https://doi.org/10.2352/CIC.2022.30.1.07

Color appearance characterization of highlight stimuli in HDR scenes across a wide range of diffuse white luminance

Hongbing Wang, Minchen Wei*
Color, Imaging and Illumination Laboratory, The Hong Kong Polytechnic University
****minchen.wei@polyu.edu.hk***

Abstract

Various uniform color spaces and color appearance models were mainly developed for characterizing stimuli under a low dynamic range condition. Real scenes in daily life, however, are commonly high dynamic range (HDR), containing highlights with luminance beyond the diffuse white, whose color appearance characterization was never investigated in the past. This study was carefully designed to investigate the color appearance characterization of highlights in HDR scenes, covering extremely wide ranges of diffuse white luminance (up to 11000 cd/m^2), stimulus luminance (up to 49000 cd/m^2), stimulus chromaticities (reach Rec. 2020 gamut), and scene luminance contrast (up to 72045). The observers viewed two stimuli, including one highlight and one dark stimulus, in a viewing booth, and were asked to adjust the color appearance of another stimulus, so that the color differences between the adjusted stimulus to each of the other two stimuli appeared the same. The results clearly showed that none of the existing models, including the one (i.e., IC_tC_P) that was recently designed for HDR scenes, has a good performance. The models using a power function to characterize the non-linear compressive responses of the human visual system (i.e., CIELAB and IPT) had a slightly better performance. The findings provided some guidance for performing tone mapping and chroma/saturation adjustments, and clearly suggest the necessity to carry out further work to develop a better model for HDR scenes.

Introduction

Color characterization is critically important in a wide range of industries. Various methods have been developed for color characterization, including color appearance models (CAM) and uniform color spaces (UCS) [1]. A CAM aims to characterize and predict the color appearance of a stimulus under a certain viewing condition, with at least three color attributes (i.e., hue, chroma, and lightness), and a UCS aims to characterize the perceptual difference between two stimuli. These models and spaces were generally developed based on three important mechanisms in the human visual system (i.e., chromatic adaptation, non-linear compressive response, and opponent responses) and the various perceptual data collected through psychophysical experiments. The psychophysical experiments, however, were mainly carried out using conventional displays and surface color samples with a clearly defined diffuse white point, which was either the color produced by the display with the RGB values of 255 or the color produced by a perfect reflector. Therefore, the luminance of the color stimuli were always lower than that of the white point, making the viewing condition low dynamic range (LDR).

The real scenes we experience every day, however, commonly contain highlights from specular reflections or self-luminous sources, with the stimulus luminance being significantly higher than the diffuse white luminance, making the viewing conditions high dynamic range (HDR). In addition, the recent development of display technologies has allowed the rendering of highlights with the luminance beyond the display diffuse white point. Unfortunately, little effort has been made to investigate the color appearance of highlight stimuli in HDR scenes.

In 2011, two experiments [1,2] were carried out to investigate the perceived brightness of a D65 stimulus at different luminance levels, with the maximal stimulus luminance being around four times the diffuse white luminance (i.e., 997 cd/m^2). The data collected from the two experiments were used to develop hdr-CIELAB and hdr-IPT color spaces for characterizing stimuli in an HDR scene, with the power functions—$f(\omega) = \omega^n$—used in CIELAB and IPT replaced with hyperbolic functions—$f(\omega) = \omega^n/(a+ \omega^n)$. It is also believed that a hyperbolic function can better characterize the non-linear compressive responses in the human visual system, and it is widely used in other recently proposed models (e.g., CIECAM02, CAM02-UCS, and CIECAM16).

Figure 1 compares the predicted lightness of a D65 stimulus at various relative luminance levels (i.e., Y/Y_w) under a D65 viewing condition using different models, which minimizes the effect of chromatic adaptation. It is clear that the power functions and hyperbolic functions introduce little difference when the stimulus has a luminance lower than the diffuse white luminance (i.e., Y/Y_w<100), as shown in Fig 1(a), but introduce significant differences when the stimulus luminance is higher than the diffuse white luminance (i.e., Y/Y_w>100), as shown in Fig 1(b). The higher the stimulus luminance, the larger the differences.

