Research Article Vol. 30, No. 10 / 9 May 2022 / Optics Express 16973
Optics EXPRESS

Effect of observer age and stimulus size on the performance of CIE color matching functions

MIN HUANG,[1] YU WANG,[1] MINCHEN WEI,[2,*] YU LI,[1] XINYUAN GAO,[1] AND XIU LI[1]

[1] *Printing and Packaging Engineering, Beijing Institute of Graphic Communication, Beijing, China*
[2] *Color, Imaging, and Illumination Laboratory, The Hong Kong Polytechnic University, Kowloon, Hong Kong*
[*] *minchen.wei@polyu.edu.hk*

Abstract: Both observer age and size of stimulus as characterized using the field of view (FOV) are two important parameters to affect the color matching functions (CMFs) of human observers. They are also included in the cone fundamental and CMFs models that were recently proposed by the International Commission on Illumination (CIE) in 2006. In contrast to the great number of studies investigating the performance of CMFs in characterizing color matches and mismatches using different primary sets, few study investigated the effect of CMFs in characterizing these two factors. In this study, we carefully designed a series of test stimuli in five colors, which had different magnitudes of calculated color difference to the corresponding reference stimulus using the CIE 2006 CMFs model. The stimuli were presented in two FOVs (i.e., 8.6° and 2.9°). A group of young and senior observers were asked to judge which of the two test stimuli appeared to be similar to the reference stimulus, which was a forced choice. The color differences calculated using the CIE 1931 2° and 1964 10° CMFs were found to have higher correlations to the visual color differences judged by the senior and young observers respectively, regardless of the actual FOVs used in the experiment. In addition, though the CIE 2006 CMFs with the different parameter settings always had better performance in predicting the perceived color differences, they failed to characterize the effect of observer age and stimulus size. Also, the experiment results did not support the CIE's recommendation of using the 10° CMFs for stimuli with an FOV greater than 4° and the 2° CMFs for those smaller than 4°.

1. Introduction

Color matching functions (CMFs) are the fundamental of colorimetry and are used in all colorimetric calculations, such as tristimulus values and color difference values. The set of CMFs describes how color stimuli with different spectra compositions match in color appearance under a same viewing condition. The most widely used CMFs—the CIE 1931 2° CMFs—were developed based on two color matching experiments that were carried out by Wright [1,2] and Guild [3] using stimuli with a field of view (FOV) around 2°. Later, Stiles and Burch [4] and Speranskaya [5] carried out color matching experiments using a 10° bipartite setup, leading to the CIE 1964 10° CMFs. The CIE recommends using the CIE 1931 2° CMFs for the stimuli with an FOV between about 1° and about 4° and the CIE 1964 10° CMFs for those with an FOV greater than 4° [6]. A CIE technical committee TC1-36 developed a model to derive cone fundamentals for different FOVs between 1° and 10° and different ages in 2006 [7], and later derived the corresponding CMFs and chromaticity diagrams in 2015, with the 2° and 10° CMFs provided directly for easy use in practice [8].

With the development of LED lighting and display technologies, the performance of various CMFs have been carefully investigated in the last 10 years. For example, Borbely and Schanda asked 10 observers between 20 and 30 years of age to perform color matching experiments using a 2° bipartite setup by adjusting two sets of LED primaries [9]. Hu and Houser asked 10 observers

#457493 https://doi.org/10.1364/OE.457493
Journal © 2022 Received 2 Mar 2022; revised 5 Apr 2022; accepted 16 Apr 2022; published 2 May 2022

Optics EXPRESS

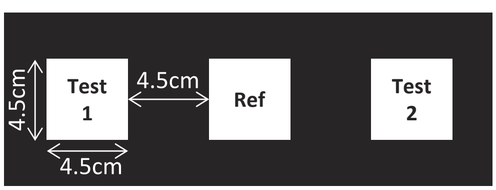
including the two authors and eight university students, with a mean age of 32.3, to complete a Maxwell color matching experiment using two FOVs (i.e., 10° bipartite and 102° × 50° bipartite) with a fixed primary set [10]. In Asano et al., 61 observers between 20 and 52 completed the color matching experiment using a bipartite field of 8.5° with different LED primary sets [11]. In Xie et al., 26 observers between 22 and 55 completed a color matching experiment using different displays with the stimuli having an FOV of 8° [12]. In Li et al., 54 observers including eight observers older than 45 years, with a mean of 33 years of age performed a series of achromatic color-matching experiments with a 10° stimulus using eight different narrowband primaries [13]. In our early study, we asked 50 and 53 observers between 19 and 38 years of age to match the color appearance of stimuli with an FOV of 4.77° and 20.2°, produced by different smartphone displays [14,15]. It is clear that most studies focused on how primary sets affected the performance of CMFs in predicting color matches and the degree of observer metamerism, and only two studies specifically investigated the effect of FOVs [10,14,15]. In contrast, the age of observer was never considered as an independent variable to investigate its effect, though it is really an important parameter and is modeled in the CIE 2006 model. In our recent work, we investigated the perceived color matches and mismatches of surface color samples between young and senior observers with an FOV of 11.4° and found significant differences [16].

In this study, we carefully investigated the performance of the CIE 2006 model in characterizing the effect of observer age and FOV using LED sources. Two levels of FOVs (i.e., 8.6° and 2.9°) and two groups of observers (i.e., young and senior) evaluated pairs of color stimuli that were carefully designed based on the CIE 2006 model.

2. Methods

2.1. Apparatus

Three spectrally tunable LED panels, with the light emitting area of 5 cm × 5 cm, were used to produce stimuli in this experiment. A black frame having a 4.5 cm × 4.5 cm opening was used to cover each panel, so that the actual light stimuli were all 4.5 cm × 4.5 cm. The three panels were placed next to each other, with the distance between the adjacent edges being 4.5 cm. Figure 1 shows the illustration and the photograph of the apparatus. During the experiment, the observer was seated 30 cm in front of the LED panels, with his or her sagittal plane aligned to the center of the reference panel, so that each stimulus had an FOV around 8.6°.



