



Reducing the CIE colorimetric matching failure on wide color gamut displays

MINJEONG KO,¹  YOUNGSHIN KWAK,^{1,*}  GUIWON SEO,²
JONGHWAN KIM,² AND YOUNGSU MOON²

¹Ulsan National Institute of Science and Technology, Department of Biomedical Engineering, 50 UNIST-gil, Ulsan 44919, Republic of Korea

²Samsung Electronics Co. Ltd., Suwon 16677, Republic of Korea

*yskwak@unist.ac.kr

Abstract: Color matching experiments were conducted for 11 pairs of displays, using 7 displays with different spectral characteristics. The color matching results between the LCD display and displays that have a narrowband spectrum, such as a laser projector, QLED, or OLED, demonstrated a significant color difference between two matched colors. The maximum difference was $18.52 \Delta E_{00}$, which indicates the white color difference between the LCD and laser projector. There was also a clear observer variability of $2.27 \Delta E_{00}$. The new cone fundamental function derived from 757 metameric pairs showed good performance compared to CIE standard observers reducing the display color mismatching significantly. This function also demonstrated a better performance when evaluating color matching in color chart image.

© 2023 Optica Publishing Group under the terms of the [Optica Open Access Publishing Agreement](#)

1. Introduction

With the development of wide color gamut display technologies having the narrowband spectrum, producing more saturated colors than conventional displays is now possible. However, when two different displays use the same CIE 1931 tristimulus values to produce the same colors, a color mismatch with different color appearances can be observed. Similarly, two colors with the same appearance on different displays were found to have different tristimulus values [1–15]. Such a color mismatch problem indicates that the CIE standard observer does not represent the real average observer.

Issues related to color mismatch have recently become critically important. Therefore, several color matching experiments have been conducted using displays with various spectral characteristics. These experiments were conducted to demonstrate the occurrence of color mismatch and observer variability when using CIE standard color matching functions (CMF).

For instance, 50 observers adjusted six test colors in the CIE $u'v'$ color space using one LCD and four OLEDs for a field size of view (FOV) of 4.77° in a study by Wu et al. [1]. The results showed that the color difference between the reference and matched colors was $0.0042 \Delta u'v'$, and there were $0.0059 \Delta u'v'$ inter-observer variations. Huang et al. conducted a color matching experiment by adjusting the intensities of one or two primaries using LED panels at an FOV of 5.7° with 70 observers [2]. The color difference between the reference and matched colors was $0.0123 \Delta u'v'$ and $0.0059 \Delta u'v'$ for inter-observer variations. In a study by Shi and Luo, 20 observers adjusted colors in the CIELAB color space using three LCDs and two OLEDs at 4° of FOV [3]. The results showed that the color difference between the reference and matched color was $4.93 \Delta E_{00}$, and there was a $2.58 \Delta E_{00}$ of inter-observer variations. Bodner et al. performed a color matching experiment with an OLED and a CRT display with a 1° FOV [4]. The 16 observers adjusted the CIELAB a^*b^* , chroma, and lightness, and the results showed that the color difference between the reference and matched color was $2.7 \Delta E_{00}$. Fan et al. conducted an experiment where 35 observers adjusted colors by manipulating the correlated color temperature

and CIELAB a^*b^* using two LCD displays at an FOV of 4° [5]. In this experiment, there were $5.5 \Delta E_{00}$ inter-observer variations.

These studies clearly demonstrated that color mismatch and the observer variability cannot be ignored when using the CIE standard color matching function. Although observers performed color matching to have the same color appearance in two displays with different spectral characteristics, the two colors actually had different CIE 1931 tristimulus values. There were also differences in color matching for each individual. Therefore, a new color matching function is required to complement the CIE metameric failure and observer variability and to reflect the characteristics of various displays, including the narrowband spectrum [16,17].

In this study, color matching experiments were performed using seven displays with different spectral characteristics, including recently developed narrowband spectrum displays such as laser projectors, QLED, and OLED. The results of the color matching experiment were used to derive new cone fundamentals using Asano's individual colorimetric observer model [16]. The performance of the new cone fundamental functions was validated by evaluating color matching in the color chart image.

2. Display color matching experiment

2.1. Apparatus and setup

Two displays with different spectral characteristics were placed side by side. Two types of experimental settings were used depending on the FOV of the stimulus. In the experimental setting A, 2° and 4° stimuli were used, whereas full-sized stimuli were used in setting B.

Figure 1(a) demonstrates the experimental setting A. The stimuli were shown in a size of $9.1 \text{ cm} \times 9.1 \text{ cm}$; excluding these parts, the remaining display was covered with a black paper. The distance between the two stimuli was set at 15 cm . The distance between the observer and the displays was set at 270 cm for a 2° viewing angle and 130 cm for a 4° viewing angle.

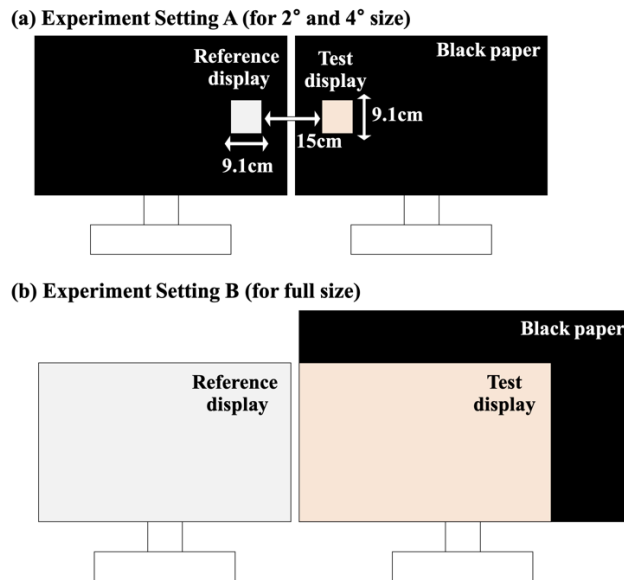


Fig. 1. Experimental settings.

Figure 1(b) demonstrates the experimental setting B. The size of the stimuli was set to fit the smaller display size between the two displays, and the remaining part of the larger display was covered with a black paper. The distance between the observer and the displays was set at 200 cm .

2.2. Test displays

Seven test displays, including three RGB LCDs with different spectral characteristics (RGB LCD1, RGB LCD2, and RGB LCD3), QLED, RGB OLED, WRGB OLED, and a laser projector, were used in the experiment. Figure 2 displays the spectral power distribution (SPD) of the test displays. Note that WRGB OLED uses RGB and white sub-pixels while other displays use RGB only.

