Journal of Imaging Science and Technology® 57(5): 050502-1–050502-9, 2013.

Color-Difference Evaluation for Digital and Printed Images*

Haoxue Liu, Min Huang, Yu Liu, Bing Wu, and Yanfang Xu
School of Printing & Packaging Engineering, Beijing Institute of Graphic Communication, Beijing 102600, China

Ningfang Liao
National Laboratory of Color Science and Engineering, Beijing Institute of Technology, Beijing 100081, China

Guihua Cui
School of Physics & Electronic Information Engineering, Wenzhou University, Wenzhou 325035, China
E-mail: guihua.cui@gmail.com

Abstract. *A series of experiments and data analyses has been performed to test the consistency between computed and perceptual color differences in images. Five International Organization for Standardization (ISO) standard color image data (SCID) images, N2, N3, N4, N5 and N7, were tested in the experiments, whose colors were altered in CIELAB lightness, chroma, and hue, either independently or simultaneously, to form the test image pairs. CIELAB, CIE94, CIEDE2000, and CMC color differences were computed by averaging the color differences pixel by pixel for the digital images or by averaging the color differences of 256 typical color patches extracted from each image for the printed images. The digital test images were displayed on an EIZO CG19 LCD monitor and the printed test images were viewed in a D50 light booth. The experimental results showed that the lightness, chroma, and hue differences behaved differently when the perceptual color differences were plotted against the computed differences. This implied that the color-difference formulas should be optimized and that different weighting factors should be applied to different visual attributes. The color-difference formulas can be optimized from the experimental data by the slope ratio of best-fit lines of lightness, chroma, and hue. The optimized formulas CIELAB(1.5:1), CIEDE2000(2.3:1), CIE94(3.0:1), and CMC(3.4:1) for digital images, or formulas CIELAB(2.4:1.5:1), CIEDE2000(2.8:1.6:1), CIE94(2.9:1.6:1), and CMC(2.7:1.5:1) for printed images, when considering hue, performed much better than the original formulas.*

[DOI: 10.2352/J.ImagingSci.Technol.2013.57.5.050502]

INTRODUCTION

Color images are one type of media that is widely used to convey visual information. The quality of a reproduced image, either digital or printed, is of interest to researchers in many fields related to vision science and engineering, such as optics and material physics, image processing, printing and media technology, and psychology. Some measures other than color difference may also be applied for assessing the perceptual image quality, such as sharpness, naturalness, contrast, graininess, and usefulness,[1–3] but these are not considered in this article. A 'high-quality' image may appear differently in color sensation when it is rendered on different media or under different viewing conditions. It is very important to evaluate the color difference between two images in practice. As an image is a non-uniform or complex color sample which consists of many pixels with different color values, its color sensation is induced by all the pixels, and it cannot be measured directly using spectroradiometers or spectrophotometers. Therefore, it is very important to appropriately calculate the color difference between two images in practice. The CIE technical committee TC08-02 was charged to address the problems involved in evaluating image color difference, and it has published its finding in a technical report, CIE 199:2011.[4] There have been many studies addressing important issues in this area. In 1991, Stokes[5] conducted experiments with six digital images on a cathode ray tube (CRT) display to derive colorimetric tolerances of displayed images. In his visual experiment, perceptibility and acceptability colorimetric tolerances for images were measured using paired comparison techniques by a panel of 32 observers. He found that the content did not significantly affect the tolerances. The CIELAB,[6] CMC,[7] and CIE94[8] color-difference formulas were shown to be inadequate for accurately modeling image tolerances. He also found that a threshold to detect a color difference in images was about 2.0 CIELAB units. Uroz et al.[9] also conducted a similar experiment using printed images, and found that the threshold was between 1.9 and 2.3 CIELAB units, depending on the experimental method used. Gibson et al.[10] tested different color-difference formulas using the data accumulated from a CRT, a liquid-crystal display (LCD), and prints, and developed perceptible color-difference tolerances for pictorial images. In their study, thresholds were also expressed colorimetrically using pixel-by-pixel comparisons with and without S-CIELAB pre-filtering.[11] They found that the overall thresholds ranged from 1.0 to 3.0 CIELAB units for each of the three perceptual color attributes. They also found experimentally that S-CIELAB pre-filtering had no significant effect. Song and Luo[12] conducted similar experiments, and found that the perceptible color difference is about 2.2 CIELAB units. Sano et al.[13] accumulated a set of experimental results for assessing perceptibility

* Parts of this article were published as part of the Proceedings of the NIP28/DF2012: International Conference on Digital Printing Technologies and Digital Fabrication 2012, Quebec, Canada, September 2012.

Received Dec. 14, 2012; accepted for publication Sep. 26, 2013; published online Sep. 1, 2013. Associate Editor: Wei Koh.
1062-3701/2013/57(5)/050502/9/$20.00

Table I. The transform functions and coefficients for CIELAB lightness and chroma used to prepare the digital images.

