Influence of viewing conditions on cross-media color matching

Min Huang[1], Ruili He[1], Chunli Guo[1], Chunjie Shi[1], Guihua Cui[2],
Ming Ronnier Luo[3], Manuel Melgosa[4*]

[1] School of Printing and Packaging Engineering, Beijing Institute of Graphic Communication,
 Beijing 102600, China

[2] School of Physics and Electronic Information Engineering, Wenzhou University,
 Wenzhou, 325035, China

[3] School of Design, University of Leeds, Leeds, LE2 9JT, UK

[4] Optics Department, University of Granada, Granada, 18071, Spain

*Corresponding author: mmelgosa@ugr.es

We investigated observer metamerism under a variety of viewing conditions, in a set of color-matching experiments using displays and printed color samples under specific light sources. A selection was made of light sources with different illuminances, spectral power distributions, and correlated color temperatures, as well as displays with different sets of primaries. A panel of 157 observers with normal color vision and ages between 20 and 59 years old performed 5465 visual color matches around 9 different color centers. The results from the simulated and real experiments were quite different. Specifically, the mean color difference from the mean changed with experimental viewing conditions, ranging from 0.73 to 1.64 CIELAB units (average 0.99 CIELAB units) in simulated experiments, and from 3.12 to 4.03 CIELAB units (average 3.55 CIELAB units) in real experiments. In real experiments, observers' variability reduced for light sources with high illuminance and high correlated color temperature. Spectral power distributions affected observer metamerism, but the role played by the primaries of the two displays employed was unclear.

Keywords: observer variability, observer metamerism, cross-media color experiment, color-matching.

1. Introduction

Observers with normal color vision may have different color perceptions because of the individual variability in their color-matching functions (CMFs). The variability among observers with normal color vision is also known as "observer metamerism", defined as "the property of specimens having different spectral characteristics and having the same color when viewed by one observer, but different colors when viewed by a different observer under the same conditions" [1]. The last few words in this defi-

nition are related to the problem considered in the current paper: the potential influence of different viewing conditions on the magnitude of observer metamerism.

The CIE 1931 and CIE 1964 standard colorimetric observers [2], also known as 2° and 10° (standard) observers, are based on the assumption that a single set of CMFs can reasonably represent a population with normal color vision. However, such an assumption may not be true, in particular for narrowband stimuli, and this failure may reduce accuracy in color reproduction [3–5]. In 1989, CIE proposed a "standard deviate observer" (SDO) to evaluate observer metamerism [6]. However, reportedly this SDO significantly underestimates observer variability [7, 8]. In 2006, the CIE also proposed a model (called CIEPO06) that provided cone fundamentals by specifying the observer's age and field size [9]. This model enables different theoretical observers to be generated and used to evaluate observer metamerism.

In 2010 and 2011, SARKAR *et al*. [10, 11] proposed different observers' categories by using a cluster analysis method, starting from 47 individual CMFs from STILES and BURCH [12] and 61 CMFs from CIEPO06. In a test experiment, 47 human observers were classified into nine categories (including the CIE 1964 standard colorimetric observer as one category) [13], requiring further investigation.

ASANO *et al*. [14] carried out a series of experiments to investigate the influence of CMFs on color discrimination. In their color-image-matching experiment, two different media (Apple Cinema HD LCD monitor and Microvision laser projector) were employed, and the results from 28 color-normal observers were analyzed. These researchers found an inter-observer variability (measured by the mean color difference from the mean, MCDM) in the range 3.2–4.4 CIEDE2000 units (roughly equivalent to 4.9–6.8 CIELAB units [15]), indicating that observer metamerism is quite high when a laser projector is used. Furthermore, using the STILES and BURCH CMFs [12], simulated color-matching experiments have been performed under different combinations of physical patches and monitors/projectors with different primaries [16]. The results showed that: *i*) the choice of spectra for the matching primaries had a significant effect on observer variability, *ii*) observer variability was large for near-neutral reference colors, and *iii*) observer variability in the lightness direction was smaller than in chromaticity directions. In a real color-matching experiment with a panel of 61 color-normal observers [16], the average inter-observer variability (measured by the MCDM) was 9.2 CIEDE2000 units (roughly equivalent to 14.2 CIELAB units [15]), spanning about 40 CIELAB units, which was much larger than reported in any previous experiment.

OICHERMAN *et al*. [17] studied observer variability of asymmetric metameric color matches using computer monitors and object color stimuli, in conditions typical for surface color industries, and reported that observer metamerism was small for all colors but neutrals, the CIE SDO [6] underestimating inter-observer variability.

To investigate the observer variability within normal color vision under a variety of viewing conditions, the current paper complements previous findings [11–14, 16, 17], by reporting the results of a set of cross-media color-matching experiments using displays and printed color samples. Specifically, we investigated the influence on observer

metamerism of light sources with different spectral power distributions (SPDs), different illuminances, and different correlated color temperatures (CCTs), as well as the use of displays with different sets of primary colors. The simulated results (*i.e.* theoretical results from CMFs) and the real results (*i.e.* experimental results achieved by real observers participating in our experiments) were compared.

