Comparing Different Colour Discrimination Data Sets

S. Y. Zhu, M. R. Luo, G. H. Cui and B. Rigg
Colour & Imaging Institute, University of Derby
Derby, United Kingdom

Abstract

Different large magnitude colour-difference (LCD) data sets were accumulated. A uniform colour space based upon modification of CIELAB was developed to fit each data set. These spaces were then compared to reveal the differences between the different data sets. The results showed that all LCD data sets have fairly similar characteristics to each other except for the Munsell data due to the incorrect balance between the lightness and chromatic data within the latter data set. The present results show that a one unit of Munsell Value appears larger than one unit of Munsell Chroma by a factor of 1.25.

Introduction

Uniform colour spaces (UCS) are essential for colour imaging applications such as colour management, gamut mapping, image compression, etc. Almost all spaces have been developed to fit some particular experimental data sets. However, it was found that there are large disagreements between different data sets from the authors' earlier studies.[1,2] In the earlier experiment,[1] five types of scales in CIELAB colour space were investigated: hue, lightness, chroma, a light series and a dark series using CRT colours. Observers were asked to adjust colours along 24 vectors in CIELAB colour space as shown in Figure 1. These included one lightness, five chromas along different hue angles, eight hues with two chromas in each of the four hue quadrants, 5 mixtures of lightness and choma from white to full colour (called light series) and five from black to a full colour (called dark series). The data were then used to test different colour difference formulae and uniform colour spaces (UCSs). The results clearly show that the models' performances are largely different according to the data sets used to develop these formulae or spaces. They can be categorised into three: those fitted to data with large magnitude of colour difference (named LCD), with small magnitude of colour difference (SCD), and the Munsell data. This was further supported in the authors' later study[2] using all available data sets.

The present study aims to test various formulae or UCSs using seven available LCD data sets and to reveal the differences between them. The SCD data sets were excluded because almost all UCSs have been developed based upon LCD data sets. Guan and Luo[3] discovered that a model with a general structure based upon CIELAB[4] can fit well to many data sets. The same structure is adopted here to fit each individual data set. Hence, the differences of lightness or chroma scales between different models represent the differences between different data sets.

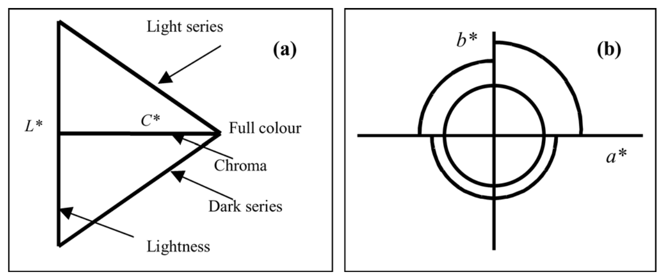

Figure 1. (a) The lightness scale, chroma scale, light series and dark series studied. (b) The hue scales studied

Testing Different Colour Models' Performance Using LCD Data Sets

Seven LCD data sets were used in this study: CII-Zhu,[1] Guan,[3] Munsell,[5,6] OSA,[7] BFB-Textile[8] and Pointer[9] including 144, 292, 844, 128, 238 and 1308 and having average ΔE^*_{ab} values of 9.9,11.4,10.2,14.3,11.7 and 8.9 respectively. Nine uniform colour spaces or colour difference formulae were tested using the above six data sets in terms of the PF/3 measure, which was developed by Luo and Rigg.[10] For example, a PF/3 of 30 indicates that there is a 30% disagreement between a model's predictions to the visual results. These spaces and formulae were selected for particular reasons. The CIEDE2000[11] represents formulae fitted to SCD data. The other spaces were developed to fit LCD data, i.e. CIELAB,[4] Kuehni,[12] SVF[13] and NC_IIIC[14] fitted to the Munsell data, OSA[7] space fitted to OSA data, GLAB[3] for the combined OSA, BFB-Textile and Guan data, and also IPT[15] fitted to a number of data sets. The Munsell data used here correspond to the physical samples used in the original experiment,[5] not the full set of the Munsell Renotation system.[6] Many extremely colourful and bright samples are included in the latter, but they were extrapolated from the original data set.[5] The colour difference pairs were formed between the neighbouring samples along Munsell Value, Chroma and Hue scales. The visual differences were set to two and one for Value and Chroma respectively (see later). The Munsell Hue differences vary according to the Chroma of given samples. CAM97s2[16] is a modified version of the CIE colour appearance model, the CIECAM97s,[17] for predicting colour appearance under different viewing conditions. For each model, two tests were carried out: the original model and the original model optimised with a lightness parametric factor (k_L), which are designated below as original and k_L models, respectively. The results in PF/3 values are summarised in Table 1 for the original and the k_L models together with optimum k_L values in brackets. The mean values for each model were also calculated.

