Colorimetric Thresholds for Printed Images

Tao Song and Ronnier Luo
Colour & Imaging Institute, University of Derby
Derby, United Kingdom

Abstract

Since Mike Stokes started his study on the colour difference evaluation metrics on pictorial images application, more and more researches in this field were investigated recently. The CIE TC 8-02 is devoted to study, develop and standardise methods to derive colour difference for images. In it's latest meeting, it is agreed that a set of test images will be produced based upon the JIS/SCID and SHIPP images. This study is associated with this purpose and color difference thresholds were invested for printed pictorial images.

Seven transforms were applied to the standard images. They are: sigmoidal compression in L*., power compression in L*, power compression in C*, linear compression in C*, hue angle shift, power compression in both L* and C*, and linear compression in L* and C*.

The effects of transforms were assessed by a panel of observers. The results were used to develop the perceptibility and acceptability for each transform. The thresholds were expressed colorimetricly using pixel by pixel comparisons between the test images and standard image using CIELAB color difference formula. In general, if the average ΔE between two images is greater than 2.5, the colour difference can be seen by the observers. The acceptability thresholds vary according to the image's content and the transform being used.

Introduction

Since the CIE recommended the CIELAB and CIELUV colour spaces and their corresponding colour-difference formulae to industries in 1976 [1], many progresses have been taken. These developments of color difference formulae were dominated by “traditional” color industries such as textiles, paint and plastics, which were primarily interested in large uniform areas of color in direct edge contact. Recently, as the capability of computer imaging systems and displays being more and more powerful, the study on the colour difference control for rendering high-quality full-color pictorial images became possible and easier comparing with the old days. The researches on such kind of applications were recently investigated and reported.

Stokes[2] did experiment on 6 complex CRT images and obtained the perceptibility threshold in terms of average CIELAB colour difference. Uroz[3] conducted a similar experiment on large size printed images and later Song[4] did another experiment on CRT using Uroz's images. At the mean time, Gibson[5] investigated the perceptibility threshold on different media including CRT, LCD and prints. These experiments all gave similar results on the perceptibility threshold except some results from Gibson on LCD and prints. But only Song's experiment investigated the acceptability tolerance. The deviation of some Gibson's result also indicates that different media may give different answer.

The CIE TC 8-02 is devoted to study, develop and standardise methods to derive colour difference for images. In it's latest meeting, it is agreed that a set of test images will be produced based upon the JIS/SCID and SHIPP images. These will be assessed at Derby and the results will be shown at the next TC meeting. These images will then be distributed between different sites for assessing their perceptibility and acceptability thresholds.

This study is associated with this purpose and color difference perceptibility and acceptability thresholds were invested for printed pictorial images. To examine these thresholds, a set of test images was prepared based on the previous studies and a psychophysical experiment was conducted.

Experimental Methods

Colour Transformations

Most causes of colour variability in Thermal Ink Jet printing (TIJ) are related to the variability in the size of printed individual dots. Because of this, colours tend to change in chroma and/or lightness. Hue changes will generally only appear as a lose of grey balance, but hue shifts as rotaions of the whole gamut are unlikely to happen [6]. However, in order to investigate the effect of solo hue change, we still employed hue shift transform in this experiment.

Colour changes can be applied in either an increasing or decreasing direction in a given perceptual colour dimension. Stokes' study demonstrated that threshold levels of colour changes in both directions were symmetrical [2]. Thus colour changes were only applied in the direction of gamut reduction, which gave the additional benefit of reducing the possibility of transformed colours being out of the gamut of the printer. Hue shift transform doesn't have such additional benefit but retain the symmetry characteristic. Therefore only the clock-wise rotation was used in the experiment.

The mathematical functions were based on those used by previous studies[2-5] and practice. Because the darkest colour that printer can generate and paper white will not change, there should be two anchor points for the lightness change. Therefore power function and sigmoidual function were the suitable choices for L*. Similar reason for the Chroma change except that we don't know the maximum Chroma that printer can produce. As the variability in the size of printed individual dots may change the colours in chroma and lightness simultaneously, we also simulated this situation in the experiment by using power function and multiplicative function. All the transforms can be found In Table 1.