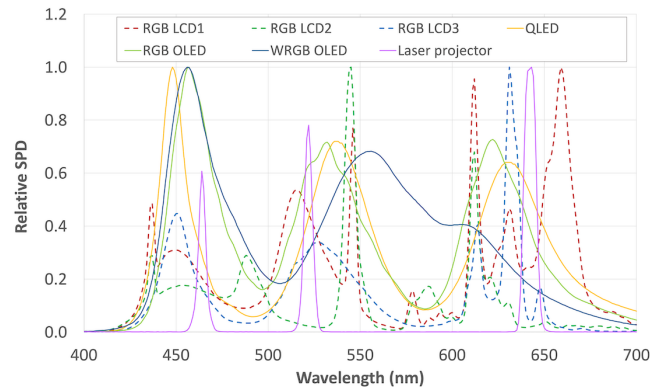


Fig. 2. Spectral power distribution of the test displays.

2.3. Experimental procedures

Before the experiment, all the test displays were warmed up for at least 30 min. The experiment was conducted in a dark room. The observers were located at the center of the two displays.

The experiment was conducted by manipulating the color of the test displays to have the same color appearance as the reference display. When the observers were making a judgement, they were allowed to freely move their eyes and postures but were guided to focus the center of the two displays to minimize the effect of color change by display non-uniformity.

Each test display was characterized using a gain-offset-gamma model [18], and the color was controlled in the CIELAB color space. The observer adjusted the brightness (L^*), redness-greenness (a^*), and yellowness-blueness (b^*) of the color on the test display using a keyboard. The changed CIELAB value was then converted to XYZ, which was subsequently converted to the corresponding RGB value and the color was observed by the observers watching the test display.

Before the main experiment, the observers performed a training session to familiarize themselves with the color control method. When color matching was performed, the color in the test display began with the color having the same luminance and different chromaticity as the reference stimulus. The observer performed the color matching without a time limit. Immediately after each color matching, the spectral data of the stimuli on the reference and test displays were measured using a spectroradiometer (Konica Minolta CS-2000). The observer was asked to press the enter key to move to the next stimulus. The same procedure was followed for each experimental session, and each experiment lasted 30 min to 90 min. The experimental protocols and procedures were approved by the Institutional Review Board of the Ulsan National Institute of Science and Technology.

2.4. Experimental session and test stimuli

Table 1 lists the experimental sessions that were conducted, which demonstrates the type of displays used, FOV, experimental setting, number of test stimuli, number of observers, and number of repetitions. Experimental session 1 was conducted using the same RGB LCD1 display to test the sufficiency of individuals in matching the colors. Experimental session 3-2 were conducted twice with different observers (6 and 7) on different days but the data were combined for further analysis. In the case of Setting B, as explained in Section 2.1, FOV was adjusted to fit the smaller display. Therefore, Laser projector's screen size was adjusted as 26° for Session 8 following RGB LCD2 size and 15° for Session 10 following RGB LCD3 size.

Table 1. Color Matching Experimental Session

	Ref. display	Test display	FOV	Setting	Test stimuli	Obs.	Repeat
1	RGB LCD1	RGB LCD1	4°	A	1	6	2
2		QLED	4°	A	1	7	2
3-1			2°	A	1	6	2
3-2	RGB LCD1	RGB LCD2	4°	A	1	6 7	2 2
4-1		RGB OLED	2°	A	1	7	2
4-2		RGB OLED	4°	A	1	7	2
5-1		QLED	4°	A	1	6	2
5-2		QLED	4°	A	15	8	1
6-1		RGB OLED	4°	A	1	7	2
6-2	RGB LCD2				25	7	1
7		WRGB OLED	4°	A	9	8	1
8		Laser projector	26°	B	14	8	1
9		RGB LCD3	15°	B	14	8	1
10		Laser projector	15°	B	14	5	1
11	RGB LCD3	RGB OLED	15°	B	10	5	1

All the test stimuli were located within the color gamut of the reference and test displays. Because white has the largest metamerism [16], it was included as a test color in all the experimental sets. Only the white color was used in the experimental sessions in which only one stimulus was tested; nine sessions belonged to this case. In the remaining seven types of experiments, color matching was performed on the white color and various colors with different hues and colorfulness, which included red, green, blue, cyan, magenta, and yellow. The CIE $u'_{10}v'_{10}$ chromaticities of the test stimuli for each experimental session are shown in Fig. 3.

2.5. Observers

Nine observers (eight females and one male) with an average age of 28 years participated in this experiment. Each observer participated in a minimum of 3 and a maximum of 11 experimental sessions. All of them had normal color vision, as tested using the Ishihara color vision test.

2.6. Data analysis method

The $X_{10}Y_{10}Z_{10}$ were calculated for the reference and matched spectra. From the tristimulus values, $u'_{10}v'_{10}$ values were calculated to represent the colors on the chromaticity coordinates and CIELAB values to calculate the color difference. The reference white for the calculation of CIELAB values was the white of RGB LCD2, and its tristimulus value is [97.62, 103.36,

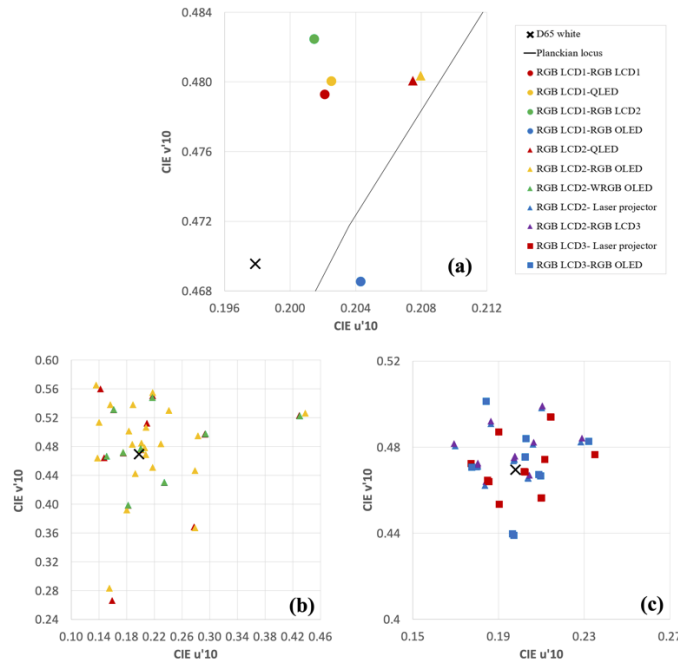


Fig. 3. Color chromaticities for test stimulus of each experimental session: (a) sessions 1, 2, 3, 4, 5-1, 6-1 (only white color), (b) sessions 5-2, 6-2, 7, (c) sessions 8, 9, 10, 11.