Functions	Coefficient k for C^*_{ab}	Coefficient k for L^*	Coefficient k for C^*_{ab} & L^*	
			C^*_{ab}	L^*
$Out = k \cdot In$	0.85	0.95	0.92	
	0.70	0.85		
$Out = k \cdot In + 100 \cdot (1 - k)$	0.95	0.95		0.92
$Out = 100 \cdot (\frac{In}{100})^k$	0.98	0.65	1.02	1.05
	1.02	0.85	0.90	0.85
	1.15	1.15	0.85	0.90
	1.25	1.25	–	–
	1.30	1.35	–	–

thresholds using printed images, and proposed using either CIEDE2000[14,15] or CIELAB with a lightness parameter of $k_L = 2$ for evaluating color differences in complex images.

The studies mentioned above evaluated the perceptibility threshold or acceptability tolerance of color difference between images, and found that the average color-difference threshold between images was about $2.0\Delta E^*ab$ units. Recently recommended color-difference formulas for images according to the experimental results are CIELAB(2:1) and CIEDE2000(2:1). However, those experiments tested only the threshold. They are not able to provide reliable information on the relationship between calculated color-difference magnitudes and their corresponding visual sensations. To test how visual color-difference sensation varies with differing magnitudes of color difference between images, two experiments were designed. A series of subjective evaluations based on complex images were performed using a categorical judgment method. The visual experiment results were assessed using CIELAB, CIE94, CMC, and CIEDE2000 color-difference formulas.

EXPERIMENTS

Two psychophysical experiments, hereafter referred to as Experiment I and Experiment II, respectively, were conducted in this study to evaluate the color difference between images for digital and printed images.

Preparation of Test Image Pairs

Five CMYK/SCID ISO 400 images,[16] N2, N3, N4, N5, and N7 (as shown in Figure 1), were used as original images in the experiment. The original images were in CMYK color mode and were transformed to CIELAB lightness (L^*), chroma (C^*_{ab}), and hue angle (h) values using the Adobe ICC (International Color Consortium) profile Japan-Color2001Coated.icc in Photoshop (see website http://www.color.org/JC200103.xalter for detail). Our own software was used to generate the serial color-difference grades by adjusting the CIELAB lightness, chroma, and hue attributes either independently or in combination. Two linear transform functions and one exponential transform function were applied to the CIELAB lightness (L^*) and chroma (C^*_{ab}) attributes using Eq. (1) and values in Table I[17–19], where In is the original CIELAB lightness or chroma value and Out is the value modified according to the equation. The parameter k changes the magnitude of the color difference in the prepared digital image, and is listed in Table I. The transformation for hue angle is given in Eq. (2),[20] where Hue_O is the modified CIELAB hue angle in degrees. It is derived from the hue angle Hue_I in the original image plus a defined *offset*.

$$Out = k \cdot In, \quad Out = k \cdot In + 100 \cdot (1 - k),$$
$$Out = 100 \cdot (In/100)^k \tag{1}$$
$$Hue_O = Hue_I + offset. \tag{2}$$

For the results, the test image pairs were made up of an original image and its transformed versions. These included eight color differences of different magnitudes for C^*_{ab} and L^*, modified individually (see the second and third columns in Table I), and four color differences of different magnitudes for C^*_{ab} and L^* modified in combination (see the fourth and fifth columns in Table I. Note that, in the first transform, the first and second linear transformations with $k = 0.92$ for C^*_{ab} and L^*, respectively, were applied) using Eq. (1). In total, 20 pairs of digital images were prepared for each of the five original images. After modification, the color of the test image pairs was transformed back from CIELAB to RGB for displaying on the monitor by a reverse sRGB transform. For the printed images, the CIELAB lightness was altered by each of two linear transformations, as shown in Eq. (1), in seven steps, individually, with k ranging from 0.85 to 0.98, and by the exponential transformation in Eq. (1) in 16 steps with k ranging from 0.77 to 1.33. The CIELAB chroma was altered in the desaturation direction by the first linear transformation in seven steps, with k ranging from 0.43 to 0.95, and by the exponential transformation in eight steps, with k ranging from 1.01 to 1.38. The hue angle was altered using Eq. (2) to give ten different magnitudes using approximately 1° intervals of hue angle. In total, 55 color-difference magnitudes for each original image and 275 test images were prepared. All test image pairs were in CIELAB color mode, and they were printed

Figure 1. The five ISO SCID images used in the experiments.

on EasyColor inkjet paper by a colorimetrically calibrated Epson 7880 eight-color inkjet printer with a resolution of 720×720 dpi using EFI Colorproof XF software (see website http://www.efi.com/products/prepress/proofing/colorproof-xf/details-color-manager.asp for detail).

The color difference of image pairs was computed pixel by pixel for the digital images and represented by the mean difference of all pixels under illuminant D50 and CIE 1931 standard colorimetric observer. Eq. (3) is the calculation for mean CIELAB difference. The subscripts 1 and 2 denote the modified and original images, i and j are the row and column number of pixels, and M and N are the width and height of the image, respectively.

$$\Delta E = \frac{\sum_{0\leqslant i\leqslant M}\sum_{0\leqslant j\leqslant N}[(L^*_{1ij}-L^*_{2ij})^2+(a^*_{1ij}-a^*_{2ij})^2+(b^*_{1ij}-b^*_{2ij})^2]^{1/2}}{M\cdot N}. \qquad (3)$$

The difficulty in preparing printed test image pairs is that the color differences cannot be calculated pixel by pixel, unlike the case for digital images. A relatively simple method was used to calculate the mean color difference in the images, in which 256 typical color samples were extracted from each original image, and they were combined with the original image to form a new original image that was modified with Eqs. (1) and (2). The modified image was then printed together with the new original image, as shown in Figure 2.[19] The 256 typical color samples were extracted as indexed colors by Photoshop software, and the printed colors were measured using an X-Rite EyeOne spectrophotometer. The 256 typical colors were selected automatically according to the colors in the original images, and differed between images. The color difference of an image pair was computed also using Eq. (3) from the 256 indexed colors. Here, $M \times N$ in the equation was equal to 256.

