



Parametric effects on the evaluation of threshold chromaticity differences using red printed samples

FERNANDO BRUSOLA,^{1,*}  IGNACIO TORTAJADA,¹  BEGOÑA JORDÁ,¹ AND MANUEL MELGOSA² 

¹Centro de Investigación en Tecnologías Gráficas, Universitat Politècnica de València, 46022 Valencia, Spain

²Departamento de Óptica, Universidad de Granada, 18071 Granada, Spain

*Corresponding author: fbrusola@upv.es

Received 1 November 2018; revised 25 January 2019; accepted 3 February 2019; posted 4 February 2019 (Doc. ID 349845); published 12 March 2019

Results from different authors showed deviations of radial orientation in the a^*-b^* plane (tilt) for the major axes of chromaticity-discrimination ellipses centered around the International Commission on Illumination (CIE) red color center [Color Res. Appl. 3, 149 (1978)], which are not considered by most of the current advanced color-difference formulas (e.g., CIEDE2000). We performed a visual experiment using red printed samples in order to test the influence of the separation between samples (gap) on the mentioned tilt. Our results confirm a counterclockwise tilt of fitted a^*-b^* ellipses with a magnitude of approximately 36° for samples with no separation, which is similar to that detected by other authors, and a reduction of the mentioned tilt owing to the separation of the samples. We detected a tilt of approximately 22° for samples with a black gap of 0.5 mm and a tilt of approximately 25° for samples with a white gap of 3 mm. Notably, the uncertainty of previous values given by the corresponding credibility intervals of 95% posterior probability is approximately $\pm 8^\circ$ of the mean values. Finally, we study the performance of the most widely used color-difference formulas in the graphic arts sector using our current experimental results, and conclude that the performance of the CAM02-SCD and CAM02-UCS color-difference formulas is significantly better than that of the CIEDE2000 formula. © 2019 Optical Society of America

<https://doi.org/10.1364/JOSAA.36.000510>

1. INTRODUCTION

The International Commission on Illumination (CIE) has requested new reliable experimental data in order to improve color-difference formulas for industrial applications [1]. CIE has also emphasized the relevance of viewing conditions (parametric effects) in color-difference evaluation [2]. Some discrepancies have been observed in previous literature regarding a potential tilt (deviation with respect to radial orientation) for a^*-b^* chromaticity-discrimination ellipses in the red area. The reason for such an effect is unknown, but it can be attributed to different factors, such as the separation between samples (gap), psychometric method, fitting methods, and material of the samples.

In the graphic arts industry, there is a particular interest in color discrimination under the usual viewing conditions.

This paper reports the results of a new experiment using printed samples in the red zone of CIELAB color space (CIE red color center [3]). Our main goal is to test the relationship between the tilt and size of the chromaticity-discrimination ellipses and the separation (gap) of the samples. Additionally, the performances of some recently proposed color-difference formulas are analyzed with the experimental data obtained.

2. BACKGROUND

Most of the current color-difference formulas (CIELAB, CIE94, CIEDE2000, CMC, CAM02-SCD, CAM02-UCS, and CAM02-LCD) have been developed from datasets published by different authors since the last half of the past century.

The adoption of new formulas or modification of the existing ones has been attempted to improve the correlation between metric differences, computed by means of a color-difference formula, and visual sensation of difference perceived by observers.

The color-difference formulas CIE94 and CMC represented an improvement with respect to CIELAB by considering the different sensitivities of the human vision system to the changes in lightness, chroma, and hue, which are considered by applying different weights in the formulas to the differences ΔL^* , ΔC_{ab}^* , ΔH_{ab}^* .

CIEDE2000 was adopted by CIE/ISO [4] with the purpose, among others, of improving the behavior of CIE94 (the previously recommended color-difference formula by the CIE) in the area of blues and grays. One of the improvements was based on the introduction of an interactive term between

chroma and hue differences to allow the rotation of the discrimination ellipses in the blue region because experimental datasets show that the major axis of the ellipses does not point toward the origin.

Luo *et al.* [5] described the development of the CIEDE2000 formula in detail. It was based on four experimental datasets: BFD-P (Luo and Rigg [6]), Leeds (Kim and Nobbs [7]), RIT-DuPont (Alman-Berns *et al.* [8,9]), and Witt [10]. From these datasets (generated mostly between 1980 and 2000), the ellipses in the red zone maintain an orientation, which responds to the pattern of the color-difference formulas CIE94 and CMC, with the major axis pointing toward the origin of CIELAB color space.

Here, we define tilt ($\Delta\theta$) as the angular difference between the hue angle (h_{ab}^*) of the color center and the angle of the major axis of the discrimination ellipses with respect to $+a^*$, in this order. Most authors have reported counterclockwise tilts for ellipses in the saturated red zone. Figure 1(a) shows the chromaticity-discrimination ellipses at the red zone obtained experimentally from the datasets BFD-P and RIT-DuPont compared with the corresponding ellipses obtained from the CIEDE2000 formula. The consistency between these two groups is evident, except for some ellipses far from the origin.

