New Uniform Colour Spaces

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

Abstract

Six new uniform colour spaces were developed based upon three types of colour discrimination data: small magnitude colour differences, large magnitude colour differences and the Munsell system. Three spaces representative of these six spaces, together with the some of the most widely used spaces, were evaluated by a psychophysical experiment in terms of space uniformity and hue constancy. The results show that IPT type spaces gave an overall better performance.

Introduction

From the authors' earlier studies,[1,2] a number of colour models including colour difference formulae and uniform colour spaces were tested by using available colour discrimination data. The results show that the colour models' performance varied considerably according to the data sets used to develop these models. They can be divided into 3 categories: those fitted to data with large magnitude colour differences (LCD), those with small magnitude colour differences (SCD) and the Munsell data.[3] This study includes the development of uniform colour spaces for each the three data categories. In this study, the basic structures of IPT[4] and GLAB[5] were used to fit these three data groups, because these two models gave overall quite satisfactory performance in fitting different large magnitude colour difference data sets in the authors' another study.[6] The GLAB model is a modified version of CIELAB as are many other colour difference formulae (CMC, CIE94, LCD, BFD, CIEDE2000). IPT represents another simple colour space and fits well to a variety of data sets.

Six new uniform colour spaces designated LCDLAB, SCDLAB, MUNLAB, LCDIPT, SCDIPT and MUNIPT were derived. These new UCSs together with CIELAB,[7] GLAB,[5] IPT[4] and DIN99d[8] and CAM97s2[9] were tested using available experimental data sets. A psychophysical experiment was also carried out to evaluate the space uniformity and hue constancy of these spaces. The results show that IPT type spaces gave a better performance overall.

Developing New Uniform Colour Spaces

New Models Based on IPT

Nine coefficients (k_1 to k_9) in the IPT model given in equation (1) were optimised to fit data sets in three categories: LCD, SCD and Munsell. A combined data set accumulated by Luo *et al.*[10] in connection with the development of the CIE 2000 colour difference formula, CIEDE2000 was used to represent the SCD category. The LCD category includes the CII-ZHU,[1] Guan,[5] OSA,[11] BADB-T,[12] Pointer[13] data sets. The Munsell data is used on its own due to its unique features. When optimising the IPT model to fit LCD data sets, the PF/3 value between the model's predictions and visual results for each individual data set was first calculated. The PF/3 values for all individual data sets were averaged and used as the measure for optimisation. Three new IPT colour models were derived and denoted as LCDIPT, SCDIPT and MUNIPT, which were fitted to the LCD, SCD and Munsell data sets respectively.

$$
\begin{aligned}
&\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.4002 & 0.7075 & -0.0807 \\ -0.2280 & 1.1500 & 0.0612 \\ 0.0 & 0.0 & 0.9184 \end{bmatrix} \begin{bmatrix} X_{D65} \\ Y_{D65} \\ Z_{D65} \end{bmatrix} \\
&L' = L^{k_1}; L >= 0 \\
&L' = -(-L)^{k_1}\, L < 0 \\
&M' = M^{k_1}; M >= 0 \\
&M' = -(-M)^{k_1}\, M < 0 \\
&S' = S^{k_1}; S >= 0 \\
&S' = -(-S)^{k_1}\, S < 0 \\
&\begin{bmatrix} I \\ P \\ T \end{bmatrix} = \begin{bmatrix} k_2 & k_3 & 1-k_2-k_3 \\ k_4 & k_5 & 0-k_4-k_5 \\ k_6 & k_7 & 0-k_6-k_7 \end{bmatrix} \begin{bmatrix} L' \\ M' \\ S' \end{bmatrix} \\
&C = \sqrt{P^2+T^2} \\
&h_{IPT} = \arctan(T/P) \\
&C_{IPT} = (k_8/k_9)\ln(1+k_9 C) \\
&a_{IPT} = C_{IPT}\cos(h_{IPT}) \\
&b_{IPT} = C_{IPT}\sin(h_{IPT})
\end{aligned}
\qquad (1)
$$

New Models Based on CIELAB

A structure based upon CIELAB and given in equation (2) was also used to fit the LCD, SCD and Munsell data sets. The lightness parametric factor k_L and weighting function β_c were optimised to achieve the smallest PF/3 value. Three new models based on CIELAB were derived. They are named LCDLAB, SCDLAB and MUNLAB respectively.

$$\begin{aligned} L' &= L^* \\ C' &= (k_L / \beta_C)\ln(1 + \beta_C C^*) \qquad (2) \\ a' &= C'\cos(h) \\ b' &= C'\sin(h) \end{aligned}$$

Evaluating Each New Model's Performance on Different Data Sets

Eleven colour spaces were tested using the CII-ZHU,[1] GUAN,[5] OSA,[11] BADB-T,[12] Pointer,[13] LCD, SCD and Munsell data sets. These were: the six new spaces together with DIN99d (the only earlier UCS developed to fit the combined SCD data sets), CIELAB (the most widely used UCS), GLAB (the UCS developed to fit some LCD data sets), IPT (the simple space developed for colour imaging) and CAM97s2 (based upon CIE colour appearance model). The results are summarised in Table 1 in terms of PF/3 values. The spaces which performed the best for each data set are underlined and printed with bold lettering, while the worst spaces printed in italic and bold.