2. Experiment

2.1. Setups

We performed two groups of cross-media color-matching experiments, where physical printed samples (12 × 12 cm) were placed in the back wall of a color-assessment cabinet, and matching samples with the same size were generated by using displays placed inside (hereafter called experiments I) or outside (hereafter called experiments II) of a color-assessment cabinet (see Fig. 1). The illuminated physical samples acted as references, and the task of the observers was to adjust the RGB channels in each display to achieve samples with a visual color match with respect to the references.

In experiments I, two different situations were analyzed: *i*) light sources (L1) with different illuminances, but very similar CCTs and SPDs; *ii*) light sources (L1, L2) with
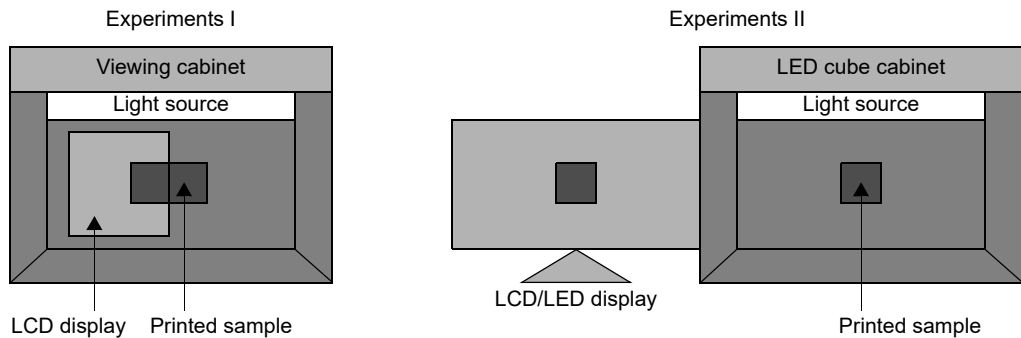

Fig. 1. Schemes of experimental setups.

T a b l e 1. Experimental information of the cross-media color-matching experiments.

| | Phase | Light sources for printed samples | | | | Displays for matching samples |
		Illuminance [lx]	Relative SPD	CCT [K]	R_a	
Experiments I	1	282	L1	6230	96.1	LCD
	2	522	L1	6302	95.9	LCD
	3	902	L1	6385	96.0	LCD
	4	922	L2	6364	94.1	LCD
Experiments II	1	585	L3	3619	94.1	LCD
	2	794	L4	6594	97.5	LCD
	3	585	L3	3619	94.1	LED
	4	794	L4	6594	97.5	LED

different SPDs, but very similar CCTs and illuminances. The LCD display employed in experiments I was an EIZO CG19. Specifically, in experiments I we developed four phases cross-media color-matching experiments (see Table 1). In experiments II, two different situations were also considered: the printed samples illuminated by two light sources with different CCTs (L3, L4), provided by a spectral tunable lighting system with 9 narrow-band and 2 broad-band LEDs [18], each one in comparison with samples provided by displays with two different sets of primaries (LCD-EIZO CG19 and LED-NEC PA242W). Therefore, four phases can be distinguished in cross-media color-matching experiments developed within experiments II (see Table 1). The dif-

T a b l e 2. Summary of the experiments with different viewing conditions.

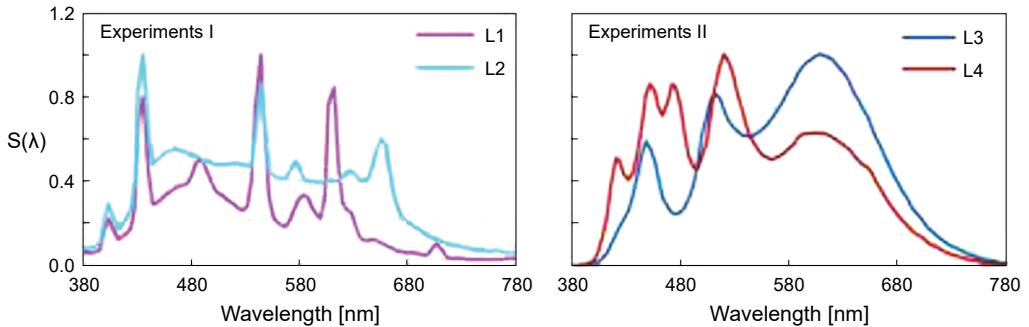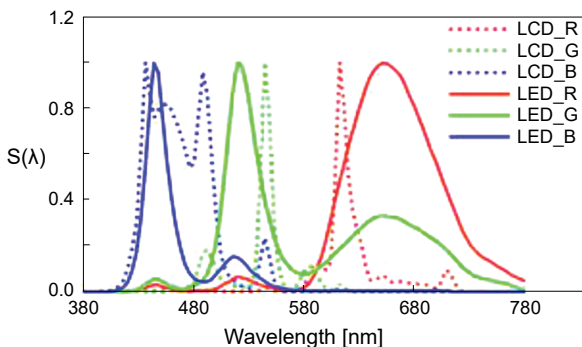| Experiments I	
Phase 1, 2 and 3	Phase 3 and 4
Different illuminance	Different SPD
Similar CCT and SPD	Similar CCT and illuminance
Experiments II	
Phase 1 and 2 (3 and 4)	Phase 1 and 3 (2 and 4)
Different CCT, illuminance, and SPD	Different displays
Identical displays	Identical CCT and illuminance, and similar SPD

Fig. 2. Normalized SPDs of the lighting sources.

