Predicting Visible Image Degradation by Colour Image Difference Formulae

Eriko Bando, Jon Y. Hardeberg, David Connah, Ivar Farup
The Norwegian Color Research Laboratory, Gjøvik University College, Gjøvik, Norway

Abstract We carried out a CRT monitor based psychophysical experiment to investigate the quality of three colour image difference metrics, the CIE ΔE ab equation, the iCAM and the S-CIELAB metrics. Six original images were reproduced through six gamut mapping algorithms for the observer experiment. The result indicates that the colour image difference calculated by each metric does not directly relate to perceived image difference.

Key words **Colour Image Difference, Image Consistency, CRT Display, iCAM, S-CIELAB**

1 Introduction

Digital imagery has become one of the major image reproduction methods, and according to the diversity of imaging methods, there is a strong need to quantify how reproduced images have been changed by the reproduction process and how much of these changes are perceived by the human eye.

A traditional colour difference equation, the CIELAB ΔE ab colour difference formula, is still the most widely used as a colour difference metric in the graphic arts industry, although it was designed to derive colour difference for a single pair of colour patches. As a result, its root mean square error is not able to predict perceived image difference. Therefore, several digital image distortion metrics, which are designed to take into account the human visual system, have been researched and developed over recent years. In this experiment, we used two perceptual image metrics as well as the CIELAB ΔE ab colour difference formula. The S-CIELAB ΔE ab metric [1] is an extension of the CIELAB ΔE ab metric, and the aim of S-CIELAB is to take into account the spatial-colour sensitivity of the human eye. One of the most recently developed image appearance models is iCAM [2 3]. It is based on the S-CIELAB spatial vision concepts, but incorporates more sophisticated models of chromatic adaptation. It is also simpler than other multiscale observer models, which are computationally complex.

2 Experimental Setup

A pair image comparison on a CRT monitor is adopted as the method of psychometric scaling for this experiment. Six sRGB images, used in previous gamut mapping research [4], were selected (see Figure 1). Each original image has been reproduced by six different gamut mapping algorithms, so that a total of 42 images (6 originals and 36 reproductions) were used for the pair image comparison. A Dell 18-inch monitor with a resolution of 1600×1200 pixels was used to display the images, and its white point was set to D65. The monitor was calibrated daily during the experiment, and the monitor's ICC profile was used to transform CRT's RGB primaries to CIE tristimulus values. The experiment was carried out in a dark room. Two pairs of images were displayed in each trial, hence, the observers judged 4 images (2 pairs) simultaneously. Each pair consisted of one original and one reproduction and the position of original and reproduction was changed randomly to avoid the observer adapting to the original image. The observer was then asked which pair showed the least difference,

This research is supported by the Norwegian research council and media technology companies through the research project, "Quality Assured Content Management"

1.Hue-angle preserving minimum _E_ab clipping,	Only out-of-gamut colours are mapped to its closest colours.
2. GAMMA, sRGB gamut,	Used to map sRGB images.
3. SGCK, image gamut 4. SGCK, sRGB gamut	Both used for mapping image gamut to reproduction gamut, but SGCK, sRGB gamut using sRGB gamut as its source gamut.
4. SGCKC, image gamut. 5. SGCKC, sRGB gamut	Using two mapping processes, firstly SGCK algorithm, and then clipped by minimum _E_ab clipping algorithm. SGCK, sRGB gamut using sRGB gamut as its source gamut.

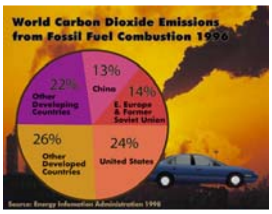
i.e. which reproduction was closest to the original.

Table 1. Gamut mapping algorithms

1. Camera

2. Girl

3. Cat



4. Pollution

5. Ski

6. Picnic

Figure 1. Sample Images

The same pairs were shown twice to confirm the repeatability. Therefore a total of 180 comparisons were made per observer. A total of 16 observers, 5 females and 11 males with ages ranging from 22 to 44, participated in the experiment. The observers were asked to take a colour deficiency test before the experiment. The observers sat on a chair which was placed 35 inches away from the monitor [5] and were asked to view the monitor's grey background for 2 minutes to adapt to the viewing conditions. The viewing angle for the monitor was about 23 degrees.