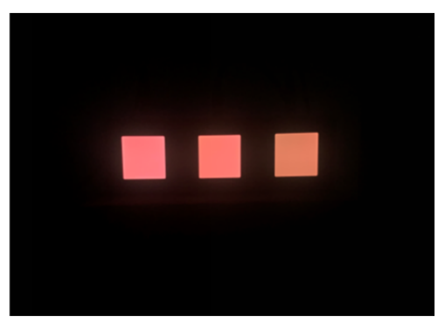


Fig. 1. Illustration and photograph of the apparatus.

The panel at the center was used to produce the reference stimulus, while the other two panels were used to produce the test stimuli. In particular, the panels producing the test stimuli had six primaries, including two sets of red (R1 and R2), green (G1 and G2), and blue (B1 and B2), and the panel producing the reference stimulus had three primaries (i.e., R1, G2, and B1) whose peak wavelengths were generally around 10 nm shorter than those used in Stiles

Optics EXPRESS

and Burch (i.e., 645.2 nm, 526.3 nm, and 444.4 nm). Table 1 summarizes the colorimetric characteristics of the primaries, with Fig. 2 showing the spectral power distributions (SPDs) measured using a PhotoResearch PR-655 spectroradiometer. The LED panels were used in our previous experiments and the stability and uniformity were characterized in our past work [17].

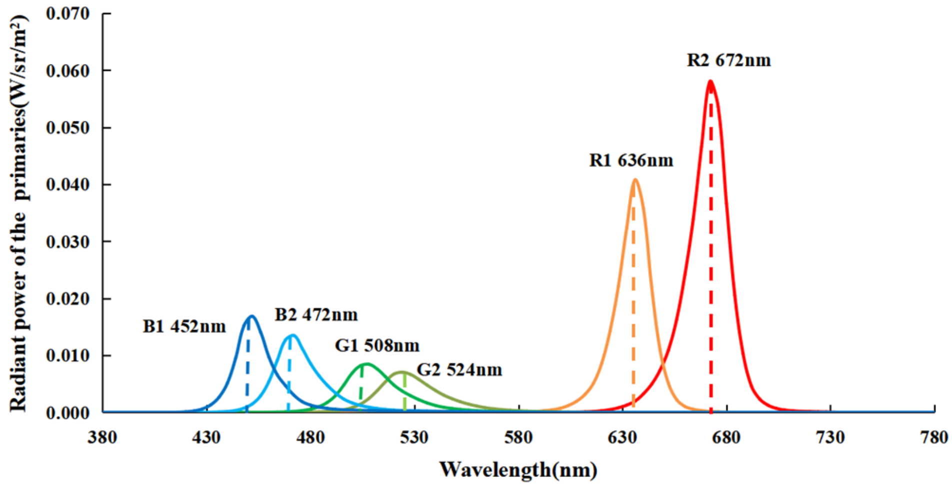


Fig. 2. SPDs of the primaries used in the LED panels, as measured using a PhotoResearch PR655 spectroradiometer.

Table 1. Summary of the colorimetric characteristics of the primaries derived from the measured SPDs (note: the chromaticities were calculated using the CIE 1964 10° CMFs).

Primary	Peak wavelength (nm)	FWHM (nm)	x_{10}	y_{10}	u'_{10}	v'_{10}
R1	636	20	0.694	0.306	0.525	0.521
R2	672	20	0.714	0.286	0.572	0.514
G1	508	28	0.116	0.705	0.041	0.565
G2	524	32	0.236	0.719	0.085	0.580
B1	452	20	0.146	0.064	0.168	0.165
B2	472	21	0.111	0.170	0.092	0.318

2.2. Color stimuli

The six primaries of the two testing panels were organized into six primary sets (i.e., L1 to L6), none of which was identical to the primary set of the reference panel. These primary sets were carefully selected to introduce different degrees of metamerism based on calculations. In particular, they were able to gradually change the magnitude of the calculated color differences between the test and reference stimuli for both young and senior observers. Five reference stimuli (i.e., grey, red, yellow, green, and blue) recommended by the CIE were designed and calibrated on the reference panel, with Fig. 3 showing the measured SPDs.

For each of the five reference stimuli, six test stimuli were carefully designed using the six primary sets, so that 30 stimuli were produced and the chromaticity distances $\Delta u'v'$ to the reference stimuli increased for a 20-year observer but decreased for a 70-year observer, as calculated using the CIE 2006 8.6° CMFs (i.e., the FOV was set to 8.6° and the age was set to 20 and 70). Table 2 summarizes the primary sets and test stimuli; Table 3 summarizes the chromaticities ($u'_{8.6}$,$v'_{8.6}$) of the reference stimuli calculated using the 20- and 70-year CIE 2006

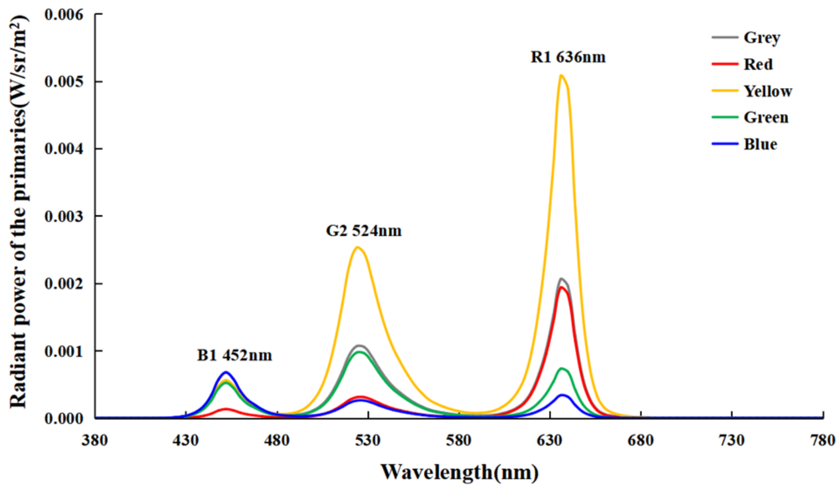


Fig. 3. SPDs of the five reference stimuli used in the LED panels, as measured using the PhotoResearch PR655 spectroradiometer.