These results show a fairly small difference between original models, i.e. the mean PF/3 values ranged from 25 to 30 except for the worst CAM97s2 (34). The difference is even smaller between k_L formulae (3 units). As expected, all k_L formulae improved from the original formulae by 1 or 2 units especially for CAM97s2 due to large difference of k_L values from unity. In addition, all models require a larger k_L value to fit the Munsell data set than those for the other data sets by a factor of 1.5. CIEDE2000, developed for fitting small colour differences, performed the worst for almost all LCD data sets. This indicates a large difference between SCD and LCD data. Overall, GLAB gave an overall better performance than the other models. This demonstrates the effectiveness of modifying CIELAB space to fit a combination of reliable data sets. It can also be found that some models performed best for a particular data set but predicted the worst for the other data set. For example, NC_IIIC developed to fit the Munsell data gave the worst fit for the CII-Zhu, OSA and Guan data sets. This indicates that the LCD data sets are inconsistent between themselves.

Developing Colour Models to Represent Each Data Set

It is valuable to understand the detailed differences between the different data sets. However, it is difficult to compare two data sets because the colour difference pairs between two data sets are different. The strategy used here is to derive colour models having the same generic structure to fit different data sets. These models can then be compared to reveal the real differences between different data sets. The first generic model developed is a modification of CIELAB having a structure similar to that of GLAB. Tolerances in colour space are commonly presented with an ellipsoid. The general form is defined as follows.

$$\Delta E = \sqrt{\left(\frac{\Delta L^*}{k_L(1+\beta_L \overline{L^*})}\right)^2 + \left(\frac{\Delta C^*}{k_C(1+\beta_C \overline{C^*})}\right)^2 + \left(\frac{\Delta H^*}{(1+\beta_H \overline{C^*})}\right)^2} \qquad (1)$$

Table 1. Testing Colour Models' Performance Using Different LCD Data Sets.

Model\Data	CII-Zhu	OSA	BFB-Textile	Guan	Munsell	Pointer	Mean
Original model (k_L=1)							
CIEDE2000	31	25	21	21	*38*	*41*	30
GLAB	26	19	23	16	33	34	25
OSA	29	20	28	18	28	34	26
IPT	27	24	32	19	25	33	27
CIELAB	29	*27*	34	27	18	35	28
Kuehni	*39*	26	33	29	18	36	25
SVF	34	22	28	19	19	33	26
NC_IIIC	38	*27*	32	27	15	36	29
CAM97s2	37	26	*35*	*37*	28	40	*34*
Optimised Formula with varying k_L							
CIEDE2000	29(0.8)	*25*(0.8)	20(0.9)	*21*(1.0)	*30*(1.5)	*38*(0.6)	*27*
GLAB	26(1.0)	19(1.0)	23(1.0)	16(1.0)	23(1.6)	34(1.0)	24
OSA	29(1.1)	20(1.0)	27(1.1)	18(1.1)	16(1.6)	34(1.1)	24
IPT	27(1.1)	24(0.9)	*32*(1.0)	19(1.0)	19(1.4)	33(0.9)	26
CIELAB	28(0.8)	24(0.6)	30(0.7)	19(0.6)	18(1.0)	34(0.7)	26
Kuehni	*35*(0.6)	22(0.6)	28(0.6)	19(0.6)	17(0.9)	34(0.6)	25
SVF	33(0.8)	21(0.8)	27(0.8)	17(0.8)	17(1.2)	33(0.8)	25
NC_IIIC	36(0.6)	*25*(0.6)	29(0.7)	*21*(0.7)	15(1.0)	35(0.7)	*27*
CAM97s2	29(0.5)	23(0.4)	26(0.5)	*21*(0.5)	27(0.8)	34(0.4)	*27*

