A novel method for evaluating color accuracy of displayed images with different formulae*

LIU Haoxue[1], HUANG Min[2,1], WU Bing[2,1], XU Yanfang[1]

(1. School of Printing and Packaging Engineering, Beijing Institute of Graphic Communications, 25 Xinghua NR., Daxing District, Beijing 102600, China)

(2. National Lab of Colour Science and Engineering, School of Information Science and Technology, Beijing Institute of Technology, Beijing 100081, China)

Abstract

A new method to evaluate color accuracy of displayed images is established, which samples the typical color values from the images and measure them on the display. The measured color difference calculated by CIEDE76, CIEDE94, CIEDE2000, CMC(1:1) and CMC(2:1) color differences formulae are compared with subjective evaluation difference and then consistency between calculated and subjective color difference is discussed in the paper. The experiment results show that the human eyes are more sensitive to the neutral color so that the small change of color value can invoke relative large difference sensation while the measured and calculated difference is relative smaller. The color difference calculated by CIEDE76 is much smaller than subjective perceived difference in the near neutral color area and that in the green area is larger than perceived difference, so that the consistency between measured and perceived color differences is very poor. Other formulae show much better result than that of CIEDE76. The results show that CMC(2:1) color difference is the best one for evaluating color images among the 5 color difference formulae, the calculated color difference is almost linear to the visual one. The color difference calculated by CIEDE2000 formula is also smaller than subjective perceived color difference in the near neutral grey area but has good consistency in other area.

Keywords: *color difference, subjective evaluation, perceptible color difference threshold, color image, soft proofing*

1. Introductions

Monitors are the necessary device of computer system and they play a very important roll in image processing process. Operators judge the color accuracy and image quality according to the displayed color. In other hands, the monitor manufacture technology is developed so fast that makes quality of displayed images very high, especially the large dimension and high resolution LCDs meet the needs of image processing industry. In graphic art and printing industry soft proofing is now developing and becoming a main method of communication between manufacture and customers, therefore some of them replace partly the hardcopy proofing system[1].

However, the color sensation of displayed images and hardcopy is somewhat different because they use different principles to render the color. How much difference they make when they render the same image and how much difference is allowable are the key factors when using soft proofing system. Tao Song and Ronnier Luo changed the displayed images to make a series of color difference to test perceived color difference for displays and check every pixel between two images[2]. Joan Uroz et al. tested perceived color difference of printed images by changing images' lightness, chroma and hue angle individually, and conclude that the perceivable color difference for prints is about $2.5\Delta E_{ab}^{*}$[3]. H. X. Liu et. al also studied the color accuracy of hardcopy proof and soft proof images and resulted in about $2.5\Delta E_{ab}^{*}$ of average perceptible color difference for hardcopy[4,5]. X.M Zhang proposed S-CIELAB method to calculated color difference of images, which processes color information with a low pass spatial filter to simulate the color perception process of eyes[6]. However, all these methods have to calculate color difference by indirect methods and can not measure the color difference of image directly. The new image evaluation method proposed in this paper samples typical color pixels from original images and direct measure these typical color on the displays. The method supposed that it is the typical colors that make the perceived difference between soft and hardcopy proofs. The validity of the method is tested by the color image evaluation experiment.

2. Experiment design

First of all, all displays and printer have to be exactly calibrated and characterized to ensure the outputted color accuracy. The testing images are then printed and served as hardcopy proofs used in the experiment. The printed color accuracy is monitored and tested by measuring color control strips. The hardcopy images are viewed in the standard light booth and compared them with displayed images. The color difference between soft proof and hardcopy is divided into 5 grades, which are no difference(grade 1), very small difference or perceivable color difference(grade 2), acceptable difference(grade 3), relative large difference but still acceptable(grade 4) and not acceptable(grade 5) respectively. The lower a grade is, the smaller perceptive color difference.

Changing the test images in Photoshop to LAB color mode with the printer's profile, which is just the same one used when the test images are printed, the target color data can be then obtained in this way. The perceived color difference is consisted of displayed and printed errors. Because printed color is easy to measure and control and the printed color error is assumed much smaller than displayed one, only displayed error is therefore considered in the experiment.