111.43]. Using the CIELAB values for the reference and the matched color, the CIEDE2000 color difference, ΔE_{00} , was calculated.

The color matching results were analyzed in terms of ‘Observer Metamerism Magnitude (OMM)’ and ‘Observer Variability (OV)’ [2,15]. The OMM indicates the average color difference between the reference and individual matching colors, as demonstrated in Eq. (1). The OV calculates the average color difference between the average matching color and the individual matching colors, as demonstrated in Eq. (2). The term i represents the observer, from 1 to N , and N represents the number of observers per experimental session.

$$OMM = \frac{1}{N} \sum_{i=1}^N (E_{00}(L^* a^* b^*_{ref,i}, L^* a^* b^*_{matched,i})) \quad (1)$$

$$OV = \frac{1}{N} \sum_{i=1}^N (E_{00}(L^* a^* b^*_{matched_{ave}}, L^* a^* b^*_{matched,i})) \quad (2)$$

2.7. Results

2.7.1. Color matching accuracy test

To evaluate the color matching accuracy, color matching was first conducted on the same RGB LCD1 display. When the color matching experiment was performed using the same display, observers should generate the same color as the reference color. The average \pm standard deviation of OMM was determined to be $0.98 \pm 0.68 \Delta E_{00}$ and the OV was $0.22 \pm 0.15 \Delta E_{00}$, indicating that the observers can match colors fairly well.

2.7.2. FOV effect on color matching

Experimental sessions 3 (RGB LCD1 and RGB LCD2) and 4 (RGB LCD1 and RGB OLED) were conducted using both 2° and 4° FOV.

Table 2 summarizes the color matching results between the 2° and 4° FOV. The results show that as the FOV increased, both the OMM and OV slightly decreased, but t-test showed that they were not significantly different ($p > .05$). This trend was also observed in the $u'_{10}v'_{10}$ chromaticity diagram.

Table 2. Color Matching results between 2° and 4° of FOV

Ref. display	Test display	FOV	OMM (Average $\Delta E_{00} \pm$ standard deviation)	OV (Average $\Delta E_{00} \pm$ standard deviation)
3-1	RGB LCD2	2°	3.87 ± 2.58	3.36 ± 1.90
		4°	2.84 ± 1.87	2.93 ± 1.98
4-1	RGB OLED	2°	5.71 ± 1.59	2.80 ± 1.08
4-2		4°	5.09 ± 1.67	3.08 ± 1.39

Figure 4 displays the white color matching results of Session 3 (RGB LCD1 and RGB LCD2), and Fig. 5 shows the results of Session 4 (RGB LCD1 and RGB OLED). The arrow direction represents the individual color matching result from the reference color coordinates. The red arrow indicates the result at 2°, and the blue arrow indicates the result at 4°. The results show that there was no significant difference between the 2° and 4° FOV of the color matching results.

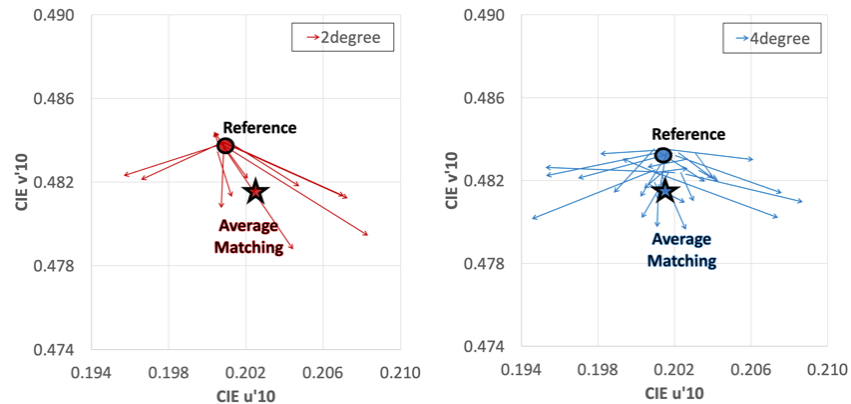


Fig. 4. Color Matching Results for RGB LCD1 vs. RGB LCD2 (left: 2°, right: 4°).

2.7.3. Relationship between the observer metamerism magnitude and observer variability

Table 3 summarizes the color matching results in terms of the OMM and OV. The maximum average OMM was $18.52 \Delta E_{00}$, which indicates the white color difference between the RGB LCD2 and laser projector. Our experimental data also confirms the significant color mismatches in wide-gamut displays having the narrowband spectrums as found in the previous studies.

Figure 6 presents the relationship between OMM and OV using the white color matching experimental data. The x-axis represents the experimental session.

Note, for OMM, the laser projector demonstrates the largest color difference, followed by the WRGB OLED and QLED when color matching was performed with the RGB LCDs. Compared to the OMM, there was a relatively smaller difference between the display pairs for OV. The

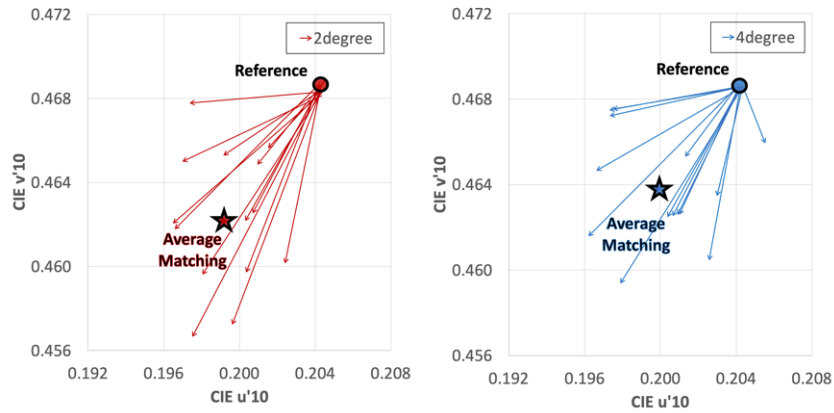


Fig. 5. Color Matching Results for RGB LCD1 vs. RGB OLED (left: 2°, right: 4°).