CMFs. In particular, the $u'_{8.6}$ values of the reference stimuli using the 20-year CIE 2006 CMFs were always smaller than those using the 70-year CIE 2006 CMFs, while the $v'_{8.6}$ values were in opposite. Figure 4 shows the chromaticity shifts of the test stimuli from the corresponding reference stimulus. It can be observed that the chromaticities of the stimuli using the 20-year CIE 2006 CMFs were all shifted towards the -u'+v' direction, while those using the 70-year CIE 2006 CMFs were all shifted towards the + u'-v' direction. Figure 5 shows the chromaticity distances $\Delta u'_{8.6}v'_{8.6}$ from the test stimuli to the reference stimuli. All the test stimuli were calibrated on both test panels.

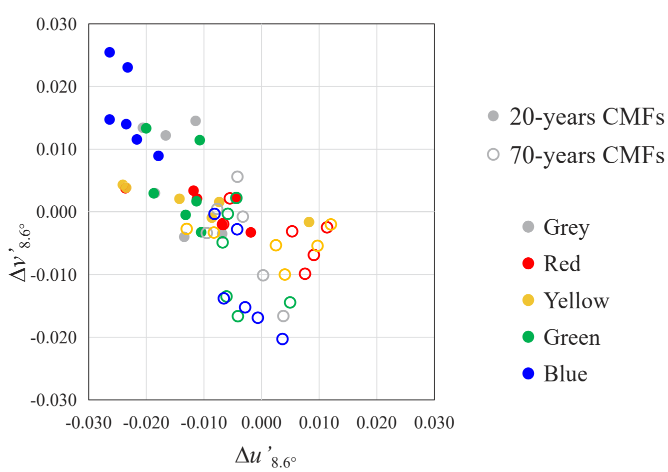


Fig. 4. Chromaticity shifts $\Delta u'$ and $\Delta v'$ between the test and reference stimuli, which were derived using the CIE 2006 8.6° CMFs for a 20- or 70-year observer.

2.3. Experimental procedures

Upon arrival, the observer completed a general information survey and the Ishihara Color Vision Test, and then was seated in front of the three panels with his or her sagittal plane aligned to

Optics EXPRESS

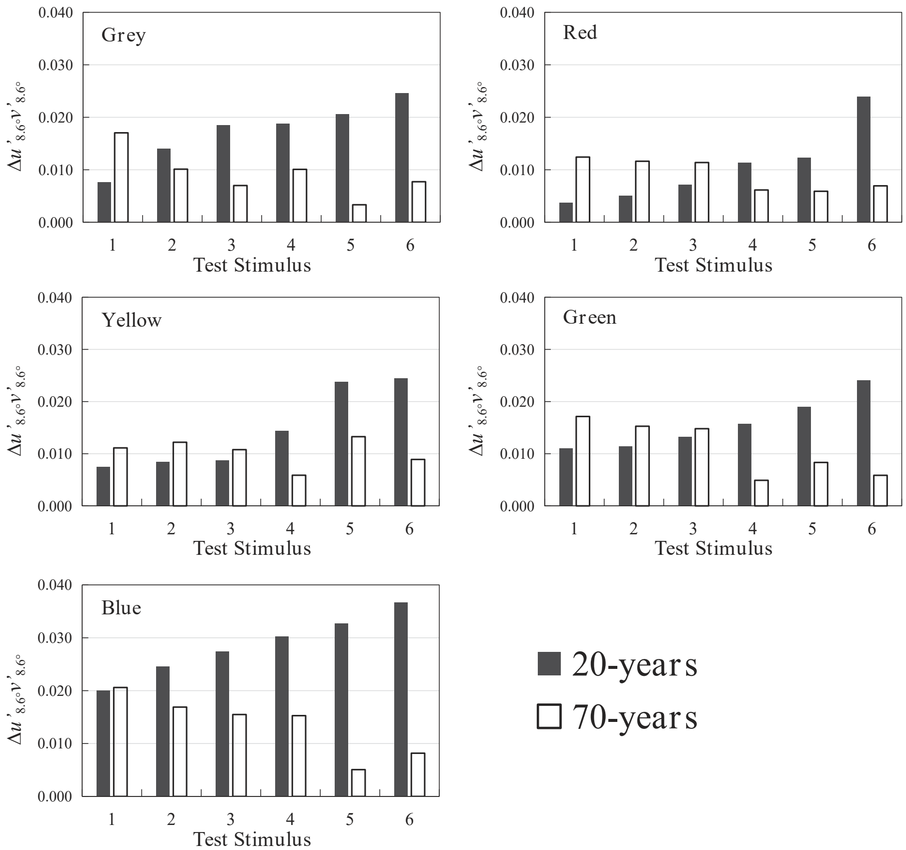


Fig. 5. Chromaticity differences $\Delta u'_{8.6}v'_{8.6}$ between the test and reference stimuli, which were calculated using the CIE 2006 8.6° CMFs for a 20- or 70-year observer.

Table 2. Summary of the primary sets used to produce the 30 test stimuli, with six stimuli for each of the five reference stimuli. (note: the primaries that were different from those used in the reference panel are underlined).