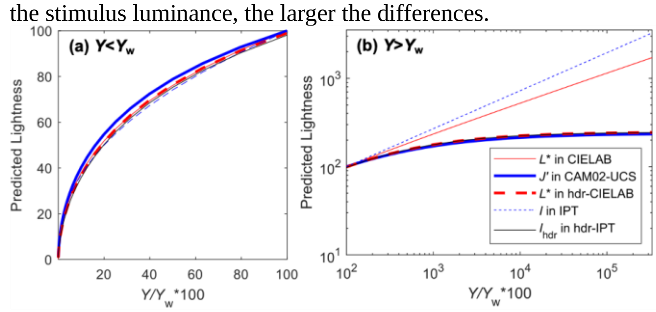

Figure 1. Comparison of the lightness of a D65 stimulus at various relative luminance levels (i.e., Y/Y_W) under a D65 viewing condition using different models. (a) when the stimulus luminance is lower than the diffuse white luminance (i.e., Y/Y_W×100<100); (b) when the stimulus luminance is higher than the diffuse white luminance (i.e., Y/Y_W×100>100).

Recently, two models—IC_tC_P [3] and $J_za_zb_z$ [4]—were specifically proposed for stimuli in HDR scenes. They use a Perceptual Quantizer (PQ) curve [5], which was developed based on the Barten contrast sensitivity function [6], to derive the perceptual attributes, aiming to quantify the threshold of difference that is perceivable at a certain luminance level, which is critically important in image encoding and bit depth. This is, however, completely different from the above models, which aim to quantify the magnitude of perceptual differences between stimuli.

With the above in mind, this study was designed to investigate the color appearance of achromatic and chromatic stimuli under a wide range of diffuse white luminance levels (i.e., 10 to 11000 cd/m^2), with the stimulus covering very wide ranges of luminance above the diffuse white luminance (i.e., 100 to 49000 cd/m^2) and chromaticities.

Method

Apparatus and experiment setup

A viewing booth, with dimensions of 70 cm (width) × 45 cm (height) × 60 cm (depth), was built for this study. The interiors of the booth were painted with Munsell N7 spectrally neutral paint. A four-channel spectrally tunable LED device was placed above the viewing booth to provide a uniform adapting field. The bottom of the front side was open, with the top portion covered to prevent observers from seeing the source of illumination.

Three 5 cm × 5 cm openings, with a 5 cm distance between edge to edge, were cut on the back wall, which were used to produce three stimuli viewed from the front side of the booth. In particular, a diffuse black sheet, with a reflectance around 4%, was attached on the back of the wall to cover the right opening to produce Stimulus 0. Two diffuse panels were attached on the back of the wall to cover the left and center openings, with each opening illuminated by a four-channel spectrally tunable LED device from the back to produce Stimulus 1 (center) and Stimulus 2 (left). The two LED devices were fixed on two tripods, so that the adjustment on one device only changed the corresponding stimulus.

During the experiment, the observer was seated in front of the viewing booth, with his or her chin fixed on a chin rest that was mounted just outside the viewing booth and centered on the opening. The viewing booth was placed on a frame, so that the three stimuli and the observer's eyes were on the same horizontal plane and each stimulus occupied a field of view (FOV) around 5°. Figure 2 shows the experiment setup.

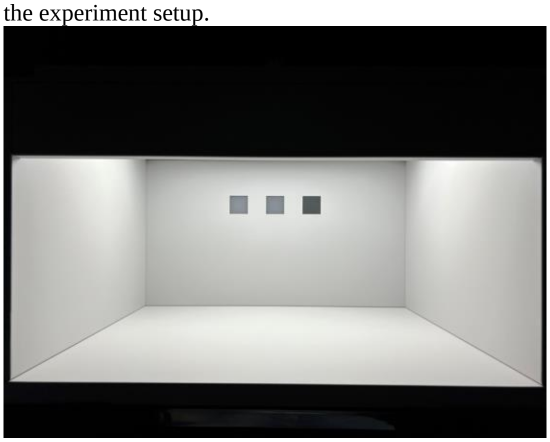

Figure 2. Photograph of the experiment setup, with the chin-rest removed. The three openings on the back wall are the three stimuli—Stimulus 0 (right), Stimulus 1 (center), and Stimulus 2 (left).