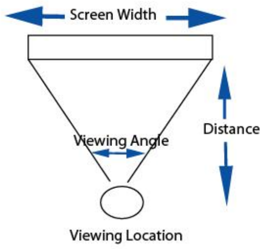


Figure 2. Viewing Conditions

3 Results

Psychophysical experiment

Except for the first (Camera) and the third (Cat) images, the hue-angle preserving minimum _E_ab clipping algorithm shows good performance. This result corresponds to a previous monitor-based gamut mapping quality survey [6], however, paper-based psychophysical experiments show different results [4, 6].

According to the overall Z score, the hue-angle preserving minimum _E_ab clipping algorithm seems to perform the best and both the SGCKC algorithms show lower z scores than other algorithms.

Image difference calculation

All the reproduced images' difference were computed using three different equations: iCAM, S-CIELAB Δ E ab and CIELAB Δ E ab. We selected the Girl image as the example for further discussion, because this image is representative of the overall Z scores (Figures 3 and 4).

The computed mean image difference values of the

Girl image are shown in Figure 5 and Table 2. We did not find any corresponding results between any of the

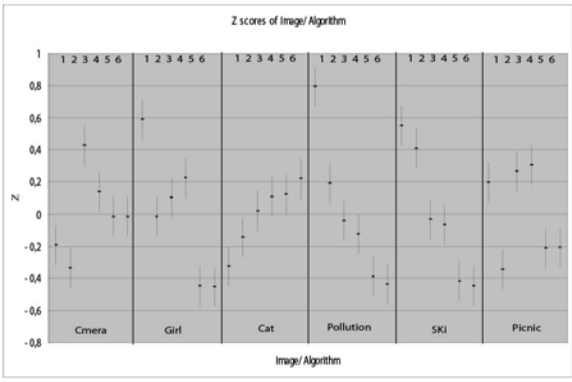


Figure 3. Z scores of each algorithm

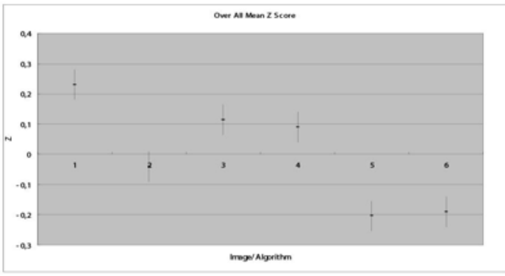
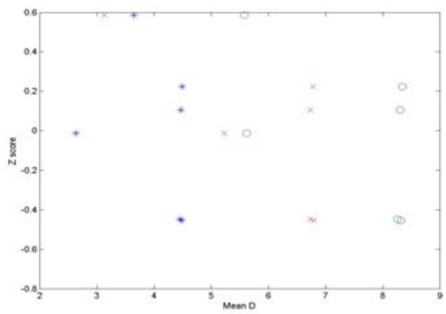


Figure 4. Overall Z score

computed results from iCAM, S-CIELAB or CIELAB ΔE ab. The overall results are plotted in Figure 6; the results from iCAM are variable, and there is no significant difference between CIELAB and iCAM or S-CIELAB. We can see that in Table 2 and 3, the mean and maximum image differences of the SGCK algorithms and the SGCKC algorithms are very close, but the difference between those two algorithm's Z scores is quite large. This suggests that the observers perceived image difference does not only depend on average or maximum pixel by pixel differences. The minimum _E_ab clipping algorithm when executed prior to the SGCK mapping algorithm does not influence pixel by pixel based image difference but does influence perceived image quality.

Image difference mapping

The following figures 7 and 8 give the visual information of image difference of the hue-angle preserving minimum _E_ab clipping (the overall best result) and SGCK, sRGB gamut mapping algorithms (the overall worst result). Note: we have to take into account that iCAM's unit of image difference is different, because it uses IPT as the colour space, instead of CIELAB.



icam:*, S-CIELAB:o, ΔE ab:x

Figure 5. Mean image difference of Girl

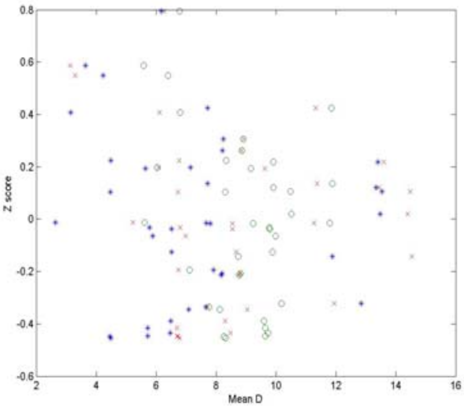
GMA	iCAM mean ΔIm	S-CIELAB Mean Δab	Mean Δab	Zscore 95% confident interval= 0,245
1	3,6531	5,5805	3,1307	0,586273
2	2,6343	5,62	5,2318	-0,01311
3	4,4727	8,3016	6,7362	0,104606
4	4,4895	8,3427	6,7766	0,223456
5	4,4583	8,2586	6,7319	-0,4473
6	4,4844	8,3214	6,796	-0,45392