Although each test image contains large amount of pixels, most of them are of almost the same or very similar color values and very few pixels can be used to represent the typical colors in the image. A very easy way to find out these typical pixels is to reduce

*Supported by the key subject construction project of Beijing Education Committee (No. 01422313010006-77)

image's pixels to 72 dpi in Photoshop with "Bi-Cube mode". This step takes effect of spatial filter similar to that X. M. Zhang used in S-CIELAB module[7]. Finally, reducing the each image to no more than 1000 pixels in Photoshop with "Nearest Neighbor" mode, the processed image contains the typical color values. With these hundreds of typical pixels the outline of the image can be still seen. Save the color values of image in a text file and import it into Evaluation window of Monaco Optix Pro with "Load Chart" button, then the typical color can be displayed on the screen and then measured(Fig.1).

We assume that it is the difference between target color and displayed color that produces the perceived difference. Therefore, calculate the average displayed color difference and compare it with visual difference, the relation between measured and visual color difference can be found.

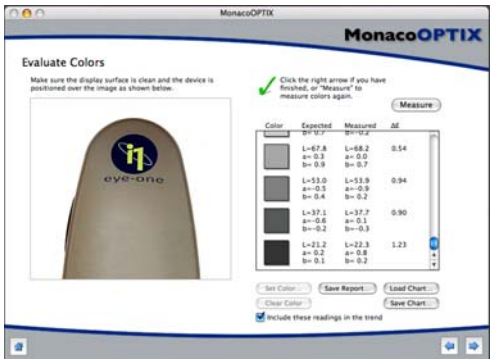

Fig.1 The measuring data can be loaded and measured in Monaco Optix Pro evaluation window.

3. Experiment conditions

The following equipments and software are used in the experiment:

LCD: Apple Cinema HD 23" (1920*1200) and Apple Cinema HD 30" (2560*1600);

Computers: Macintosh MacPro (Mac OS X 10.4.8);

Light booth: GretagMacbeth The Judge II;

Spectral photometer: GretagMacbeth Eye One;

Software: Monaco Optix Pro 2.0, Photoshop CS2, GretagMacbeth ProfileMaker 5.

6 test images are used in the experiment and the outlines of them are shown in Fig.2. Among the images, LX.tif is an oil painting, which is consisted of mainly reddish and yellowish colors but of rich texture. 5517.tif is a Chinese painting, which is almost a grey image but has reddish flavor. Others are photo works, which are of different colors and tone. Most of colors in Parrot.tif are dark green and cyan, and that in the Girl.tif are lighted grey and flesh color, in the Flower.tif are saturated colors and in the BDLG.tif are sky blue and golden color. The test images represent different kind of one and the paintings are known as very difficult to reproduce. The original images are of very large amount of pixels and very high resolution, their original parameters are listed in Table 1.

20 observers took part in the subjective evaluation experiments. All observer are engaged with printing or something to do with printing, such as painter, designer, and therefore have experience of evaluating color. They are aged from 22~70.

4 Experiment results

4.1 Subjective evaluation

Table 2 lists the results of subjective evaluation. The viewing experiments are carried out on 2 LCDs, Apple Cinema HD 23" and 30". The row tagged as MEAN corresponding to the average estimated color difference grade of 20 observers, STDEV corresponding to the standard deviations. It is can be seen that the results for two displays have very similar rank for 6 images and the standard deviations are relative small, but the grades of 30" LCD are a little lower than that of 23" LCD's indicating 30" LCD has smaller perceived color difference. Because the results of two displays are consistent very well, we only use the results of 30" to compare with measured results in the paper.

