Research Article Vol. 29, No. 8 / 12 April 2021 / *Optics Express* 12292

Optics EXPRESS

Color mismatch and observer metamerism between conventional liquid crystal displays and organic light emitting diode displays

JIALU WU,[1] MINCHEN WEI,[1,*] YANG FU,[2] AND CHUNHUI CUI[2]

[1] *Color and Illumination Laboratory, The Hong Kong Polytechnic University, Kowloon, Hong Kong, China*
[2] *Huawei Technologies Co., Ltd., China*
[*] *minchen.wei@polyu.edu.hk*

Abstract: Organic light emitting diode (OLED) displays use red, green, and blue primaries with a higher saturation level to produce larger color gamuts than conventional liquid crystal displays (LCD). No past study, however, experimentally investigated how such a difference between these two display types causes color mismatch and observer metamerism using the most widely used color matching functions (CMFs)—the CIE 1931 2° CMFs—for color calibration and specification. In this study, 50 human observers performed color matching tasks for six color stimuli with a field-of-view of 4.77° between four test displays (i.e., one LCD and three OLED) and a reference OLED display. The color gamuts of the LCD and OLED displays were similar to the sRGB and P3 standard color gamuts. It was found the CIE 1931 2° CMFs cannot accurately characterize the color matches between the LCD and OLED displays, with different chromaticities required to produce matched color appearance. Particularly, when the stimuli had matched color appearance, the chromaticities of the stimuli produced by the LCD display were all shifted towards the $-u'+v'$ direction in the CIE 1976 $u'v'$ chromaticity diagram in comparison to those produced by the OLED display. This suggested that using the CIE 1931 2° CMFs for display calibration would cause the colors shown on OLED displays to have a yellow-green tint if those on LCD displays appear neutral. In addition, a larger degree of observer metamerism was found between the LCD and OLED displays, while little differences, in terms of color mismatch and observer metamerism, were found between the OLED displays. The CIE 2006 2° CMFs were found to have better performance than the CIE 1931 2°, 1964 10°, and 2006 10° CMFs, which could be partially due to the size of the stimulus used in the experiment.

1. Introduction

Displays produce different colors by mixing the three primaries (i.e., red, green, and blue) with different intensities, with the color gamut defined as the triangle enclosed by the chromaticities of the three primaries. The larger the color gamut, the more colors that can be produced by a display. Organic light emitting diode (OLED) displays are recently becoming more popular, replacing conventional liquid crystal displays (LCD) in different applications, such as smartphones. As OLEDs can produce more saturated colors, OLED displays have much larger color gamuts than conventional LCD displays. For example, conventional LCD displays typically have color gamuts similar to the sRGB standard color gamut, while OLED displays can have color gamuts similar to the P3 standard color gamut, with the sRGB gamut covering 33.5% and the P3 gamut covering 45.5% of the CIE 1931 chromaticity diagram.

In order to produce similar color appearance of an image on different displays, color consistency is critically important in display calibration and specification. It refers to not only the consistency between individual units using the same technology, but also the consistency between units using different technologies. Display calibration allows the displays to produce the colors at the target chromaticities at the given digital counts. For example, both OLED and LCD displays are

#418675 https://doi.org/10.1364/OE.418675
Journal © 2021 Received 22 Jan 2021; revised 30 Mar 2021; accepted 31 Mar 2021; published 6 Apr 2021

typically calibrated to have a white point (i.e., the digital counts of RGB channels set to 255) with the same chromaticities as D65, no matter whether P3 or sRGB color gamut is used. To derive the chromaticities of a color, the spectral power distribution (SPD) of the color is weighted by a set of color matching functions (CMFs)—$\bar{r}(\lambda)$, $\bar{g}(\lambda)$, $\bar{b}(\lambda)$, which describes the amounts of three lights (i.e., primary stimuli) needed to make the mixture of the three lights have a same appearance, in terms of color and brightness, as a monochromatic light with a fixed radiance at wavelength λ.

Currently, the set of CMFs that was recommended by the International Commission on Illumination (CIE) in 1931, known as the CIE 1931 2° CMFs, is the most widely used for display color calibration and specification. It was derived based on two sets of 2° bipartite color matching data collected by Wright [1,2] and Guild [3], with Wright using monochromatic lights at 460, 530, and 630 nm and Guild using broadband lights as the primary stimuli. Since the color matching data can be transformed between different sets of primary stimuli based on Grassmann's laws [4], these two datasets were transformed to an imaginary set of primary stimuli and resulted in the CIE 1931 2° CMFs [5]. In 1964, the CIE recommended another set of CMFs—CIE 1964 10° CMFs—for larger field of views (FOV) based on the 10° bipartite color matching data collected by Stiles and Burch [6] and Speranskaya [7]. The CIE 1931 2° CMFs are recommended to be used for an FOV between 1° and 4°, while the CIE 1964 10° CMFs are recommended to be used for an FOV beyond 4° [8]. In 1991, CIE established a technical committee TC1-36 to "establish a fundamental chromaticity diagram of which the coordinates correspond to physiological significant axes". In 2006, the committee derived a model to estimate the cone fundamentals for normal observers with an FOV ranging from 1° to 10° [9]. In 2015, the committee further derived the corresponding CMFs and chromaticity diagrams based on the cone fundamentals, with the CMFs for the 2° and 10° FOVs being directly provided for easy use in practice [10]. We refer to these two sets as the CIE 2006 2° and 10° CMFs. Thus, the CIE 1931 2°, 1964 10°, 2006 2°, and 2006 10° CMFs are the four standard CMFs.

Though Grassmann's laws suggest that stimuli having the same chromaticities should match in color appearance regardless of the spectral compositions and primary stimuli, psychophysical experiments found that the performance of CMFs was significantly affected by the primary stimuli. In other words, stimuli having the same chromaticities but different spectral compositions were found to have different color appearance [11,12] or stimuli having matched color appearance were found to have different chromaticities [13–17]. Moreover, the variations of cone fundamentals from person to person were found to cause observer metamerism that a pair of stimuli having matched color appearance to one observer have different appearance to another observer. Observer metamerism was found to vary with the primary stimuli [17–23], especially when narrow-band primaries were used. For example, Hu et al. [17] carried out a color matching experiment using 16 three-primary sets and found that the shift of the blue primary wavelength significantly affected the observer metamerism.