Figure 2. An example of a printed test pair of images (N7: original at the top, modified at the bottom) and their 256 typical color patches; these patches were used to measure the color difference between the two test images.

Experiment I

The test image pairs were displayed on a colorimetrically calibrated EIZO CG19 LED monitor in random order under typical normal office environment conditions (i.e., illuminance level of 300–500 lx, correlated color temperature of 5300 K). The resolution of the monitor was set to 1280×1024, and the observation distance was approximately 450 mm. The image resolution was either 448×560 or 560×448 pixels, giving displayed image sizes of 130×165 mm^2 or 165×130 mm^2 and approximately 27 pixels per viewing angle.[11] Each modified image was paired with its original at a separation of approximately 10 mm between them to form a pair of test images.

Twelve observers (six male, six female, with an average age of 26.4 years and normal color vision according to the Ishihara test) took part in the experiment. Each observer assessed each image pair three times. All 300 assessments (5 images $\times$ 20 grades $\times$ 3 replications) were arranged in a random order and divided into ten sessions for each observer; the duration of each session was limited approximately to 15 min to avoid visual fatigue. In total, 3600 judgments were obtained. Prior to the experiment, all observers were first trained to assess image differences using

Table II. Color-difference categories for visual experiment.

Color perception	Category
No difference	0
Just perceptible difference	1
Small difference, which is ensured to exist	2
A small obvious difference, which is deemed acceptable	3
An obvious difference, which is not acceptable	4

the category judgment method. The observers were asked to judge the image color difference (ΔV) using five categories of nearly uniform color-difference sensation scales as listed in Table II. Use of a decimal value was encouraged if the difference was considered to lie between any adjacent levels, and a value greater than 4 was permitted for differences judged greater than level 4. The observers were encouraged to point out the location of the color difference between images.

Experiment II

Experiment II had a similar design to Experiment I, but used the printed images. The printed test image pairs were viewed in a D50 light booth with an illumination level of approximately 930 lx in a dark room. The viewing distance for the test images was 600 mm, with 45°/0° viewing geometry. In this experiment, 12 observers, five male and seven female, with an average age of 23 and normal color vision, took part. The observers adapted to D50 light for approximately 1 min before their assessments, and they assessed each test pair twice in the same way as in Experiment I. In total, 6600 (12 observers × 5 images × 55 grades × 2 replications) assessments were obtained in Experiment II.

RESULTS AND DISCUSSION

In general, a color-difference formula based on CIELAB, e.g., CIELAB, CIEDE2000, CIE94 and CMC, may be rewritten into a generic form as

$$\Delta E = \sqrt{\left(\frac{\Delta L^*}{k_L S_L}\right)^2 + \left(\frac{\Delta C^*}{k_C S_C}\right)^2 + \left(\frac{\Delta H^*}{k_H S_H}\right)^2}, \qquad (4)$$

where ΔL^*, ΔC^*, and ΔH^* are the CIELAB lightness, chroma, and hue differences, respectively, and S_L, S_C, and S_H are the weighting functions for the correction of CIELAB lightness, chroma, and hue differences. $S_L = S_C = S_H = 1$ for CIELAB; k_L, k_C, and k_H are the parametric factors for the above three color-difference components. The parametric factors are usually applied in color-difference evaluation to consider the deviation of actual viewing conditions from the reference conditions recommended by CIE,[6,8,15] for example, the difference between complex images and uniform colors in this study. Previous studies found the lightness parametric factor (k_L) to be the most influential of the three weighting factors.[19] This is possibly due to changes in viewing conditions having the greatest effect on lightness differences. In all forms of the generic equation, k_H is often set to 1. It is worth mentioning here that CIELAB was not defined with parametric factors for lightness (k_L), chroma (k_C), and hue (k_H), but it was suggested by CIE[6] that in different practical applications it may be necessary to use different weightings for ΔL^*, ΔC^*, and ΔH^*. To distinguish a modified color-difference formula from its original form, it is usual to append the ratio k_L:k_C or k_L:k_C:k_H in parentheses to the name of original formula; e.g., CMC(2:1) indicates a modified CMC formula with $k_L = 2$ and $k_C = 1$.

Experiment I

The raw visual data from the observers for a pair of images was averaged as the true visual color difference of the pair. The intra-observer and inter-observer variability[21] was measured by the *STRESS* index[22] to demonstrate the observers' uncertainty. The intra-observer variability ranged from 10.4 to 19.6, with an average of 15.6, and the inter-observer variability varied from 17.7 to 28.4, with an average of 24.3. Considering that the current assessments were made by a category judgment method without a reference pair, the observers' variability was judged to be acceptable.