The three LED devices were connected in series and a control program was developed to control the intensity of each channel through a DMX controller.

Adapting conditions and stimuli

The intensities of the four channels (i.e., red, green, blue, and white) in the three LED devices can be adjusted individually with a bit depth of 16, with very good channel independence. A gain-offset-gamma model was developed for each LED device to derive the relationship between the input signal and the tristimulus values *XYZ* calculated using the CIE 1931 2° Color Matching Functions (CMFs).

The adapting condition was calibrated to have chromaticities around D65 with a wide range of luminance levels L_w using a PhotoResearch PR-655 spectroradiometer with a standard reflectance placed at the center opening. Stimulus 2 was also carefully calibrated to have a very wide range of chromaticities and luminance. In order to minimize the effect of observer metamerism, both the adapting illumination and Stimulus 2 were always produced using two color channels and the white channel in the LED device, so that they had broadband spectra.

For investigating achromatic stimuli, 15 levels of L_w (i.e., 10, 50, 100, 500, 1000, 2000, 3000, 5000, 5621, 6000, 7000, 8000, 9000, 10000, and 11000 cd/m^2) were calibrated. The chromaticities of Stimulus 2 were fixed around the D65 chromaticities. This was purposely designed to make the stimulus appear white regardless of the stimulus luminance and adapting luminance, which was based on the findings in a recent study investigating the white appearance of a stimulus in an HDR scene [7]. The luminance of Stimulus 2 L_2 was calibrated to a total of 151 levels under the various L_w, ranging between 500 and 49000 cd/m^2, with L_2 generally above L_w. Considering the luminance of Stimulus 0 L_0, the luminance ratio of the scene (i.e., L_2/L_0) ranged between 15 and 72045.

For investigating chromatic stimuli, five levels of L_w (i.e., 50, 100, 1000, 5000, and 10000 cd/m^2) were calibrated. The chromaticities of Stimulus 2 were calibrated to cover three hues (i.e., red, green, and blue) and different levels of saturation. Based on the color gamut of the spectrally tunable LED device, ten chromaticities were designed, with three for red stimuli, three for green stimuli, and four for blue stimuli. The luminance of Stimulus 2 was calibrated to cover a range as large as possible based on the chromaticities of Stimulus 2 and L_w, ranging between 90 and 38562 cd/m^2. In total, 200 Stimulus 2 were calibrated with different combinations of luminance and chromaticities. Similarly, the luminance ratio of the scene (i.e., L_2/L_0) ranged between 12 and 13960.

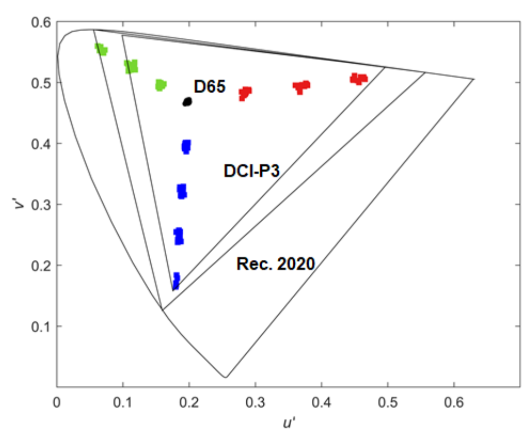

Figure 3. Measured chromaticities of Stimulus 2 under the different adapting conditions, with the achromatic stimuli shown in black and chromatic stimuli shown in red, green, and blue, respectively, together with the DCI-P3 and Rec. 2020 gamuts.

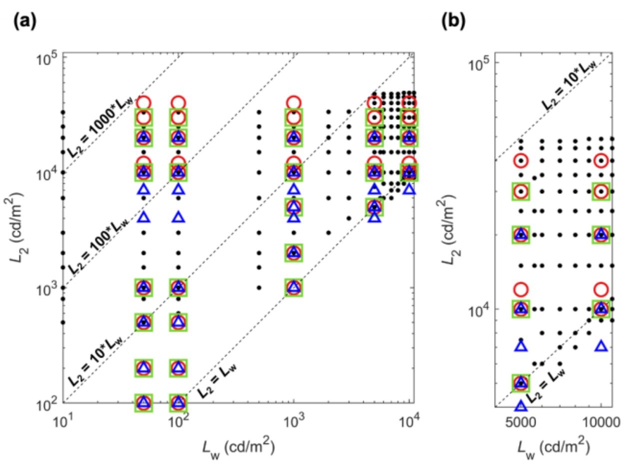

Figure 4. Combinations of L_2 and L_w, which were measured using the spectroradiometer under the corresponding adapting condition, with the achromatic stimuli shown in black dots and chromatic stimuli shown in different colors. (a) All combinations under all the adapting conditions; (b) A close-up in the region with L_w of 5000 and 10000 cd/m^2.