Note: For each data set, the best model is underlined and bold, and the worst model is in italic and bold.

where k_L, k_C, β_L, β_C and β_H are five coefficients optimized to give the best fit to each data set in terms of the PF/3 measure. Equation (1) is named as 5CLAB below, indicating the five coefficients were optimised in modifying CIELAB. In addition, a three coefficients model, 3CLAB, was also developed by setting k_C=1 and β_L=0 for each data set. The optimised coefficients together with PF/3 value for each data set are given in Tables 2 and 3 respectively. The results can be summarised below:

- Equation (1) (5CLAB) fits the data quite well with an average of about 22 PF/3 units. The largest error occurs for the Pointer data set due to the larger noise in this data set compared to the others.
- The PF/3 measures for the 5CLAB and 3CLAB models are very similar (within 2 units). This indicates that k_C and β_L are not as important as the other coefficients.
- The model-coefficients for the different data sets are very different, i.e. for the 3CLAB models, the k_L, and β_C values vary by factors about 2 and 6 respectively.

Table 2. 5CLAB Coefficients and PF/3 Values for Each Data Set

Data	PF/3	k_L	β_L	k_C	β_C	β_H
CII-Zhu	22	1.3	0.0016	2.2	0.0036	0.0186
OSA	19	0.9	-0.0022	1.1	0.0117	0.0014
BFB-Textile	21	0.4	0.0210	0.8	0.0389	0.0029
Guan	15	0.5	0.0100	0.9	0.0123	-0.0012
Munsell	17	1.2	-0.0018	0.9	0.0083	0.0039
Pointer	33	0.8	-0.0028	0.9	0.0079	-0.0022

Table 3. 3CLAB Coefficients and PF/3 Values for Each Data Set

Data	PF/3	k_L	β_C	β_H
CII-Zhu	24	0.75	0.0092	-0.0003
OSA	19	0.74	0.0155	0.0005
BFB-Textile	22	0.93	0.0281	0.0061
Guan	15	0.72	0.0107	-0.0007
Munsell	17	1.19	0.0068	0.0057
Pointer	33	0.74	0.0048	-0.0004

Comparing Scales between Different Models

A quantitative method was used to compare the different colour models corresponding to different data sets. Each 3CLAB model's coefficients were used to construct an UCS as given in equation (2). The β_H coefficient in equation (1) was excluded due to fact that their values are quite small and also by removing them the UCS becomes much simpler.

$$\begin{aligned} L' &= L^* \\ C' &= (k_L / \beta_C)\ln(1 + \beta_C C^*) \\ a' &= C'\mathrm{Cos}\,(h) \\ b' &= C'\,\mathrm{Sin}\,(h) \end{aligned} \quad (2)$$

The k_L values in Table 3 indicate the differences in lightness weighting (compared to chroma and hue differences) between different data sets. The k_L value of 1.19 for Munsell is the largest, followed by BFB-Textile (0.93) with the others (around 0.75). This implies that for a pair of samples exhibiting only lightness differences, the UCSs derived from the Munsell and BFB-Textile data sets predict a smaller difference than the other UCSs.

Comparing different UCSs' chroma scales, the *C'* scale in equation (2) was used. They are plotted against the CIELAB C* scale as shown in Figure 2. The results clearly show that the *C'* scales from all models had a reasonably good agreement except for that of Munsell, which agrees best to the CIELAB *C** scale, i.e. closest to the 45o line. The chroma scale shows the largest discrepancy between the Munsell and the other data sets. Note that the equation for *C'* involves k_L as well as β_C. Much of the difference between the different *C'* scales is caused by the lightness weighting factor k_L.