Table 2. Mean image difference and Z score of Girl

GMA	iCAM Maximum ΔIm	S-CIELAB Maximum Δab	Maxmum Δab	Zscore 95% confident interval= 0,245
1	121,9071	138,7801	23,6005	0,586273
2	110,4462	91,8321	29,5777	-0,01311
3	121,1202	120,5012	24,5047	0,104606
4	121,0488	120,1001	24,217	0,223456
5	121,2534	115,5833	24,0233	-0,4473
6	121,0938	116,1212	23,7349	-0,45392

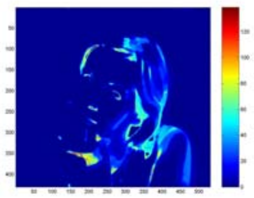
Table 3. Maximum image difference and Z score of Girl

The overall trend between the best image and the worst image is the difference of the background colour. Consequently, the image background might have a strong influence on the perceived image difference. For instance, the Cat image shows an atypical result (Figure 3), and one of the characteristics of the image is its weak texture on the dark background.

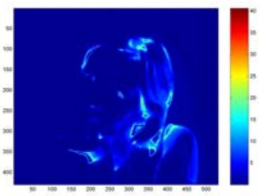


icam:* S-CIELAB:o, ΔE ab:x

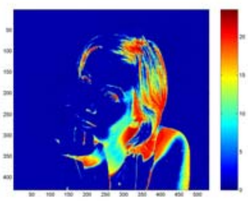
Figure 6. Mean image difference of all reproductions



S-CIELAB

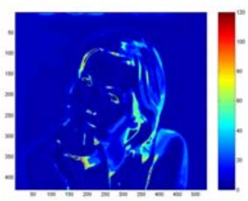


iCAM

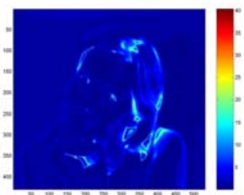


CIELABΔE

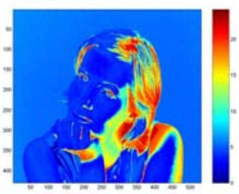
Figure 7. Hue-angle preserving minimum _E_ab clipping



S-CIELAB



iCAM



CIELABΔE ab

Figure 8. SGCK, sRGB gamut

4 Conclusions

We do not find a correlation between the perceived image differences and pixel by pixel image difference calculation values, but, it also means that there are potential improvements for iCAM and S-CIELAB. Further research will be carried out the subjects of image difference and image quality metric, perceptual difference in different media, and background effect of perceptual image difference.

Acknowledgements

The authors would like to thank Professor Lindsay MacDonald for his advice, Garrett Johnson for providing iCAM and S-CIELAB MATLAB codes, Morten Amsrud who kindly helped to reproduce images, and Øivind Kolloen for Java programming help.

References

1 X. Zhang, B. A. Wandell, "A spatial extension of CIELAB for digital color image reproduction", SID Journal , (1997).

2 M. D. Fairchild, G. Johnson, "Meet iCAM: A Next-Generation Color Appearance Model", IS&T/SIT Tenth Color Imaging Conference, pp. 33-38, Arizona, USA (2002).

3 M. D. Fairchild and G.M. Johnson, "The iCAM framework for image appearance, image differences, and image quality," Journal of Electronic Imaging, in press (2004).

4 I. Farup, J. Y. Hardeberg, and M. Amsrud, "Enhancing the SGCK Colour Gamut Mapping Algorithm", CGIV 2004, pp.520-524, Aachen, Germany (2004).

5 D. R. Ankrum, "Viewing Angle And Distance In Computer Workstations", http://www.combo.com/ergo/vangle.htm

6 P. J. Green, L. W. MacDonald, "Colour Engeneering", Jon Wiley & Sons, ltd, West Sussex, England (2002).