Fig. 3. Normalized SPDs of the 3 primaries in the two displays (dotted line – LCD, solid line – LED).

ferent comparatives made from experimental results found in experiments I and II are summarized in Table 2.

The relative spectral power distributions of the lighting sources and the spectral curves of the primary colors of the displays used in the experiments I and II are shown in Figs. 2 and 3, respectively. All the light sources used in our experiments had CIE general color-rendering indices [19] above 94.1. Light sources L1 and L2 were selected because they are typical simulators of the main CIE illuminant (*i.e.* D65), available in commercial color-assessment cabinets currently employed in industry and research laboratories. Light sources L3 and L4 were selected as representatives of modern and solid-state light sources increasingly employed in our everyday life.

2.2. Reference color samples

CIE has recommended 17 color centers for industrial color difference evaluation [20]. For all phases in experiment I, we selected as reference colors 9 of these 17 centers (the remaining 8 centers had high chroma values outside the gamut of our displays). We produced physical samples with the colors of these centers (illuminant D65, CIE 1964 standard colorimetric observer) using an Epson Stylus Pro 7908 inkjet printer. For all phases in experiment II, to have a reasonable workload, we only considered 5 of these
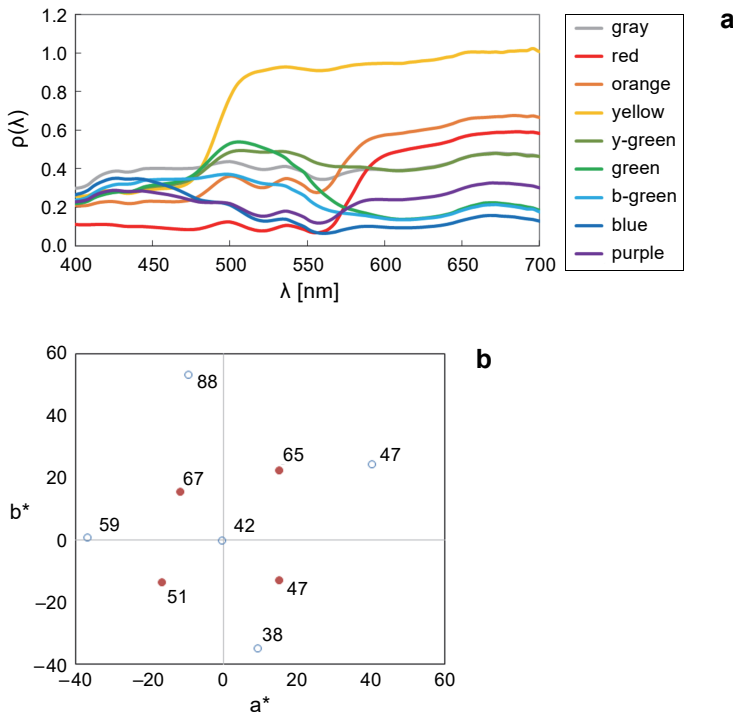

Fig. 4. Measured spectral reflectance (**a**) and CIELAB coordinates in experiment I, phase 3 (**b**) for the 9/5 printed samples employed in experiments I/II (solid/open symbols), with L^* values indicated by integers.

9 color centers. The spectral reflectance for each physical sample was measured by an X-Rite SpectroEye spectrophotometer with geometry 0°:45° (Fig. 4**a**), and the tristimulus values of the samples were calculated under the different light sources shown in Fig. 2, assuming the CIE 1964 standard colorimetric observer [2]. As an example, the CIELAB coordinates of the reference samples employed in phase 3 of experiment I are shown in Fig. 4**b**.

2.3. Observers

The four phases in experiments I were performed by 27–35 observers aged from 20 to 59 years old, up to 9 observers made replications, and there were nearly 700 color matches for each phase in experiments I. Four phases in experiments II were performed by 117–120 observers aged from 17 to 25 years old, 8 observers made replications, and there were over 670 color matches for each phase in experiments II. All observers participating in our experiments had normal color vision, and were students and teachers with some background in color science, because they had majored in printing engineering at Beijing Institute of Graphic Communication (BIGC) and had participated previously in similar color experiments. The experiments I and II were completed within a three-month period, the number of performed color matches being 2745 and 2720, respectively. So a total of 5465 matches were made in this paper.

2.4. Visual assessments

The cross-media color-matching experiments were conducted in a dark room, and the displays and viewing cabinet were allowed to warm-up for 90 min before the experiments started. It was found that the displays had a reasonable stable condition after 90 min, because at this time the change of CIELAB color differences was below 1.0 CIELAB unit. The height of each observer's eyes was adjusted so that they were perpendicular to the display and the distance to the color samples was about 80 cm. The experiments I and II were conducted in two stages. First, observers were asked to adjust the background of the display to match the gray board of the back wall of the color assessment cabinet under the selected light source. Second, observers were asked to perform the color matches with the illuminated reference samples, presented in random order, by modifying the intensities of the RGB channels in the display from 0 to 255, using the program interface. After each color match, we measured the color of the reference and matched samples using a PR-655 SpectraScan spectroradiometer, placed at the same position as the observers' eyes.

