https://doi.org/10.2352/CIC.2023.31.1.23

Novel methods of brightness and saturation testing for high-dynamic-range images

Luke Hellwig[1,2], Dale Stolitzka[2], Mark D. Fairchild[1]; [1] Munsell Color Science Lab, Rochester Institute of Technology, Rochester, NY; [2] Samsung Display American Lab, San Jose, CA

Abstract

A psychophysical experiment was conducted in which observers compared the saturation and brightness between high-dynamic-range images that had been modulated in chroma and achromatic lightness. Models of brightness which account for the Helmholtz-Kohlrausch effect include both chromatic and achromatic inputs into brightness metrics, and this experiment was an exploration of whether these metrics could be expanded to images. The observers consistently judged saturation in agreement with the predictions of our color appearance modeling. However, some unexpected results and differences between observers in their methods for judging brightness indicates that further modeling, including spatial effects of color perception, need to be included to apply our model of the Helmholtz-Kohlrausch effect to images.

Introduction

The perceived brightness of a visual stimulus increases with its chromatic saturation, even if its luminance is held constant. This phenomenon, known as the Helmholtz-Kohlrausch effect, is due to the contribution of the chromatic visual pathways to brightness perception, whereas luminance is solely a measure of the achromatic visual channel. Recently, interest in the Helmholtz-Kohlrausch effect has been renewed by the spread of high-dynamic-range (HDR) imaging and display and the increased perceptual brightness offered by HDR, wide-color-gamut (WCG) displays [1, 2, 3, 4, 5].

A new model of the Helmholtz-Kohlrausch (H-K) effect based on a key experimental innovation and involving new techniques in modeling has recently been proposed.[6] The proposed H-K model uses the Commission Internationale de l'Eclairage recommended CIECAM16 color appearance model as a substrate [7, 8] with the inclusion of several recently proposed corrections to fundamental flaws in the CIECAM16's formulas for measures of achromatic and chromatic perceptual attributes [9]. In the proposed model, revised CIECAM16 (achromatic) lightness, J, and revised CIECAM16 chroma are combined by the following formula to predict H-K-compensated perceived lightness, J_{HK}:

$$J_{HK} = \sqrt{J^2 + 66C} \tag{1}$$

Perceived brightness is then calculated by scaling to the overall achromatic white reference signal in the scene. Figure 1 shows how Equation 1 can be used to predict lines of equal perceived brightness in color space.

In this study, we sought to test whether this model of brightness could be applied to predict how observers would perceive modulations in the color of HDR, WCG imagery, simulating how differences in display color gamut and peak luminance may effect ratings of brightness and saturation. Due to spatial effects, color appearance in images is more complex when compared to the simplified geometric stimuli used in the direct brightness matching studies for the H-K effect [10, 11, 12, 13]. However, directly applying this model to images serves as a useful first study in how the Helmholtz-Kohlrausch effect functions for images.

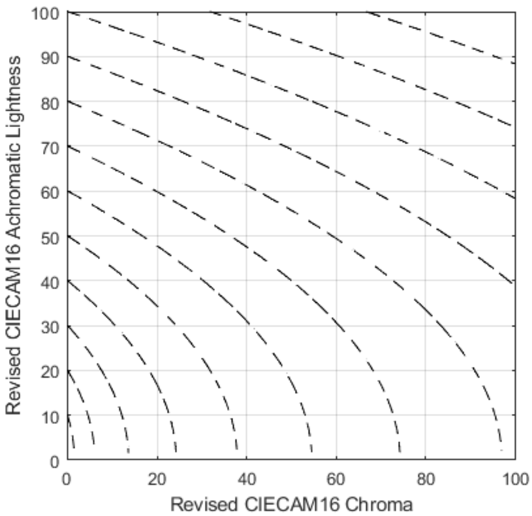

Figure 1. *Lines of equal perceived brightness (equal J_{HK} from Equation 1) in the achromatic lightness-chroma plane [9]. The lines slope downward because more colorful stimuli appear equally bright to less colorful stimuli with greater achromatic lightness.*

Methods

Image Modulation

A novel method of image color modulation based on color appearance modeling was used to generate stimuli for perceptual evaluation. Nine frames of HDR video were selected from a set of publicly available HDR videos developed previously by the Video Electronics Standards Association for visual experiments (see Figure 2). Each frame was mapped to the gamut of the Sony PVM-X3200 reference monitor used in the experiment with the diffuse white point set to D65 at 100 cd/m^2 and the peak luminance of the images clipped to the white point of the display, 790 cd/m^2. The CIE XYZ values of each image pixel were then used as input to the revised CIECAM16 color appearance model. The diffuse white point was used as the reference white point for the model with the degree of adaptation set to 1 and the surround specified as "dark."