Figure 3 shows the results of Experiment I. The abscissa in Fig. 3 is the calculated color difference from CIELAB(1:1),[6] CIEDE2000(1:1),[14,15] CIE94(1:1),[8] and CMC(1:1)[7] formulas, respectively, and the ordinate is the perceptual difference (ΔV). The experimental data for each image is plotted in different dot shapes, and the lines are the best-fit linear trend lines. Table III lists the coefficients, including slope, intercept, and Pearson's correlation coefficient (R), of the fitted lines. From Fig. 3 and Table III it can be seen that the perceptual color difference was roughly proportional to the calculated difference with reasonably higher Pearson's correlation coefficients (R), although the dots were somewhat scattered. This means that there is some kind of relevant relationship between the perceptual and calculated color differences. The color-difference formulas which are derived from uniform color samples may also be used to approximate the color difference between images.

In order to investigate the scatter patterns between different formulas, the experimental data was fitted according to CIELAB lightness and chroma, respectively, as shown in Table IV. It was found that the fitted lines for lightness and chroma were separated significantly from each other[19,23]. It can be seen from Table IV that there was a large difference in slope between the two fitted lines for each of the four tested formulas, i.e. the slope of the chroma (C^* in Table IV) is greater than that of the lightness (L^* in Table IV). This indicates that the trends of lightness and chroma sensations were very different to the calculated differences. The slope ratios of fitted lines for chroma to lightness (C^*/L^*), 1.5, 2.3, 3.0, and 3.4 for CIELAB, CIEDE2000, CIE94, and CMC, respectively, are listed in Table IV, and were used as the lightness parametric factors k_L for optimizing the color-difference formulas.

Figure 4 shows the results of Experiment I with the optimized color-difference formulas. The coordinates are

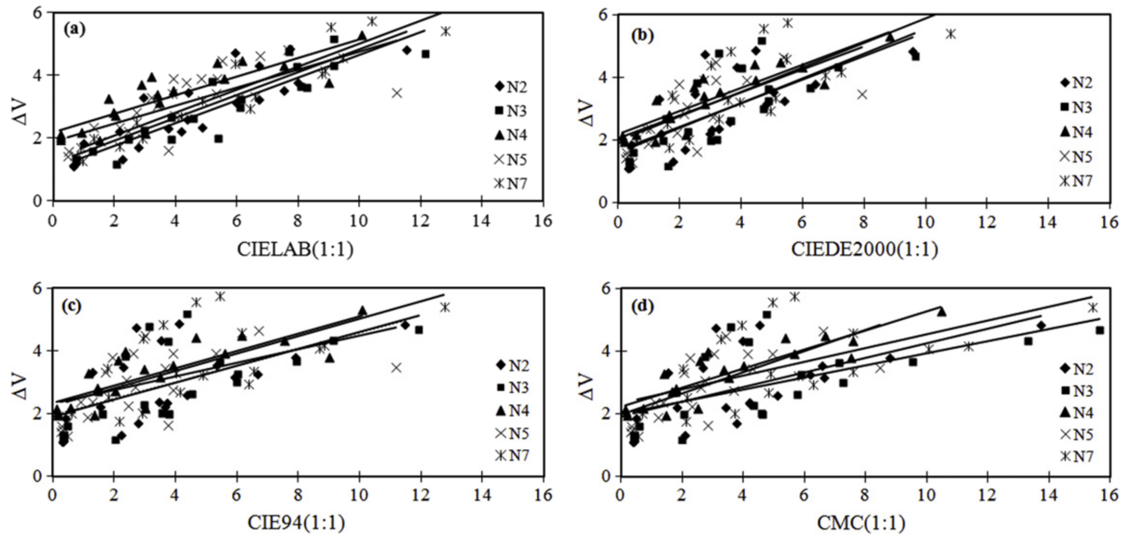

Figure 3. The relationship between visual difference (ΔV) and calculated color difference for (a) CIELAB(1:1), (b) CIEDE2000(1:1), (c) CIE94(1:1), and (d) CMC(1:1) from Experiment I data.

Table III. The coefficients and correlation factors of the fitted lines in Fig. 3.

	CIELAB(1:1)			CIEDE2000(1:1)			CIE94(1:1)			CMC(1:1)		
	Slope	Intercept	*R*	Slope	Intercept	*R*	Slope	Intercept	*R*	Slope	Intercept	*R*
N2	0.37	1.17	0.86	0.37	1.68	0.70	0.27	1.92	0.60	0.23	1.93	0.61
N3	0.36	1.02	0.92	0.39	1.61	0.74	0.27	1.88	0.65	0.19	2.01	0.61
N4	0.30	2.19	0.86	0.37	2.19	0.88	0.28	2.34	0.82	0.30	2.21	0.88
N5	0.29	1.90	0.71	0.37	2.04	0.67	0.22	2.30	0.55	0.34	1.96	0.70
N7	0.37	1.34	0.93	0.39	1.98	0.75	0.28	2.23	0.65	0.22	2.33	0.62
Mean	0.34	1.52	0.85	0.38	1.90	0.75	0.26	2.13	0.65	0.26	2.09	0.69

Table IV. The coefficients and correlation factors of the lightness and chroma fitted lines.