2.5. Observer metamerism indices for simulated and real experiments

In this section we will define the indices we have employed to measure observer metamerism in our simulated and real experiments. Specifically, a computational color-matching simulation was performed to examine the potential magnitude of observers' variability in the color-matching experiments with real observers. The simu-

lation workflow was the one proposed by ASANO *et al.* [16], using as reference samples the 9 (or 5) printed samples (Fig. 4), illuminated by the light sources in the cabinets (Fig. 2), and as matching colors those from primaries of a selected display (LCD or LED, Fig. 3). Moreover, the Stiles and Burch 10° CMFs for 47 observers were considered for these simulations, as representative of a population with normal color vision [12]. As a result of this simulation, the SPDs of matching colors were established from adjusted scalar factors of the 3 primaries in each display, for each of the 47 observers [4]. Then, assuming the CIE 1964 standard colorimetric observer, it was possible to compute the X_{10}, Y_{10}, Z_{10} tristimulus values, and CIELAB color differences between the reference and matched colors. On the other hand, in the so-called real experiments we computed CIELAB color differences, from the spectroradiometric measurements of the color coordinates of the two color stimuli matched by different real observers.

To quantify observers' variability (*i.e.* observer metamerism), for each color center (illuminated printed sample P) the next four indices were defined. The indices termed MCDM (mean color difference from the mean), OM_{max} (maximum observer metamerism) and OM_{mean} (mean observer metamerism), were previously proposed [16], and are defined as follows:

$$ \text{MCDM} = \frac{1}{N} \sum_{i = 1, N} \Delta E(V_i, V_{\text{ave}}) \tag{1} $$

where V_i and V_{ave} represent the CIELAB coordinates of matched color for a given observer i and for the average of all the observers, respectively, N is the number of observers, and ΔE indicates the CIELAB color-difference, $\Delta E^*_{ab, 10}$. The mean color difference from the mean (MCDM) is often used to evaluate measurement precision, higher MCDM values indicating larger observer variability in the current paper. Next,

$$ \text{OM}_{\text{max}} = \max(\overline{\Delta E_{P, i}}) \tag{2} $$

$$ \text{OM}_{\text{mean}} = \overline{\Delta E_{P, i}} \tag{3} $$

where $\Delta E_{P, i}$ is the mean CIELAB color difference between the 9 (or 5) printed samples and their matched colors for a given observer i, and the word "max" in Eq. (2) and the overline in Eqs. (2) and (3) mean the maximum and the average values of such color differences considering all the observers, respectively. In simulated experiments, OM_{max} and OM_{mean} indicate the maximum and average observer metamerism from STILES and BURCH CMFs [12], averaging our 9 (or 5) color centers. Finally, CIELAB color coordinates for a reference stimulus and simulate matched colors for each observer i were computed and fitted to ellipses in the a^*b^* plane, using covariance analysis and 90% statistical significance (Fig. 5). The areas of these ellipses (average of all centers) were considered to be another appropriate index of observer metamerism, named OM_{area},

$$ \text{OM}_{\text{area}} = \text{Area}(a^*, b^*)_P \tag{4} $$

3. Results and discussion

3.1. Observers' variability in real experiments

As stated in the previous section, over 8 observers made replications of the experiments, allowing the computation of intra-observer variability from results of each observer's replication and the mean of all replications made by that observer. The intra-observer variability in an experiment is defined as the average intra-observer variability of all the observers performing replications of that experiment. On the other hand, the inter-observer variability for a given experiment is defined as the average deviation between average results of individual observers and the mean results from all the observers. Table 3 shows the intra- and inter-observer variability, in terms of CIELAB color differences, for each of our experiments. From the average of the 8 experiments, the inter-observer variability was 3.55 CIELAB units, which is larger than the intra-observer variability (2.23 CIELAB units), and in agreement with results found in previous similar color-matching experiments [14, 17].

3.2. Simulated and real results

To quantify observers' variability (*i.e.* observer metamerism), we employed the average color difference of the 9 (or 5) values provided by Eqs. (1)–(4) for each color center P. Table 4 shows the results of simulated and real color-matching experiments, with the

T a b l e 3. Intra- and inter-observer variability in terms of CIELAB color-difference values for each of our experiments.

	Experiment I				Experiment II			
	1	2	3	4	1	2	3	4
Intra-observer	2.94	2.46	2.32	2.62	2.02	1.83	1.94	1.71
Inter-observer	4.03	3.79	3.12	3.67	3.69	3.12	3.71	3.27

T a b l e 4. Results of simulated and real color-matching experiments (average of all color centers), using observer metamerism indices defined in Eqs. (1)–(4).