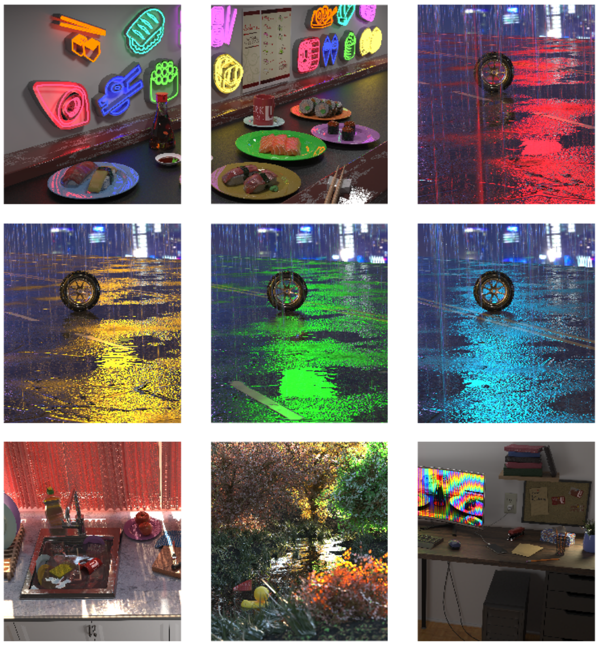

Figure 2. SDR approximation of HDR images used in the experiment. Images were used under a CC by 4.0 license from VESA and York University.

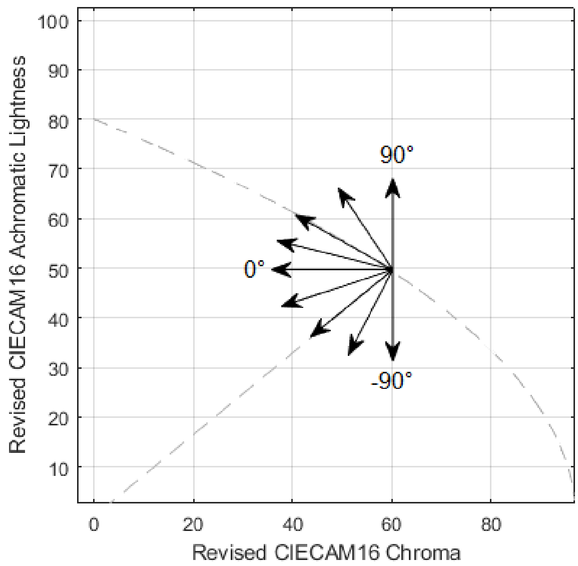

Figure 3. Example directions of image pixel color modulation in the achromatic lightness-chroma plane. The upper dashed line represent the equal-perceived-brightness line predicted by our model of the H-K effect (Equation 1) [14] and the lower dashed line represents the equal saturation line.

The color of each image pixel was then modulated in one of nine directions in the CIECAM16 lightness-chroma plane to generate nine new images for each start image. The directions of modulation were as follows (ordered counter-clockwise) (Figure 3):

1. The positive lightness direction (+90° in Figure 3).
2. The angle halfway between directions 1 and 3.
3. The direction of the equal brightness line as predicted by Equation 1.
4. The angle halfway between directions 3 and 5.
5. The negative chroma direction (0° in Figure 3).
6. The angle halfway between directions 5 and 7.
7. The direction of the equal saturation line (towards the origin of the lightness-chroma plane).
8. The angle halfway between directions 7 and 9.
9. The negative lightness direction (−90° in Figure 3).

Given that the derivative of the equal brightness line and the direction of the equal saturation line depend on achromatic lightness and chroma, these directions and the intermediate directions were calculated for each pixel. The magnitude of the modulation was set to either 5% or 10% of the chroma of each pixel, generating 18 modulated images for each reference image. No pixels with chroma below 10 were modulated so as to encourage the viewer to only use the chromatic areas of the image to judge brightness. This expectation was confirmed via post-experiment interviews with the participants.