	CIELAB(1:1)			CIEDE2000(1:1)			CIE94(1:1)			CMC(1:1)		
	Slope	Intercept	*R*	Slope	Intercept	*R*	Slope	Intercept	R	Slope	Intercept	*R*
C^*	0.44	1.56	0.94	0.85	1.52	0.96	0.90	1.52	0.96	0.84	1.42	0.96
L^*	0.29	1.48	0.85	0.37	1.47	0.87	0.30	1.48	0.85	0.25	1.64	0.83
C^*/L^*	1.5			2.3			3.0			3.4		

the same as in Fig. 3 but the abscissa is the calculated color differences from the optimized color-difference formula, CIELAB(1.5:1), CIEDE2000(2.3:1), CIE94(3.0:1), and CMC(3.4:1), respectively. Comparing Fig. 4 to Fig. 3, the improvement of the optimized color-difference formulas is clear; that is, the data dots were less scattered and the fitted lines for the five images were almost parallel to each other. This result shows that the calculated color differences with the original formulas had different sensations with different visual attributes and the color-difference formulas should be weighted with proper k_L and k_C in evaluating the color difference for complex images.

The significance of the difference between the optimized form and its original formula with $k_L = 1$ was tested using an F-test[22] (with a critical value $FC = 0.673$ in this case). The F-test values for the formulas CIELAB, CIEDE2000, CIE94, and CMC from the original ($k_L = 1$) to the optimized version were 0.88, 0.48, 0.38, and 0.43, respectively. The F-test values were less than 1, which means that the modified formulas performed better than their original forms. Furthermore, the F-test values were less than FC for the three advanced formulas, CIEDE2000, CIE94, and CMC, and this indicates that the improvement was statistically significant for them. It is worth noting that $k_L = 1.5$ for CIELAB and $k_L = 2.3$ for CIEDE2000 found in this study were in almost total agreement with the values given in the CIE recent recommendations made in CIE 199:2011 technical report. The current finding suggests that

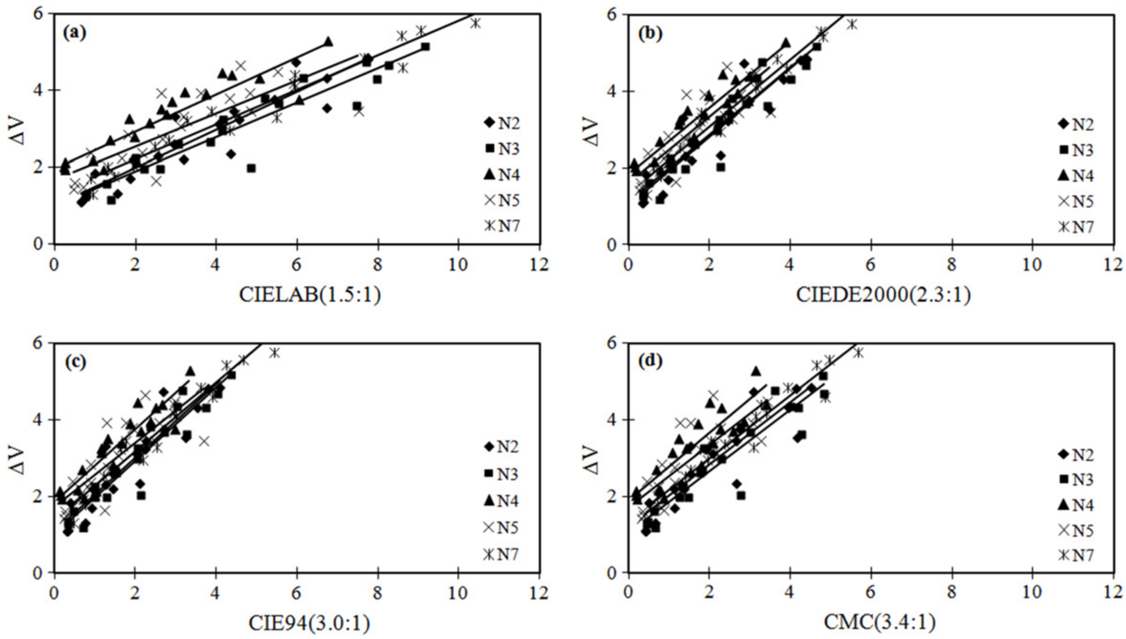

Figure 4. The relationship between visual difference (ΔV) and calculated color difference for color-difference formulas (a) CIELAB(1.5:1), (b) CIEDE2000(2.3:1), (c) CIE94(3.0:1), and (d) CMC(3.4:1) optimized from Experiment I data.

the lightness parameter (k_L) should be further investigated to evaluate the color difference for complex images.

Experiment II

As with Experiment I, the true color difference for each tested image pair was computed by averaging the panel's visual data. The *STRESS* index was calculated again for checking intra-observer and inter-observer variability. The intra-observer variability varied from 12.0 to 30.7, with an average of 18.9, and the inter-observer variability varied from 17.6 to 41.4, with an average of 23.1 for Experiment II.

The correlation between perceptual and calculated lightness, chroma, and hue differences of Experiment II are shown in Figure 5. The results are very similar to those of Experiment I, even though they came from printed images. The results show different trends for lightness, chroma, and hue (see Fig. 5). The best-fit lines indicate the linear relationships between different perceptual color differences and corresponding calculated color differences for lightness, chroma, or hue difference. The parameters of the regression lines in Fig. 5 are listed in Table V. The slopes of h, C^*, and L^* in Table V indicate quantitatively the diversity of the different perceptual attributes in Fig. 5. The larger the slope, the more sensitive the attribute will be. The slope in Table V suggests that the human visual system is much more sensitive to the hue difference than to chroma and lightness differences when compared with the calculated differences of all four color-difference formulas.