		Simulated experiments				Real experiments			
		MCDM	OM_{max}	OM_{mean}	OM_{area}	MCDM	OM_{max}	OM_{mean}	OM_{area}
Experiments I	1	0.74	1.48	0.76	0.86	<u>4.03</u>	<u>8.07</u>	<u>4.62</u>	<u>3.67</u>
	2	0.75	1.51	0.77	0.87	3.79	7.86	4.18	3.15
	3	**0.73**	**1.46**	**0.76**	**0.86**	**3.12**	**6.82**	**3.44**	**2.95**
	4	<u>0.82</u>	<u>1.60</u>	<u>0.86</u>	<u>1.02</u>	3.67	7.25	4.04	3.34
Experiments II	1	**0.76**	2.97	**1.82**	**0.95**	3.69	<u>9.53</u>	<u>5.74</u>	<u>3.67</u>
	2	0.84	**2.61**	1.88	1.00	**3.12**	**7.17**	**4.33**	**2.89**
	3	<u>1.64</u>	<u>4.80</u>	1.69	1.61	<u>3.71</u>	8.24	5.00	3.65
	4	1.60	3.89	<u>2.09</u>	<u>1.87</u>	3.27	7.37	4.84	3.20

arithmetical mean of values of indices MCDM, OM_{max}, OM_{mean}, and OM_{area} (see Eqs. (1)–(4)), considering the 9 (or 5) centers. The minimum/maximum values within experiments I and II were distinguished by using bold/underlined fonts in Table 4. The results for MCDM and real experiments in Table 4 are identical to the inter-observer variability shown in the last row of previous Table 3.

Table 4 shows pronounced differences between simulated and real results, in such a way that observer metamerism in real experiments is considerably higher than in simulated experiments, in agreement with previous literature [16]. An explanation of this result is that in our previous simulations we considered only the observer metamerism caused by the variability in a given set of CMFs. However, in real color-matching experiments, in addition to the variability of CMFs of real observers, other additional factors like experimental errors inherent to the color-matching task performed by observers, or specific viewing conditions, may also play a role. For example, it can be noted that in experiment I the separation between the samples (gap) was lower than in experiment II, although it can be considered that observers performed a "simultaneous color-matching task" in both experiments and it was not reported by the observers that experiment II was more difficult than experiment I. However, the separation between the samples may have induced different adaptations states in the observers performing the color-matching experiments, and may have also generate different degrees of accuracy in the color-matching task, because, as reported in previous literature, the separation between samples reduces visual sensitivity to color differences [21, 22]. Unfortunately, from current results it is not possible to quantify the influence of these effects. The viewing conditions had different effects: in experiments I, the phase 4 had the largest observer metamerism for simulated experiments, while for real experiments, the phase 1 had the largest observer metamerism. Similarly, in simulated experiments II, the phase 3 and 4 (both with the same display primaries) had the largest observer metamerism, while this was not true for real experiments. In real experiments, the inter-observer variability ranged from 1.18 to 7.70 and from 1.53 to 7.31 CIELAB units in experiments I and II, respectively, with average values shown in Table 3. These ranges are wider than those reported by Asano *et al*. in their color-image matching experiments [14] (around 4.9–6.8 CIELAB units), although in another color-matching experiment, these authors also reported a substantially higher value around 14.2 CIELAB units for a neutral color [16].

As stated before, the OM_{max} and OM_{mean} indices are the maximum and average CIELAB color differences between a reference sample (illuminated printed sample) and the color match with this sample provided by different observers using specific visual displays. For more details concerning these CIELAB color differences, Table 5 shows the percentage of observers registering color differences below three specific tolerance limits (3.0, 5.0 and 7.0 CIELAB units) in each of our experiments. From results in Table 5, in experiments I and II, the highest color-matching accuracies were found in the phase 3 of experiment I and the phase 2 of experiment II, respectively,

Table 5. Percentages (%) of observers achieving color matches bellow three different CIELAB values.

	Phase	CIELAB < 3.0	CIELAB < 5.0	CIELAB < 7.0
	1	3.9	59.7	94.8
Experiments I	2	20.5	74.0	97.3
	3	39.7	96.2	100.0
	4	20.8	84.4	96.1
	1	0.0	29.4	81.6
Experiments II	2	5.8	75.9	98.5
	3	2.2	53.0	97.0
	4	1.5	55.9	99.3

with 39.7%/96.2% and 5.8%/75.9% of the observers achieving values below 3.0/5.0 CIELAB units.