Table 3. Color matching results based on the OMM and OV

Ref. display	Test display	Test color	OMM (Average $\Delta E_{00} \pm$ standard deviation)	OV (Average $\Delta E_{00} \pm$ standard deviation))	
1	RGB LCD1	1 (White)	0.98 ± 0.68	0.22 ± 0.15	
	QLED	1 (White)	8.73 ± 2.49	2.61 ± 1.47	
3	RGB LCD1	RGB LCD2	1 (White)	3.16 ± 2.14	3.22 ± 1.92
	RGB OLED	1 (White)	5.40 ± 1.63	2.96 ± 1.30	
5	QLED	1 (White)	6.53 ± 2.52	2.14 ± 1.13	
		13	3.72 ± 2.59	1.61 ± 0.96	
6	RGB OLED	1 (White)	5.88 ± 1.98	1.86 ± 0.87	
		25	3.14 ± 2.01	1.57 ± 0.92	
7	RGB LCD2	WRGB OLED	1 (White)	9.21 ± 2.14	1.85 ± 1.22
		9	3.53 ± 2.71	1.12 ± 0.69	
8	Laser projector	1 (White)	18.52 ± 3.55	4.39 ± 3.87	
		14	13.99 ± 4.34	4.22 ± 2.79	
9	RGB LCD3	1 (White)	6.90 ± 2.08	2.97 ± 1.55	
		14	4.31 ± 1.66	2.22 ± 1.11	
10	RGB LCD3	Laser projector	1 (White)	11.82 ± 2.11	3.01 ± 1.41
		14	10.68 ± 3.41	3.24 ± 2.13	
11	RGB OLED	1 (White)	3.92 ± 1.71	2.66 ± 1.21	
		10	3.39 ± 1.40	1.92 ± 0.92	

average OV was $2.27 \pm 1.81 \Delta E_{00}$ for the entire experiment. For instance, when Exp8 (RGB LCD2 and laser projector) and Exp3 (RGB LCD1 and RGB LCD2) are compared for the white test color, the OMMs are 18.52 and 3.16 ΔE_{00} , respectively, while the OVs are 4.39 and 3.22 ΔE_{00} , respectively, showing relatively similar results.

This result indicates that the OV is not significantly affected by the type of display, while there can be a significant color mismatch between the LCD and new types of displays, such as a laser projector when the colorimetric matching was performed using the CIE XYZ values. Additionally, when comparing the results of color matching of only white to those of matching various colors including white, color mismatch is demonstrated to be more significant for the white color.

The color matching data were further analyzed using the $u'_{10}v'_{10}$ chromaticity diagram. Figure 7 shows the white color matching result of Session 8 (RGB LCD2 and Laser projector) in the $u'_{10}v'_{10}$ color space. The blue arrow direction represents the reference color (CIE standard observer) to the individual matching result (each observer) indicating OMM. On the other hand, red arrow connects the average matching data point with the individual matching data indicating the variability within the observers in this study (OV).

The results show that although deviations of individual matchings exist, the chromaticities of all the matched colors on the laser projector are generally located in the $+u'_{10}-v'_{10}$ direction

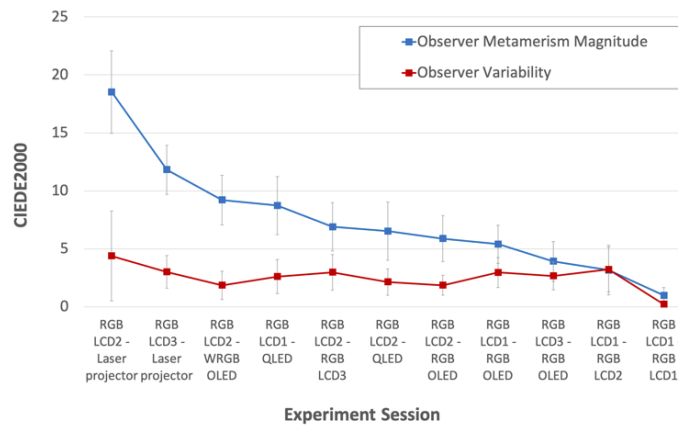


Fig. 6. Relationship between the OMM and OV (ΔE_{00}).

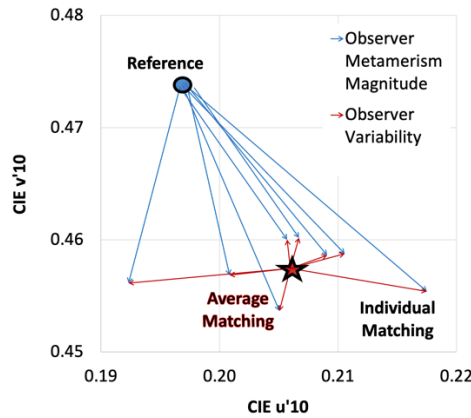


Fig. 7. Color Matching result of RGB LCD2 vs. Laser projector.

from that of the reference color on RGB LCD2. This indicates that when the RGB LCD2 and laser projector were set to have the same $u'_{10}v'_{10}$ coordinates, all the observers perceived the laser projector to appear greener than the LCD display, although there were slight differences in the color perception.

3. New LMS cone fundamentals derivation

A total of 829 metameric pairs were collected through the color matching experiment, and the color matching data using RGB channels are used for new LMS cone fundamentals derivation excluding WRGB OLED (Session 7) matching data. Using these 757 color matching data, the LMS cone fundamentals were derived by optimizing the parameters based on an individual colorimetric observer model by Asano [16,19,20]. The Asano model is based on the CIE 2006 cone fundamentals. This model proposes cone fundamentals while considering individual variability. To propose the cone fundamentals, the following 10 parameters were optimized: age (a), field size (v), deviation from the average lens density (d_{ens}), deviation from the average macula density (d_{macula}), deviation from the average LMS cone density (d_L, d_M, d_S), and deviation from the average LMS cone shift (s_L, s_M, s_S). Among the 10 parameters, age and field size were obtained from the CIE 2006 model, and the remaining eight physiological parameters were obtained from the Asano model.

3.1. LMS cone fundamentals parameter optimization process

The overall process of the LMS cone fundamentals parameter optimization is as follows.

First, the tristimulus LMS value of the reference stimulus was calculated by multiplying the reference stimulus spectral data of the ($n \times 1$) matrix and the test cone fundamentals (test CMF) of the ($n \times 3$) matrix. In this case, n is 79, which is 390 nm to 780 nm with a 5 nm step. The initial parameters for the test CMF were 38 for age, 2° for field size, and 0 for the remaining.