Table 1 summarizes the characteristics of the datasets used for the development of the CIEDE2000 formula. The last two

columns indicate the CIELAB coordinates of the color centers closest to the CIE red color center ($L^* = 44$, $a^* = 37$, $b^* = 23$) and the corresponding tilt ($\Delta\theta$). The tilts in Table 1 for BFD-P and RIT-DuPont have been computed from Melgosa *et al.* [11,12], and those corresponding to the Witt [10] and Leeds [7] datasets have been calculated from the data collected in [13].

As shown in Table 1, the tilts of the ellipses near the CIE red color center are very small, except that of the Leeds dataset. However, the mentioned tilts at the CIE red center contrast with those reported by other authors [14–19], as shown in Fig. 1(b) and Table 2, and by Melgosa *et al.* [20] around the red–orange region.

It can be observed that most of the differences between Tables 1 and 2 are related to the psychometric method and the material of the samples. Most of the datasets of Table 1 have been collected from painted or textile samples using the gray scale method, whereas those from Table 2 have been collected from CRT or printed samples using the threshold method.

Wen [17] proposed an alternative color-difference formula to CIEDE2000 that considers the orientation pattern of MacAdam [21] chromaticity-discrimination ellipses. Wen concluded in his paper that CIEDE2000 is not suitable for predicting color differences at threshold because most of the datasets

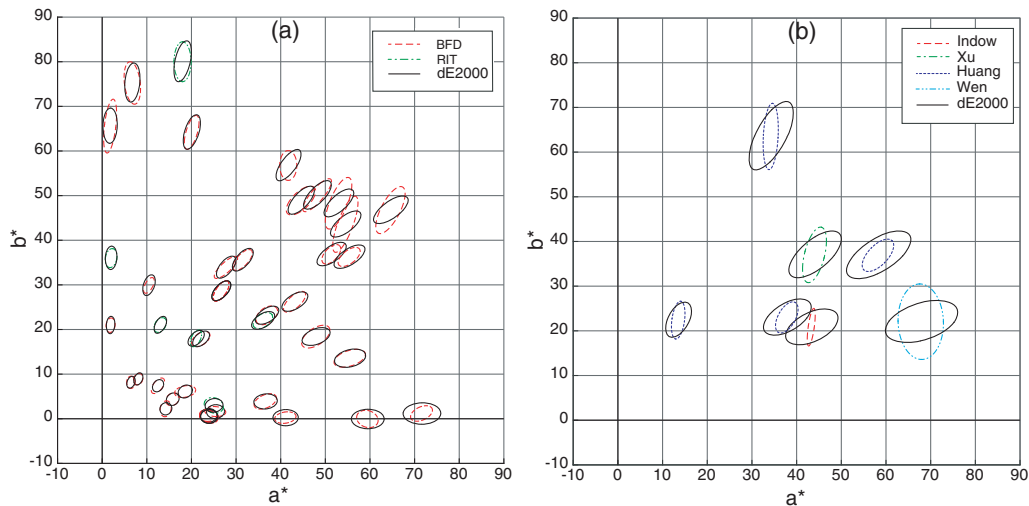


Fig. 1. (a) BFD (in red) and RIT-DuPont (in green) experimental chromaticity-discrimination ellipses around the CIE red color center compared with the corresponding ellipses from the CIEDE2000 color-difference formula (black lines) for $\Delta E_{00} = 1$; CIELAB parameters of the ellipses are obtained from Melgosa *et al.* [11,12]. (b) Indow *et al.* [14], Xu and Yaguchi [15], Huang *et al.* [16], and Wen [17] experimental chromaticity-discrimination ellipses (scaled by a factor of 2) surrounding the CIE red color center compared with the corresponding ellipses from CIEDE2000 color-difference formula (black lines) for $\Delta E_{00} = 1$.

Table 1. Experimental Datasets Used for the Development of CIEDE2000 and Computed Tilt around CIE Red Color Center [3]

Dataset	Pairs	Material	Method	Fitting Model	Color Center [$L^* a^* b^*$]	Tilt ($\Delta\theta$) (°)
BFD-P [6]	2776	painted surfaces and textiles	various	various	[44 37 24]	3
Leeds [7]	307	textile	pair comparison gray scale	probit least squares	[45 37 22]	13
RIT-DuPont [8,9]	156	glossy paint	pair comparison	probit	[44 36 22]	2
Witt [10]	418	glossy paint	gray scale	least squares	[44 37 23]	1

Table 2. Tilt of the Chromaticity-Discrimination Ellipses near the CIE Red [44 37 23] Color Center Reported by Other Authors Whose Datasets Have Not Been Considered for the Development of the CIEDE2000 Color-Difference Formula

Dataset	Material	Method	Fitting Model	Color Center	
				[L* a* b*]	Tilt ($\Delta\theta$) (°)
Indow <i>et al.</i> [14]	CRT	constant stimuli threshold	probit least squares	[52 43 21]	56
Xu and Yaguchi [15]	CRT	threshold constant stimuli	probit maximum likelihood	[44 37 23]	44
Huang <i>et al.</i> [16]	printed on glossy paper	threshold	probit (z-score) least squares	[47 38 23]	30
Wen [17]	CRT	matching experiment threshold	standard deviation	[76 68 22]	74
Huang <i>et al.</i> [18]	printed on glossy paper	gray scale	least squares minimum STRESS	[47 38 23]	41
Rich and Billmeyer [19]	glossy paint	threshold	Gaussian psychometric max likelihood	[68 19 19]	13

used for the development of the formula were obtained using the gray scale method or the constant stimuli method.