Figure 3 shows the measured chromaticities of Stimulus 2, together with the DCI-P3 and Rec. 2020 gamuts in the CIE 1976 *u*'*v*' chromaticity diagram; Figure 4 shows the combinations of L_2 and L_w.

Observers

A total of 30 observers (17 males and 13 females) between 24 to 34 years of age (mean = 27.17, std. dev. = 2.54) completed the experiments. All the observers had a normal color vision as tested using the Ishihara Color Vision Test.

Experimental procedure

Upon arrival, the observer completed the general information survey and the Ishihara Color Vision Test. Then the general illumination in the experiment space was switched off, and the observer was seated in front of the viewing booth, with his or her chin being fixed on the chin-rest.

Under each adapting luminance level, the observer was asked to look into the viewing booth for around 90 seconds for adaptation, then the observer was asked to adjust the color appearance of Stimulus 1, so that it appeared in the middle of Stimulus 0 and Stimulus 2, with the color difference between Stimulus 1 and Stimulus 0 being the same as that between Stimulus 1 and Stimulus 2.

In the experiment session when Stimulus 2 was achromatic, the chromaticities of Stimulus 1 were fixed around the D65 chromaticities and the observer was only allowed to adjust the luminance of Stimulus 1 using two keys on a remote controller, with each adjustment corresponding to 0.5% of L_2. The initial luminance of Stimulus 1 was set to zero or L_2. In contrast, in the experiment session when Stimulus 2 was chromatic, the observer was allowed to adjust the color appearance of Stimulus 1 along the L^*, C^*, and h axes in the CIELAB (or L^*C^*h) color space using six keys on a remote controller, with each adjustment corresponding to 0.5° hue angle and 0.5% of L^* and C^* of Stimulus 2. Stimulus 1 was initially set to be the same as Stimulus 2.

The observer was allowed to take as much time as he or she needed. Once the observer was satisfied with the adjusted color appearance of Stimulus 1, he or she pressed a key to confirm the adjustment, with the signals of the LED devices being recorded, and Stimulus 2 and Stimulus 1 automatically switched to the next setting. The order of Stimulus 2 under each adapting condition was randomized, and the order of the adapting luminance was also randomized. The adjustments under some Stimulus 2 settings were repeated for evaluating the intra-observer variations.

Result and discussion

Data analyses

After the experiment, the spectral power distributions (SPDs) of Stimulus 1 adjusted by the observers were measured by sending the recorded signals under the corresponding adapting condition.

In general, the inter- and intra-observer variations, as characterized using the mean difference from the mean values, were similar to past studies. Also, the two initial luminance of Stimulus 1 (i.e., $L_1 = L_2$ and $L_1 = 0$) did not introduce significant differences. Therefore, the experiment results were believed to be reliable.

Luminance and chromaticities of Stimulus 1 adjusted by observers

The luminance and chromaticities of Stimulus 1 adjusted by the observers were calculated based on the measured SPDs with the CIE 1931 2° CMFs. Figure 5 shows the average chromaticities of the adjusted Stimulus 1 for each Stimulus 2 under the corresponding adapting condition in the CIE 1976 *u*'*v*' chromaticity diagram. It can be observed that the chromaticities of Stimulus 1 were generally between those of Stimulus 2 and D65 when Stimulus 2 was in red or green. In contrast, when Stimulus 2 was in blue, the chromaticities of the adjusted Stimulus 1 were shifted clockwise. This suggested the poor hue linearity of the CIE 1976 *u*'*v*' chromaticity diagram.