Table 1. The original parameters of test images

	Parrot.tif	Girl.tif	Flower.tif
Pixels	1696*2545	2912* 4368	3156* 2173
Size(mm)	144*215	211*317	250*172
	BDLG.tif	LX.tif	5517.tif
Pixels	3150*1203	3755*3755	10197*5110
Size(mm)	200*76	318*318	740*310

Table 2. Average results of subjective evaluation

Image:		5517	LX	Girl	Flower	BDLG	Parrot
23"	MEAN	3.45	3.3	2.8	2.5	2.6	2.55
	STDEV	0.76	0.73	0.52	0.61	0.68	0.89
30"	MEAN	3.15	2.85	2.65	2.55	2.45	2.4
	STDEV	0.75	0.67	0.49	0.6	0.51	0.68

4.2 Measured results

The measured results on Apple Cinema HD 30" are listed in Table 3. The rows tagged as M and S are the same meaning as Table 2. In addition to CIELAB color difference CIEDE94, CIEDE2000, CMC(1:1) and CMC(2:1) color difference for each image are also

Table 3. Measured color difference of 6 test images calculated with different formulae

Image:		5517	LX	Girl	Flower	BDLG	Parrot
DE76	M	1.54	2.36	1.75	2.01	1.73	2.04
	S	0.31	0.35	0.50	0.59	0.70	1.14
DE94	M	1.41	1.76	1.42	1.34	1.09	1.40
	S	0.29	0.28	0.41	0.34	0.50	0.64
DE 2000	M	1.47	1.94	1.82	1.28	1.05	1.34
	SV	0.30	0.34	0.49	0.36	0.52	0.64
CMC (1:1)	M	2.07	2.33	1.93	1.45	1.32	1.76
	S	0.48	0.57	0.71	0.52	0.69	0.88
CMC (2:1)	M	2.02	2.10	1.83	1.31	1.24	1.51
	S	0.47	0.45	0.53	0.44	0.60	0.78

calculated and listed in Table 3, with which the effect to represent

Parrot.tif

Flower.tif

BDLG.tif

5517.tif

LX.tif

Girl.tif

Fig.2 Test images used in the experiments.

color difference in each formula can be compared. It can be seen that calculated color differences for the same image but in different formula are also different, some of them are of relative large difference.

5. Discussions

Just as mentioned above, the painting images(5517.tif and LX.tif) are difficult to reproduce, very small difference can trigger off relative larger difference sensation, the evaluated color difference grades for these two images are therefore 3.15 and 2.85 respectively. In the image, Girl.tif, there are large areas of white, highlighted grey and flesh color. Our eyes are also sensitive to these colors and therefore get relative large difference grade(2.65).

The subjective estimated grades are between 2.5 to 3.2 and average grade for all images and all observers is 2.8. This means that the displayed color images are very close to the hardcopy, the average perceived difference is between perceivable and acceptable grades. This result shows that if the image system has been well color managed the displayed images can simulate hardcopy very accurately and therefore can be used as soft proofing.

To establish the relation between subjective evaluated difference grades and directly measured difference on the display, put the perceived grade as x-coordinate, measured difference as y-coordinate and plot the results in a chart as shown in Fig 3. In Fig.3 the results of different formulae are plotted in different dot shape and different colors. From Fig.3 we can find out their relation easily.

From the data in Table 2 and Table 3 and referring to the test images in Fig.2, it can be seen that the 5517.tif has the largest perceived color difference among the test images but it has very small corresponding measured difference because it is almost a grey image. This means that the color difference calculated by CIELAB formula is much smaller than perceived one. The CMC(1:1) and CMC(2:1) formulae have improved greatly.

Parrot.tif contains much dark green and cyan, some of such color is located out of the display gamut so that it makes large measured color difference. The largest measured color difference is 12.39 ΔE^*_{ab} made by out of gamut color. However, because it is

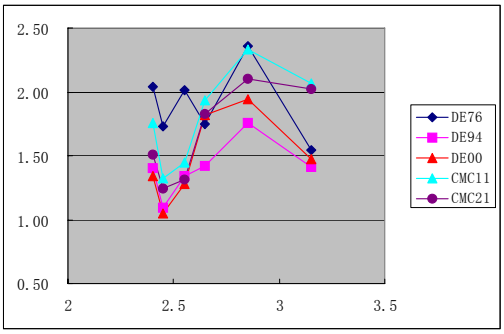

Fig.3 Subjective grades vs. measured color differences.

colorful and the eyes are relatively not sensitive to such color so that the average perceived color difference grade is very low, only 2.4. The average color difference in CIELAB is therefore 2.04, larger than 5517.tif. It is evident that this result does not accord with the visual result. The calculated color differences in CIEDE94, CIEDE2000 and CMC(2:1) , on the contrary, are much smaller and therefore closer to the subjective result.