Though the issues of color mismatch and observer metamerism were not recently found, they become critically important to the display community due to the wider usage of OLED displays. Some users have recently reported on some online forums about the appearance of green tint on OLED displays of a specific smartphone brand, whose reason has not been identified. In addition, one study investigated the possible effect of using the CIE 1931 2° CMFs on color calibration and specification for OLED and LCD displays through mathematical calculations [11]. No study, however, has experimentally investigated the degrees of color mismatch and observer metamerism between OLED and LCD displays and how the standard CIE CMFs (i.e., CIE 1931 2° CMFs) can characterize the color matches between these two display types. In this study, the human observers adjusted the color appearance of the stimuli produced by four smartphone displays, including one LCD display and three OLED displays, to match the color appearance of the six color stimuli produced by another OLED display. The results of the color matching experiment allow us to quantify the performance of the four standard CIE CMFs, especially the CIE 1931 2° CMFs, in characterizing the color matches between the LCD and OLED displays.

2. Methods

The experiment was carried out in Color and Illumination Laboratory at The Hong Kong Polytechnic University. The protocols and procedures of the experiment were approved by the Institutional Review Board at The Hong Kong Polytechnic University.

2.1. Apparatus and setup

A viewing booth with dimensions of 60 cm (width) × 60 cm (depth) × 60 cm (height) was built for the experiment. The interiors of the booth were painted using Munsell N7 spectrally neutral paint. Two 5 cm × 5 cm square openings were cut at the center of the back panel, with a distance of 6 cm between the center edges of the two openings. Two tripods were placed behind the back panel, with two smartphones being fixed on the tripods and the displays being placed at the openings. A chin rest was mounted just outside the front of the viewing booth, so that an observer can simultaneously view the center of the displays of the two smartphones through the openings, with each stimulus occupying an FOV around 4.77°. Figure 1 illustrates the experiment setup.

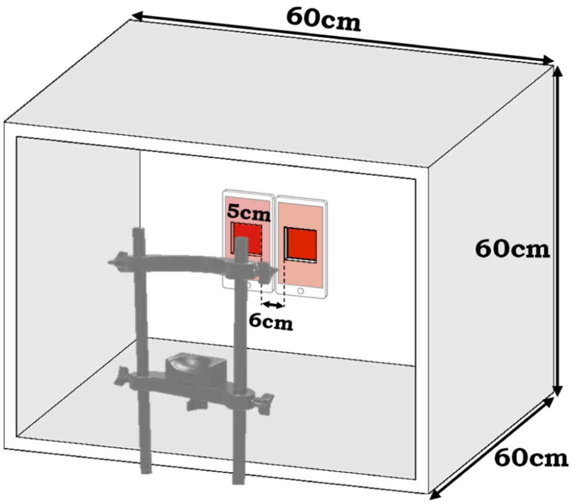

Fig. 1. Schematic illustration of the experiment setup. The chin rest was mounted just outside the booth, so that the viewing distance was 60 cm.

It is worthwhile to mention that though using the entire displays as the stimuli would be similar to how the displays are viewed in reality, we only used a small area at the center of the displays to produce the stimuli for avoiding the possible problems caused by the non-uniformity of the displays. In addition, though placing the two stimuli next to each other without any distance between them would make the color matching task easier, the 6 cm used in the current setup was the smallest distance for allowing the openings to present the center of the displays.

2.2. Displays and color stimuli

Five smart phone displays, including one LCD display and four OLED displays, were used in this experiment. These smartphones, whose dimensions were around 70 mm (width) × 140 mm (height) with small variations, were prototypes for testing display characteristics and algorithms, instead of commercially available products. The four OLED displays were carefully selected from eight OLED displays, so that they had slight differences in their primaries, in terms of peak wavelengths and shape of the SPDs. Though the differences among the OLED displays were small, noticeable color differences were observed by us. The OLED display having the smallest color gamut was selected as the reference display, so that the comparisons can be made between the LCD and OLED displays and also among the OLED displays. Figure 2 shows the color gamuts of the displays derived using the CIE 1931 2° CMFs, with the SPDs of the primaries shown in Fig. 3.

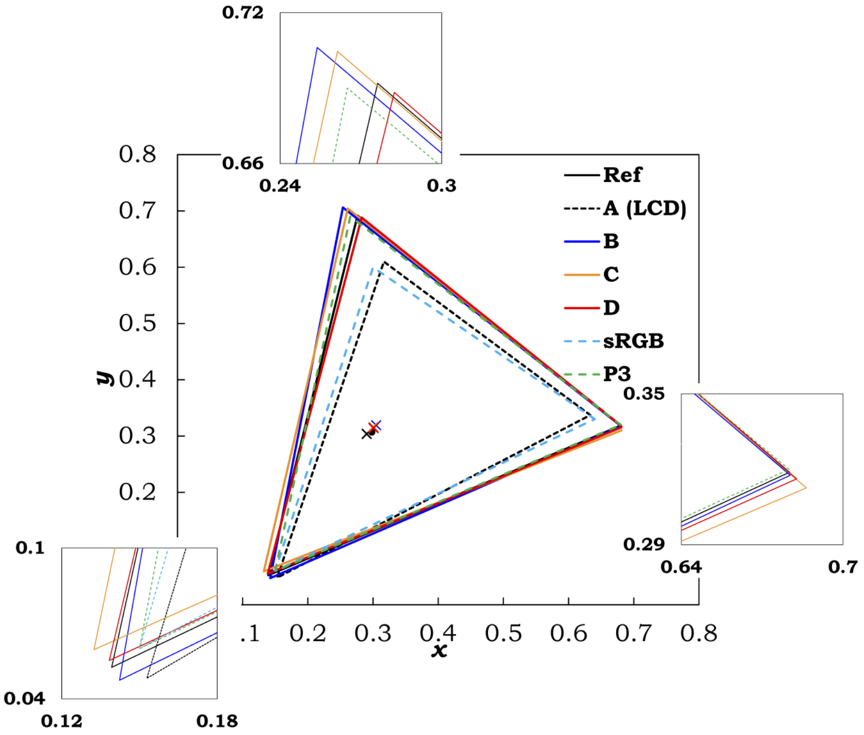

Fig. 2. Color gamuts and the white points of the five displays derived using the CIE 1931 CMFs, together with the standard sRGB and P3 color gamuts. Note: the white point of the reference display is labeled with a circle and those of the other four displays are labeled with crosses.

Six color stimuli were carefully selected to be presented on the reference display in the MacLeod - Boynton chromaticity diagram [24] for 2° FOV [10], which is a physiologically based chromaticity diagram, with the rectangular axes—l and s—formed using the cone sensitivity functions to better represent the relative cone excitation (i.e., $l = L/(L + M)$ and $s = S/(L + M)$). As shown in Fig. 4, all the stimuli were selected within the color gamuts of all the displays, so that they can be matched by human observers. Specifically, Stimuli 1 and 2 had similar s values but a large difference along the l axis, Stimuli 2 and 3 had similar l values but a large difference along the s axis, Stimuli 2, 4, and 5 simultaneously varied the s and l values. Stimulus 6 was selected to be a neutral color, and the CIE D70 illuminant was employed to calculate the target l and s values. Based on the luminance ranges that can be achieved by the displays at these six chromaticities, the luminance of the six stimuli were fixed at 93 cd/m^2, so that the possible influence of adaptation was minimized under the dark viewing condition.