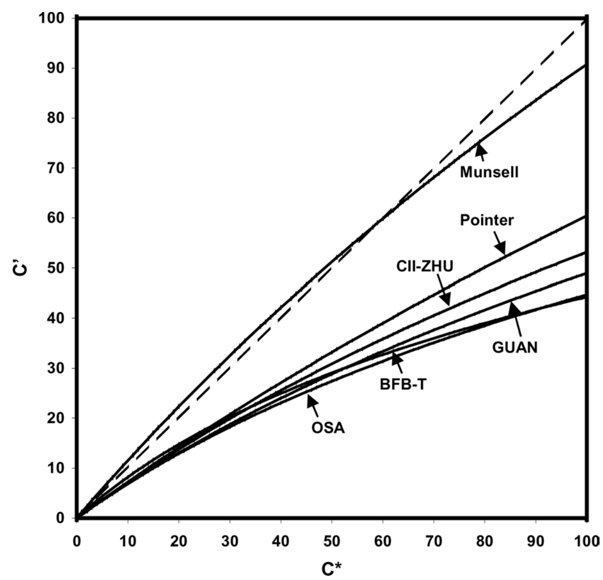

Figure 2. The C' scales developed from different data sets are plotted against the CIELAB C scale.*

Munsell Data Set

Figure 2 shows that the *C'* scale of the Munsell model is the most different from those from the other models. This is caused by a larger k_L value than those of the others. However, its β_C value is not very much different from those of the other models as shown in Figure 3, in which *C''* scale for each model calculated using $(1/\beta_C)\ \ln\ (1 + \beta_C C^*)$ are plotted. It clearly shows that Munsell *C''* is quite similar to the other *C''* scales. The discrepancy between Munsell *C''* and *C'* is caused by the balance between lightness and chromatic differences within the data. Although a careful literature survey was conducted, very little information was found regarding this balance. Almost all researchers assume that one unit of Munsell Value appears twice of one unit of Munsell Chroma.[18,19] Hence, the models developed to fit the Munsell data had a k_L value close to one (see CIELAB, Kuehni, SVF and NC_IIIC in Table 1). The present results show that the characteristics of the Munsell data are similar to those of the other LCD data sets. Considering the k_L value of 1.19 and 0.74 for the Munsell data and the average from the other data sets excluding BFB-Textile, one unit of Munsell Value should be preceived 1.25 times of one unit of Munsell Chroma in order to have a good agreement between the Munsell and the other LCD data sets.

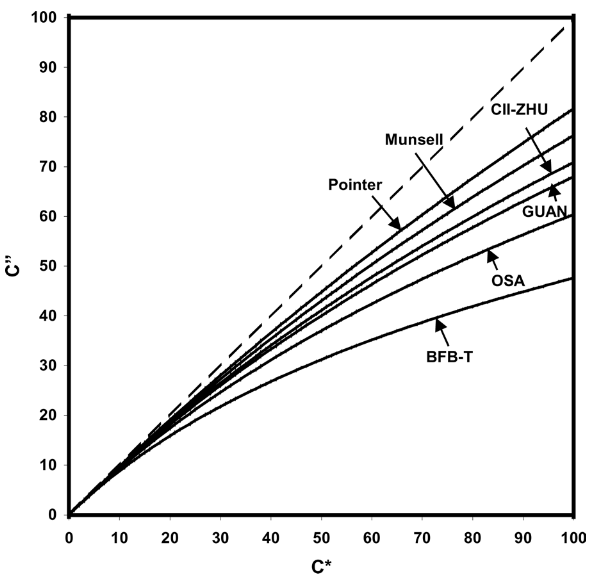

Figure 3. The C'' scales developed from different data sets are plotted against the CIELAB C scale.*

Conclusions

Various large colour-difference data sets were accumulated. A uniform colour space based upon modification of CIELAB was developed to fit each data set. These spaces were then compared to reveal the differences between the different data sets. The results showed that the Munsell data disagreed with the other data sets due to the incorrect balance between the lightness and chromatic data within the data set. The present results show that a one unit of Munsell Value is equivalent to 1.25-units of Munsell Chroma. In conclusion, all LCD data sets have fairly similar characteristics to each other.