Second, the predicted spectrum, which has the same LMS tristimulus values on the test display as the reference display, was calculated using Eq. (1). The input values were reference LMS tristimulus values of the (3×1) matrix obtained from the first step, the test display maximum RGB primary spectrum of the ($n \times 3$) matrix, and the test CMF of the ($n \times 3$) matrix.

$$P = M[[C^T M]^{-1} R]. \quad (3)$$

$$P = \text{predicted spectrum}(n \times 1)$$

$$M = \text{Test display max RGB spectrum}(n \times 3)$$

$$C = \text{Test CMF}(n \times 3)$$

$$R = \text{Reference LMS}(3 \times 1)$$

Third, the tristimulus values of the matched and predicted spectra were calculated using the CIE 1964 color matching function. The predicted spectrum from the second step and the matched spectrum of the actual color matching result were used.

Fourth, using the tristimulus values, the $u'_{10}v'_{10}$ values of the matched and predicted colors were calculated; the color difference between them was calculated using $\Delta u'_{10}v'_{10}$. The 10 parameters were optimized using an unconstrained nonlinear minimization method, such that the average $\Delta u'_{10}v'_{10}$ was minimized. During the parameter optimization process, the age was limited to 20-80 years, the field size was limited to 1° - 10° , and the remaining parameters were optimized without any restrictions.

To derive the new cone fundamentals, 757 metameric pairs were grouped into 26 training datasets such as using the data from each experimental session only, each observer's data and using all the data etc. For instance, when obtaining the cone fundamentals of each experimental session, 10 parameters were optimized using all the observer data of each experimental session as the input values.

3.2. *New cone fundamentals performance test results*

The performance test was conducted to determine the sufficiency of the new cone fundamentals in explaining our datasets. It was analyzed using the color difference, ΔE_{00} , between the matched colors from the actual color matching data and the predicted colors using optimized cone fundamentals in each experimental session. The performance of the CIE 1931 2°, CIE 1964 10°, CIE 2006 2° color matching functions, 3-by-3 matrix correction method, and 10 categorical observers by Asano [16] as well as the 26 new LMS cone fundamentals were tested. The 3-by-3 matrix correction method means applying 3×3 matrix to $X_{10}Y_{10}Z_{10}$ values of the test display color to correct the color mismatch [3,4]. In this study, the optimized 3×3 matrix was obtained by minimizing the $\Delta u'_{10}v'_{10}$ values between the references and predicted colors.

The results are summarized in Table 4 showing the average data without distinguishing white and other colors for Sessions 5-11. The new LMS cone fundamentals were numbered from 1 to 26 in order of good performance. Then, those 26 functions were grouped into 9 categories showing the similar performances.

The best-performing method was the cone fundamentals function, CMF1, derived using all the experimental data. The performance of the CIE CMFs was good in the order of CIE 1964 10°, CIE 2006 2°, and CIE 1931 2°, on average. In the case of Asano's categorical observer, Category2 observer showed the better performance than CIE CMFs. For the color matching experiments with broadband displays, the CIE CMF, Asano categorical observers, and new cone fundamentals presented a similar performance; however, for experimental sessions 8 and 10, which included the narrowband display, new cone fundamentals outperform the others.

It is notable that the performance of 3×3 matrix correction method is fairly good indicating that applying 3×3 matrix to current CIE XYZ values can be an easy and practical solution to reduce the display color mismatch. However, applying the same 3×3 matrix to all types of display will cause the errors when applied to the display that CIE colorimetry should work such as Session 1. The average observer function that works well both for broadband and narrowband spectrums are needed to solve the display color mismatching problem.

Table 4. LMS cone fundamentals performance test

OMM (ΔE_{00})		Experimental Session										
CAT	CMF	1	2	3	4	5	6	8	9	10	11	Average
	CIE1931	0.98	3.17	4.23	6.73	1.83	3.49	21.49	3.96	20.12	3.64	6.96 ± 7.46
	CIE1964	0.98	8.73	3.16	5.40	3.72	3.14	13.99	4.31	10.68	3.39	5.75 ± 4.07
	CIE2006	0.98	3.62	4.86	5.19	2.11	3.10	20.69	4.77	17.93	3.04	6.63 ± 6.84
3 × 3 Matrix Correction		5.31	3.97	4.55	5.59	3.05	1.75	10.05	4.41	7.17	5.39	4.80 ± 3.65
	Asano_CAT1	0.98	3.62	4.86	5.19	2.11	3.10	20.69	4.77	17.93	3.04	6.63 ± 6.84
	Asano_CAT2	0.99	3.32	3.21	4.15	1.93	1.88	14.21	3.53	12.74	2.68	4.86 ± 4.64
	Asano_CAT3	0.97	4.27	7.62	6.95	2.29	4.93	26.54	6.58	23.67	3.95	8.78 ± 8.87
	Asano_CAT4	0.98	6.38	3.47	5.14	3.24	2.78	15.44	4.67	12.02	3.38	5.75 ± 4.52
	Asano_CAT5	0.98	4.63	6.76	5.22	1.76	2.94	23.40	4.95	21.27	2.97	7.49 ± 8.03
	Asano_CAT6	0.98	4.80	4.93	5.76	3.09	3.90	21.53	5.90	17.56	3.58	7.20 ± 6.73
	Asano_CAT7	1.00	3.40	4.54	4.10	1.74	1.81	17.70	4.19	15.96	2.68	5.71 ± 5.99
	Asano_CAT8	0.98	4.98	7.42	6.41	1.81	4.16	25.62	5.44	23.41	3.46	8.37 ± 8.75
	Asano_CAT9	0.99	4.10	3.40	4.48	2.21	2.15	14.91	3.80	12.83	2.86	5.17 ± 4.73
	Asano_CAT10	0.97	5.11	9.51	8.40	2.31	6.03	29.32	7.61	27.21	4.74	10.12 ± 9.92
1	CMF1	1.00	6.42	3.28	3.88	2.96	1.91	5.60	2.98	4.20	3.96	3.62 ± 1.60
1	CMF4	0.99	5.51	2.97	3.86	2.57	1.83	7.70	2.97	6.45	3.66	3.85 ± 2.10
1	CMF5	0.99	7.10	3.81	3.75	2.69	2.08	6.50	2.71	5.86	3.15	3.86 ± 2.00
1	CMF6	0.99	5.27	2.93	3.55	2.78	1.85	8.43	3.31	6.20	3.52	3.88 ± 2.19
1	CMF7	0.99	6.87	3.04	4.32	2.96	1.93	7.51	3.11	5.97	3.86	4.06 ± 2.13
1	CMF12	0.98	7.47	3.54	4.52	2.71	2.04	8.95	3.02	8.19	3.12	4.46 ± 2.77
2	CMF2	1.00	5.91	3.34	3.77	2.58	2.00	6.11	2.73	5.34	3.72	3.65 ± 1.70
2	CMF3	1.03	5.14	5.04	4.16	2.11	2.79	4.60	4.80	5.26	2.94	3.79 ± 1.47
2	CMF8	1.01	5.89	4.97	5.04	2.30	2.60	4.82	4.85	6.02	3.26	4.08 ± 1.68
3	CMF9	1.00	4.80	3.50	3.78	2.39	2.41	10.33	2.80	9.06	2.38	4.25 ± 3.06
3	CMF11	0.99	7.18	4.49	4.51	2.40	2.38	8.38	2.89	8.16	2.67	4.41 ± 2.64
3	CMF13	1.00	8.11	5.60	4.41	2.48	2.53	7.71	3.08	7.44	2.56	4.49 ± 2.56
4	CMF10	0.99	4.65	4.25	2.97	3.27	2.78	10.93	4.48	6.35	3.22	4.39 ± 2.70
4	CMF14	0.99	4.65	3.79	4.29	2.93	1.74	10.47	4.15	7.51	4.70	4.52 ± 2.75
4	CMF15	0.99	5.15	3.60	4.48	2.88	1.70	10.54	3.98	7.85	4.48	4.56 ± 2.83
4	CMF16	0.99	4.40	3.77	3.87	2.65	1.78	12.39	4.04	9.15	3.84	4.69 ± 3.47
4	CMF17	0.98	6.72	2.92	3.99	3.30	2.25	11.44	3.98	8.22	3.30	4.71 ± 3.16
4	CMF18	0.98	6.59	3.22	4.59	2.99	1.81	10.93	3.70	8.76	3.69	4.73 ± 3.13
5	CMF19	1.00	3.21	3.59	4.93	2.13	1.67	12.35	4.35	11.18	3.27	4.77 ± 3.88
5	CMF20	1.00	2.69	3.85	5.38	2.26	1.71	12.64	4.95	12.51	3.28	5.03 ± 4.20
5	CMF21	0.99	3.51	3.33	5.44	1.94	2.29	15.01	3.63	14.10	3.04	5.33 ± 5.01
5	CMF22	0.98	4.54	4.25	4.32	2.71	2.03	15.60	4.60	11.94	3.56	5.45 ± 4.62
6	CMF23	1.01	2.45	4.50	8.68	2.71	2.54	12.03	4.99	14.80	4.54	5.83 ± 4.55
7	CMF24	0.98	7.33	4.65	5.73	4.17	4.85	22.13	6.52	17.13	3.61	7.71 ± 6.62
8	CMF25	0.98	4.21	6.18	9.03	1.90	5.21	27.12	5.25	25.73	4.88	9.05 ± 9.42
9	CMF26	1.01	20.36	10.01	9.34	6.00	5.46	17.01	3.67	15.98	8.79	9.76 ± 6.25