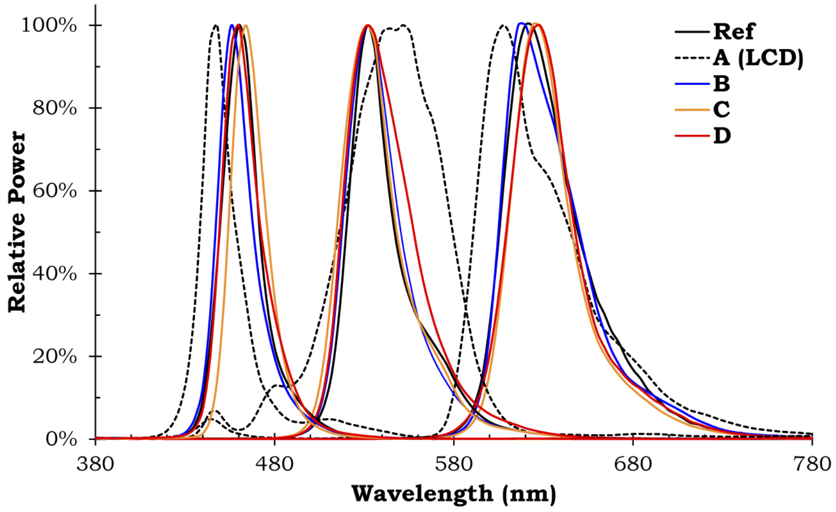

Fig. 3. Relative SPDs of the primaries of the five displays.

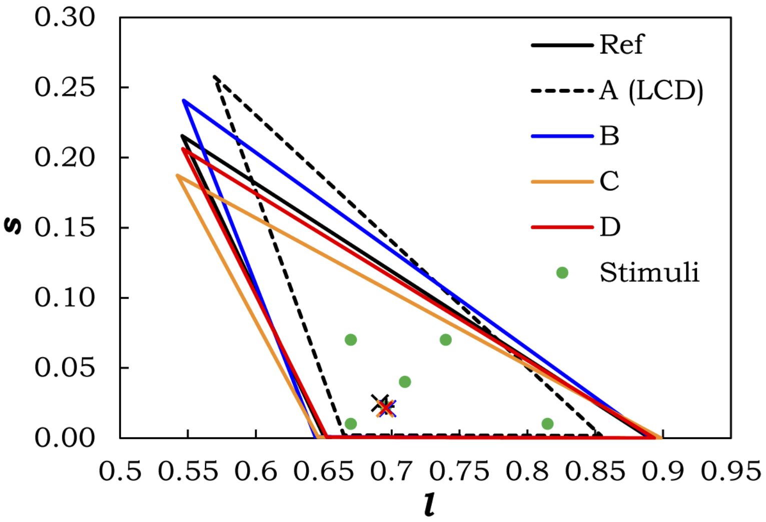

Fig. 4. Display color gamuts and white points in the MacLeod-Boynton chromaticity diagram for 2° FOV [10] and the chromaticities of the stimuli shown on the reference display for color matching.

2.3. Display calibration and control program

The six stimuli presented on the reference display were calibrated using a calibrated JETI Specbos 1811 spectroradiometer.

For the four displays that were used to match the color appearance of the stimuli shown on the reference display, a customized control program was developed. It allowed the observers to change the color appearance of the stimulus by adjusting its chromaticities along the *u*' and *v*' directions in the CIE 1976 *u*'*v*' chromaticity diagram, with + *u*' for red, -*u*' for green, +*v*' for yellow, and -*v*' for blue. Such an adjustment method was used in a past study [25] and was found understandable to naïve observers.

For each display, the spectroradiometer was used to measure the SPDs of 1000 colors, which were all possible combinations of 10 digital count levels for the RGB channels (i.e., 0, 28, 56, 85, 113, 141, 170, 198, 226, and 255). They were measured by placing the display at where they would be placed in the experiment and the spectroradiometer was placed at the observer's eye position. These data were used to build a 10×10×10 3D look-up-table (LUT) between XYZ and RGB combinations. Then, the color gamut at the luminance of 93 cd/m^2 was estimated by finding the stimuli with the luminance between 88.35 and 97.65 cd/m^2 (i.e., within ±5% of 93 cd/m^2), as illustrated in Fig. 5. To estimate the RGB combination for producing a target stimulus with chromaticities (u',v') within the gamut, the color having the smallest color difference ΔE_{ab} from the target stimulus was identified from the above 1000 colors. Its RGB combination, together with the all possible combinations of the adjacent 6 digital count levels (i.e., three above and three below) for the RGB channels, and their corresponding XYZ values were used to build another 3D LUT with an interval of 1 digital count for the RGB channels through a trilinear interpolation. The RGB combination having the smallest color difference ΔE_{ab} from the target stimulus was considered as the estimated RGB combination for producing the stimulus with chromaticities (u', v'). Such a procedure was used to estimate the RGB combinations for all the chromaticities within the color gamut for each display at the luminance of 93 cd/m^2, with a step of 0.0015 along the u' and v' directions. During the experiment, when the observer adjusted the color appearance of the stimulus by adjusting its chromaticities, the control program simply found the corresponding RGB combination to produce the color on the display. Though a gain-offset-gamma (GOG) model or a piecewise linear interpolation assuming constant chromaticity coordinates (PLCC) model is commonly used to develop control program to change the color of a display [17,25,26], we found these models did not have good performance for the OLED displays used in the experiment.

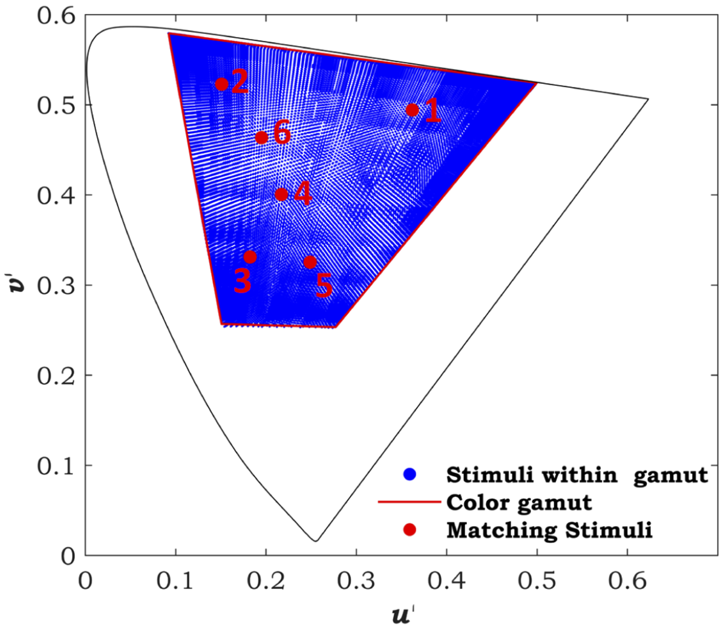

Fig. 5. Illustration of using the 3D LUT for identifying color gamut at the luminance around 93 cd/m^2 for Display B.