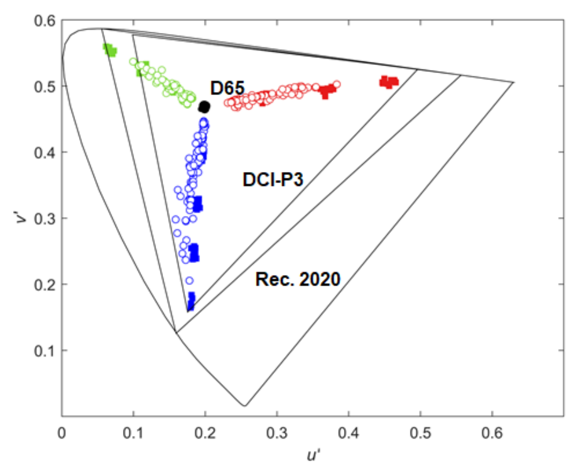

Figure 5. Average chromaticities of the adjusted Stimulus 1, together with those of Stimulus 2.

Figure 6 shows the scatter plot of the average luminance of the adjusted Stimulus 1 (L_1) versus the sum of the luminance of the corresponding Stimulus 2 and Stimulus 0 (L_2+L_0). If luminance level can characterize the perceptual attributes (i.e., perceived brightness) of a stimulus, the data points should be around the ½ line, especially for the achromatic stimuli, since the perceptual difference between Stimulus 1 and 0 should be the same as that between Stimulus 1 and 2. Such a result was not surprising.

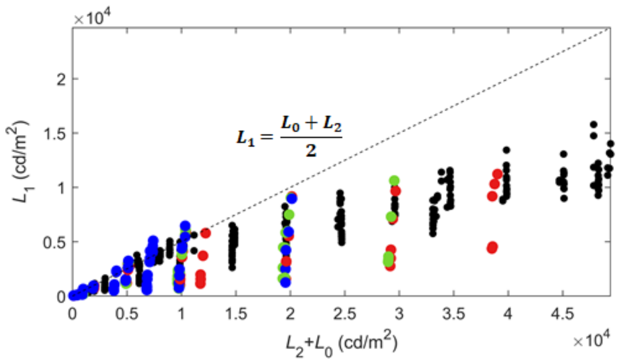

Figure 6. Scatter plot of the average luminance of the adjusted Stimulus 1 (L_1) versus the sum of luminance of the corresponding Stimulus 2 and Stimulus 0 (L_2+L_0), with the color of the data point representing the color of Stimulus 2.

Performance of the existing UCSs in characterizing perceived lightness, chroma, and hue

Since the observers were asked to adjust the color appearance of Stimulus 1 to make the perceptual color difference between Stimulus 1 and Stimulus 0 to be the same as that between Stimulus 1 and Stimulus 2, the analyses below mainly focus on the performance of various UCSs.

In particular, for a certain UCS, the perceived lightness of Stimulus 1 is characterized using the average lightness values of Stimulus 2 and Stimulus 0 that are calculated based on the luminance of Stimulus 2 and Stimulus 0. Figure 7 shows the scatter plot of the lightness values of Stimulus 0, Stimulus 1, and the perceived lightness of Stimulus 2 versus the (relative) luminance levels of the stimuli in six UCSs. In addition, the dotted lines represent the predicted lightness of a stimulus at various luminance levels. If a UCS is able to accurately characterize the perceived lightness of the stimuli in the experiment, the perceived and predicted lightness should be similar for Stimulus 1, with the data points of Stimulus 1 being close to the dotted line. Table 1 summarizes the minimum, maximum, and mean values of the differences between the perceived and predicted lightness of Stimulus 1. It can be observed that CIELAB and IPT, the two UCSs using a power function to characterize the non-linear compressive response in the human visual system, had better performance.

Table 1. Summary of the performance of various UCSs in characterizing the perceived lightness of stimuli, in terms of the minimum, maximum, and mean of the differences between the perceived and predicted lightness of Stimulus 1.