Primary sets	Primaries	Number of test stimuli produced using the primary set				
		Grey	Red	Yellow	Green	Blue
L1	R2-G2-B1		1	1		
L2	R1-G1-B1	1		1		
L3	R1-G2-B2				1	
L4	R1-G1-B2	2		1	2	2
L5	R2-G2-B2	2	1		2	4
L6	R2-G1-B1	1	4	3	1	

Table 3. Chromaticities ($u'_{8.6}$,$v'_{8.6}$) of the reference stimuli calculated using the 20- and 70-year CIE 2006 CMFs

Reference stimuli	($u'_{8.6}$,$v'_{8.6}$) 20-year CMFs	($u'_{8.6}$,$v'_{8.6}$) 70-yaer CMFs
Grey	(0.1988,0.5045)	(0.2050,0.4969)
Red	(0.3138,0.5220)	(0.3206,0.5202)
Yellow	(0.2071,0.5414)	(0.2141,0.5367)
Green	(0.1459,0.5118)	(0.1446,0.5019)
Blue	(0.1692,0.3915)	(0.1730,0.3691)

Optics EXPRESS

the center of the reference panel. The general illumination in the experiment space was then switched off, so that the observer only saw the three stimuli in his or her field of view. For each of the five reference stimuli, all possible pairs of the six test stimuli were presented using the two test panels, with a total of 15 pairs of test stimuli. The locations of the two stimuli within each pair and the order of the 15 pairs were randomized. For each pair of the test stimuli, the observer was asked to compare their color appearance to the reference stimulus, and to judge which one had a larger color difference to the reference stimulus, which was a forced choice. Then the stimuli were switched to the next pair, during which the observer's eyes were covered by an eye mask. The evaluations of all the stimuli pairs were repeated 12 times for evaluating the intra-observer variations and for deriving the probability statistics. The order of the stimuli pairs was randomized, and the locations of the stimuli within each pair were also randomized.

2.4. Observers

Sixty observers, including 30 young and 30 senior observers, were recruited for the experiment. In particular, the 30 young observers (15 males and 15 females) were between 21 to 25 years of age (mean = 23, std. dev. = 0.9) and the 30 senior observers (13 males and 17 females) were between 60 and 86 years of age (mean = 69, std. dev.=4.9). All the observers had normal color vision, as tested using the Ishihara Color Vision Test. In total, 54,000 evaluations (i.e., 15 stimuli pairs × 5 colors × 12 repetitions × 60 observers) were made.

3. Results

3.1. Intra- and inter-observer variations

The intra-observer variations were characterized based on the repeated judgements made by each observer for the same pair of stimuli. In particular, the judgment that was made for more than 50% for the same pair of the stimuli was considered as the final judgment made by the observer. Thus, 3.8% to 19.3% of the repeated judgments were different for the young observers, with a mean of 9.7%; 9.9% to 27.5% of the repeated judgments were different for the senior observers, with a mean of 18.1%. It is obvious that the young observers were able to make consistent judgements.

The inter-observer variations were characterized based on all the judgments made by each individual observer, in comparison to those made by an average observer (i.e., the judgments made by more than half of the observers). For the young observers, 5.8% to 21.0% of the judgements were different from those made by an average observer, with a mean of 11.7%; for the senior observers, 13.4% to 76.6% of the judgments were different from those made by an average observer, with a mean of 31.5%. It can be observed that the senior observers had much larger inter-observer variations, which may be due to the wide range of the ages of the senior observers and their color vision conditions. In particular, the judgements made by five senior observers were very different from those made by an average observer, with more than 50% judgements being different (i.e., 51.9%, 54.1%, 63.3%, 67.6%, and 76.6%). As none of these five observers reported any known vision problems, their data were not discarded.

3.2. Calculation of the visual color differences

The 12 judgements made by each observer for each pair of the stimuli were used to derive a probability score, which was then converted to a Z-score. This Z-score was used to represent the visual color difference ΔV between each test stimulus to the reference stimulus by adding a value of 2.0, with a higher Z-score for a larger visual color difference. Figure 6 summarizes the average visual color differences for each stimulus for the young and senior observers. It can be seen that the visual color differences judged by the young and senior observers were negatively correlated, as shown in Fig. 7, which corroborated the calculated color differences shown in Fig. 5.

Optics EXPRESS

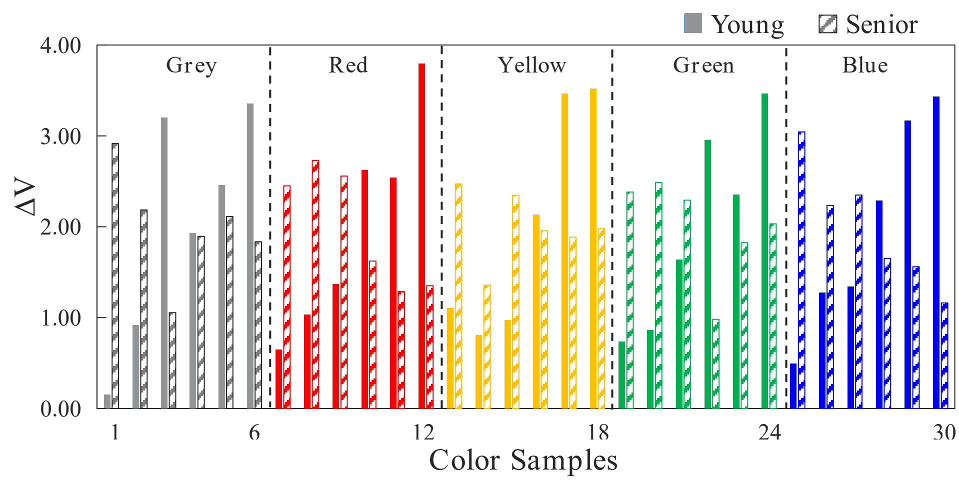


Fig. 6. Average visual color differences ΔV between the test and reference stimuli judged by the young and senior observers.