2.4. Observers

Fifty observers (14 females and 36 males) between 19 and 38 years of age (mean = 22, std. dev. = 3.01) participated in this experiment. All the observers had normal color vision, as tested using the Ishihara Color Vision Test.

2.5. Experimental procedure

Upon arrival, the observer completed the general information survey and the Ishihara Color Vision Test. Then he or she was escorted to the viewing booth, and the general illumination in the space was switched off. The experimenter explained how the observer can use the four arrow keys on the keyboard (right: +*u*', left: -*u*', up: +*v*', down: -*v*') to adjust the color appearance of a stimulus to make the color appearance of the two stimuli appear the same. The observer was instructed to keep his or her chin on the rest during the experiment. Before the beginning of the recorded trials, the observer completed two practice trials to get familiar with the control program. When performing the color match, the stimulus on the left was always produced by the reference display, while the stimulus on the right was produced by one of the four test displays. Though such a setup may introduce positional bias, we did not change the positions for display stability. For each test display, the observer made seven adjustments, with Stimulus 2 being adjusted twice to evaluate the intra-observer variations, in a random order. The order of the four test displays was randomized for each observer. In total, each observer made 28 adjustments, which took around one hour. All the displays were warmed up for at least 30 minutes for stabilization before the experiment. The displays were fixed on different tripods and two lines were marked on the back of the booth, so that the experimenter simply replaced the tripods and aligned the top and left edges of the displays to the markers when switching between different displays during the experiment.

3. Results and discussions

3.1. Verification of the control program accuracy

Though the observers adjusted the color appearance of the stimulus by changing its chromaticities using the customized control program, the RGB combinations that were used to produce the

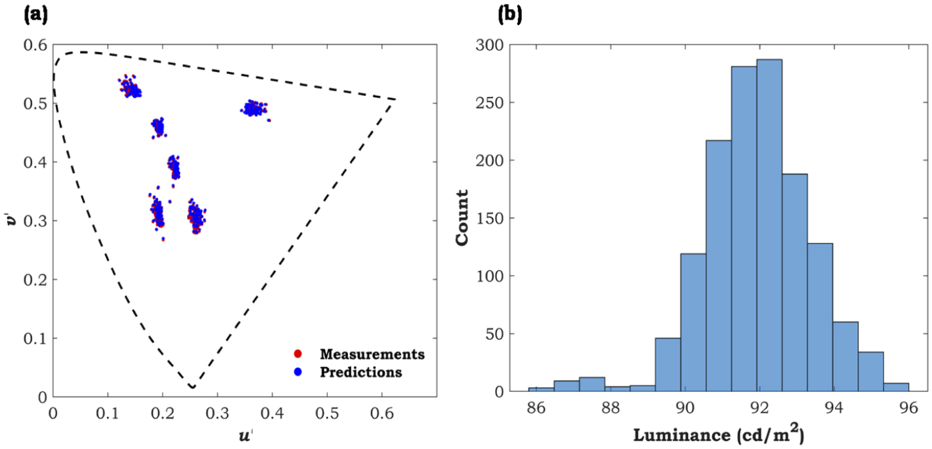

Fig. 6. Accuracy of the control program, in terms of chromaticities and luminance. (a) Chromaticities of the adjusted stimuli derived from the measured SPDs and the predictions using the control program in the CIE 1976 *u*'*v*' chromaticity diagram; (b) Histogram of luminance of the adjusted stimuli derived from the measured SPDs.

stimuli were recorded during the experiment. After the experiment, these RGB combinations were used to reproduce the stimuli on the corresponding displays, with the SPDs being measured using the spectroradiometer. The chromaticities (u', v') of all the adjusted stimuli derived from the measured SPDs and from the predictions using the control program are shown in Fig. 6(a), with $\Delta u'v'$ values between 0 and 0.0053 and the average $\Delta u'v'$ of 0.0017. Figure 6(b) shows the histogram of the luminance levels of the adjusted stimuli derived from the measured SPDs, with an average of 91.9 cd/m^2. Therefore, the control program was believed reliable.

3.2. Intra- and inter-observer variations

Both intra- and inter-observer variations were characterized using the mean color difference from the mean (MCDM) in the CIE 1976 $u'v'$ chromaticity diagram. Specifically, the intra-observer variations were characterized based on the average color difference between the chromaticities of the two matches for Stimulus 2 and the average chromaticities of the two matches. Figure 7 shows the chromaticities, together with the 95% confidence error ellipses, of the two matches performed by each observer for Stimulus 2 using the four test displays, with the histograms of the MCDM values being shown in Fig. 8. The average MCDM values were 0.0020, 0.0031, 0.0026, and 0.0024 for Display A, B, C, and D respectively, which were smaller than 0.004 units of $u'v'$ (i.e., ≈ 1 unit of just-noticeable color difference, JND) [4].

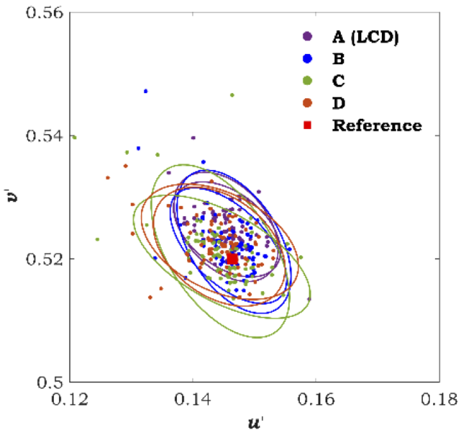

Fig. 7. Chromaticities, together with the 95% confidence error ellipses, of the two matches performed by each observer for Stimulus 2 using the four test displays.

The inter-observer variations were characterized by calculating the chromaticity differences between the 28 adjustments made by each observer and the 28 average adjustments made by the 50 observers (i.e., an average observer), with the MCDM values between 0.0027 and 0.0122 and the average MCDM value of 0.0059. Both the intra- and inter-observer variations, in terms of the MCDM values, were comparable to a recent color matching experiment [17], suggesting the high reliability of the experiment results.