Table 1. Colour Transform Functions

Function	Apply on	Description
Power	Lightness	$L_{out}=100 \times (L_{in} / 100)^a$ a>1
Sigmoidual	Lightness	$L_{out}=50 \times (L_{in} /50)^a$ when $L_{in}<50$ $L_{out}=50 + 50 \times [(L_{in} - 50)/50]^{1/a}$ when $L_{in}>=50$
Power	Chroma	$C_{out}=100 \times (C_{in} / 100)^a$ a>1
Multiplicative	Chroma	$C_{out}=k\,C_{in}$ $k<1$
Power	Lightness & Chroma	$L_{out}=100 \times (L_{in} / 100)^a$ $C_{out}=100 \times (C_{in} / 100)^a$ a>1
Multiplicative	Lightness & Chroma	$L_{out}=k\,L_{in}$ $C_{out}=k\,C_{in}$ $k<1$
Offset	Hue angle	$h_{out}=h_{in} + off$ $off<0$

Colour Reproduction System

The colour reproduction system used to generate the standard images and the test images was composed of an HP laptop with MS-Window 95 operation system and LCD screen. The images were then printed using an HP DesignJet large format thermal inkjet printer, on HP glossy Photo Paper, using HP dye-based CP inks.

Printer characterisation is to build up the relationship between the input Lab colour and the actual printed L*a*b* values so that the colours on the prints can be predicted. This is necessary because, unlike the uniform single colour patches, no instrument can measure all the colours in a pictorial image.

In order to characterise the printer, a colour chart with 1331 colour patches was printed and measured. Another colour chart with 396 random colour patches was printed and measured as well for evaluating the characterisation performance. After the analysis of STD images, we found that the a values in the images are within the range from 90 to 192, the b values are within the range from 81 to 212. Thinking of that colour rendering might cause some values out of these ranges, we decided to enlarge the ranges a little bit and the sample values were chosen within the new ranges. Therefore, finally the samples of L values are: [0, 25, 50, 75, 100, 125, 150,175, 200, 225, 255]; samples of a values are: [72, 86,100,114,128,142,156,170,184,198,212]; samples of b values are: [60, 77, 94, 111, 128, 145, 162, 179,196, 213, 230]. The 1331 colour patches were the combination of these three groups of values.

Tetrahedral method[7] was used in the charaterisation model to predict the colours that falled outside the measured samples. The performance data are given below (in CIELAB ΔE unit):

Min: 0.12 Max: 6.51 Mean: 1.64 Std: 1.05

From the data, we can see that the charaterisation model performs reasonably well.

Image Selection and Generation

Two images (See Figure 1) were selected for use in this study to keep the number of stimuli manageable.

Girl

Threads

Figure 1. The STD Images

The first image, denoted as “Girl”, comes from the JIS XYZ/SCID standard.[8] This is the only image that contains skin tone in the ISO working draft proposal [9], which is the combination of JIS XYZ/SCID images and SHIPP XYZ/SCID images [10]. It is suitable for the assessment fresh tone reproduction and the blue area reproduction. The second image, denoted as “Threads”, comes from the SHIPP images, which contains several saturated colours and has a large colour gamut. These two images were encoded in TIFF format using Lab colour space.

Once the images were selected, in order to make sure the image gamut is within the printer gamut, they were firstly transferred from the Lab values of each pixel in the image into CMYK values of the printer using ICC profile, which was created by the printer manufacturer. And after that, the CMYK values were converted back to the Lab value using the reverse ICC profile. These two final Lab (8 bits per channel) format files are used as our standard images.

By doing so, all the colours in the original images was clipped on the printer gamut boundary and the colours in STD images are assured to be within the printer gamut.

All the colour renderings were based on these STD images and then printed in A3 size with 600 dpi resolution and with all the printer built-in colour management controls being disabled. Figure 2 illustrated the overall working flow of the image generation.

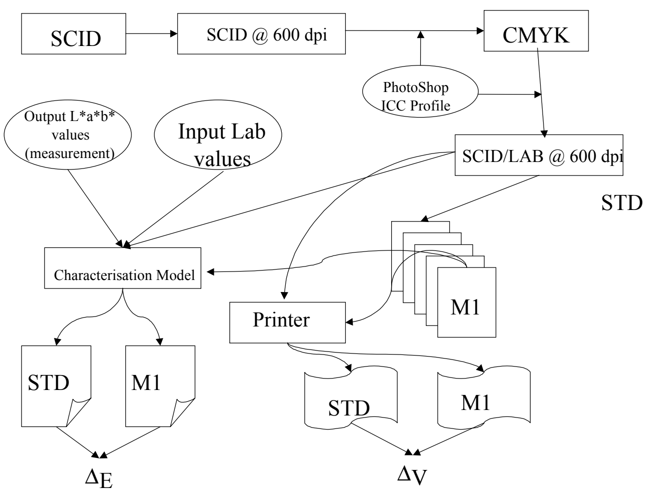

Figure 2. Working Flow of Image Generation

Viewing Environment

The pairs of test images were judged inside a Macbeth SpectraLight II viewing cabinet placed in a dark room. The light source was simulating the CIE Standard Illuminant D65. We used D65 instead of D50, which is recommanded in the Graphic Arts industry [11], is because the XYZ/SCID images were taken under the D65 illuminant.