UCS	Min	Max	Mean	Type of non-linear compression
CIELAB	74%	137%	111%	Power
CAM02-UCS	107%	180%	133%	Hyperbolic
hdr-CIELAB	107%	179%	131%	Hyperbolic
IPT	58%	133%	103%	Power
hdr-IPT	107%	180%	131%	Hyperbolic
IC_tC_p	106%	158%	119%	PQ

Figure 8 shows the predicted chroma of Stimulus 1 C_1^* that were calculated based on the measured SPDs versus the perceived chroma of Stimulus 1, which is characterized as the average of the chroma of Stimulus 0 and Stimulus 2 $(C_0^* + C_2^*) / 2$ in the six UCSs For a UCS that can accurately characterize the perceived chroma of the stimuli, the predicted and perceived chroma should be similar, with the data points distributed around the diagonal line. It can be observed that none of these six UCSs have very good performance in characterizing the chroma of stimuli in such HDR scenes. In comparison to the other UCSs, CIELAB and IPT had slightly better performance, with the data points forming a linear trendline.

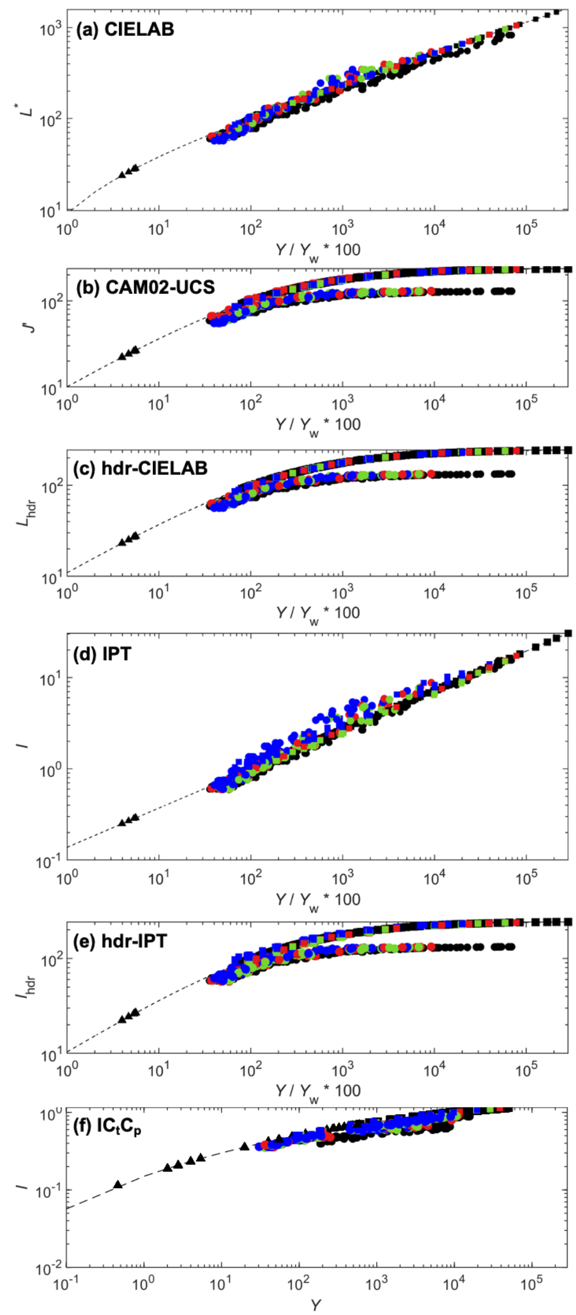

Figure 7. Lightness of Stimulus 2 (labeled using squares) and Stimulus 0 (labeled using triangles), and perceived lightness of Stimulus 1 (labeled using circles) versus the luminance of the stimuli in the six UCSs, with the color of the data point representing the color of Stimulus 2. The perceived lightness of Stimulus 1 are calculated using the average of lightness of Stimulus 2 and Stimulus 0. The dotted lines represent the predicted lightness of a stimulus at various luminance levels. The closer the data points of Stimulus 1 to the dotted line, the better the performance of the UCS. (a) CIELAB; (b) CAM02-UCS; (c) hdr-CIELAB; (d) IPT; (e) hdr-IPT; (f) IC_tC_p.