3.3. Characterization of the color matching results using the four CMFs

The chromaticities of the stimuli adjusted by the observers using the four displays, together with the 95% confidence error ellipses, are plotted in Fig. 9 using the four CMFs, with Fig. 10 showing the chromaticity shifts by moving the chromaticities of the reference stimuli to the origin. It can be observed that most ellipses had similar orientations when different CMFs were used. Figure 11 shows the chromaticities of the stimuli shown on the reference display and the average chromaticities of the stimuli adjusted by the observers on the four displays, which were calculated

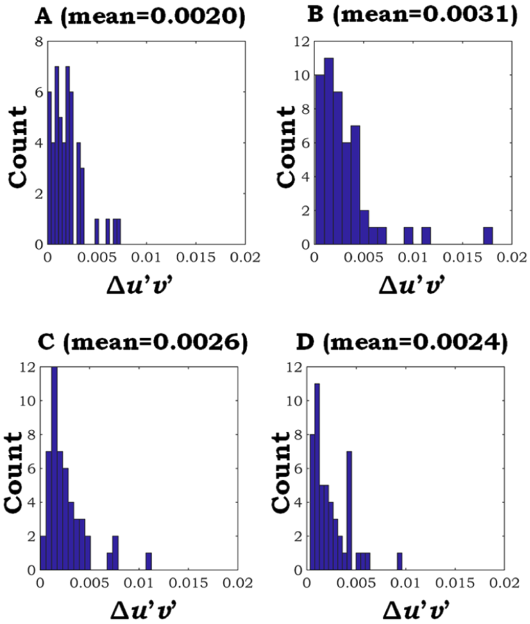

Fig. 8. Histogram of MCDM values of the intra-observer variations for each display.

using the four CMFs. The distances between the average chromaticities and the chromaticities of the corresponding reference stimulus and the areas of the ellipses are shown in Figs. 12 and 13 respectively. It is interesting to see that the longer axes of the ellipses generally oriented towards the center of the chromaticity diagram, and the ellipses of Stimulus 6 (white color) were closest to circles. Also, the chromaticities derived using the two 10° CMFs were close, and the effect of the four CMFs on the LCD display seemed to be smaller than that on the OLED displays.

When using the CIE 1931 2° CMFs, the most widely used CMFs for display calibration, Display A (LCD) had the largest chromaticity distances and ellipses areas compared to the other three OLED displays. The chromaticity distances of the six stimuli were all greater than 0.004 units of *u*'*v*' (i.e., 1 JND) [4], suggesting obvious color differences. In addition, all the chromaticities of Display A (LCD) shifted towards the -*u*'+*v*' direction, in comparison to the stimuli shown on the reference OLED display, for producing matched appearance between the two displays. This suggests that using the CIE 1931 2° CMFs to calibrate Display A (LCD) and the reference OLED display to the same chromaticities will introduce serious color mismatches. Specifically, a stimulus that appears neutral on an LCD display will appear to have a yellow-green tint on an OLED display with the same chromaticities, as its chromaticities are shifted towards the -*u*'+*v*' direction from those producing a matched color appearance. In contrast, a stimulus that appears neutral on an OLED display will appear to have a pinkish tint on an LCD display with the same chromaticities, as its chromaticities are shifted towards the + *u*'-*v*' direction from those producing a matched color appearance. This seems to also corroborate the findings in a recent study that color mismatches are more likely to appear pinkish or greenish [27]. The shifts did not only happen to neutral colors, but also to other colors, as shown in Fig. 11. For the other three OLED displays (i.e., Display B, C, and D), most chromaticity distances were below 0.004 units of *u*'*v*' (i.e., 1 JND).

It is also worthwhile to point out that the above comparisons between the LCD display (Display A) and the OLED display (Reference) were especially meaningful by comparing these calculated chromaticity distances with the MCDM values for the intra-observer variations. The chromaticity

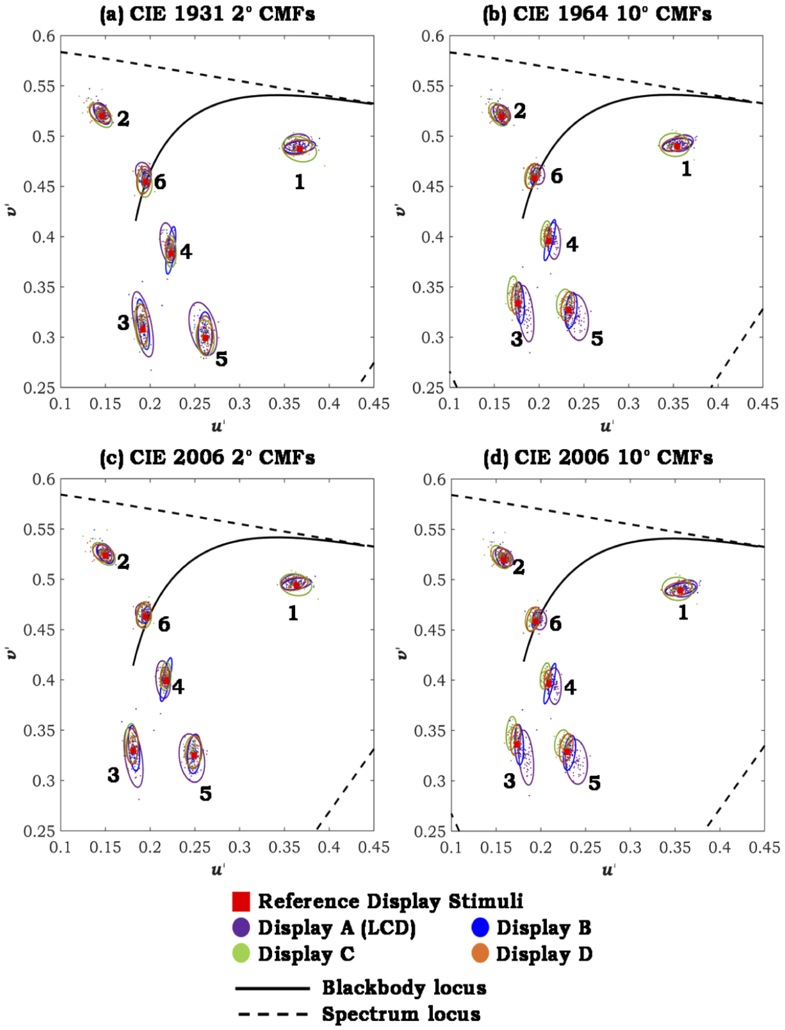

Fig. 9. Chromaticities, together with the 95% confidence error ellipses, of the stimuli adjusted by the 50 observers using the four displays to match the color appearance of the stimuli shown on the reference display calculated using the four CMFs. (a) CIE 1931 2° CMFs; (b) CIE 1964 10° CMFs; (c) CIE 2006 2° CMFs; (d) CIE 2006 10° CMFs. (Note: the numbers shown on the figures represent Stimulus 1 to 6).

