Fig. 3 at various spatial extents, where the observer has limitations on discriminating the spatial Gabor pattern target. In such experiments, a maximum value of 3.3 log units was found for perception of a 10% contrast pattern in such bright-dark alternating patterns. Thus, more dynamic range on a display is wasted during any instantaneous viewing of a still image. Furthermore, the simultaneous dynamic range was found to be stimulus size-dependent, as well as maximum-luminance level dependent. The practical implication of this is that a bright display with more than 3 log units of luminance dynamic range, and some clever processing to manipulate adaptation state, is all that is needed to reasonably simulate most perceptual experiences. Additional display capability could actually degrade viewer experience by introducing other artifacts.

That is where the important concept of floating adaptation levels comes in. In temporally changing images, such as video and motion pictures, there is not a single adaptation state, as there might be time for viewers to adapt to very dark scenes and move their simultaneous dynamic range down to that level and then later adapt to a very bright scene and move their simultaneous dynamic range to a higher luminance. Thus, we have HDR displays with much more than 3 log units of luminance and the capability to encode those 3 log units at both high and low luminance levels. While the visual system cannot perceive all of that information at once, there might be scenes where the dark areas are critical and other scenes where the bright areas are critical. This can also be encountered, to a lesser degree, in very large displays. Thus system engineers are left with either encoding and displaying extremely large dynamic ranges beyond what can be seen at any one time, or developing algorithms to simulate the adaptation process on hardware with a lesser dynamic range, something with a long history in image reproduction.¹⁻⁴

This distinction between threshold/encoding models and appearance models is why multiple types of color systems exist, such as ICTCP models, for encoding HDR image content without visible artifacts and appearance-related models, such as CIELAB and CIECAM02, for describing what the stimuli look like to observers and for important specifications, such as display color-gamut volumes. Both types of models have important applications in HDR display technology.

Appearance Scales and Color-Gamut Volume

At fixed adaptation levels, color appearance models find their most appropriate use in display evaluation. (Alternatively, they can be implemented in a dynamic way, where the adaptation parameters in the models are set in spatially and temporally localized ways.) Color spaces such as CIELAB and CIECAM02 can be used to accurately and appropriately represent display appearance.¹¹ Their luminance transfer functions (predictors of lightness and brightness) have been shown to extend smoothly to HDR content and levels above diffuse white, or $L^* = 100$. There is no limitation in using these spaces in such situations, and it is incorrect to assume that they only apply to reflecting materials. In fact, they have been developed and rigorously tested using self-luminous and HDR displays. Recently, researchers are looking to fine-tune the brightness and lightness functions and develop HDR versions of these spaces and scales.⁴

Fig. 4 illustrates an application of such appearance spaces via simulation of a simple change in display technology: the difference between an RGB display and an RGBW display. The visual differences (Fig. 4) are fully predicted and quantified in appearance spaces such as CIELAB and CIECAM02. The initial image data were captured in HDR and then visually rendered into a more limited dynamic range for illustration purposes. The top image is a

simulation of a typical RGB display in which the maximum white luminance is the sum of the maximum luminance in each RGB channel. The lower image is a simplified simulation of an RGBW display in which half the luminance is introduced through an additional white channel. In the RGBW display, colorful and high-luminance regions cannot be reproduced. While both displays have the same luminance dynamic range and chromaticity gamut, the image appearance is quite different because the relationship between diffuse white and various image regions is significantly different. This illustrates why the triangular area of chromaticities (on an xy or $u'v'$ diagram) that can be produced by a display does not define the color gamut. Color is a three-dimensional (at least) perception, and the relationship between luminance and chromaticity must be specified in an appearance space.

Fig. 5 shows color-gamut volumes in the CIELAB color space for both RGB and RGBW displays. First, note that there is no difficulty expressing meaningful CIELAB values above $L^* = 100$ for highlight regions of the display, and second, note how this 3D gamut illustrates the lost colorfulness of bright areas in the RGBW display (recall that the chromaticity gamuts, such as RGB primaries, are identical for the two displays.) Masaoka et al. recently explored the ramifications of various methods for measuring and specifying color-gamut volumes for HDR displays.¹²

Color-gamut volume is a metric used to describe a display's color reproduction capability. Traditionally, the 2D area coverage in chromaticity diagram, xy or $u'v'$, was used. Because at least three dimensions are needed to describe color perception, it is more meaningful to compute the 3D color-gamut volume. Additionally, the 3D color-gamut volume should be computed in a relatively color uniform space (i.e., CIELAB or CIECAM02). Both gamuts in Fig. 5 are DCI/P3 primaries with 1000 cd/m² peak luminance level and 200 cd/m² diffuse white setting. The brownish gamut is the RGBW system, where the R+G+B = 500 cd/m², and the white optical channel provides another 500 cd/m². The greenish gamut is the traditional RGB system, where R+G+B can achieve 1000 cd/m² peak luminance.