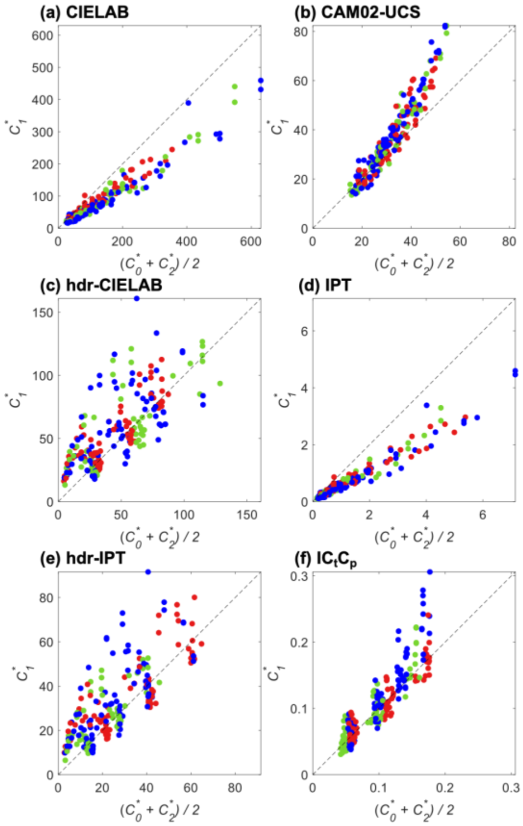

Figure 8. Predicted chroma of Stimulus 1 C_1^* that are calculated using the measured SPDs versus perceived chroma of Stimulus 1 that are characterized as the average of the chroma of Stimulus 0 and Stimulus 2 (C_0^* + C_2) / 2 in different UCSs, with the color of the data point labeling the hue of Stimulus 2. The closer the data points to the diagonal line, the better the performance of the corresponding UCSs. (a) CIELAB; (b) CAM02-UCS; (c) hdr-CIELAB; (d) IPT; (e) hdr-IPT; (f) IC_tC_p.*

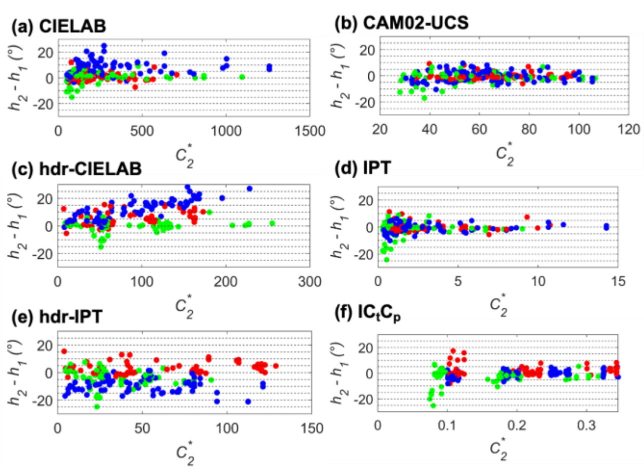

Figure 9. Hue angle differences between Stimulus 2 and Stimulus 1 h_2-h_1 versus the chroma of Stimulus 2 C_2 in the six UCSs. (a) CIELAB; (b) CAM02-UCS; (c) hdr-CIELAB; (d) IPT; (e) hdr-IPT; (f) IC_tC_p.*

Figure 9 shows the hue angle difference between Stimulus 2 and Stimulus 1 h_2-h_1 versus the chroma of Stimulus 2 C_2^*. Since the observers were asked to adjust the color appearance of Stimulus 1 to be in the middle of Stimulus 2 and Stimulus 0, the perceived hue of Stimulus 1 should be identical to that of Stimulus 2, with the hue angle difference close to 0. It can be found that CAM02-UCS has the best hue linearity. The poor performance of CIELAB in the blue hue corroborated the past studies.

Lightness-chroma relationship

Figure 10 shows the lightness and chroma of Stimulus 0, Stimulus 2, and Stimulus 1, with the lines showing the shifts from Stimulus 2 to Stimulus 1, in the six UCSs. The lightness and chroma values of Stimulus 1 are calculated using the measured SPDs. A UCS that can accurately characterize the perceptual color differences should result in both the lightness and chroma values of Stimulus 1 to be the average of those of Stimulus 0 and Stimulus 2, with the lines converging towards Stimulus 0 labeled using the black triangles. It can be observed that CIELAB and IPT have better performance than the other four UCSs, but the lines converge towards a point on the lightness axis with the lightness value greater than that of Stimulus 0. Such a result can be used in tone mapping adjustments, helping the adjustment of chroma (or saturation) after the adjustment of the lightness.