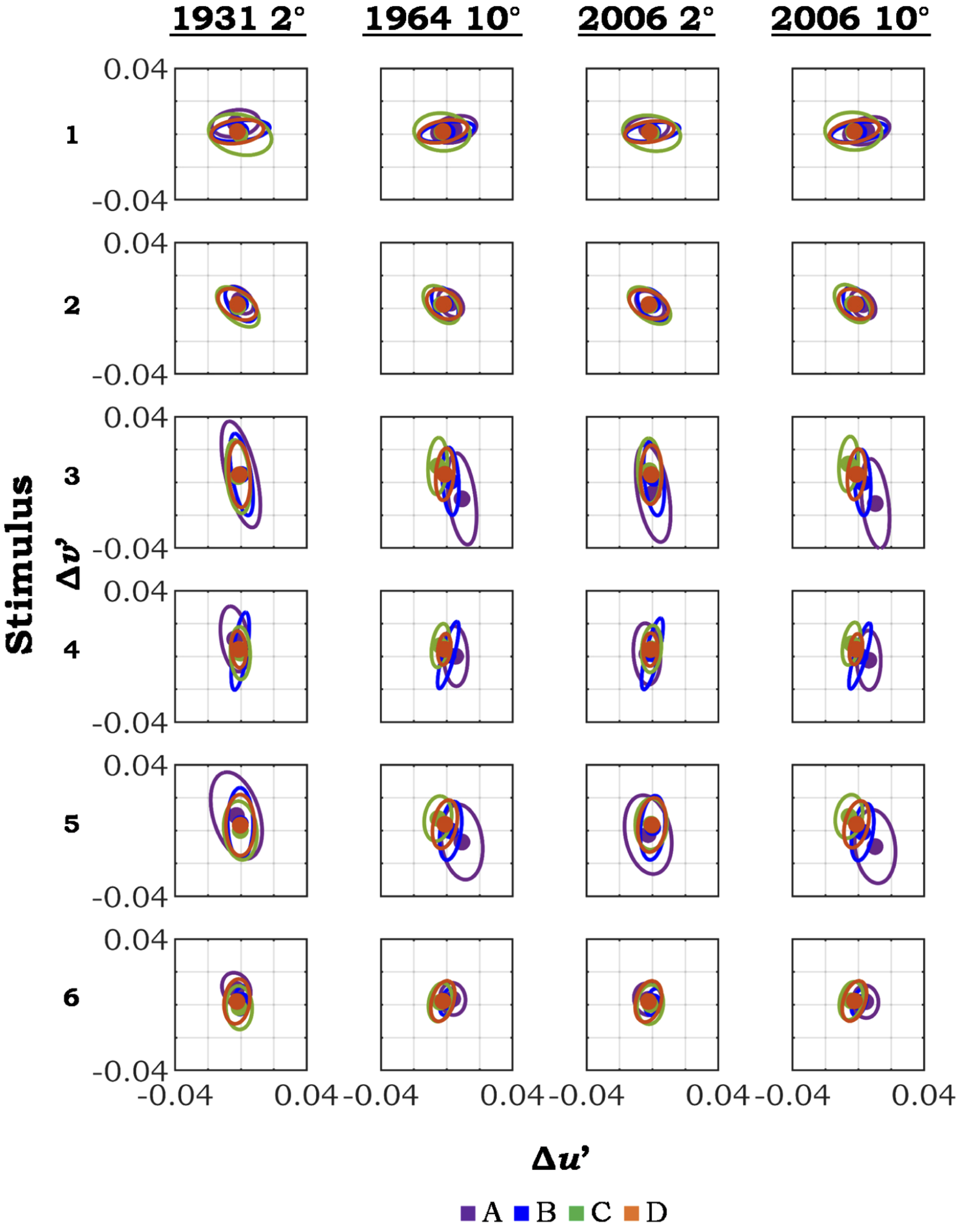

Fig. 10. 95% confidence error ellipses of the chromaticities adjusted by the observers, in relative to the chromaticities of the corresponding reference stimuli, plotted using the four CMFs.

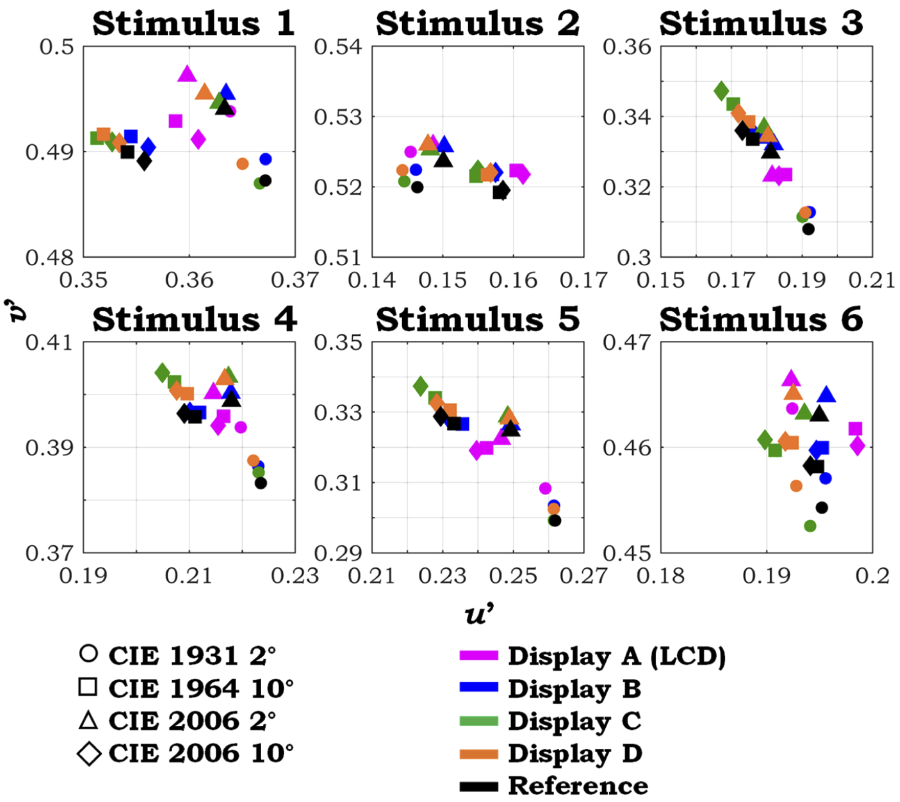

Fig. 11. Average chromaticities of the stimuli adjusted by the observers using the four displays and the corresponding stimuli shown on the reference display calculated using the four CMFs. (note: for illustration purposes, the axis ranges are different, but each gird corresponds to 0.01 unit in *u*' and *v*')

differences between the OLED displays (i.e., Display B, C, and D) and the reference display were generally similar to the intra-observer variations, while the difference between the LCD display and the reference display was significantly larger than the intra-observer variation. The ratio of the chromaticity distance of Stimulus 2 to the MCDM value was 2.55, 0.81, 0.81, and 1.33 for Display A, B, C, and D respectively, with the chromaticity distances being 0.0051. 0.0025, 0.0021, 0.0032 and the MCDM values being 0.0020, 0.0031, 0.0026, 0.0024 for the four displays.