Experimental Design

The experiment followed a two-alternative forced-choice method where observers were presented one of the 18 modulated images next to the corresponding reference image (separated by a vertical bar) and asked one of two questions:

- Which image is brighter?
- Which image is more saturated?

This was repeated in random order for all modulated images, all reference images, and both brightness and saturation, leading to 324 total judgements per observer. 18 observers participated in the experiment, 13 of which had previously received formal color science education and 5 of which were naïve. They were instructed in the definition of saturation by the demonstration of pages from the DIN color order system.

Stimuli were displayed on a Sony PVM-X3200 reference monitor in a dark room. Colorimetric characterization was performed with a Colorimetry Research CR-100 colorimeter. A $10 \times 10 \times 10$ test grid in Rec. 2020 RGB space was used to test the color accuracy of the display and optimize a 3×3 transform matrix, which improved the color accuracy of the display from 2.3 ΔE_{00} to 1.3 ΔE_{00}.

Results

General Trends

The direction of image pixel modulation in color space should predict the response of observers. When then angle of modulation is in the positive achromatic lightness dimension (90° in the notation from Figure 3), the expectation is that the observers will always perceive the modulated image as brighter and less saturated than the reference image. The converse is expected for the −90° dimension: the modulated image should now appear less bright but more saturated than the reference image. Figure 4 shows that these predictions hold true when our data is averaged across all observers, images, and modulation intensities.

As the modulation angle decreases in our notation (Figure 3), observers should be less likely to rate the modulated image as brighter, as can be seen in Figure 4. For modulated images

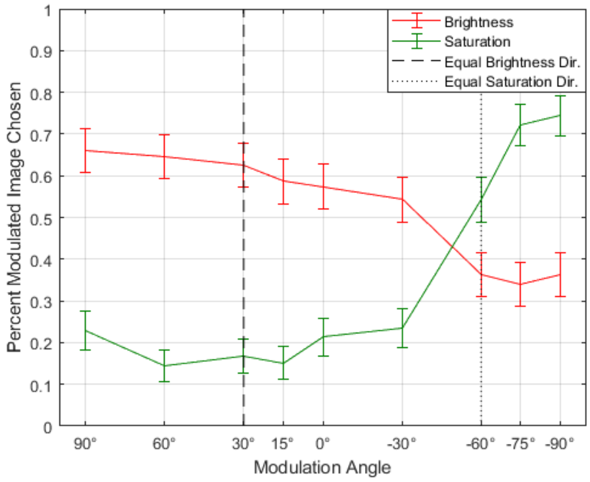

Figure 4. Percent of trials in which the modulated image was chosen as brighter or more saturated as a function of modulation angle, averaged across all observers and images.

between 90° and 0°, chroma is reduced in the modulated images and achromatic lightness is increased in the modulated images. In our model of H-K-compensated brightness, both chroma and achromatic lightness contribute to perceived brightness. Thus, according to the prediction of our model, reducing the chroma should make the image appear less bright to observers but increasing the achromatic lightness should make the image appear brighter to observers. So, there is an inherent brightness tradeoff in this quadrant. If the achromatic lightness is boosted much more than the chroma is decreased (e.g., modulation angle 2), then we would expect that the small loss in perceived brightness due to decreased chroma to be offset by the increase in achromatic lightness in the modulated image, which would make the modulated image still appear brighter than the reference image. Conversely, if the chroma is decreased much more than the achromatic lightness is boosted (e.g., modulation angle 4), we would expect that the decrease in chroma to cause the modulated image to appear less bright than the reference image, even though there was a small boost to the (physical) achromatic lightness of the image. The point at which the loss in chroma is exactly offset by a gain in achromatic lightness should be when observers were equally likely to choose either the modulated image or the reference image as brighter. Surprisingly, though, when the only modulation was to reduce the chroma (modulation angle of 0°), observers did not perceive the modulated images as less bright! Even when the achromatic lightness was slightly reduced along with the chroma (the first negative modulation angle), observers were equally likely to choose the modulated image as the reference image. Possible explanations for this unexpected result are discussed below.

Observers should be more likely to rate the modulated image as more saturated as the modulation angle decreases in our notation (Figure 3). Saturation is typically defined as chroma relative to lightness, so decreasing the chroma or increasing the lightness should cause the perceived saturation to decrease. Thus, for the positive angles of modulation where both chroma is decreased and lightness is increased, the perception of the observers that the reference image was more saturated agrees with our expectations.