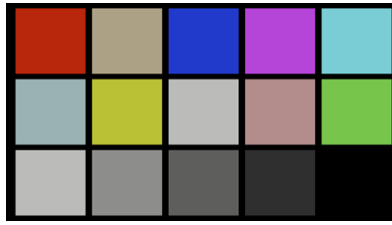


Fig. 8. Color chart (15 colors).



Fig. 9. Actual experimental scene (left: Laser projector, right: RGB LCD2)

4. New LMS cone fundamentals verification

A new cone fundamentals evaluation experiment was conducted to determine whether the new cone fundamentals work sufficiently with color chart images.

4.1. Apparatus and setup

The experimental setting was the same as that displayed in Fig. 1(b) of the color matching experiment. The image size was set based on a smaller display. A larger display presents as black on the remaining part and an additional black paper is used to prevent light observed from the oversized area.

4.2. Test displays and test stimuli

Two test displays were used, an RGB LCD2 and a laser projector. The test stimuli were represented by a color chart, as displayed in Fig. 8. The color chart consisted of 15 colors, including two white, one black, six low chromatic, and six high chromatic colors. All the colors were located within the color gamut of the reference and test displays. Figure 9 demonstrates the actual experimental scene.

4.3. Test color matching function

From Table 4 in the previous section, 10 test color matching functions were selected, including CIE 1931 2° , and CIE 2006 2° color matching functions. A representative function from each category was chosen as the test color matching function to have various chromaticity coordinates when predicting colors using them.

Figure 10 presents the white color prediction results using ten test color matching functions in RGB LCD2 and a laser projector. Each circle indicates the chromaticity of the color predicted by each color matching function. Note that CMF7 is from Category1 showing the good performance

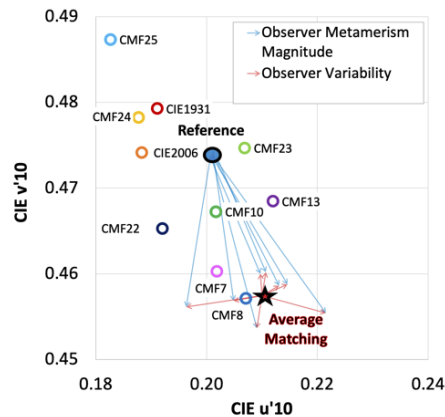


Fig. 10. Color prediction of RGB LCD2 vs. Laser projector.]

on average and CMF8 is from Category2 showing the good performance for Session 8 (RGB LCD2-Laser projector) experiment.

In the experiment, a laser projector was used as the reference display, and RGB LCD2 was used as the test display. The stimuli for the experiment were designed to have the same LMS tristimulus values on the two test displays according to each color matching function. In this case, the color difference between the actual measured color and the target color of the RGB LCD2 stimulus was set to less than $1 \Delta E_{00}$.

4.4. Observers

Fourteen observers (ten females and four males) with an average age of 27 years participated in this experiment. Seven of them also participated in the color matching experiment. All observers had normal color vision, as tested using the Ishihara color vision test.

4.5. Experimental procedures

Before the experiment, all test displays were warmed up for at least 30 min. The experiment was conducted in a dark room. The observers were seated in the center of the two displays and were guided to focus the center of the two displays. The experiment was conducted using two-interval forced-choice method. The stimulus was set to appear on the entire screen, and arrow keys were used to alternately present two stimuli. The observers selected one of the two stimuli that looked more similar to the stimulus on the reference display. They were instructed to not just look at one color, but look at a whole color during their selection. The observers started the experiment after clearly setting their own selection criteria through the training session and were able to see two stimuli as many times as they wanted in each trial. Following the selection, the next stimuli were presented in a random order when the enter key was pressed. For each trial, 45 pairs were evaluated, three repetitions were performed, and the experiment lasted for approximately 40 min.