The viewing distance was fixed to about one meter so that the angle subtended by a 600 dpi printed dot is about 9 seconds of arc. This is equivalent to a spatial frequency of 200 cpd, which is far beyond the highest detectable spatial frequency (at high luminance levels) of 50-6- cpd [12]. Observers were asked to adapt to viewing conditions for a few minutes before starting the test. They can adjudge the viewing angle slightly to avoid the gloss. The viewing geometry was 45/0. These are shown in Figure 3.

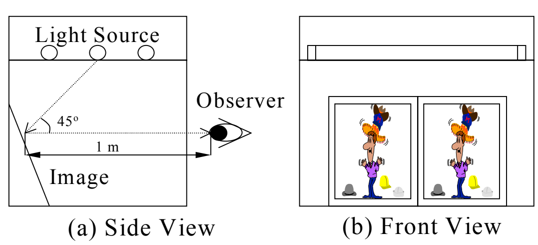

Figure 3. Viewing Condition

Psychophysical Experiment Method

A modified version of the method of single stimulus was used. The test images were shown to observers in random order together with the STD image. Observers were asked to answer the questions: a) whether or not they could see difference between both images, and b) if they could see difference, whether or not they could accept such reproduction. The assessments were in term of colour fidelity and differences due to other season such as physical damage were ignored. The duration of the judgement sessions was controlled within 45 minutes to avoid observer fatigue. The number of pairs per session was 49, which is 7 transforms by 7 parametric levels.

For each visual stimulus (pair of images), the following data were collected:

The average colour difference between two images using the CIELAB colour difference formula pixel by pixel (computed on a HP Visualize workstation with HP-UNIX operation system).

The number of times that the colour difference between two images can be perceived.

The number of times that the colour difference between two images was not accepted.

Data from threshold experiments is often analysed using Probit analysis [2-4]. This method was used again and the threshold stimulus intensity corresponds to 50% of positive responses.

Results & Discussion

Perceptibility

In order to be able to access the acceptability thresholds, the maximum colour differences between the trial images and the STD images were purposely chosen large. In the other hand, in order to limit the amount of trial image, for the reasons of avoiding changing ink cartridge and controlling the total assessment time, we only choose seven parametric levels for each transform.

Table 2. The Minimum Difference for Each Transform and Response

	Girl		Threads	
	ΔE^*_{ab}	Positive Response	ΔE^*_{ab}	Positive Response
PL	2.71	83%	2.60	78%
SL	2.71	83%	2.57	78%
PC	2.69	83%	2.50	67%
MC	2.58	92%	2.54	67%
HO	3.04	83%	3.56	78%
PLC	2.71	100%	2.60	67%
MLC	2.74	92%	2.56	89%

It turns out later that the minimum differences are all greater than 2.5 ΔE^*_{ab}. According to previous studies, the perceptibility of pictorial images is around 2 ΔE^*_{ab}. Therefore perceptibility can not be obtained directly from this set of test images. Table 2 lists the minimum difference for each transform of every image and their corresponding percentage positive response. From the data, we can see the current results agree the previous studies in some degree. The perceptibility should be around 2 ΔE^*_{ab}.

Acceptability

Acceptability is obviously trickier than perceptibility. The criteria may vary according to the background of the observers such as age, experience, profession and so on. In stokes' study,[2] the acceptability is larger than 6 ΔE^*_{ab} while in Song's study,[4] it's only 4.4 ΔE^*_{ab}. This situation is partly due the development of technique. For example about 10 years ago, a computer with 4M RAM was thought as acceptable or even excessive, but now definitely unacceptable. Also whether or not image contents influences in the acceptability tolerance of colour changes is an unsolved question. Observers make the judgement with more or less preference and also their criteria may vary from time to time.

Table 3 tabulates the acceptability results of current study and Figure 4 shows the percentage of rejections. From the data, we can see the acceptability for 'Girl' image is smaller than that of 'Threads' image. This might be because observers were more sensitive to the fresh tone change. Most of the observers stated that they couldn't accept the difference because the girl in reproductions looks being ill. For the 'Threads' image, the reasons for unacceptable are much more then those for 'Girl' image. Therefore, the image content did influence the judgements of observers in this study. The transforms affected acceptability as well but not significantly. The solo hue shift transformation having the largest tolerance indicates that hue change was the most tolerated. Lightness and Chroma changes gave similar results.