Huang *et al.* [16] investigated the performance of color-difference formulas for predicting threshold color differences surrounding 17 color centers recommended by the CIE [22]. After an analysis using the STRESS index [23], they concluded that most of the current color-difference formulas performed similar to each other and outperformed CIELAB for small color differences.

Thus, the question of whether the differences of tilt and size of chromaticity-discrimination ellipses at the CIE red color center could be due to parametric effects caused by the psychometric method used (threshold, gray scale, or pair comparison) or by the separation between samples. As shown in Table 2, the tilt is much higher when samples without a contact hairline are displayed on a CRT.

Regarding the influence of the method on the tilt of the ellipses, several studies [24,25] have been conducted to compare the gray scale method with the method of constant stimuli, and there is no clear conclusion. Some authors [24] concluded that both methods are equivalent, whereas others [25] showed that the constant stimuli method provided a tolerance threshold approximately twice that obtained using the gray scale method. Given the high fiducial limits reported by these authors, a clear conclusion could not be established. In [16], the authors indicate that, at the CIE red center, the threshold and gray scale methods provide similar results, given that they observed a difference of 12.8 STRESS units. Additionally, if we compare

Table 6 of [16] and Table VII of [18], the discrimination ellipses differ by a factor size of 0.9 and a tilt of 9° (difference between the tilt obtained using the gray scale method and that obtained using the threshold method), which can be considered small, given the usual variability of observers. Thus, taking into account the considerations made by the authors at the end of Section E of [16], it does not seem evident that the threshold method would be the main cause of the tilt detected at the CIE red center by the authors shown in Table 2.

Table 3 summarizes the results obtained by different authors regarding the effect of the separation of the samples (Factor column) on the size (K_G column) and tilt ($\Delta\theta_F/\Delta\theta_R$ column) of the discrimination ellipses. More specifically, K_E represents the average value of the ratio between the visual differences under the factor conditions to the visual differences under the reference conditions. Similarly, K_G is the ratio factor between the volumes or areas of the ellipsoids or ellipses of discrimination under factor conditions to reference conditions. The smaller the effect of the factor, the closer to 1 K_E and K_G are. Finally, K_L , K_C , and K_H are the ratios between the principal axes of the discrimination ellipsoids or ellipses under factor conditions to reference conditions.

As shown in the second row of Table 3, the effect determined by Strocka *et al.* [26] by observing pairs with a 1-mm-thick black line separation generates a deviation coefficient of approximately 7% with respect to the reference condition. Therefore, the authors concluded that the separation line does not generate significant effects in terms of the evaluation of the

Table 3. Parametric Effects of the Separation of the Samples on the Size and Tilt of the Chromaticity-Discrimination Ellipses Published by Several Authors

Author	Support	Method	Factor	K_E^a	K_G^b	K_L^c	K_C	K_H	$\Delta\theta_F/\Delta\theta_R^d$ (°)
Guan and Luo [24]	textile	gray scale	gap 3"	0.97	0.97	NA	0.96	0.98	2/1
Strocka <i>et al.</i> [26]	textile	pair comparison	1-mm-thick black line	NA ^c	$V_{FR} = 0.068 \approx 7\%^f$				NA/9
Witt [27]	acrylicpainted	threshold	gap 3 mm	NA	2.1	1.7	2.4	2	-17/-17
Xin <i>et al.</i> [28]	acrylicpainted	gray scale	gap 3"	0.91	1.12	1.27	1.0	1.10	NA
Cui <i>et al.</i> [29]	CRT	gray scale	gap 2" (140 pixel)	1.21	0.96	1.44	0.93	0.79	10/8
Cui <i>et al.</i> [29]	CRT	gray scale	0.35-mm-thick black line (1 pixel)	1.05	0.89	1.1	0.9	0.71	10/8

^a K_E , visual color-difference ratio [28].

^b K_G , size factor.

^c K_L , K_C , and K_H , ellipsoid weighting factors.

^d $\Delta\theta_F/\Delta\theta_R$, tilt under factor/tilt without gap.

^eNA, not available.

^f V_{FR} , coefficient of deviation [26].

color differences with respect to the reference conditions. Witt [27] analyzed the effect of a gap of 3 mm between painted samples using the threshold method. The corresponding coefficients in Table 3 are calculated from the results reported in Table V of [27] considering the reference condition in this case to be the one corresponding to a lightness $L_{10}^* = 41$. As shown in Table 3, the tilt $\Delta\theta_F$ under the factor conditions (3 mm gap), and the tilt $\Delta\theta_R$ under the reference condition (dark surrounding) are different from the ones reported by other authors and even from the tilt reported by the author himself in [10] for the reference conditions, as shown in Table 1. The size factor $K_G = 2.1$ shown in Table 3, which indicates that the ellipsoid of chromatic discrimination determined with a 3 mm gap is approximately twice as large as that corresponding to the reference conditions, is very high compared with that in the other studies shown in the same table.