4.6. Data analysis method

Assuming Thurstone's Case V model [21], the z-score of the test color matching functions was calculated from the pair-comparison data. In the calculation process, proportion 0 was converted to 0.01, and proportion 1 was converted to 0.99. A higher z-score indicates a color matching function that the observer chooses is similar to a reference stimulus, whereas a lower z-score indicates a color matching function that looks different from the reference stimulus.

4.7. Results

Figure 11 demonstrates the results of the pair-comparison with 10 test color matching functions. The x-axis represents the test color matching functions, and the y-axis represents the z-score. When seven observers who participated in the previous color matching experiment were marked as experienced and the remaining seven observers were marked as naïve, t-test result showed that the difference between the two groups was insignificant ($p > .05$). The results indicate that CMF10 shows the best performance followed by CMF7 and CMF8.

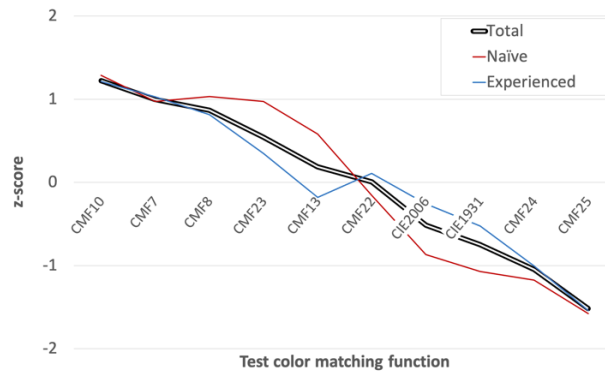


Fig. 11. Pair-comparison with 10 test color matching functions.

The pair-comparison data were further analyzed using the $u'_{10}v'_{10}$ chromaticity diagram, as illustrated in Fig. 10. The arrow direction represents the individual color matching result from the reference color coordinates. As indicated in the color matching experiment, the chromaticities of all the matched colors on the laser projector were generally located in the $+u'_{10}-v'_{10}$ direction from that of the reference color on RGB LCD2. It is notable that the color prediction results by the best performance CMFs, CMF10, 7 and 8 are also located in the same direction as the individual matching results. However, CMF8 showing the best performance for the color matching data for white color was not the best CMF for the chart image experiment.

5. Proposed LMS function

The results of our two experiments demonstrated that the optimal LMS function, CMF7, obtained from the color matching experiment performed well, even in images containing various colors. In other words, two different experiments, which are color matching experiment and new cone fundamentals verification experiment, presented similar results with a high performance in the same LMS cone fundamentals. The optimal LMS function obtained in this study is depicted in Fig. 12 and tabulated in Table 5 to be used for evaluation in the future studies.

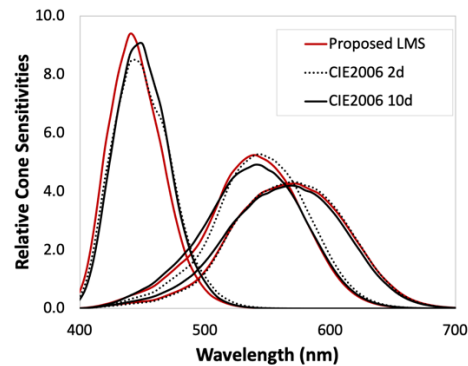


Fig. 12. Cone sensitivities of Proposed LMS cone fundamentals

6. Conclusion

A color matching experiment was designed to collect 829 metameric pairs using seven displays with different spectral characteristics, including displays with a narrowband spectrum. The results demonstrated that the color matching accuracy was reliable, and there was no significant difference between the 2° and 4° FOV. The OMM was the highest for the laser projector, a display having the narrowest spectrum among all the test displays. A significant OV that cannot be ignored always appeared, which was $2.27\Delta E_{00}$ on average. The color matching results indicate that the CIE standard observer fails to explain the color matching results, requiring more accurate color matching functions.

Based on the 757 color matching data and the individual colorimetric observer model by Asano, the parameters of cone fundamentals were optimized and new cone fundamentals were derived. The performance differed depending on the training datasets used. For color matching with broadband displays, both CIE CMFs and new cone fundamentals demonstrated a good performance. However, when narrowband displays such as laser projector were used, the performance was more than three times better in the new cone fundamentals.

As a result of the new cone fundamentals verification experiment using a color chart stimulus, the response of the observers indicated that the new cone fundamentals are similar to the reference stimulus compared to the existing CIE 1931 2° and CIE 2006 2° color matching functions. The cone fundamentals with high z-scores predict the color in the direction of the actual individual matching results. Additionally, although individual differences existed for each observer, it was possible that an observer was satisfied to a certain extent despite using the same color matching function without using different color matching functions for each observer. Therefore, the experimental results indicated that the optimal LMS demonstrated a good performance, even in images with various colors, and it is likely to explain the observations of several individuals using one common color matching function – the real average observer.

In this study, the necessity and possibility for a “new average colorimetric observer function” was presented. The new LMS cone fundamentals could solve the display color mismatch problem caused by current CIE standard observers. Further intensive and global cooperative researches are required to obtain new standard average observer. The proposed LMS functions obtained in this study are based on limited number of Korean observers and the limited number of metameric color pairs obtained from display color matching experiment. More diverse observers’ data covering wide range of ages, gender and ethnic groups should be obtained using more diverse experimental methods and setting including the size effect, color patch vs. natural images etc. Also, more researches are needed for the colorimetric observer models.