Table 3. The Acceptability for Each Transform in ΔE^*_{ab}

	Girl	Threads
PL	2.76	4.25
SL	3.00	4.00
PC	2.91	4.10
MC	2.80	3.50
HO	3.60	5.36
PLC	2.65	5.44
MLC	2.65	4.33

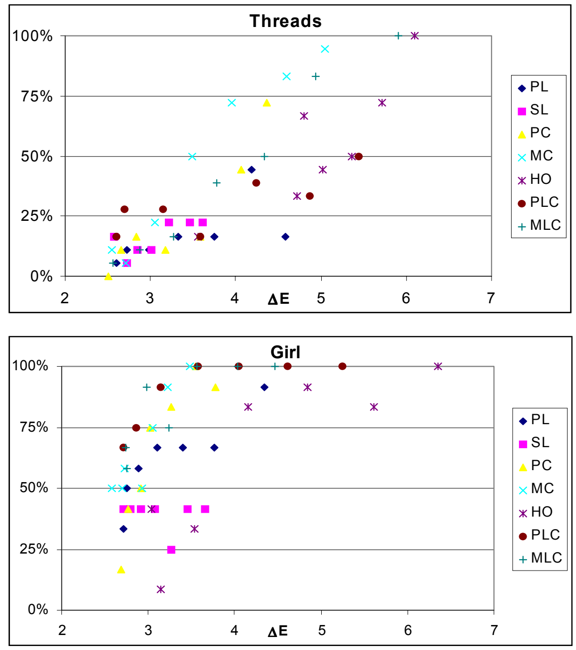

Figure 4. Percentage of Unacceptable Responses

Conclusion

An experiment associating with CIE TC8-02 task was conducted. In this experiment, a set of test images was created and assessed to obtain the perceptibility and acceptability for pictorial images. From the experimental results, conclusions can be drawn as follow:

1) The perceptibility from this study agrees with those previous studies in some degree.
2) Acceptability is image content dependent.
3) The functions for transform influence acceptability but not significantly.
4) Among the Lightness, Chroma, and Hue transform, Hue change was most tolerated. Lightness and Chroma changes gave similar results.

Unfortunately perceptibility thresholds could not be obtained from this study. But somehow the study results indirectly support the previous researches. Generally, for the printing industry, if the average ΔE^*_{ab} between original and reproductions can be controlled within 2 units, most of the customers would be satisfied with the reproduction quality. Also because the perceptibility thresholds of pictorial images for printing and CRT monitor are similar,[2-4] it is possible to use CRT monitor to simulate prints, which will give more accurate results.

In the future, more images should be investigated in order to achieve more general results.

Acknowledgements

The authors thank Joan Uroz and Hewlett-Packard Barcelona for technical advice and image preparation. Time and patience of observers who took part in the experiments is highly appreciated.

References

1. Colourimetry. CIE Publication 15.2. Vienna: CIE Central Bureau, (1986).
2. M. Stokes, *Colourimeteric tolerances of digital images,* MSc Thesis, RIT, University of Rochester, New York, USA. (1991).
3. Joan Uroz, Colour Difference Perceptibility for Large-Size Printed Images, MSc Thesis, University of Derby, Derby, (1999)
4. T. Song and R. Luo, Testing Color-Difference Formulae on Complex Images Using a CRT Monitor, Proceeding of 8th Colour Imaging Confrerece, (2000).
5. Jason E. Gibson, Mark D. Fairchild, Steven L. Wright, Colorimetric Tolerances of Various Digital Image Displays, Proceeding of 8th Colour Imaging Confrerece, (2000).
6. Joan Uroz, Private communication, (2001).
7. P. Hung, Colorimetric Calibration in Electronic Imaging Devices Using a Look-up-table Model and Interpolations, Journal of Electronic Imaging, Vol 2(1), pg. 53, (1993).
8. JIS X 9204:2000, Graphic technology - Prepress digital data exchange- XYZ/sRGB standard colour image data (XYZ/SCID), (2000).
9. ISO 12640-2, Graphic technology - Prepress digital data exchange- XYZ/sRGB standard colour image data (XYZ/SCID), Working draft, (2000).
10. K. Sakamoto, H. Urabe Standard high Precision Pictures: SHIPP, The Fifth Color Imaging Conference: Color Science, Systems, and Applicatins, (1997).
11. ISO 3664:1998(E), Viewing Conditions - for Graphic Technology and Photography. Working draft (2000).
12. B. Wandell, Foundations of vision, Sinauer Associates, Inc. (1995).