In the case of negative angles of modulation, the chroma is still being decreased, which would cause saturation to decrease, but the lightness is now also being decreased, which would cause saturation to increase. The modulation angle at which the decrease in lightness offsets the decrease in chroma and observers are equally likely to rate either image as more saturated should fall along the line of equal saturation which points from the stimulus to the origin of the lightness-chroma plane. In our results, the average percentage of saturation ratings did cross the 50% point at the angle that pointed to the origin, in agreement with the expectations outlined above.

Differences between observers

The 18 observers who participated in the experiment were consistent in their judgements of saturation, with all but two following the trend from the average results. However, there was significantly more variation between observers in their ratings of brightness. The majority (12) of observers' results followed the average trend from Figure 4. However, several observers had results curves that were relatively flat or lacked a clear trend relative to modulation angle (Figure 5). Additionally, two observers appeared to conflate brightness and saturation, with their brightness results following the same trend as saturation (Figure 5). Interestingly, all five of the naïve observers fell into the first category of observer; the observers that conflated the meaning of brightness and saturation had all received formal color science training. These severe differences between observers that we observed are a cautionary tale for researchers of brightness. Given the fluid linguistic meaning of the term "brightness," we cannot assume that the interpretation of brightness by observers exactly aligns with the technical color science definition [15].

Discussion

The perceptual effect on saturation of our pixel-by-pixel modulation of images was well predicted by the direction of modulation in the revised CIECAM16 achromatic lightness-chroma plane. This indicates that our color appearance modeling pipeline was accurate in predicting the perceptual color attributes of image content and that observers were consistent in their ratings of saturation.

However, our results for observers' ratings of brightness were more mixed. Observers had differing interpretations of brightness or differing implementations in how they made brightness comparisons (Figure 5). Thus, brightness judgments may be less reliable in image psychophysics than saturation judgments.

Furthermore, the overall trend in observers' brightness comparisons did not follow the expected transition point predicted by the Helmholtz-Kohlrausch effect as the direction of image pixel color modulation in the lightness-chroma plane changed. This failure of our color appearance modeling pipeline to predict brightness judgments could be partially due to our exclusion of spatial effects in our calculations. For example, in many images, the background of the image was not modulated because its chroma fell below our chroma cutoff. So, the interaction between the constant background and the shifting image content could have created spatial brightness effects for which we did not account.

Spatial effects alone, though, would not be able to account

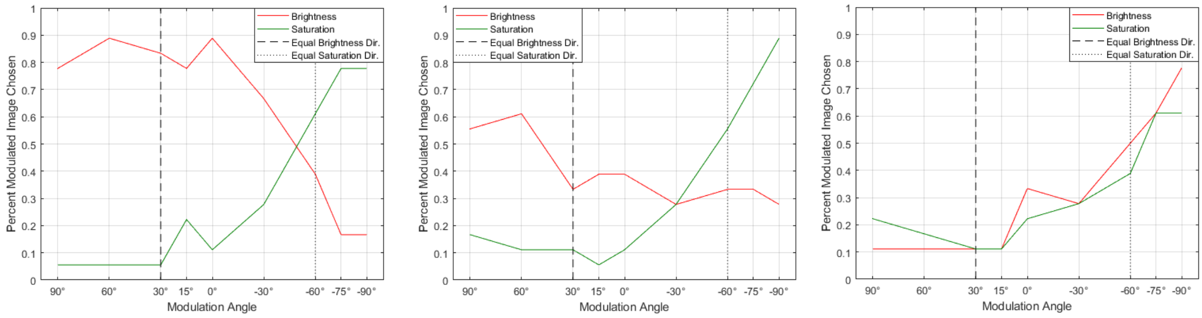

Figure 5. *Representative examples of brightness curves generated by individual observers in the experiment.*

for the most surprising result, which is that when image was modulated to only reduce chroma and keep achromatic lightness constant, observers on average rated the modulated image as brighter than the reference image. Even if the Helmholtz-Kohlrausch effect did not exist, then observers should still have only been equally likely to rate either image as brighter, but with the well-established validity of the Helmholtz-Kohlrausch effect (more colorful stimuli are brighter at equal luminance), this result is completely unexpected.