Appendix: information of proposed LMS cone fundamentals**Table 5. Cone sensitivities of Proposed LMS cone fundamentals**

Wavelength (nm)	l	m	s	Wavelength (nm)	l	m	s
390	0.0024	0.0030	0.1261	590	3.9437	2.2177	0.0002
395	0.0055	0.0069	0.2810	595	3.7684	1.8194	0.0001
400	0.0122	0.0145	0.6076	600	3.5469	1.4479	0.0001
405	0.0249	0.0302	1.2906	605	3.3043	1.1282	0.0001
410	0.0447	0.0570	2.4210	610	3.0190	0.8622	0.0000
415	0.0688	0.0951	3.9060	615	2.7073	0.6477	0.0000
420	0.0938	0.1468	5.5163	620	2.3823	0.4773	0.0000
425	0.1160	0.2014	6.7320	625	2.0735	0.3466	0.0000
430	0.1425	0.2757	7.9137	630	1.7545	0.2489	0.0000
435	0.1696	0.3520	8.7531	635	1.4389	0.1765	0.0000
440	0.1995	0.4403	9.3936	640	1.1655	0.1229	0.0000
445	0.2188	0.5149	9.1268	645	0.9377	0.0866	0.0000
450	0.2408	0.5885	8.3067	650	0.7376	0.0605	0.0000
455	0.2639	0.6736	7.3663	655	0.5612	0.0414	0.0000
460	0.3036	0.7927	6.4423	660	0.4193	0.0282	0.0000
465	0.3727	0.9629	5.5024	665	0.3101	0.0194	0.0000
470	0.4564	1.1390	4.3849	670	0.2268	0.0134	0.0000
475	0.5417	1.3238	3.3685	675	0.1638	0.0093	0.0000
480	0.6306	1.5196	2.5051	680	0.1167	0.0064	0.0000
485	0.7375	1.6979	1.7937	685	0.0821	0.0043	0.0000
490	0.8619	1.9294	1.3125	690	0.0566	0.0030	0.0000
495	1.0355	2.2177	0.9439	695	0.0391	0.0021	0.0000
500	1.2829	2.6029	0.6532	700	0.0273	0.0014	0.0000
505	1.5786	3.0651	0.4391	705	0.0190	0.0010	0.0000
510	1.9293	3.5613	0.3078	710	0.0131	0.0007	0.0000
515	2.3072	4.1090	0.2126	715	0.0089	0.0005	0.0000
520	2.6838	4.5418	0.1421	720	0.0062	0.0003	0.0000
525	3.0239	4.8173	0.0925	725	0.0043	0.0002	0.0000
530	3.3023	5.0407	0.0600	730	0.0030	0.0002	0.0000
535	3.5229	5.1928	0.0376	735	0.0021	0.0001	0.0000
540	3.7522	5.2425	0.0234	740	0.0014	0.0001	0.0000
545	3.9434	5.1471	0.0144	745	0.0010	0.0001	0.0000
550	4.0369	5.0334	0.0089	750	0.0007	0.0000	0.0000
555	4.1182	4.8540	0.0055	755	0.0005	0.0000	0.0000
560	4.1951	4.6068	0.0033	760	0.0004	0.0000	0.0000
565	4.2431	4.3072	0.0021	765	0.0003	0.0000	0.0000
570	4.2664	3.9376	0.0013	770	0.0002	0.0000	0.0000
575	4.2439	3.4886	0.0008	775	0.0001	0.0000	0.0000
580	4.1408	3.0537	0.0005	780	0.0001	0.0000	0.0000
585	4.0423	2.6404	0.0003				

Funding. Samsung Electronics Co. Ltd.; National Research Foundation of Korea (NRF-2021R1A2C1013610).

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. J. Wu, M. Wei, Y. Fu, and C. Cui, "Color mismatch and observer metamerism between conventional liquid crystal displays and organic light emitting diode displays," *Opt. Express* **29**(8), 12292–12306 (2021).
2. M. Huang, Y. Li, Y. Wang, X. Li, and M. Wei, "Effect of primary peak wavelength on color matching and color matching function performance," *Opt. Express* **29**(24), 40447–40461 (2021).
3. K. Shi and M. R. Luo, "Methods to improve colour mismatch between displays," in *Proceedings of the 29th Color and Imaging Conference* (2021), pp. 37–41.
4. 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).
5. H. Fan, Y. Hu, and M. R. Luo, "Observer Metamerism for Assessing Neutrality on Displays," in *Proceedings of 2020 11th Academic Conference on Printing and Packaging*, 754 (2021), pp. 47–53.
6. M. Shaw and M. D. Fairchild, "Evaluating the 1931 CIE color-matching functions," *Color Res. Appl.* **27**(5), 316–329 (2002).
7. SONY Corporation, *Colour matching between OLED and CRT* (SONY, 2013).
8. D. L. Long and M. D. Fairchild, "Reducing observer metamerism in wide-gamut multiprimary displays," *Proc. SPIE* **9394**, 93940T (2015).
9. D. L. Long and M. D. Fairchild, "Observer metamerism models and multiprimary display systems," *SMPTE Motion Imag. J.* **125**(3), 18–29 (2016).
10. H. Xie, S. P. Farnand, M. J. Murdoch, J. S. Park, and B. Min, "Color Mismatches across Commercial Displays: Modeling the Effect of Observer Metamerism," *SID Int. Symp. Dig. Tech. Pap.* **50**(1), 1374–1377 (2019).
11. J. Li, P. Hanselaer, and K. A. G. Smet, "Pilot Study on Color Matching Accuracy using Different Primaries," in *Proceedings of 29th Quadrennial session of the Commission Internationale de l'Eclairage* (2019), pp. 1010–1018.
12. H. Xie, S. P. Farnand, and M. J. Murdoch, "Observer metamerism in commercial displays," *J. Opt. Soc. Am. A* **37**(4), A61–A69 (2020).
13. 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).
14. Y. Park and M. J. Murdoch, "Effect of color gamut and luminance on observer metamerism in HDR displays," in *Proceedings of the 28th Color and Imaging Conference* (2020), pp. 237–243.
15. Y. Park and M. J. Murdoch, "Efficiently evaluating the effect of color gamut and spectral bandwidth on observer metamerism in high dynamic range displays," *J. Soc. Inf. Disp.* **29**(9), 704–722 (2021).
16. Y. Asano, "Individual colorimetric observers for personalized color imaging," PhD dissertation at Rochester Institute of Technology (2015).
17. C. Y. H. Bai and L. C. Ou, "Observer variability study and method to implement observer categories for novel light source projection system," *Color Res. Appl.* **46**(5), 1019–1033 (2021).
18. R. S. Berns, R. J. Motta, and M. E. Gorzynski, "CRT Colorimetry. Part 1: Theory and Practice," *Color Res. Appl.* **18**(5), 299–314 (1993).
19. CIE, "Fundamental chromaticity diagram with physiological axes - part 1," in *CIE 170-1:2006* (CIE, 2006).
20. CIE, "Fundamental chromaticity diagram with physiological axes - part 2: Spectral luminous efficiency functions and chromaticity diagrams," in *CIE 170-2:2015* (CIE, 2015).
21. G. A. Gescheider, *Psychophysics: the fundamentals*, 3rd ed. (Psychology Press, 1997), pp. 198–206.