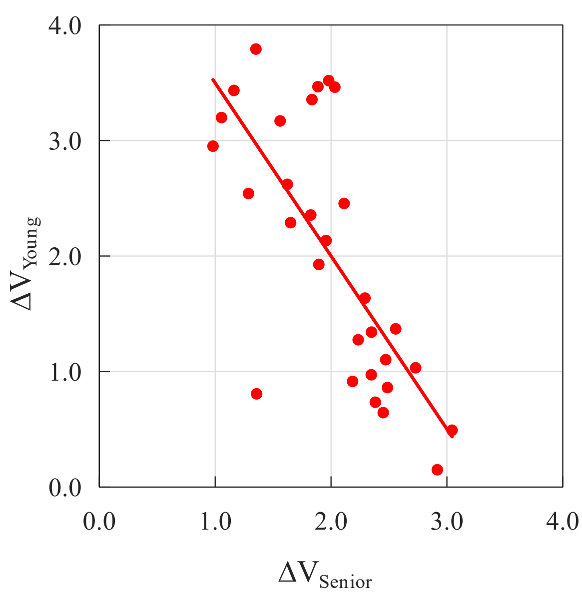


Fig. 7. Relationship between the average visual color differences ΔV judged by the young and senior observers.

Optics EXPRESS

3.3. Performance of the CIE CMFs

The performances of six CIE CMFs, including CIE 1931 2°, CIE 1964 10°, and four CMFs derived using the CIE 2006 model, were compared. In particular, four FOVs (i.e., 2°, 4°, 8.6°, and 10°) were used for the CIE 2006 model, as 2° and 10° are widely used in practice, 4° is the critical size for recommending 2° or 10° CMFs, and 8.6° was the FOV used in the experiment. In addition, the mean ages of the observers (i.e., 23 and 69) were used to derive the CMFs for the young and senior observers respectively. These six CIE CMFs were used to calculate the color difference between each test stimulus and the reference stimulus. Figures 8(a) and (b) show the scatter plots of the calculated color differences $\Delta u'v'$ versus the visual color differences for the young and senior observers respectively.

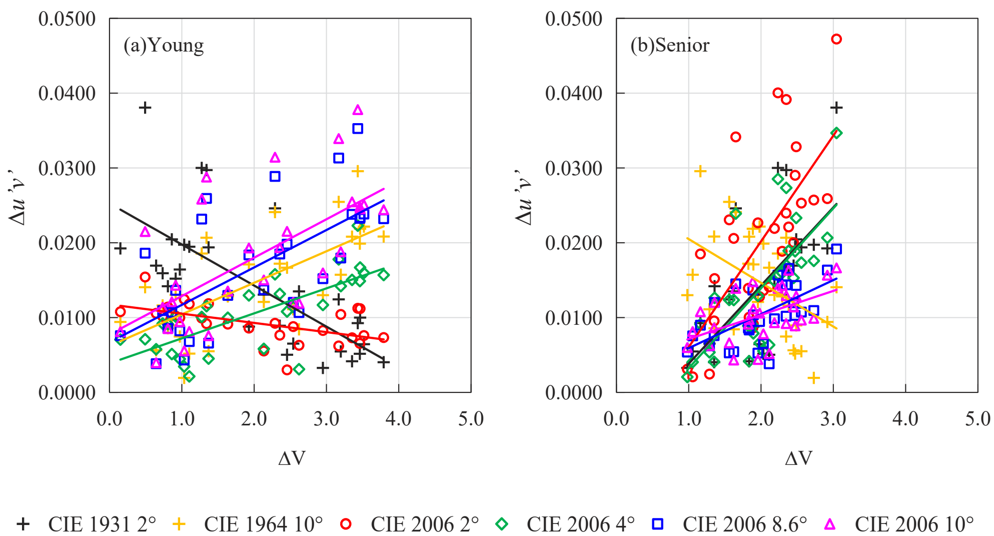


Fig. 8. Scatter plots of the calculated color differences $\Delta u'v'$ derived using various CMFs versus the average visual color differences ΔV judged by the young and senior observers. (a) Young observers; (b) Senior observers.

For the young observers, the color differences calculated using the CIE 1964 10°, 2006 4°, 8.6°, and 10° were positively correlated to the visual color differences; for the senior observers, the color differences calculated using the CIE 1931 2°, 2006 2°, 4°, 8.6°, and 10° were positively correlated to the visual color differences.

The performance of the CIE CMFs in characterizing the visual color differences were further evaluated and compared using the Standardized Residual Sum of Squares (STRESS) [18] values, with a smaller STRESS value suggesting a higher correlation between the visual color differences and the calculated color differences. Table 4 summarizes the STRESS values between the visual and calculated color differences using the six CIE CMFs.

Table 4. Summary of the STRESS values between the visual color differences ΔV and the calculated color differences $\Delta u'v'$ using the various CIE CMFs, with the smallest STRESS value underlined.

Observer Group	CIE 1931 2°	CIE 1964 10°	CIE 2006 (age = 23 or 69)			
			2°	4°	8.6°	10°
Young	81.3	35.5	62.4	**32.7**	36.8	38.2
Senior	39.4	56.0	36.2	37.4	**27.8**	28.7

Optics EXPRESS

Though all the stimuli having an FOV of 8.6°, the CIE 2006 8.6° CMFs did not have the best performance for the young observers. In contrast, the CIE 2006 4° CMFs had the best performance, while the CIE 1931 2° CMFs had the worst performance. For the senior observers, however, the CIE 2006 8.6° CMFs had the best performance, while the CIE 1964 10° CMFs had the worst performance. Moreover, the CIE 2006 10° CMFs resulted in smaller STRESS values than the CIE 2006 2° CMFs for both the young and senior observers, which supported the CIE's recommendation of using the 10° CMFs for stimuli with an FOV beyond 4°.