Guan *et al.* [24], Xin *et al.* [28], and Cui *et al.* [29] reported K_G factors close to unity and very small tilts.

However, the last four rows of Table 3 contradict the hypothesis defended by some authors by means of which an increase in the uncertainty, represented to a certain extent by an increase in the visual differences ($K_E > 1$), should necessarily be accompanied by an increase in the size of the ellipse or the ellipsoid of chromatic discrimination ($K_G > 1$). It can be observed that, as deduced from the data reported by Xin *et al.* [28] for painted samples, the effect of a 3" gap corresponds to $K_E = 0.91$ and $K_G = 1.12$, whereas from the data reported by Cui *et al.* [29], for samples displayed on a screen monitor, the effect of a gap of 2" corresponds to $K_E = 1.21$ and $K_G = 0.96$, which contradicts the mentioned hypothesis. Thus, an increase in uncertainty to perceive a color difference between a sample and a reference by observers need not necessarily be associated with an increase in the size of the chromaticity-discrimination ellipse.

It can be deduced from Table 3 that most of the parametric effects studied correspond to painted, textile, or CRT-displayed samples but not to printed samples on paper (very important for the graphics arts industry). Hence, in this paper, we will study the possible parametric effects owing to the separation of the samples on the evaluation of chromaticity differences at the CIE red color center [3,22], using the threshold method for printed samples on a white support.

3. METHODS

Color samples were printed on a 250 g/m² semigloss coated paper with an Epson SC 7000 inkjet plotter, equipped with 12 ink cartridges. Samples around the color center were printed in an attempt to follow an ideal pattern (similar to that used by Alman *et al.* [8]) of CIELAB color differences in eight vector directions angularly separated by 45° from 0 to 360° in the a^*-b^* plane, and 10 pairs of samples in each direction with an increasing color difference of 0.4 ΔE_{ab}^* units from 0 to 4, following an algorithm developed in Brusola *et al.* [30].

Three sets of color samples pairs were printed. The first set was printed with no separation between samples, the second was printed with samples separated by a 0.5-mm-thick black line, and the third was printed with a 3 mm white gap between samples (the same as that used by Witt [27]). The thickness of the black line was selected to lie between 1 mm proposed by

Strocka *et al.* [26] and 0.35 mm (1 pixel displayed at 72 dpi) proposed by Cui *et al.* [29]. Therefore, for each set, $10 \times 8 = 80$ pairs of samples were printed.

Samples were printed with a size of 40 × 40 mm, subtending an angle of 6° (according to the reference conditions for the CIE generalized color-difference equations [4]) to an observation distance of 40 cm, and were measured using a X-Rite I1 spectrophotometer. The CIELAB values were determined for D65 illuminant and 10° observer. The actual distribution of the measured color differences for pairs of printed samples with no separation is shown in Fig. 2. Similar results were obtained for the other two sets. Figure 2 shows the distribution of measured pairs of samples compared with the fitted values in eight vector directions obtained via linear regression passing through the origin.

The samples were presented to the observers in a Verivide viewing cabinet equipped with a Philips MASTER TL-D 90 Graphica 18 W/965 lamp, which has a correlated color temperature of 6500 K and color rendering index $R_a = 98$. The illuminance level at the test position was approximately 1000 lux. The background of the viewer corresponds to a neutral gray of coordinates CIELAB = [62 0 0].

Every sample was presented to a panel of 10–18 observers with normal color vision according to the Farnsworth–Munsell 100 Hue Test. The observers had to judge whether they could perceive a color difference between each pair of samples (threshold method). The observers repeated the judgments 4–10 times each, randomly in different sessions. Sets of pairs separated by a white line or black line were assessed 100 times, and the sets of pairs in contact were assessed 69 times.

To verify the performance of the most widely used color-difference formulas in the graphic arts sector with the data obtained, we have computed STRESS values and have studied the

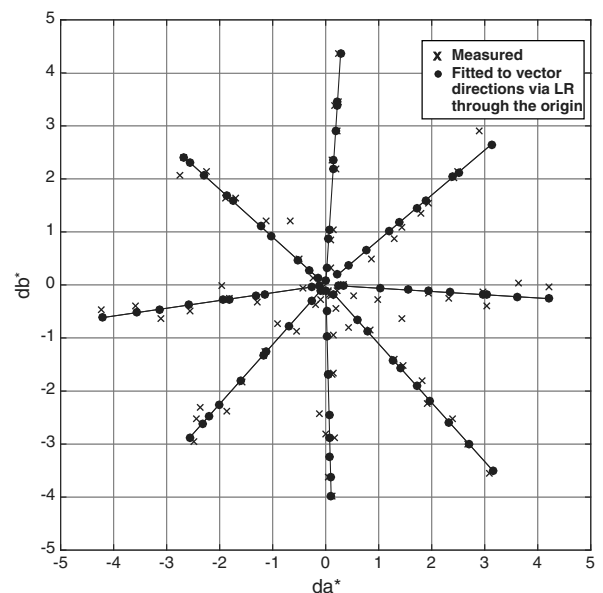


Fig. 2. Distribution of measured color differences for pairs of printed samples with no separation compared with the color differences fitted in eight vector directions obtained via linear regression (LR) passing through the origin.

performance of color-difference formulas selected as indicated in [16,23].