3.3. Chromaticity ellipses and lightness differences

From color-matching results in each of the 8 experiments performed by real observers, the average size of fitted a^*b^* ellipses at different centers was the metamerism index OM_{area} in previous Eq. (4). However, the OM_{area} index provides partial information, lacking the orientation and shape of the fitted ellipses, which are also useful. Figure 5 allows a comparison of the orientations, shapes, and sizes of the ellipses found in our visual experiments at the different centers. Figure 5 also shows the coordinates of the 4/2 reference samples used in experiments I/II at each color center (see Table 1 and Fig. 4), which are relatively close to the centers of the ellipses. Anyway, some differences between reference stimuli and color-matching ellipses have been reported in previous literature using visual colorimeters [23], although the reasons of such phenomenon, generically qualified as an "additivity failure", are not well understood.
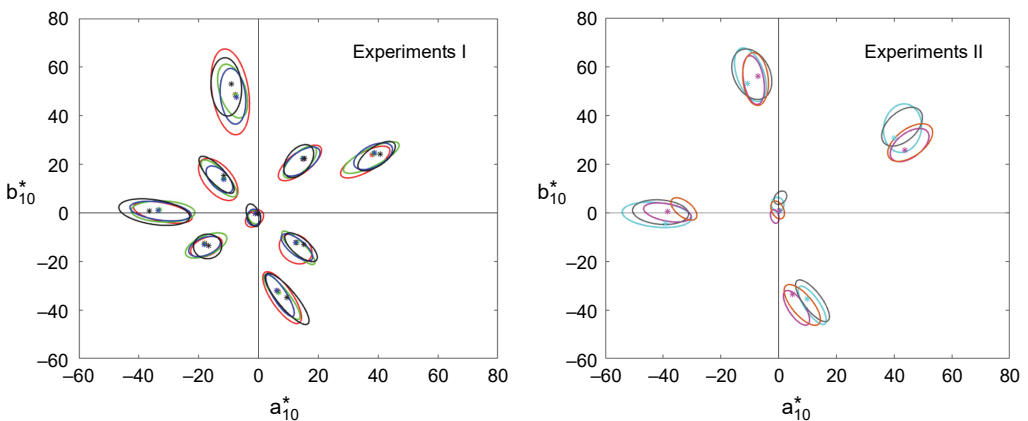

Fig. 5. Experimental 90% confidence color-matching ellipses fitted to results in our real experiments I-1 (red), I-2 (green), I-3 (blue), I-4 (black), II-1 (cyan), II-2 (magenta), II-3 (orange) and II-4 (gray). The dots represent the coordinates of the reference samples in our experiments.

As we can see (Fig. 5), the orientation, shape, and size of the ellipses in experiments I and II follow similar trends. Except at the gray center, where ellipses were close to circles and with smallest size, the major axes of the ellipses are oriented approximately in radial directions, and have higher values at the most chromatic centers. In fact, the ellipses shown in Fig. 5 agree with conventional color discrimination ellipses in a^*b^* reported in previous literature [24–26], which is not surprising, because current color-matching experiments can be considered a particular case of color-discrimination experiments.

Because chromaticity ellipses in Fig. 5 just focus on observers' variability in the a^*b^* plane, this piece of information must be complemented by considering the results found in the third dimension L^*. In ASANO *et al.* [16], the variability in L^* was minor in their simulation experiments, so that they made L^* set constant for their color-matching visual experiments, and their observers only adjusted the a^* and b^* values. For comparison of the magnitude of lightness and chromaticity differences in our simulated and real color-matching experiments, Fig. 6 shows CIELAB coordinates using the 47 STILES and BURCH CMFs [12] considered at the simulation experiment (green points), and the visual results achieved by real observers in phase 4 of experiment II (blue points), for
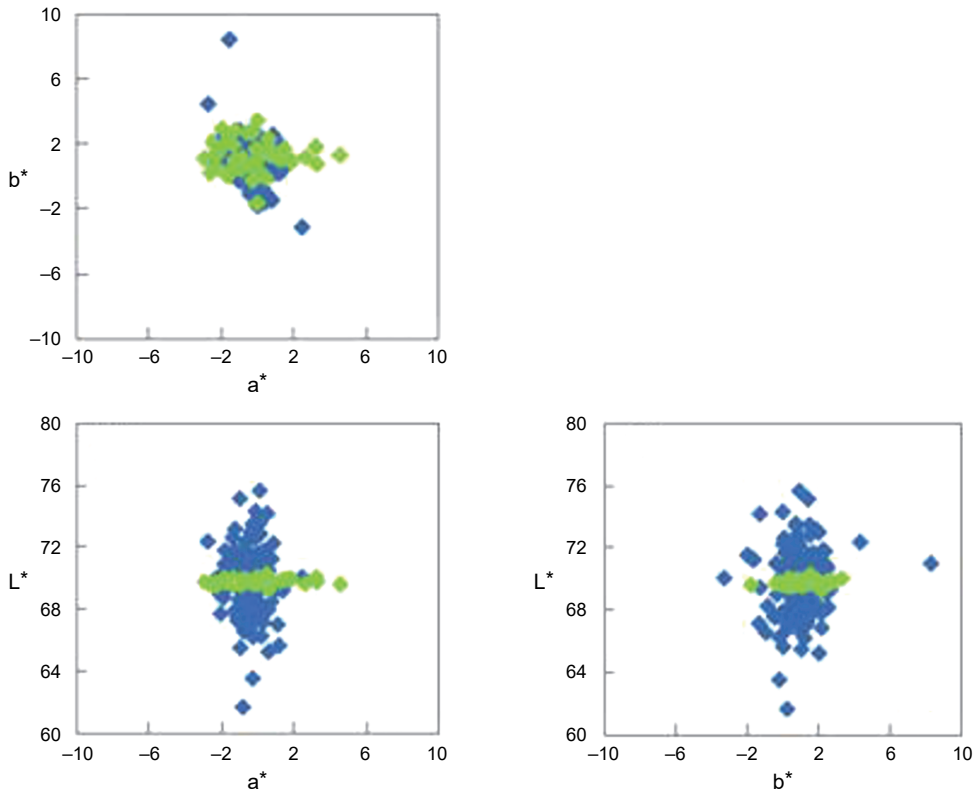