The same optimized method used in Experiment I was applied again to Experiment II's data, and the k values obtained are listed in Table V. The last row in Table V lists the optimized k values calculated by the slope ratio. For example, the slope ratio for CIELAB(1:1:1) was L^*:C^*:h = 0.41:0.65:1.00 (normalized by the slope of h); this means that the slope of the h fitted line was 1.5 times larger than that of C^* and 2.4 times larger than that of L^*. In order to make the fitted lines of the three visual attributes have the same slope, the slope of the C^* fitted line should be multiplied by 1.5 and that of the L^* fitted line by 2.4. Figure 6 shows the optimized results, and the regression coefficients after optimization are listed in Table VI. The coordinates of Fig. 6 are the same as those of Fig. 5, but the abscissa is plotted in optimized color-difference units.

Comparing Figs. 5 and 6, the fitted lines of the three visual attributes were closer to each other after optimization, that is, nearly parallel to each other, and the data points were crowded around the fitted lines. Comparing Tables V and VI, the slopes of the fitted lines for h, C^*, and L^* in Table VI are much closer to each other than those in Table V, indicating that the optimized color-difference formulas performed better than the original formulas, because the slopes of chroma and lightness were also increased and almost overlapped each other. Also, the Pearson's correlation coefficients (R) for h and C^* showed significant increase. However, it can be also seen that the slope of the optimized hue lines of all four formulas were visibly different from those of the lightness and chroma. This is probably because the test samples did not consist purely of a single attribute difference, that is, they included combinations of ΔL, ΔC, and ΔH to some extent. This is due to the difficulty in accurately controlling the color of a printed image that changes only in a single color attribute. For example, modifying only the hue attribute in the test sample images using Eq. (2), their chroma or lightness also changed in the printed image. Mathematically, if the test samples vary only in the hue attribute, the slopes of all three attributes should remain unchanged after optimization, because k_H is equal to 1 and k_L and k_C have no effect on hue difference (since

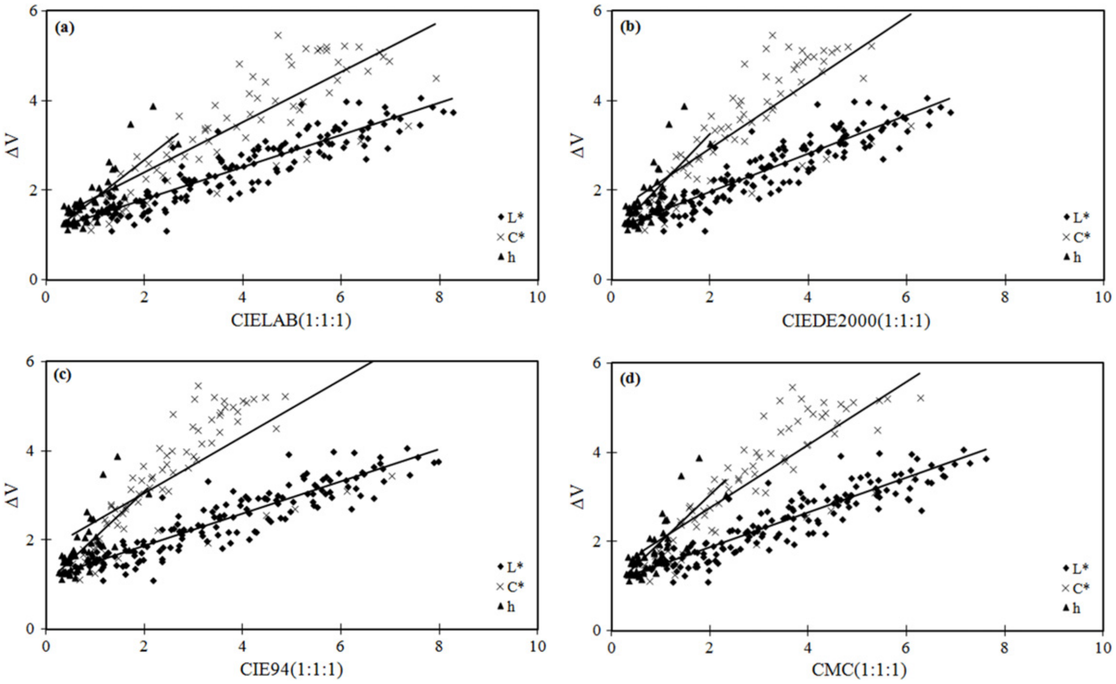

Figure 5. The relationship between visual difference (ΔV) and calculated color difference in lightness, chroma, and hue for (a) CIELAB(1:1:1), (b) CIEDE2000(1:1:1), (c) CIE94(1:1:1), and (d) CMC(1:1:1) from Experiment II data.

Table V. The coefficients and correlation factors of the lightness, chroma, and hue fitted lines in Fig. 5.