The hypothesis test described in [23] is applied to the differences of quality of the adjustment provided by two different formulas of color difference with respect to the experimental data. We set the null hypothesis, H_0 , such that there is no significant difference between color-difference formulas A and B in terms of the quality of the fit provided, and the alternative hypothesis, H_1 , such that there are significant differences. Hence, if H_0 is true, the quotient, F , calculated using Eq. (1) will be distributed according to a distribution F with degrees of freedom $df_A = N - 1$ and $df_B = N - 1$, where N is the number of color pairs in the dataset.

$$F = \frac{\text{STRESS}_A^2}{\text{STRESS}_B^2}, \quad F \sim F(df_A, df_B). \quad (1)$$

In this case, H_0 will be rejected if F takes values outside the 95% confidence region defined by the interval $[1/F_C, F_C]$, where the limits of the indicated interval are the values of the cumulative inverse distribution of F corresponding to the cumulative probability values of 2.5% and 97.5%, respectively.

Finally, we determine the tolerance that minimizes the number of wrong classification decisions with respect to a color-difference formula, according to Berns [31], which in turn allows the determination of the minimum percentage of wrong classification decisions with each formula selected.

4. RESULTS

The parameters of chromaticity-discrimination ellipses were determined using the logistic psychometric curve and observed fail frequency (number of times observers perceived a difference of color between the standard and the sample) using the Bayesian method described in [32]. Table 4 shows the parameters of the chromaticity ellipses obtained for the three situations under study using the cited Bayesian procedure with the actual measured color differences without the linear fitting shown in Fig. 2.

As shown in Table 4, pair separation generates a reduction of the tilt between 11° and 14° . The 95% credibility intervals of the orientation of the ellipses are $\pm 8^\circ$ for samples in contact and separated with a gap of 3 mm. However, the credibility interval for samples separated by a black line is slightly lower ($\pm 6^\circ$). Nevertheless, given the magnitude of the credibility intervals, we can conclude that the separation induces some kind of tilt reduction, but we cannot conclude that tilt reduction increases with the gap. The values of S (standard deviation of the psychometric curve) shown in Table 4 indicate that the zone of uncertainty of the observers, in which they are doubtful about whether they can perceive color differences, is greater for the samples separated with a gap (black or white) and maximum for pairs with a 3-mm-thick white gap. However, as in some cases shown in Table 3, K_G seldom increases with a 3 mm gap and even decreases with 0.5 mm gap, which indicates that 50% probability chromaticity-discrimination ellipse need not necessarily increase in size with the separation of the samples, as in [24] and [29], where K_G decreased with the gap.

Table 5 shows the STRESS [23] values obtained when applying the CIELAB, CIE94, CIEDE2000, CMC, CAM02-SCD, CAM02-UCS, and CAM02-LCD color-difference formulas to the datasets of this work. We have chosen these color-difference formulas because they are the most widely used formulas in the graphics arts industrial sector. The parameters used for CAM02-UCS and CAM02-LCD color-difference formulas correspond to the following default values: adapting luminance, $L_a = 100 \text{ cd/m}^2$; relative CIE Y Tristimulus of the background, $Y_b = 20\%$ (mid gray background), surround condition, $c = 0.69$ (average), although the calculated values hardly vary for $L_a = 250 \text{ cd/m}^2$ and $Y_b = 92\%$ (white background).

It can be observed that the values of STRESS when applying the formulas CAM02-SCD and CAM02-UCS are very similar (Table 5), which indicates that these formulas fit very well with the datasets obtained in this work. The results of application of the hypothesis test mentioned in Section 3 are shown in Tables 6–8.

Table 4. Chromaticity Ellipse Parameters in $\Delta a^* - \Delta b^*$ Plane^a

Case	g_{11}^b	g_{22}	g_{12}	A^c	B	θ ($^\circ$)	Tilt ($^\circ$)	Size ^d	K_G	S^e	$2\sigma^f$ ($^\circ$)
No separation	1.34	0.91	-0.22	1.11	0.84	68	36	2.92	1.00	0.71	8
Black line 0.5 mm thick	1.7	1.38	-0.46	0.98	0.70	54	22	2.16	0.74	0.92	6
White line 3 mm thick	1.14	0.92	-0.25	1.15	0.88	57	25	3.17	1.09	1.36	8

^a θ : angle of the major axis with respect to $+a^*$.

^b g_{11} , g_{22} , and g_{12} , metric coefficients of the ellipse.

^c A and B , major semiaxis and minor semiaxis, respectively.

^dSize, πAB .

^e S , standard deviation of the psychometric curve.