Fig. 6. CIELAB plots (a^* *vs.* b^*, a^* *vs.* L^*, and b^* *vs.* L^*) of color matches achieved for the gray color center using the CMFs in [12] (green dots), and the visual results from observers in experiment II-4 (blue dots).

the gray color center. For easier comparison, in the plots in Fig. 6 we used the same range (20 CIELAB units) for all axes. Figure 6 shows that results from individual observers have large discrepancies. More specifically, we can see that in simulated experiments the variability in the a^* and b^* directions was much larger than the one in the L^* direction, in agreement with results from ASANO *et al.* [16], using also the CMFs of the 47 STILES and BURCH observers [12] and a gray color center. This result can be explained by considering that the differences among the $\bar{y}(\lambda)$ CMFs in [12] are very small, which leads to almost identical L^* values in the simulations. However, in our experiment with real observers the variability in the L^* direction cannot be disregarded, because it is even larger than the one in the a^* and b^* directions (see blue dots in the bottom plots in Fig. 6). This result can be attributed to the fact that in our color-matching experiments the real observers may have different CMFs, and factors related to general color appearance and uncertainty in the visual task may have also played a role, as explained in previous Section 3.2.

3.4. Effect of different viewing conditions

Light sources with different illuminances, but similar correlated color temperature and SPDs, were used in phases 1, 2, 3 of experiment I. From Table 4, the light source with the highest illuminance (phase 3 of experiment I) resulted in the lowest MCDM values (3.12), OM_{max} (6.82), OM_{mean} (3.44), and OM_{area} (2.95), and phase 1 of experiment I, with the lowest illuminance, led to the largest observer metamerism values. We can conclude that light sources with high/low illuminances reduce/increase observer metamerism.

The comparison of results in phases 3 and 4 of experiment I allows the study of the influence of light sources with only different SPDs on observer metamerism. Table 4 shows that light sources with different SPDs cause different observer metamerism, because in phase 3 of experiment I, the values are lower than those in phase 4 of experiment I. Thus, we can conclude that light sources with different SPDs alter the observer's metamerism of the same physical samples.

The results from phases 1 and 2 (or phases 3 and 4) of experiment II involve light sources with different CCTs, but relatively similar illuminances and identical displays. From results in Table 4, light sources with high CCTs (phases 2 or 4 of experiment II) led to lower observer metamerism values than those with low CCTs (phases 1 or 3 of experiment II).

A comparison of results in phases 1 and 3 of experiment II (or phases 2 and 4 of experiment II) indicates the influence of the primaries of the displays (Fig. 3) on observer metamerism. Our results for real experiments indicated slightly higher observer metamerism for the LED than for the LCD employed, which is inconsistent with the results reported for simulation in ASANO *et al.* [16]. Figure 3 shows that the bandwidths of the R and G channels in our LCD display (phases 1 and 2 of experiment II) were narrower than those in our LED display (phases 3 and 4 of experiment II), the opposite happening for the B channels of these displays.

4. Conclusions

Simulated and real cross-media color-matching experiments were conducted under different viewing conditions, and observer metamerism was measured by using the MCDM, OM_{max}, OM_{mean} and OM_{area} indices as well as considering a^*, b^* fitted ellipses. The main conclusions are as follows:

1) The results in simulated and real experiments are quite different, which can be partly explained by considering that in real experiments other different factors than the CMFs of the observers may also be relevant.

2) Light sources producing high/low illuminances or with high/low CCTs reduce/increase observer metamerism. The SPDs of the primary colors of the displays employed have no obvious influence on magnitude of observer metamerism, perhaps because the difference between the SPDs of our light sources is not large enough.

3) The observer metamerism in our experimental visual results can be expressed by fitted a^*, b^* 90% confidence ellipses, which show very similar trends to those in color-discrimination experiments reported in previous literature [24–26].

References

[1] *ASTM. Standard Terminology of Appearance*, E 284 – 95a, 1995, pp. 246.

[2] *CIE, Colorimetry (4th Edition)*, CIE Publication 015:2018, Central Bureau of the CIE, Vienna, 2018.

[3] FAIRCHILD M.D., WYBLE D.R., *Mean observer metamerism and the selection of monitor primaries*, [In] *Final Program and Proceedings IS&T/SID Color Imaging Conference*, 2007, pp. 151–156.

[4] RAMANATH R., *Minimizing observer metamerism in display systems*, Color Research and Application **34**(5), 2009, pp. 391–398, DOI: 10.1002/col.20523.

[5] BURGOS-FERNÁNDEZ F.J., VILASECA M., PERALES E., HERRERA-RAMÍREZ J.A., MARTÍNEZ-VERDÚ F.M., PUJOL J., *Spectrally tunable light source based on light-emitting diodes for custom lighting solutions*, Optica Applicata **46**(1), 2016, pp. 117–129, DOI: 10.5277/oa160111.

[6] *CIE, Special metamerism index: Change in observer*, CIE Publication 80, Central Bureau of the CIE, Vienna, 1989.