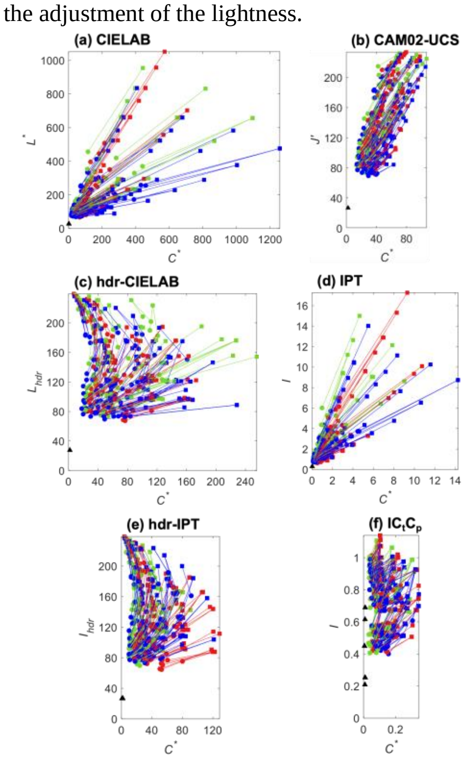

Figure 10. Lightness and chroma of Stimulus 0 (labeled with black triangles), Stimulus 2 (labeled with squares), and Stimulus 1 (labeled with circles), with the lines showing the shifts from Stimulus 2 to Stimulus 1 and the color of the data point labeling the hue of Stimulus 2. (a) CIELAB; (b) CAM02-UCS; (c) hdr-CIELAB; (d) IPT; (e) hdr-IPT; (f) IC_tC_p.

Conclusion

Two experiments were carefully designed to investigate the perceived color attributes of highlights in HDR scenes, with very wide ranges of diffuse white luminance, stimulus luminance, stimulus chromaticities, and scene luminance contrasts. The observer viewed two stimuli (i.e., a highlight stimulus and a dark stimulus), and was asked to adjust the color appearance of the other stimulus, so that the color differences between the adjusted stimulus and each of the other two stimuli appeared the same. The experiment results were then used to test the performance of six different UCSs.

It was found that none of the existing models can accurately characterize the color appearance of highlights in an HDR scene. In particular, the two models using a power function to characterize the non-linear compressive response of the human visual system (i.e., CIELAB and IPT) seemed to have a slightly better performance in characterizing the perceived lightness, chroma, and lightness-chroma relationship than those using a hyperbolic function or the PQ function. The experiment results can also be used for adjusting chroma (or saturation) in tone mapping adjustments.

The findings of this work clearly suggested the necessity to further investigate how to characterize the color appearance of highlights in an HDR scene, with an objective to develop better models (e.g., CAM and UCS).

References

[1] Fairchild MD. Color Appearance Models. 3rd ed. John Wiley & Sons (2013).

[2] Fairchild MD and Chen PH. Brightness, lightness, and specifying color in high-dynamic-range scenes and images. Electronic Imaging. California, USA. (2011).

[3] Dolby. What is ICtCp? White Paper (Version 7.1). (2016).

[4] Safdar M, Cui G, Kim YJ, and Luo MR. Perceptually uniform color space for image signals including high dynamic range and wide gamut. Optics Express. 25(13):15131-15151. (2017).

[5] SMPTE. SMPTE ST2084:2014 High dynamic range electro-optical transfer function of mastering reference display. (2014).

[6] Barten PGJ. Contrast sensitivity of the human eye and its effects on image quality. SPIE. WA, USA. (1999).

[7] Li Y and Wei M. Perception of white for stimuli with luminance beyond the diffuse white. Journal of Imaging Science and Technology. 65(5):050405. (2021).

Author Biography

Hongbing Wang is a PhD student at Color, Imaging, and Illumination Laboratory at The Hong Kong Polytechnic University. He received his master degree in at Shenzhen University in 2020. His research interests are color appearance and HDR imaging.

Minchen Wei is an associate professor, and the director of Color, Imaging, and Illumination Laboratory at The Hong Kong Polytechnic University. His research interests are color science, imaging science, and illumination technologies. He is currently an associate editor of the Journal of the Optical Society of America A, and Color Research & Application.