^f 2σ twice the standard deviation of the posterior distribution of θ (equivalent to a credibility interval of 95% probability).

Table 5. STRESS Values Obtained by Applying Different Color-Difference Formulas to the Datasets of This Work^a

Case	CIELAB	CIE94	CMC (1:1)	CIEDE2000	CAM02-SCD	CAM02-UCS
No separation	21	25	26	26	19	19
White line 3 mm thick	18	19	20	21	16	16
Black line 0.5 mm thick	21	19	20	21	16	17

^a ΔV values have been obtained from frequency data following the procedure described in [16].

Table 6. F Values Corresponding to Pairs of Color-Difference Formulas^a

	CIELAB	CIE94	CMC (1:1)	CIEDE2000	CAM02-SCD	CAM02-UCS
CIELAB		0.88	0.83	0.75	1.29	1.26
CIE94	1.14		0.94	0.85	1.47	1.43
CMC (1:1)	1.20	1.06		0.90	1.56	1.52
CIEDE2000	1.34	1.18	1.11		1.73	1.68
CAM02-SCD	0.77	0.68	0.64	0.58		0.97
CAM02-UCS	0.80	0.70	0.66	0.60	1.03	

^a3-mm-thick white line dataset ($1/F_C = 0.64$, $F_C = 1.56$); $N = 79$. $1/F_C$ and F_C , interval limits of the confidence region; N , number of pairs in the dataset whose observed relative frequency is different from 0 and 1. Bold, results outside the confidence region.

Table 7. F Values Corresponding to Pairs of Color-Difference Formulas^a

	CIELAB	CIE94	CMC (1:1)	CIEDE2000	CAM02-SCD	CAM02-UCS
CIELAB		1.13	1.06	0.98	1.72	1.51
CIE94	0.89		0.94	0.87	1.53	1.34
CMC (1:1)	0.95	1.06		0.93	1.63	1.42
CIEDE2000	1.02	1.14	1.08		1.75	1.53
CAM02-SCD	0.58	0.65	0.62	0.57		0.88
CAM02-UCS	0.66	0.75	0.71	0.65	1.14	

^a0.5-mm-thick black line separation dataset ($1/F_C = 0.58$, $F_C = 1.72$); $N = 55$. $1/F_C$ and F_C , interval limits of the confidence region; N , number of pairs in the dataset whose observed relative frequency is different from 0 and 1. Bold, results outside the confidence region.

Table 8. F Values Corresponding to Pairs of Color-Difference Formulas^a

	CIELAB	CIE94	CMC (1:1)	CIEDE2000	CAM02-SCD	CAM02-UCS
CIELAB		0.71	0.67	0.63	1.15	1.18
CIE94	1.40		0.94	0.88	1.62	1.65
CMC (1:1)	1.49	1.06		0.94	1.72	1.72
CIEDE2000	1.59	1.14	1.07		1.84	1.88
CAM02-SCD	0.87	0.62	0.58	0.54		1.02
CAM02-UCS	0.85	0.61	0.58	0.53	0.98	

^aNo separation dataset ($1/F_C = 0.57$, $F_C = 1.74$); $N = 52$. $1/F_C$ and F_C , interval limits of the confidence region; N , number of pairs in the dataset whose observed relative frequency is different from 0 and 1. Bold, results outside the confidence region.

Tables 6–8 show that, in general, there are no significant differences with respect to the quality of fit between most of the color-difference formulas used for the three datasets, except in the cases of the CIEDE2000 formula with respect to CAM02-SCD and CAM02-UCS formulas, in which there are significant differences in favor of the latter two formulas.

Thus, it can be concluded that the color-difference formulas CAM02-SCD and CAM02-UCS are the best ones with respect to the experimental data of this work.

A summary of the application of the procedure to determine the tolerances, with respect to the color-difference formulas evaluated, which minimized the number of wrong classification decisions, according to the procedure described by Berns [31], is shown in Table 9. As shown in Table 9, CAM02-SCD and CAM02-UCS formulas again produce the best results, with the lowest classification error rate. It can also be inferred from Table 9 that the optimal tolerance for minimizing the classification error should be 1 unit for the CIELAB color-difference formula and approximately 0.5 units for the remaining color-difference formulas. This seems to indicate that the acceptance limit should be approximately 0.5 units when attempting to make color comparisons at threshold with color-difference formulas other than CIELAB and 1 unit when using CIELAB.

Table 9. Optimization Tolerance That Minimizes the Number of Wrong Classification Decisions with Respect to a Color-Difference Formula, According to Berns [31]

Formula	Dataset	Tolerance	Wrong Decisions (%)
CIELAB	No separation	1	6
	3-mm-thick white line	1	6.8
	0.5-mm-thick black line	1	8.0
CIE94	No separation	0.5	8.4
	3-mm-thick white line	0.5	9.3
	0.5-mm-thick black line	0.5	8.0
CMC	No separation	0.6	8.7
	3-mm-thick white line	0.7	9.6
	0.5-mm-thick black line	0.6	8.4
CIEDE2000	No separation	0.5	8.9
	3-mm-thick white line	0.6	10.3
	0.5-mm-thick black line	0.5	8.4
CAM02-SCD	No separation	0.5	6.2
	3-mm-thick white line	0.5	6.2
	0.5-mm-thick black line	0.4	5.1
CAM02-UCS	No separation	0.5	6.3
	3-mm-thick white line	0.6	6.0
	0.5-mm-thick black line	0.5	5.2

Moreover, this result seems to suggest that the tolerance of 0.5 could be the constant factor to add up when computing ΔV values from the frequency data in threshold experiments after applying the logit or probit transformation.