Conclusions

Studying color perception and trying to model it mathematically for practical colorimetry and image reproduction is a fascinating field. While new HDR and WCG display technologies provide new challenges, they are not anything that has not been considered previously when studying perception of the real world. Regardless, there remains much to learn, both about human perception and the technology we use to simulate the world.

There is a classic idiom that the devil is in the details. In HDR displays, it might be more appropriate to consider that the devil is in the shadows, the highlights, and on the gamut boundaries. These are the areas that require special attention (and sometimes benign neglect). It should also be noted that this devil is quite good at adapting to change, perceptual or otherwise. 

References

- Evans, R.M. (1974). *The Perception of Color*. John Wiley & Sons, New York.
- Hunter, R.S. & Harold, R.W. (1987). *The Measurement of Appearance*, 2nd Ed., Wiley, New York.
- Hunt, R.W.G. (2004). *The Reproduction of Colour*, 6th Ed. Wiley-IS&T Series in Imaging Science and Technology, Chichester, UK.
- Fairchild, M.D. (2013). *Color Appearance Models*, 3rd Ed. Wiley-IS&T Series in Imaging Science and Technology, Chichester, UK.
- Asano, Y., Fairchild, M.D., & Blondé, L. (2016). Individual colorimetric observer model. *PLOS ONE* 11, e014567.
- Asano, Y. & Fairchild, M.D. (2020). Categorical observers for metamerism. *Color Res. Appl.* 45, 576–585.

Res. Appl. 45, 576–585.

⁷ Xie, H., Rodríguez-Pardo, C.E., & Sharma, G. (2017). Multiobjective optimization for color display primary designs. *J. Electron. Imaging* 26, 063013.

⁸ CIE. (1989). *Special Metamerism Index: Change in Observer*, CIE Publ. No. 80, Central Bureau of the CIE, Vienna.

⁹ Long, D.L. & Fairchild, M.D. (2014). Modeling observer variability and metamerism failure in electronic color displays. *J. Imaging Sci. & Technol.* 58, 030402.

¹⁰ Xie, H., Farnand, S.P. & Murdoch, M.J. (2020). Observer metamerism in commercial displays. *J. Opt. Soc. Am. A* 37(4), A61–A69.

¹¹ Fairchild, M.D. & Chen, P.-H. (2011). Brightness, lightness, and specifying color in high-dynamic-range scenes and images. *SPIE/IS&T Electronic Imaging Conference*, Vol. 7867, San Francisco, 786700-1-786700-14.

¹² Masaoka, K., Jiang, F., Fairchild, M.D. & Heckaman, R.L. (2020). Analysis of color volume of multi-chromatic displays using gamut rings. *J. Society for Information Display* 28, 273–286.



Mark D. Fairchild is professor and founding head of the Integrated Sciences Academy at the Rochester Institute of Technology's College of Science and is director of the Program of Color Science and

Munsell Color Science Laboratory. He received his B.S. and M.S. degrees in imaging science from RIT and a Ph.D. degree in vision science from the University of Rochester. He is a Fellow of the Society for Imaging Science and Technology and the Optical Society of America.

Fu Jiang just defended his Ph.D. dissertation at Munsell Color Science Laboratory, Rochester Institute of Technology. Jiang's research focuses on HDR display perception in several aspects:



color gamut volume, "perfect" black on display, and the simultaneous dynamic range. Jiang is interested in psychophysics, applied vision, and computer graphics.



Hao Xie is a Ph.D. student in the Munsell Color Science Lab at the Rochester Institute of Technology. He received his B.Eng. degree in optical engineering from Zhejiang University, China,

and his M.S. degree in electrical engineering from the University of Rochester. His research interests include applied vision science and computational photography.