References

1. S. Y. Zhu, M. R. Luo and G. H. Cui, New experimental data for investigating uniform colour spaces, Proc. 9th Session of the Association Internationale de la Couleur (AIC Color 2001), Rochester, USA (2001) 626-629.
2. M. R. Luo, S. Y. Zhu and G. H. Cui, Testing the uniformity of colour spaces, CIE Expert Symposium, CIE x021, 54-58, 2001.
3. S. S. Guan and M. R. Luo, A Colour-Difference Formula for Assessing Large Colour Differences, *Col. Res. Appl.,* **24**, 344-355, 1999.
4. Commission Internationale de l'Éclairage: Colorimetry, *CIE* 15.2-1986.
5. S. M. Newhall, Preliminary Report of the O.S.A. Subcommittee on the Spacing of the Munsell Colors, *J. Opt. Soc. Am.*, **30**, 617-645,1943.
6. S. M. Newhall, D. Nickerson, D. Judd, Final report of the O.S.A. Subcommittee on spacing of the Munsell colors, *J. Opt. Soc. Am.*, **33**, 385-418, 1940.
7. D. L. MacAdam, Uniform Color Scales, *J. Opt. Soc. Am.*, **55**, 1619-1702,1974.
8. A. S. Badu, Large Colour Differences between Surface Colours, Ph.D. thesis, University of Bradford, 1986.
9. M. R. Pointer and G. G. Attridge, Some Aspects of the Visual Scaling of Large Colour Differences, *Col. Res. Appl.,* **22**, 298-307, 1997.
10. M. R. Luo and B. Rigg, BFD(l:*c*) Colour-Difference Formula, Part II—Performance of the Formula, *J. Soc. Dyers Col.,* **103**, 126-132, 1987.
11. M. R. Luo, G. Cui and B. Rigg, The development of the CIE 2000 colour difference formula, *Color Res. Appl.* **26,** 340-350, 2001.
12. R. G. Kuehni, Towards an Improved Uniform Color Space, *Col. Res. Appl.,* **24**, 253-265, 1999.
13. T. Seim and A. Valberg, Towards a Uniform Color Space: A Better Formula to Describe the Munsell and OSA Color Spaces, *Col. Res. Appl.,* **11**, 11-24, 1986.
14. K. Ikeda and K. Obara, Improvement of Uniformity in Color Space in Terms of Colour Specification and Colour Difference Evaluation, *CIE Proc. 22nd*, 1991.
15. F. Ebner and M. D. Fairchild, Development and Testing of a Color Space (IPT) with Improved Hue Uniformity, *Proc. of The Sixth Color Imaging Conference: Color Science, System, and Applications*, 8-13, 1998.
16. C. J. Li, M. R. Luo and R.W. G. Hunt, A Revision of the CIECAM97s Model, *Col. Res. Appl.,* **25**, 260-265, 2000.
17. M. R. Luo and R. W. G. Hunt, The structures of the CIE 1997 colour appearance model (CIECAM97s), *Color Res. Appl.*, **23**, 138-146, 1998.
18. F. W. JR. Billmeyer, Survey of color order system, *Col. Res. Appl.,* **12**, 173-186, 1987.
19. D. Nickerson, The specification of colour tolerances, *Text. Res.,* **6**, 505-514, 1936.

Biography

Dr. M. Ronnier Luo is the Director of the Colour & Imaging Institute and Professor of Colour Science at University of Derby, UK. He received his B.Sc. in Fibre Technology from the National Taiwan Institute of Technology in 1981 and his Ph.D. in Colour Physics from the University of Bradford in 1986. He has published over 140 papers in the field of colour science. He is the chairman of the CIE TC 8-02 on colour differences in images and CIE1-52 on chromatic adaptation transforms.