Finally, another interesting point that may be considered is the relationship between visual mechanisms and chromaticity-discrimination ellipses for samples with and without separation. In a recent paper, a specific experimental design was used to investigate the possible relationship between the separation of the samples and the “directions of color differences” in CIELAB color space [33]. These authors reported that the effect of separation of samples for “tritan pairs” (two samples in which the difference between them is only due to the response of S cones) was lower than for “deutan pairs” (pairs in which the ratio of responses of L to M cones signals is different for the two samples), which could be attributed to the negligible contribution of S signal difference in producing a naturally occurring border between two abutting samples. This kind of perceptual anisotropy in CIELAB color space may be considered consistent with our current results, and the differences between tritan and deutan mechanisms may be a potential explanation of tilt in a^*-b^* chromaticity-discrimination ellipses. Specifically, if separation of samples increases visual differences in the deutan direction (a^* axis) more than in the tritan direction (b^* axis), as indicated in Fig. 5 of Ref. [33], the corresponding result should be a decrease in the orientation and tilt of the a^*-b^* ellipses, which is in agreement with values shown in Table 4. Generally speaking, the relationship between confusion lines and chromaticity-discrimination ellipses is not new in the literature. For example, Le Grand [34] already pointed out that the major axes of MacAdam’s ellipses [21] are almost parallel to the axis corresponding to the blue fundamental response, and some relationship between tilt of a^*-b^* ellipses in the blue region and tritan confusion lines has been also reported in previous literature [35].

5. CONCLUSION

Three datasets have been generated to analyze parametric factors for printed samples around the CIE red color center using the threshold method. From the analysis, it is concluded that the chromaticity-discrimination ellipses have an average tilt of 36° for samples with no separation, a tilt of approximately 22° for samples with a black gap of 0.5 mm, and a tilt of approximately 25° for samples with a white gap of 3 mm. The results for samples with no separation are consistent with those reported by Huang *et al.* [16] and are approximately the average of the results reported by Melgosa *et al.* [20] for the red–orange region.

The results also show that the parametric factors related to the separation between samples have a slight effect on the reduction of the tilt (approximately 10°) with respect to the tilt of the chromaticity-discrimination ellipse for samples with no separation.

A minimum wrong classification error rate not greater than 10% is generated by most of the color-difference formulas evaluated in this work. Nevertheless, CAM02-SCD and CAM02-UCS provide the best values, approximately 5%.

The results show that all color-difference formulas performed similarly, but CAM02-SCD and CAM02-UCS performed significantly better than CIEDE2000.

In summary, as it was requested by CIE [1], the authors consider that future studies should be performed in other color centers near the red zone in order to finally make a possible decision of improving CIEDE2000 or recommend the use of other color-difference formulas such as CAM02-SCD and CAM02-UCS, when evaluating color differences at threshold in the red zone.