4. Follow-up experiment changing the FOV

In order to further investigate the effect of FOV on the performance of the CIE CMFs, a follow-up experiment was carried out to reduce the size of the stimulus to 2.9° by reducing the size of the stimulus to 3.5 cm × 3.5 cm and increasing the viewing distance to 70 cm. The stimuli and the experimental procedures remained the same.

Seven young observers (four females and three males) between 23 and 25 years (mean = 24) and seven senior observers (five females and two males) between 60 and 70 years (mean = 66) completed the experiment. The same method was used to derive the visual color differences ΔV. Since the same stimuli were used, the differences should only be caused by the smaller FOV. Figure 9 shows the comparison of the visual color differences evaluated under the two FOVs. It can be observed that the FOVs generally did not cause a very significant impact on the color difference evaluations. In particular, the two FOVs almost did not cause any difference to the young observers, with the trendline being close to the diagonal line. In contrast, the larger FOV (i.e., 8.6°) seemed to reduce the range of the visual color difference evaluations to the senior observers.

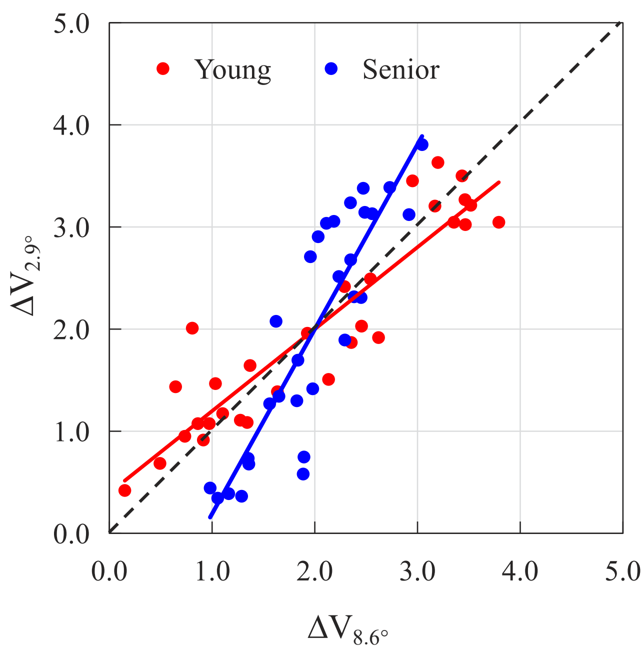


Fig. 9. Scatter plots of the visual color differences ΔV judged by the young and senior observers for the two FOVs.

Table 5 summarizes the STRESS values between the visual color differences and the calculated color differences. For the young observers, the CIE 2006 4° CMFs still had the best performance; for the senior observer, the CIE 2006 2° CMFs had the best performance. The performance of the CIE 2006 2° and 10° CMFs, however, were not consistent. Though the stimulus was 2.9°, the CIE 2006 10° CMFs resulted in a smaller STRESS value for the young observers, which did not support the CIE's recommendation.

Table 5. Summary of the STRESS values between the visual color differences ΔV and the calculated color differences $\Delta u'v'$ using the various CIE CMFs in the follow-up experiment (FOV of 2.9°), with the smallest STRESS value underlined.

Observer Group	CIE 1931 2°	CIE 1964 10°	CIE 2006 (age = 24 or 66)			
			2°	2.9°	4°	10°
Young	78.3	37.0	57.0	34.1	**29.8**	39.8
Senior	43.4	71.8	**36.2**	39.2	54.9	59.6

5. Discussions

5.1. *Performance of color difference formulas—ΔE^*_{ab} and ΔE_{00}*

Though the test stimuli were designed to produce different color differences, as characterized using $\Delta u'v'$, the CIE 1976 $u'v'$ chromaticity diagram is known to have a poor uniformity. In comparison, two color difference formulas—ΔE^*_{ab} [19] and ΔE_{00} [20]—are more widely used for color difference characterizations, which need a defined adapted white point. Though the color stimuli were viewed as unrelated colors with a black background, D65 was employed as the white point based on the findings of an early study [21]. Table 6 summarizes the STRESS values between the visual color differences and the calculated color differences, using the various CMFs. In general, ΔE^*_{ab} always had smaller STRESS values for each observer group, regardless of the CMFs used. This was likely due to the fact that the ΔE^*_{ab} color differences of the 30 stimuli ranged between 2.44 and 25.01, with a mean of 10.97 using the 23-year CIE 2006 CMFs, and ranged between 3.25 and 13.21, with a mean of 8.14 using the 69-year CIE 2006 CMFs, and most values were greater than ΔE^*_{ab} of 5.0, which was not suitable for using ΔE_{00}.

Table 6. Summary of the STRESS values between the visual color differences ΔV and the calculated color differences ΔE^*_{ab} and ΔE_{00} using the various CIE CMFs for the two FOVs, with the smallest STRESS value underlined. (a) FOV of 8.6°; (b) FOV of 2.9°.

(a) FOV of 8.6°

ΔE	Observer Group	CIE 1931 2°	CIE 1964 10°	CIE 2006 (age = 23 or 69)			
				2°	4°	8.6°	10°
ΔE^*_{ab}	Young	70.6	**35.0**	56.6	36.7	35.3	36.1
	Senior	34.4	56.5	34.8	34.4	**34.0**	36.2
ΔE_{00}	Young	72.5	42.3	58.4	**42.1**	**42.1**	42.8
	Senior	39.1	57.0	37.7	**37.6**	43.9	45.9

(b) FOV of 2.9°

ΔE	Observer Group	CIE 1931 2°	CIE 1964 10°	CIE 2006 (age = 24 or 66)			
				2°	2.9°	4°	10°
ΔE^*_{ab}	Young	67.1	33.4	50.7	37.4	**31.3**	36.2
	Senior	40.4	69.8	**38.6**	39.7	41.2	57.7
ΔE_{00}	Young	68.4	39.8	52.2	41.0	**36.9**	41.6
	Senior	44.1	69.6	**37.8**	40.4	46.1	64.0