Similar to 5517.tif, the image LX.tif and Girl.tif contain much unsaturated color or large white and light grey area, the subjective estimate gives again higher difference grade but the measured CIELAB color difference is relative small. The Fig.3 shows that the measured CIELAB color difference changes around constant, 1.8, and a little decrease with the increasing of perceived color difference, so we can conclude that there is no evident correlative relation between measured and perceived color difference. The measured color differences in CIEDE2000, CMC(1:1) and CMC(2:1) formulae, on the other hands, are much better, the measured differences are increased roughly with the perceived grade, among which CMC(2:1) formulae behaves best, the slope of fitted line is nearly 45 degrees.

6 Conclusions

A new method of measuring and calculating color accuracy of displayed color images is put forward, which samples typical color valves from the test images and directly measures them on the screen. The experiment results prove that the new method is feasible and practicable. The measured color differences do reflect the perceived color difference in color images.

The experiment results also show that CIELAB color difference formula has very poor consistency with visual results, the calculated differences are much smaller than perceived one in the near grey area, but large calculated color differences in green and cyan area. This is perhaps the reason why current digital proof system reproduces neutral grey not very well because it uses CIELAB color difference formula.

It can be concluded from the experiment results that CMC(2:1) formula behaves very well when color difference of images is evaluated by it, the calculated color difference is almost linear to the visual one. The color difference calculated by CIEDE2000 formula is also smaller than subjective perceived color difference in the near neutral grey area.

Because only 6 images are tested by the new method in the experiment, the conclusions should be test further.

Acknowledgements

The project is supported by the signal and information processing key construction project of Beijing education committee, and the experiments are carried out partly in Beijing Sun Rainbow company and Beijing C & C printing corporation, the authors would like express their thankfulness to them and people who attended the experiments as observers. The authors also owe a lot of thanks to colleagues working in Color Laboratory, who provided us great support in providing experimental instruments and conditions.

References

[1] Bob Chung, Soft Proofing, The Last Digital Frontier in Printing and Publishing, The Softproofing Era Conference, Sep. 27, 2005, Hong Kong. www.softproofing.com.hk/presentation/ 4RobertChung.pdf.

[2] T. Song and M. R. Luo, Testing colour difference formulae on complex images using a CRT monitor, The eighth Colour Imaging Conference, IS&T and SID, Scottsdale, Arizona, 7-10 November, (2000) pp44-48.

[3] J. Uroz, M. R. Luo and J. Morovic, Perception of colour difference between printed images, L. W. MacDonald and M. R. Luo (Eds), Colour Science: Exploiting digital media, John Wiley & Sons Ltd, (2002) pp49-73.

[4] Liu Haoxue, LV Yuxiang, XU Yanfang, LIANG Jiong. Subjective Evaluation for Printing Image's Colour Difference，2005 International Conference on Printing and Publication，Shanghai, 2005.11

[5] Haoxue Liu, Min Huang, Bing Wu, Yanfang Xu, Evaluation of Printing Soft Proofing，2006中国科协年会——数字成像技术及影像材料科学学术交流会论文集，2006.09Beijing, p188.

[6] X. M. Zhang, et al, A spatial extension of CIELAB for digital color image reproduction, *SID Journal*.

[7] X. M. Zhang, Farrell, & Wandell, Application of a spatial extension to CIELAB, *IS&T/SPIE Electronic Imaging 97* (1997).

Author Biography

LIU Haoxue, professor of school of printing and packaging engineering, Beijing institute of graphic communication. In 1987, he received Msc. degree from Xi'an Institute of Optical Precise Machine, Chinese Academy of Science. His research interests focus on the color science and imaging technology. He has published more than 50 papers.