3.4. *Comparisons among the four CMFs*

As shown in Fig. 11, the average chromaticities of the stimuli produced by the three OLED displays were very close to each other when the two sets of 2° CMFs were used. In contrast, the average chromaticities, especially for Stimulus 3 and 5, had large differences among the three OLED displays. To quantitatively compare the performance of the CMFs, both chromaticity distance between the average adjusted stimulus and reference stimulus and ellipse size fitted using the chromaticities of the adjusted stimuli are commonly used [13,17]. The chromaticity distance can be used to evaluate how a set of CMFs can accurately characterize the color match performed by an average observer, with a smaller distance suggesting a higher accuracy; the area of the ellipse can be used to evaluate how a set of CMFs can characterize the variations among the observers, with a smaller area suggesting a smaller variation. As summarized in Table 1, the CIE 2006 2° CMFs had the smallest chromaticity distance, followed by the CIE 1931 2°,

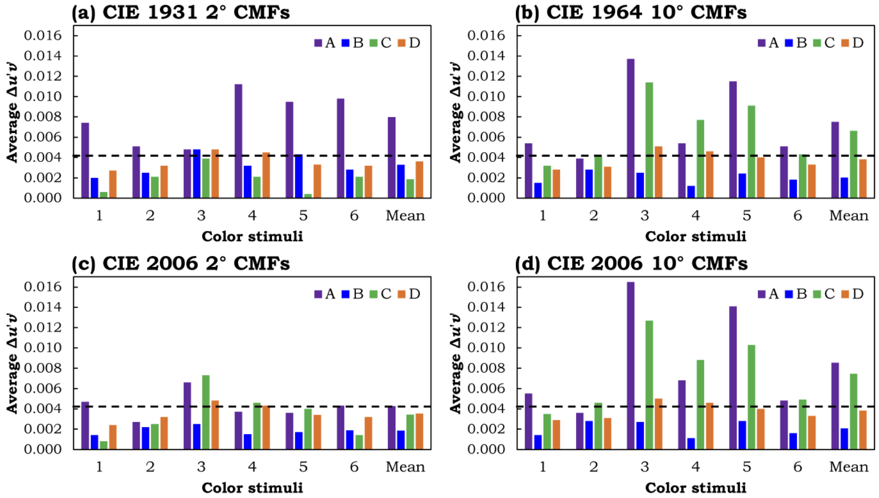

Fig. 12. Average Δ*u*’*v*’ between the chromaticities of the stimuli adjusted by the 50 observers using the four displays and the chromaticities of the stimuli shown on the reference display calculated using the four CMFs. (a) CIE 1931 2° CMFs; (b) CIE 1964 10° CMFs; (c) CIE 2006 2° CMFs; (d) CIE 2006 10° CMFs. (note: the dash line labels 1 just-noticeable color difference, JND. However, it was developed only based on the CIE 1931 CMFs and may not be applicable to the other CMFs).

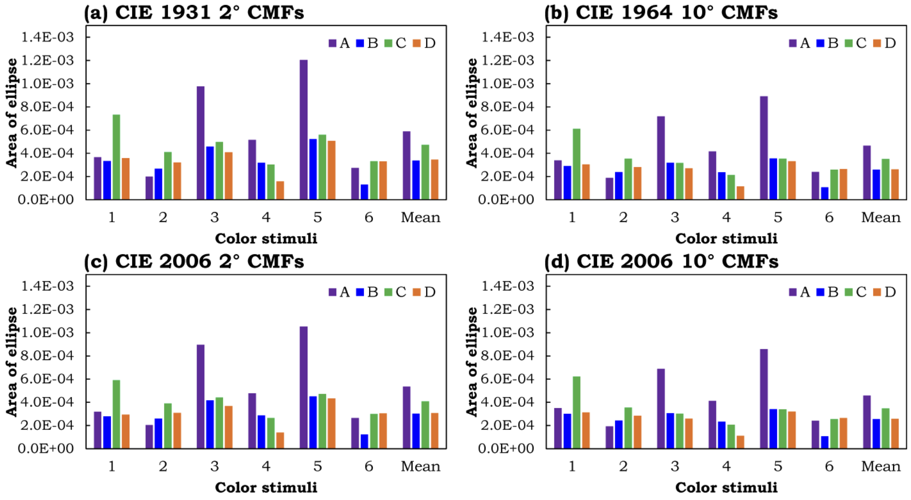

Fig. 13. Area of the 95% confidence error ellipses for the chromaticities of the stimuli adjusted by the 50 observers using the four displays calculated using the four CMFs. (a) CIE 1931 2° CMFs; (b) CIE 1964 10° CMFs; (c) CIE 2006 2° CMFs; (d) CIE 2006 10° CMFs.

CIE 1964 10°, and CIE 2006 10° CMFs. This did not corroborate the recommendations that the 10° CMFs should be used for an FOV beyond 4° [8]. In terms of observer variations, the two 10° CMFs had the best performance, followed by the CIE 2006 2°, and the CIE 1931 2° had the worst performance. Therefore, using the CIE 1931 2° CMFs would cause serious observer metamerism.

When comparing the chromaticity distances or ellipse areas calculated using the different CMFs, the values cannot be directly compared given the different scales. For example, though

Table 1. Summary of the average chromaticity distances and areas of the 95% confidence error ellipses using the four CMFs.

	CIE 1931 2°	CIE 1964 10°	CIE 2006 2°	CIE 2006 10°
Original				
Chromaticity distance in *u*'*v*' units ($\times 10^{-3}$)	4.17	4.98	3.29	5.47
Area of the ellipses in *u*'*v*' units ($\times 10^{-4}$)	4.37	3.34	3.89	3.30
Scaled				
Chromaticity distance in *u*'*v*' units ($\times 10^{-1}$)	7.32	8.77	5.53	9.60
Area of the ellipses in *u*'*v*' units ($\times 10^{-2}$)	7.67	5.89	6.55	5.79

standard D65 and D70 illuminants have a fixed perceived color difference, the chromaticity distances vary with the CMFs, with the values of 0.005701, 0.005675, 0.00594, and 0.005698 *u*'*v*' units for the CIE 1931 2°, 1964 10°, 2006 2°, and 2006 10° CMFs. Thus, to better compare the performance of the four CMFs, both the chromaticity distances and the area of the ellipses were scaled using the chromaticity distance between the standard D65 and D70, as summarized in Table 1, which resulted in the same trend.