Even for ΔE^*_{ab}, it can be observed that the CMFs resulting the smallest STRESS values were not those with the FOVs used in the experiments, and they did not support the CIE's recommendation of using 4° as the threshold. Also, regardless of the FOVs used in the experiments, the 10° CMFs always had better performance than the 2° CMFs for the young observers (i.e., CIE 1964 versus CIE 1931, CIE 2006 10° versus 2°), while the 2° CMFs always had better performance than the 10° CMFs for the senior observers. In particular, changing the FOV from 2° to 10° resulted in the similar changes to the STRESS values as changing the observers from young to senior. This was mainly because of the similar effects introduced by changing the age and the FOV size in the CIE 2006 model, with the former increasing the lens pigment optical density and the latter increasing the macular pigment optical density.

5.2. Analyses on individual colors

When the FOV was 2.9°, the CMFs having the smallest STRESS value were the same for both $\Delta u'v'$ and ΔE^*_{ab}. In contrast, when the FOV was 8.6°, $\Delta u'v'$ and ΔE^*_{ab} resulted in different CMFs having smallest STRESS value. Therefore, the STRESS values were calculated for each of the five colors used in the experiments, as summarized in Tables 7 and 8.

Table 7. Summary of the STRESS values between the visual color differences ΔV and the calculated color differences $\Delta u'v'$ and ΔE^*_{ab} using the various CIE CMFs for the FOV of 8.6°, with the smallest STRESS value underlined.

ΔE	Color	CIE 1931 2°	CIE 1964 10°	CIE 2006 (age = 23 or 69) 2°	4°	8.6°	10°
			Young observers				
$\Delta u'v'$	Grey	86.6	33.4	65.1	32.0	**27.6**	27.7
	Red	78.2	24.2	64.3	37.0	16.6	**15.1**
	Yellow	71.6	12.2	59.6	28.2	9.4	**9.0**
	Green	81.4	29.4	58.3	26.0	**24.6**	25.2
	Blue	77.4	27.3	62.6	**16.1**	29.4	31.3
ΔE^*_{ab}	Grey	82.7	28.8	59.7	31.0	**22.5**	22.6
	Red	68.4	37.8	56.8	37.8	25.4	**24.5**
	Yellow	67.5	12.8	63.3	26.9	8.5	**8.2**
	Green	58.5	21.6	37.0	**21.0**	22.4	23.5
	Blue	61.8	**21.0**	52.0	21.6	25.3	27.6
			Senior observers				
$\Delta u'v'$	Grey	35.8	42.9	28.3	35.5	29.5	**25.9**
	Red	18.3	74.9	24.8	17.9	**13.1**	21.0
	Yellow	27.5	52.8	27.3	**26.3**	33.1	39.5
	Green	31.3	38.9	27.9	34.4	21.7	**17.9**
	Blue	15.4	49.7	**10.0**	15.1	19.3	20.7
ΔE^*_{ab}	Grey	28.6	47.5	24.5	29.2	18.2	**17.0**
	Red	**23.0**	58.8	25.2	23.5	31.9	34.8
	Yellow	19.0	47.4	**17.7**	19.6	19.3	22.4
	Green	28.8	53.5	**22.8**	24.1	33.5	36.2
	Blue	**15.9**	54.1	17.7	18.9	32.5	37.7

It can be noted that for the FOV of 8.6°, $\Delta u'v'$ and ΔE^*_{ab} had much larger differences, in terms of the CMFs having the smallest STRESS values, especially to the senior observers. Also, the CMFs having the smallest STRESS values varied with colors. In contrast, for the FOV of 2.9°, the CMFs having the smallest STRESS values were almost the same when characterized using either $\Delta u'v'$ or ΔE^*_{ab}. More importantly, the performance of the CMFs did not significantly vary

Optics EXPRESS

Table 8. Summary of the STRESS values between the visual color differences ΔV and the calculated color differences $\Delta u'v'$ and ΔE^*_{ab} using the various CIE CMFs for the FOV of 2.9°, with the smallest STRESS value underlined.

ΔE	Color	CIE 1931 2°	CIE 1964 10°	CIE 2006 (age = 24 or 66)			
				2°	2.9°	4°	10°
			Young observers				
$\Delta u'v'$	Grey	84.6	35.4	61.4	40.2	**31.1**	32.1
	Red	65.1	32.5	46.7	31.6	31.5	**29.3**
	Yellow	63.8	17.3	47.1	21.2	**16.6**	19.1
	Green	80.9	35.6	58.1	38.7	32.1	**31.4**
	Blue	78.0	29.6	64.8	31.1	**18.4**	33.2
ΔE^*_{ab}	Grey	81.3	32.4	57.8	41.2	31.6	**29.5**
	Red	54.9	25.0	40.8	27.8	23.0	**21.8**
	Yellow	62.5	16.9	55.4	33.6	**14.2**	20.5
	Green	54.5	22.7	29.0	18.4	**16.3**	24.5
	Blue	62.1	22.6	51.7	36.0	**20.5**	29.3
			Senior observers				
$\Delta u'v'$	Grey	46.1	61.3	**27.4**	38.5	50.0	55.2
	Red	14.4	86.8	**9.3**	11.3	18.1	60.1
	Yellow	37.5	78.8	**23.9**	29.7	38.9	77.6
	Green	39.1	49.5	**32.1**	39.0	44.3	38.2
	Blue	**18.7**	66.8	22.8	20.2	19.7	52.5
ΔE^*_{ab}	Grey	32.6	61.7	**21.3**	28.6	35.5	45.8
	Red	31.5	75.8	**28.6**	32.4	37.6	59.4
	Yellow	28.3	74.7	**21.8**	24.2	29.5	61.6
	Green	33.9	54.2	**29.1**	30.4	31.7	44.9
	Blue	**33.4**	70.3	35.9	35.8	37.2	61.6

with the colors, with the CIE 2006 4° and 2° CMFs having the best performance for the young and senior observers respectively, which was the same as those suggested in Table 5.