	CIELAB(1:1:1)			CIEDE2000(1:1:1)			CIE94(1:1:1)			CMC(1:1:1)		
	Slope	Intercept	*R*	Slope	Intercept	*R*	Slope	Intercept	*R*	Slope	Intercept	*R*
h	0.86	0.95	0.76	1.16	0.96	0.73	1.03	1.02	0.68	1.06	0.94	0.76
C^*	0.56	1.28	0.85	0.73	1.46	0.80	0.64	1.78	0.69	0.71	1.35	0.82
L^*	0.35	1.10	0.91	0.42	1.13	0.91	0.36	1.46	0.91	0.39	1.09	0.91
$L^* : C^* : h$	0.41/0.65/1.00			0.36/0.63/1.00			0.35/0.61/1.00			0.36/0.67/1.00		
$k_L : k_C : k_H$	2.4:1.5:1			2.8:1.6:1			2.9:1.6:1			2.7:1.5:1		

Table VI. The coefficients and correlation factors of the fitted lines after optimization.

	CIELAB(2.4:1.5:1)			CIEDE2000(2.8:1.6:1)			CIE94(2.9:1.6:1)			CMC(2.7:1.5:1)		
	Slope	Intercept	*R*	Slope	Intercept	*R*	Slope	Intercept	*R*	Slope	Intercept	*R*
h	1.17	0.98	0.81	1.61	0.95	0.80	1.67	0.96	0.78	1.41	0.92	0.83
C^*	0.87	1.13	0.93	1.11	1.31	0.90	1.20	1.33	0.88	0.93	1.38	0.88
L^*	0.86	0.88	0.91	1.19	0.82	0.92	1.06	0.92	0.91	1.08	0.74	0.92

$\Delta L = \Delta C = 0$). However, here this was found not to be the case: the slopes of the fitted hue lines also changed together with changes in the fitted lightness and chroma lines. This is due to the color differences of all pairs of test images being a combination of ΔL, ΔC, and ΔH to a certain extent, and interactive effects exist between ΔL, ΔC, and ΔH. This implies that the proposed optimization method is limited by the preparation of experimental samples, i.e. color-difference samples should be prepared with a change in only one of the color attributes, either a lightness difference or a chroma difference or a hue difference. Otherwise, the optimized result will be affected by the interaction between different attributes.

The significance of the difference between the optimized form and its original formula ($k_L = 1$) was tested again using the F-test (with a critical value $FC = 0.789$ in Experiment II). The F-test values for the formulas CIELAB, CIEDE2000, CIE94, and CMC from the original ($k_L = 1$) to the optimized forms were 0.49, 0.41, 0.35, and 0.42; that is, the improvement was statistically significant for all optimized formulas (i.e., F-test values were less than FC).

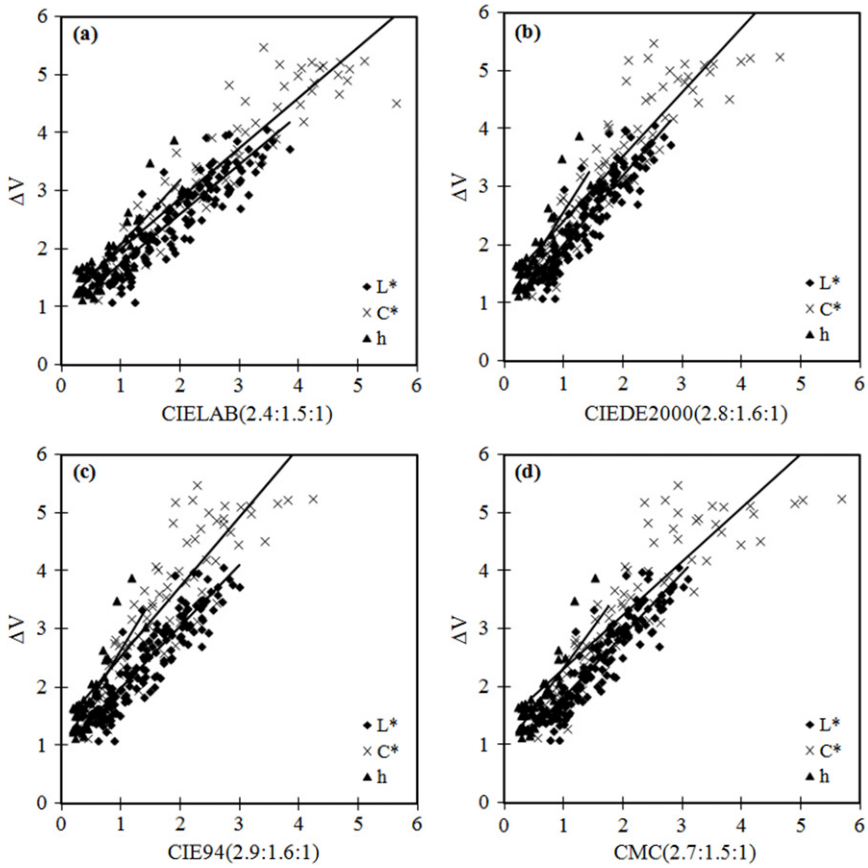

Figure 6. The relationship between visual difference (ΔV) and calculated color difference for color-difference formulas (a) CIELAB(2.4:1.5:1), (b) CIEDE2000(2.8:1.6:1), (c) CIE94(2.9:1.6:1), and (d) CMC(2.7:1.5:1) optimized from Experiment II data.