REFERENCES

1. M. Melgosa, “Request for existing experimental datasets on color differences,” *Color Res. Appl.* **32**, 159 (2007).
2. CIE, “Parametric effects in colour-difference evaluation,” CIE 101-1993 (CIE Central Bureau, 1993).
3. A. R. Robertson, “CIE guidelines for coordinated research on colour-difference evaluation,” *Color Res. Appl.* **3**, 149–151 (1978).
4. ISO/CIE, “Colorimetry—Part 6: CIEDE2000 colour-difference formula,” ISO/CIE 11664-6:2014 (2014).
5. M. R. Luo, G. Cui, and B. Rigg, “The development of the CIE 2000 colour difference formula,” *Color Res. Appl.* **26**, 340–350 (2001).
6. M. R. Luo and B. Rigg, “Chromaticity-discrimination ellipses for surface colours,” *Color Res. Appl.* **11**, 25–42 (1986).
7. D. H. Kim and J. H. Nobbs, “New weighting functions for the weighted CIELAB colour difference formula,” in *8th International Colour Association*, Kyoto, Japan, 1997, Vol. 1, pp. 446–449.
8. D. H. Alman, R. S. Berns, G. D. Snyder, and W. A. Larsen, “Performance testing of color-difference metrics using a color tolerance dataset,” *Color Res. Appl.* **14**, 139–151 (1989).
9. R. S. Berns, D. H. Alman, L. Reniff, G. D. Snyder, and M. R. Balonon-Rosen, “Visual determination of suprathreshold color-difference tolerances using probit analysis,” *Color Res. Appl.* **16**, 297–316 (1991).
10. K. Witt, “Geometric relations between scales of small colour differences,” *Color Res. Appl.* **24**, 78–92 (1999).
11. M. Melgosa, E. Hita, J. Romero, and L. J. Del Barco, “Color-discrimination thresholds translated from the CIE (x, y, Y) space to the CIE 1976 (L^*, a^*, b^*),” *Color Res. Appl.* **19**, 10–18 (1994).
12. M. Melgosa, E. Hita, A. J. Poza, D. H. Alman, and R. S. Berns, “Suprathreshold color-difference ellipsoids for surface colors,” *Color Res. Appl.* **22**, 148–155 (1997).
13. <https://sites.textiles.ncsu.edu/color-science-lab/resources>.
14. T. Indow, A. R. Robertson, M. Grunau, and G. H. Fielder, “Discrimination ellipsoids of aperture and simulated surface colors by matching and paired comparison,” *Color Res. Appl.* **17**, 6–23 (1992).
15. H. Xu and H. Yaguchi, “Visual evaluation at scale of threshold to supra-threshold color difference,” *Color Res. Appl.* **30**, 198–208 (2005).
16. M. Huang, H. Liu, G. Cui, M. R. Luo, and M. Melgosa, “Evaluation of threshold color differences using printed samples,” *J. Opt. Soc. Am. A* **29**, 883–891 (2012).
17. S. Wen, “A color difference metric based on the chromaticity discrimination ellipses,” *Opt. Express* **20**, 26441–26447 (2012).
18. M. Huang, H. Liu, G. Cui, and M. R. Luo, “Testing uniform color spaces and colour-difference formulae using printed samples,” *Color Res. Appl.* **37**, 326–335 (2012).
19. R. M. Rich and F. W. Billmeyer, “Method for deriving color-difference-perceptibility ellipses for surface-color samples,” *J. Opt. Soc. Am.* **65**, 956–959 (1975).
20. M. Melgosa, R. Huertas, A. Yebra, and E. Hita, “A comparison between the rotation terms proposed by recent CIELAB-based color-difference formulas,” in *1st European Conference on Colour in Graphics, Image and Vision (CGIV)* (2002), pp. 136–139.
21. D. L. MacAdam, “Visual sensitivities to color differences in daylight,” *J. Opt. Soc. Am.* **32**, 247–274 (1942).
22. T. Maier, “CIE guidelines for coordinated future work on industrial color-difference evaluation,” *Color Res. Appl.* **20**, 399–403 (1995).
23. P. A. García, R. Huertas, M. Melgosa, and G. Cui, “Measurement of the relationship between perceived and computed color differences,” *J. Opt. Soc. Am. A* **24**, 1823–1829 (2007).
24. S. S. Guan and M. R. Luo, “Investigation of parametric effects using small colour differences,” *Color Res. Appl.* **24**, 331–343 (1999).

25. E. D. Montag and D. C. Wilber, "A comparison of constant stimuli and gray-scale methods of color difference scaling," *Color Res. Appl.* **28**, 36–44 (2003).
26. D. Strocka, A. Brockes, and W. Paffhausen, "Influence of experimental parameters on the evaluation of color-difference ellipsoids," *Color Res. Appl.* **8**, 169–175 (1983).
27. K. Witt, "Parametric effects on surface color-difference evaluation at threshold," *Color Res. Appl.* **15**, 189–199 (1990).
28. J. H. Xin, C. C. Lam, and M. R. Luo, "Investigation of parametric effects using medium colour-difference pairs," *Color Res. Appl.* **26**, 376–383 (2001).
29. G. Cui, M. R. Luo, B. Rigg, and W. Li, "Colour-difference evaluation using CRT colours. Part II: parametric effects," *Color Res. Appl.* **26**, 403–412 (2001).
30. F. Brusola, I. Tortajada, B. Jordá, I. Lengua, and L. Dunai, "Método para la reproducción precisa de muestras de color impresas a partir de los datos de caracterización necesarios para la generación de perfiles de color ICC," in *11th Congreso Nacional de Color*, Ourense, Spain, 2016, pp. 109–112.
31. R. S. Berns, "Deriving instrumental tolerances from pass-fail and colorimetric data," *Color Res. Appl.* **21**, 459–472 (1996).
32. F. Brusola, I. Tortajada, I. Lengua, B. Jordá, and G. Peris, "Bayesian approach to color-difference models based on threshold and constant-stimuli methods," *Opt. Express* **23**, 15290–15309 (2015).
33. H. Saeedi and S. G. Kandi, "How anisotropy of CIELAB color space affects the separation effect: an experimental study," *J. Opt. Soc. Am. A* **36**, 51–60 (2019).
34. Y. Le Grand, "Les seuils différentiels de couleurs dans la théorie de Young," *Revue d'Optique* **28**, 261–278 (1949) [R. M. Boynton and K. Knoblauch, "Color difference thresholds in Young's theory," *Color Res. Appl.* **19**, 296–309 (1994)].
35. A. Yebra, R. Huertas, M. M. Pérez, and M. Melgosa, "On the relationship between tilt of a^*b^* tolerance ellipses in blue region and tritanopic confusion lines," *Color Res. Appl.* **27**, 180–184 (2002).