Despite the fact that all the CIE 2006 CMFs generally had better performance than the CIE 1931 2° and 1964 10° CMFs, the differences between CIE 1931 2° and 1964 10° CMFs were very obvious and were consistent across different colors. The CIE 1964 10° CMFs always had much smaller STRESS values for the young observers, while the CIE 1931 2° CMFs always had much smaller STRESS values for the senior observers, regardless of the colors and FOVs. Also, though the CIE 1931 2° CMFs are always considered to underestimate the contribution in the short-wavelength region (i.e., the blue color) [22,23], the much smaller STRESS values for the senior observers were likely due to the fact that the transmittance of the lens at the short wavelengths decreases when an observer is older.

6. Conclusion

Given the great efforts on investigating the performance of different CMFs in characterizing color matches, little attention has been made to investigate how observer age affect the performance of CMFs. In this study, a series of test stimuli, in five colors (i.e., red, yellow, green, blue, and grey), were carefully designed and calibrated to have different magnitudes of calculated color differences to young and senior observers using the CIE 2006 CMF models. The young and senior observers were asked to compare two test stimuli to the reference stimulus each time and to judge which test stimulus had a larger color difference, which was a forced choice. The experiment was carried out using two stimulus sizes (i.e., FOVs of 8.6° and 2.9°). The judgements made by the observers were converted to z-scores and to derive the visual color differences ΔV, and the performance of the various CMFs were evaluated by comparing the calculated color difference (i.e., $\Delta u'v'$ and ΔE^*_{ab}) and visual color differences ΔV using the STRESS values.

In general, the visual color differences were highly consistent under the two FOVs. The CMFs derived using the CIE 2006 model were found to have better performance in characterizing the visual color differences than the CIE 1931 2° and 1964 10° CMFs. It was found that the CIE 1931 2° and 1964 10° CMFs had better performance for the senior and young observers respectively, regardless of the FOVs and colors. This was likely due to the fact that the transmittance of the lens at the short wavelengths were generally lower for senior observers, which compensated the underestimations of the CIE 1931 2° CMFs. Also, for the CIE 2006 2° and 10° CMFs, the results did not support the CIE's recommendation of using the 2° CMFs for stimuli with an FOV between 1° and 4° and the 10° CMFs for those with an FOV beyond 4°. For the CMFs based on the CIE 2006 model, those with the FOV and observer age being set according to the experiment conditions, however, did not have the smallest STRESS values, which suggested the necessity to further investigate how FOV and age affect the color matching mechanism.

Funding. College's Scientific Research Project (BIGCEc202003, BIGCEc202102); Research Grant Council, University Grants Committee (PolyU R5006-21).

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. **30**(4), 141–164 (1929).
2. W. D. Wright, "A re-determination of the mixture curves of the spectrum," Trans. Opt. Soc. **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. W. S. Stiles and J. M. Burch, "N. P. L color-matching investigation: Final report," Opt. Acta **6**(1), 1–26 (1959).
5. N. I. Speranskaya, "Determination of spectral color coordinates for twenty-seven normal observers," Optics Spectrosc. **7**, 424–448 (1959).
6. CIE, "Colorimetry, 4th edition," in *CIE 015:2018* (CIE, 2018).
7. CIE, "Fundamental chromaticity diagram with physiological axes – part 1," in *CIE 170-1:2006* (CIE, 2006).
8. CIE, "Fundamental chromaticity diagram with physiological axes – part 2: spectral luminous efficiency functinos and chromaticity diagrams," in *CIE 170-2:2015* (CIE, 2015).
9. A. Borbely and J. Schanda, "Colour matching using LEDs as primaries," Color Res. Appl. **29**(5), 360–364 (2004).
10. X. Hu and K. W. Houser, "Large-field color matching functions," Color Res. Appl. **31**(1), 18–29 (2006).
11. 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).
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. J. Li, P. Hanselaer, and A. G. Smet K, "Impact of Color-Matching Primaries on Observer Matching: Part I–Accuracy," Leukos. **18**(2), 104–126 (2022).
14. 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).
15. J. Wu and M. Wei, "Color mismatch and observer metamerism between conventional liquid crystal displays and organic light emitting diode displays, Part II: adjacent stimuli with a larger field of view," Opt. Express **29**(25), 41731–41744 (2021).
16. M. Huang, Y. Xi, J. Pan, R. He, and X. Li, "Colorimetric observer categories for young and aged using paired-comparison experiments," IEEE Access **8**, 219473–219482 (2020).
17. 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).
18. P. García, R. Huertas, M. Melgosa, and G. Cui, "Measurement of the relationship between perceived and computed color differences," J. Opt. Soc. Am. A **24**(7), 1823–1829 (2007).
19. A.R. Robertson, "The CIE1977 colour-difference formulae," Color Res.Appl. **2**(1), 7–11 (1977).
20. M. R. Luo, G. Cui, and B. Rigg, "The development of the CIE2000 colour difference formula: CIEDE2000," Color Res. Appl. **26**(5), 340–350 (2001).

Optics EXPRESS

21. W. Bao, M. Wei, and K. Xiao, "Investigating unique hues at different chroma levels with a smaller hue angle 363 step," J. Opt. Soc. Am. A **37**(4), 671–679 (2020).
22. D. B. Judd, "Report of US secretariat committee on colorimetry and artificial daylight," *Proceedings of the 12th Session of the CIE*, 1–11(1951).
23. J. J. Vos, "Colorimetric and photometric properties of a 2 fundamental observer," Color Res. Appl. **3**(3), 125–128 (1978).