A failure of the achromatic lightness dimension in CIECAM16 to be completely isoluminant is one possible explanation for why the less colorful image was perceived as brighter by observers, for in this case changing the chroma could have also led to an increase in luminance that was not accounted for. Additionally, the image content could have played a role in this surprising result. In post-experiment interviews, observers focused on the brightest and most colorful elements in the images when making their judgments, which in six of the nine images were colored lights. It is possible that decreasing the chroma of the colored lights made them look whiter, which observers may have interpreted as brightness in the context of lighting. Perhaps the use of a color appearance model tailor-made for self-luminous stimuli, such as CAM18sl [16], may be more appropriate for this type of image content than CIECAM16.

The color modulations (Figure 3) performed in this study could approximate the differences between displays with different peak luminance or color gamut. As expected, increasing the simulated color gamut of an image by increasing its chroma led the observers to view the image as more saturated. While they did not necessarily view such images as brighter, another recent study indicated that wider color gamut and more saturated colors are more important to the brightness and vividness of HDR images than peak luminance [17]. Thus, further exploration of this topic is necessary to address the potential shortcomings in our modeling described above, including the incorporation of spatial effects on the perception of HDR content.

References

[1] H. Xie and M. D. Fairchild, "g_0 revisited as equally bright reference boundary," in *29th Color and Imaging Conference*, pp. 247–252, 2021.

[2] Y. Park, "Determining the color appearance of Helmholtz-Kohlrausch effect for self-emissive displays," *J. Inf. Disp.*, vol. 23, no. 3, pp. 235–242, 2022.

[3] L. Hellwig, D. Stolitzka, Y. Yi, and M. D. Fairchild, "75-1: Brightness and vividness of high dynamic range displayed imagery," in *SID 2022 Digest*, pp. 1009–1012, 2022.

[4] G. High, P. Green, and P. Nussbaum, "The Helmholtz-Kohlrausch effect on display-based light colours and simulated substrate colours," *Color Res. Appl.*, vol. 48, 2023.

[5] G. Seong, Y. Kwak, and H. Kim, "Determining the color appearance of Helmholtz-Kohlrausch effect for self-emissive displays," *Electronic Imaging*, vol. 35, no. 15, pp. 186–1–186–4, 2023.

[6] L. Hellwig and D. Stolitzka, "An advanced color model for evaluating new display technologies," *Inf. Disp.*, vol. 39, no. 3, pp. 11–15, 2023.

[7] C. Li, Z. Li, Z. Wang, Y. Xu, M. R. Luo, G. Cui, M. Melgosa, M. Brill, and M. Pointer, "Comprehensive color solutions: CAM16, CAT16, and CAM16UCS," *Color Research and Applications*, vol. 42, pp. 703–718, 2017.

[8] "CIE 248:2022 the CIE 2016 colour appearance model for colour management systems: CIECAM16," tech. rep., CIE, Vienna, 2022.

[9] L. Hellwig and M. D. Fairchild, "Brightness, lightness, colorfulness, and chroma in CIECAM02 and CAM16," *Color Res. Appl.*, vol. 47, no. 5, pp. 1083–1095, 2022.

[10] J. Lubin, *Vision Models for Target Detection and Recognition*, ch. A visual discrimination model for imaging system design and evaluation, pp. 245–283.

[11] X. M. Zhang and B. A. Wandell, "A spatial extension to CIELAB for digital color image reproduction," in *SID 1996 Digest*, pp. 731–734, 1996.

[12] M. D. Fairchild and G. M. Johnson, "iCAM framework for image appearance, differences, and quality," *Journal of Electronic Imaging*, vol. 13, no. 1, pp. 126–138, 2004.

[13] J. J. McCann, "Retinex at 50: color theory and spatial algorithms, a review," *Journal of Electronic Imaging*, vol. 26, no. 3, p. 031204, 2017.

[14] L. Hellwig, D. Stolitzka, and M. D. Fairchild, "The brightness of chromatic colors," *Color Res. Appl.*, 2023 [submitted].

[15] R. W. G. Hunt, "Colour terminology," *Color Res Appl*, vol. 3, no. 2, pp. 79–87, 1978.

[16] S. Hermans, K. A. G. Smet, and P. Hanselaer, "Color appearance model for self-luminous stimuli," *JOSA A*, vol. 35, no. 12, pp. 2000–2009, 2018.

[17] W. Jung, Y. Kwak, K. J. Kang, and M. Y. Lee, "Image-quality change by white-boost function of WRGB OLED display," *SID Digest*, vol. 54, pp. 1117–1120, 2023.