CONCLUSIONS

The magnitudes of perceptual color differences for complex images are roughly proportional to those of calculated color differences even though the color-difference formulas were developed from homogeneous color samples. However, the trends of lightness, chroma, and hue sensation were very different to those of the calculated differences. A method to optimize color-difference formulas for computing the color difference in images was proposed in this article, by which weighting factors k could be calculated simply by using the slope ratio of best-fit lines of the lightness, chroma, and hue attributes. The accuracy of the optimized values of k depends on the sample's color-difference independence from three perceptual attributes.

ACKNOWLEDGMENT

This research is supported by the National Natural Science Foundation of China (contract grant number: 61078048, 61040066, 61308081), and key research project of BIGC (No.: E-a-2003-13, 23190113049).

REFERENCES

[1] F. H. Imai, N. Tsumura, and Y. Miyake, "Perceptual color difference metric for complex images based on Mahalanobis distance," J. Elec. Imag. **10**, 385–393 (2001).

[2] T. Eerola, "Computational quality of digitally printed images," Ph.D. thesis (Lappeenranta University of Technology, Lappeenranta, Finland 2010).

[3] M. Pedersen, "Image quality metrics for the evaluation of printing workflows," Ph.D. thesis (University of Oslo, Oslo, Norway 2011).

[4] CIE Publication 199:2011, *Methods for Deriving Colour Differences in Images* (CIE Central Bureau, Vienna, 2011).

[5] M. Stokes, "Colorimetric tolerances of digital images," MSc thesis (Rochester Institute of Technology, Rochester, 1991).

[6] CIE Publication 15:2004, *Colorimetry* (CIE Central Bureau, Vienna, 2004).

[7] F. J. J. Clarke, R. McDonald, and B. Rigg, "Modification to the JPC79 Colour-difference Formula," J. Soc. Dyers Colourists **100**, 128–132 (1984); 281–282.

[8] CIE Publication 116:1995, *Industrial Colour-difference Evaluation* (CIE Central Bureau, Vienna, 1995).

[9] J. Uroz, M. R. Luo, and J. Morovic, "Perception of colour difference between printed images," *Colour Science: Exploiting Digital Media*, edited by L. W. MacDonald and M. R. Luo (John Wiley & Sons, New York, 2002), pp. 49–73.

[10] J. E. Gibson, M. D. Fairchild, and S. L. Wright, "Colorimetric tolerances of various digital image displays," *Proc. IS&T/SID Eighth Color Imaging Conf.* (IS&T, Springfield, VA, 2000), pp. 295–200.

[11] X. Zhang and B. A. Wandell, "A spatial extension of CIELAB for digital color image reproduction," J. Soc. Information Display **5**, 61–63 (1997).

[12] T. Song and M. R. Luo, "Testing color-difference formulae on complex images using a CRT monitor," *Proc. IS&T/SID Eighth Color Imaging Conf.* (IS&T, Springfield, VA, 2000), pp. 44–48.

[13] C. Sano, T. Song, and M. R. Luo, "Colour differences for complex images," *Proc. IS&T/SID Eleventh Color Imaging Conf.* (IS&T, Springfield, VA, 2003), pp. 121–126.

[14] M. R. Luo, G. Cui, and B. Rigg, "The development of the CIE 2000 colour-difference formula: CIEDE2000," Color Res. Appl. **26**, 340–350 (2001).

[15] CIE Publication 142-2001, *Improvement to Industrial Colour-Difference Evaluation* (CIE Central Bureau, Vienna, 2001).

[16] ISO 12640-1: 1997, *Graphic Technology - Graphic Technology - Prepress Digital Data Dxchange - CMYK Standard Colour Image Data (CMYK/SCID)* (ISO, Geneva, 1997).
[17] M. Xie, "Study on methods of image quality evaluation based on the color-difference formula," MSc thesis (Beijing Institute of Graphic Communication, Beijing, 2010) (in Chinese).
[18] M. Xie, H. Liu, and M. Huang, "Subjective Evaluation experiment and data analysis of color difference for digital images," *Proc. ICIS's 31st Int'l. Congress on Imaging Science* (ICIS, Beijing, 2010), pp. 180–183.
[19] H. Liu, M. Huang, G. Cui, M. R. Luo, and M. Melgosa, "Color-difference evaluation for digital images using a categorical judgment method," J. Opt. Soc. Am. A **30**, 616–626 (2013).
[20] Y. Wang, "Color difference evaluation for printed images based on the color-difference formulae," MSc thesis (Beijing Institute of Graphic Communication, Beijing, 2011) (in Chinese).
[21] M. Melgosa, P. Garcia, L. Gómez-Robledo, R. Shamey, D. Hinks, G. Cui, and R. Luo, "Notes on the application of the standardized residual sum of squares index for the assessment of intra- and inter-observer variability in color-difference experiments," J. Opt. Soc. Am. A **28**, 949–953 (2011).
[22] P. A. García, R. Huertas, M. Melgosa, and G. Cui, "Measurement of the relationship between perceived and computed colour difference," J. Opt. Soc. Am. A **24**, 1823–1829 (2007).
[23] H. Liu, M. Xie, and M. Huang, "Image color-difference evaluation based on color-difference formula," *Proc. IEEE's 4th Int'l. Congress on Image and Signal Processing* (IEEE, Shanghai, 2011), pp. 1801–1804.