4. Conclusion

OLED displays are becoming popular in recent years. They can produce larger color gamuts than conventional LCD displays using primaries with higher saturation levels. Though past studies suggested that such types of primaries were likely to cause serious color mismatch and observer metamerism, no experiment was carried out to specifically investigate how the standard CIE CMFs affected the color matches between these two types of displays. In this study, 50 human observers performed a color matching experiment using four smartphone displays (i.e., one LCD display having an sRGB color gamut and three OLED displays having P3 color gamuts) to match the color appearance of six color stimuli produced by another OLED smartphone display, with the stimuli being carefully selected in a physiologically based chromaticity diagram. Considering the non-uniformity of the displays, the color stimuli were produced by the center of the displays with a field of view of 4.77°. When the CIE 1931 2° CMF were used to characterize the color matches, the chromaticities of the stimuli produced by the LCD display were significantly different from those produced by the OLED displays, with the chromaticities being shifted towards the -*u*'+*v*' direction in the CIE 1976 *u*'*v*' chromaticity diagram. This suggests that if the LCD and OLED displays are calibrated to produce the same chromaticities using the CIE 1931 2° CMFs, the colors on the OLED displays will have a green-yellow tint if those on the LCD appear neutral. In contrast, all the OLED displays had similar chromaticities though their primaries had slight differences. Among the four sets of CMFs—CIE 1931 2°, CIE 1964 10°, CIE 2006 2°, and CIE 2006 10°, the CIE 2006 2° CMFs were found to have the best performance in characterizing the color matches, which did not support the recommendation to use 10° CMFs for an FOV beyond 4°. Moreover, the observer metamerisms between the LCD and OLED displays were larger than those between the OLEDs, regardless of which of the four CMFs were used.

Funding. Huawei Technologies Co. Ltd.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

1. W. D. Wright, "A re-determination of the trichromatic coefficients of the spectral colors," Trans. Opt. Soc., London **30**(4), 141–164 (1929).

2. W. D. Wright, "A re-determination of the mixture curves of the spectrum," Trans. Opt. Soc., London **31**(4), 201–218 (1930).
3. J. Guild, "The colorimetric properties of the spectrum," Phil. Trans. R. Soc. Lond. A **230**(681-693), 149–187 (1931).
4. G. Wyszecki and W. S. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulae*, 2nd Edition (Wiley, 1982).
5. CIE, "Colorimetry, 4th edition," in *CIE 015:2018* (CIE, 2018).
6. W. S. Stiles and J. M. Burch, "N.P.L. color-matching investigation: Final report," Opt. Acta **6**(1), 1–26 (1959).
7. N. I. Speranskaya, "Determination of spectral color co-ordinates for twenty-seven normal observers," Optics Spectrosc. **7**, 424–428 (1959).
8. CIE, "Colorimetry, 4th Edition," in *CIE 15:2018* (CIE, 2018).
9. CIE, "Fundamental chromaticity diagram with physiological axes – part 1," in *CIE 170-1:2006* (CIE, 2006).
10. CIE, "Fundamental chromaticity diagram with physiological axes – part 2: spectral luminous efficiency functions and chromaticity diagrams," in *CIE 170-2:2015* (CIE, 2015).
11. J. Fang and Y. J. Kim, "A matrix-based method of color correction for metamerism failure between LCD and OLED," SID Int. Symp. Dig. Tech. Pap. **49**(1), 1044–1047 (2018).
12. B. Bodner, N. Robinson, R. Atkins, and S. Daly, "Correcting metameric failure of wide color gamut displays," SID Int. Symp. Dig. Tech. Pap. **49**(1), 1040–1043 (2018).
13. J. Li, P Hanselaer, and K. A. G. Smet, "Impact of color matching primaries on observer matching: Part I – Accuracy," Leukos, published online, DOI: 10.1080/15502724.2020.1864395.
14. J. Li, P Hanselaer, and K. A. G. Smet, "Impact of color matching primaries on observer matching: Part II – Observer variability," Leukos, published online, DOI: 10.1080/15502724.2020.1864396.
15. B. Oicherman, M. R. Luo, B. Rigg, and A. R. Robertson, "Adaptation and color matching of display and surface colors," Color Res. Appl. **34**(3), 182–193 (2009).
16. A. Sarkar, L. Blonde, P. L. Callet, F. Autrusseau, P. Morvan, and J. Stauder, "A color matching experiment using two displays: design considerations and pilot test results," In *Proceedings of the 5th European Conference on Color in Graphics, Imaging and Vision Optics* (McGraw Hill, 2010).
17. Y. Hu, M. Wei, and M. R. Luo, "Observer metamerism to display white point using different primary sets," Opt. Express **28**(14), 20305–20323 (2020).
18. Y. Asano, M. D. Fairchild, L. Blonde, and P. Morvan, "Color matching experiment for highlighting interobserver variability," Color Res. Appl. **41**(5), 530–539 (2016).
19. M. D. Fairchild and R. L. Heckaman, "Measuring observer metamerism: the nimeroff approach," Color Res. Appl. **41**(2), 115–124 (2016).
20. D. L. Long and M. D. Fairchild, "Modeling observer variability and metamerism failure in electronic color displays," in *Proceedings of the 22nd Color and Imaging Conference* (2014), pp. 14–27.
21. H. Xie, S. P. Farnand, and M. J. Murdoch, "Observer metamerism in commercial displays," J. Opt. Soc. Am. A **37**(4), A61–A69 (2020).
22. H. Xie, C. E. Rodrigues-Pardo, and G. Shama, "Multiobjective optimization for color display primary design," J. Electro. Imaging **26**(6), 063013 (2017).
23. M. D. Fairchild and D. R. Wyble, "Mean observer metamerism and the selection of display primaries," in *Proceedings of the 15th Color and Imaging Conference* (2007), pp. 151–156.
24. D. A. MacLeod and R. M. Boynton, "Chromaticity diagram showing cone excitation by stimuli of equal luminance," J. Opt. Soc. Am. A **69**(8), 1183–1186 (1979).
25. M. Wei and S. Chen, "Effects of adapting luminance and CCT on appearance of white and degree of chromatic adaptation," Opt. Express **27**(6), 9276–9286 (2019).
26. Y. Zhu, M. Wei, and M. R. Luo, "Investigation of effects on adapting chromaticities and luminance on color appearance on computer displays using memory colors," Color Res. Appl. **45**(4), 612–621 (2020).
27. Y. Park, M. J. Murdoch, and M. D. Fairchild, "Observer metamerism: why do [mis]matches of neutral appear pinkish or greenish," in *Proceedings of the 28th Color and Imaging Conference* (2020), pp. 7–12.