[7] RICH D.C., JALIJALI J., *Effects of observer metamerism in the determination of human color-matching functions*, Color Research and Application **20**(1), 1995, pp. 29–35, DOI: 10.1002/col.5080200106.

[8] ALFVIN R.L., FAIRCHILD M.D., *Observer variability in metameric color matches using color reproduction media*, Color Research and Application **22**(3), 1997, pp.174–188, DOI: 10.1002/(SICI)1520 -6378(199706)22:3<174::AID-COL6>3.0.CO;2-O.

[9] *CIE, Fundamental Chromaticity Diagram with Physiological Axes–Part I*, CIE 170-1:2006, Central Bureau of the CIE, Vienna, 2006.

[10] SARKAR A., BLONDÉ L., LE CALLET P., AUTRUSSEAU F., MORVAN P., STAUDER J., *Toward reducing observer metamerism issue in industrial applications: colorimetric observer categories and observer classification*, [In] *Color and Imaging Conference, 18th Color and Imaging Conference Final Program and Proceedings*, 2010, pp. 307–313.

[11] SARKAR A., *Identification and Assignment of Colorimetric Observer Categories and Their Applications in Color and Vision Sciences*, PhD Thesis, Université de Nantes, 2011.

[12] STILES W.S., BURCH J.M., *N.P.L. colour-matching investigation: final report*, Journal of Modern Optics **6**(1), 1959, pp. 1–26, DOI: [10.1080/713826267](10.1080/713826267).

[13] MORVAN P., SARKAR A., STAUDER J., BLONDÉ L., KERVEC J., *A handy calibrator for color vision of a human observer*, 2011 IEEE International Conference on Multimedia and Expo, pp. 1–4, DOI: 10.1109/ICME.2011.6012090.

[14] ASANO Y., FAIRCHILD M.D., BLONDÉ L., MORVAN P., *Observer variability in color image matching on a LCD monitor and a laser projector*, [In] *Color and Imaging Conference, 22nd Color and Imaging Conference Final Program and Proceedings*, 2014, pp. 1–6.

[15] MELGOSA M., CUI G., OLEARI C., PARDO P.J., HUANG M., LI C., LUO M.R., *Revisiting the weighting function for lightness in the CIEDE2000 colour-difference formula*, Coloration Technology **133**(4), 2017, pp. 273–282, DOI: 10.1111/cote.12294.

[16] ASANO Y., FAIRCHILD M.D., BLONDÉ L., MORAVAN P., *Color matching experiment for highlighting interobserver variability*, Color Research and Application **41**(5), 2016, pp. 530–539, DOI: [10.1002/col.21975](10.1002/col.21975).

[17] OICHERMAN B., LUO M.R., RIGG B., ROBERTSON A.R., *Effect of observer metamerism on colour matching of display and surface colours*, Color Research and Application **33**(5), 2008, pp. 346–359, DOI: [10.1002/col.20429](10.1002/col.20429).

[18] http://www.thouslite.com/PRODUCTS/16.html (accessed August 15, 2018).

[19] *CIE, Method of measuring and specifying colour rendering properties of light sources*, CIE13.3-1995, CIE, Central Bureau Vienna, Austria, 1995.

[20] WITT K., *CIE guidelines for coordinated future work on industrial colour-difference evaluation*, Color Research and Application **20**(6), 1995, pp. 399–403, DOI: [10.1002/col.5080200609](10.1002/col.5080200609).

[21] WITT K., *Parametric effects on surface color-difference evaluation at threshold*, Color Research and Application **15**(4), 1990, pp. 189–199, DOI: [10.1002/col.5080150404](10.1002/col.5080150404).

[22] SHING–SHENG GUAN, LUO M.R., *Investigation of parametric effects using small colour differences*, Color Research and Application **24**(5), 1999, pp. 331–343, DOI: 10.1002/(SICI)1520-6378(199910)24:5<331::AID-COL5>3.0.CO;2-9.

[23] HITA E., JIMÉNEZ DEL BARCO L., ROMERO J., *Differential color thresholds from metameric matches: experimental results concerning failures of colorimetric additivity*, Journal of the Optical Society of America A **3**(8), 1986, pp. 1203–1209, DOI: 10.1364/JOSAA.3.001203.

[24] HUANG M., LIU H., CUI G., LUO M.R., MELGOSA M., *Evaluation of threshold color differences using printed samples*, Journal of the Optical Society of America A **29**(6), 2012, pp. 883–891, DOI: 10.1364/JOSAA.29.000883.

[25] MELGOSA M., HITA E., ROMERO J., JIMÉNEZ DEL BARCO L., *Color-discrimination thresholds translated from the CIE (x, y, Y) space to the CIE 1976 (L^*, a^*, b^*)*, Color Research and Application **19**(1), 1994, pp. 10–18, DOI: [10.1111/j.1520-6378.1994.tb00054.x](10.1111/j.1520-6378.1994.tb00054.x).

[26] LUO M.R., RIGG B., *Chromaticity-discrimination ellipses for surface colours*, Color Research and Application **11**(1), 1986, pp. 25–42, DOI: [10.1002/col.5080110107](10.1002/col.5080110107).