The results showed that all newly developed spaces fit well to the data sets which were used to develop them. The spaces can be divided into three categories according to the goodness of fit to the three data groups: LCD, SCD and Munsell. The spaces which fitted best to the LCD data sets are LCDLAB, GLAB, and LCDIPT. For Munsell group, MUNLAB and MUNIPT gave equal and best performance followed by IPT. None of the spaces gave a good fit to the SCD combined data set except for DIN99d. Although SCDLAB and SCDIPT were intended to fit this data set, their structures are not adequate. This suggests that there is large fundamental difference between results for small as opposed to large colour differences.

A psychophysical experiment was then followed to further evaluate the performances of the various spaces.

Experimental

An experiment was conducted to compare lightness, chroma and light series scales between the 8 out of the 11 spaces discussed above, in terms of space uniformity and hue constancy. The visual results of the three equivalent scales obtained by the authors, which were included in the CII-Zhu data set, were also investigated and designated as VISUAL. Hence, nine spaces in total were studied here: three new spaces (LCDLAB, LCDIPT, MUNIPT) and CIELAB, GLAB, DIN99d, CAM97s2, IPT and VISUAL. Only three of the new spaces were tested here since they gave a better performance than the other new spaces. The experiment was conducted in a darkened room using CRT colours. A total of 10 observers between 25 and 35 years old took part in the experiment. Each observer did same experiment twice. All of them had normal colour vision. They were familiar with the colour perceptual attributes such as lightness, chroma and hue and had previous experience in performing psychophysical experiments.

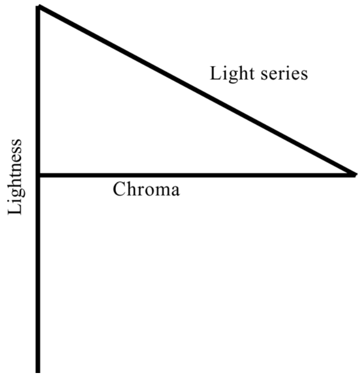

Figure 1. The lightness, chroma and light series scale studied.

Table 1. Summary of Colour Spaces' Performance Using PF/3 Measure (Optimised k_L Spaces)

Model\Data	CII-ZHU	GUAN	OSA	BADB-T	Pointer	Average	SCD	Munsell
DIN99d	***32***	***25***	***26***	24	***39***	***29***	**35**	31
GLAB	26	16	**19**	23	34	24	44	23
CIELAB	28	19	24	30	34	26	52	18
LCDLAB	24	**15**	**19**	24	34	**23**	44	31
SCDLAB	30	19	22	24	36	26	42	***38***
MUNLAB	28	*26*	25	32	34	***29***	***54***	**17**
IPT	27	19	24	32	**33**	27	53	19
LCDIPT	**23**	16	**19**	24	**33**	**23**	44	21
SCDIPT	26	18	20	**22**	35	24	38	28
MUNIPT	25	17	21	***29***	**33**	***25***	50	**17**
CAM97s2	29	21	23	26	34	27	40	27

The three attributes studied are illustrated in Figure 1. There were six chroma scales and three light series scales together with a lightness scale for each space. The six chroma scales had the CIELAB hue angles of 42°, 139°, 305°, 102°, 197° and 330° corresponding to the first and secondary primaries of the CRT used (red, green, blue, yellow, cyan and magenta). The three light series scales only had hue angles at the red, green and blue primaries. In total, 110 colour scales were compared (10 scales for each model).

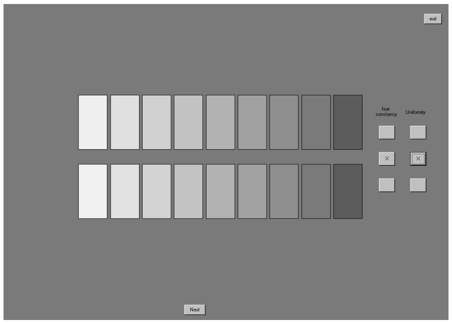

Figure 2. Experimental conditions

A paired comparison method was used. Figure 2 illustrates the experiment situation. A randomly selected pair of scales was displayed on the screen for observers to judge which scale had a better hue constancy first and which appeared to be more uniform. The uniformity was defined as equal distances in a colour space corresponding to equal perceived colour differences. For a good scale, all the pairs of neighbouring colours should have equal colour difference. Hue constancy was defined as all the colours in one scale should have the same hue as the most colourful colour displayed at the end of scale (see Figure 2). Each patch occupied an area of 37 by 20 mm^2 with a gap of 3 mm between patches. The two scales compared were displayed against a grey background with a L* of 50 and chromaticity close to illuminant D65 and CIE 1931 colorimetric observer conditions.

The uniformity of CRT was investigated using a white, a grey and six first and secondary primaries. Each colour was displayed to fill whole screen and nine regions (3 by 3) were divided and measured by a Minolta CS1000 telespectroradiometer. The results showed a satisfactory uniformity with a maximum of 3.8 and a mean of 2.2 ΔE^*_{ab} units. These differences are small compared to the colour differences from one end of a scale to the others, typically 70 to 90 units (see Table 2). The Gain-Offset-Gammar (GOG) model developed by Berns *et al.*[14,15] was used to convert a CRT's RGB signals into CIE tristimulus values. A typical model accuracy was 0.6 ΔE^*_{ab} units from a test set including 34 colours.

The anchor colours at both ends of each scale are given in Table 2 as CIELAB L* C* and hue angle. For establishing a non-CIELAB scale, the two anchor colours were first converted from L*, C and h values to the corresponding values for an appropriate colour space, say the *J, C* and *h* values of CAM97s2. The intermediate values of *J* and *C* were then linearly interpolated for the 7 steps in between two ends of the scale. The inverse form of CAM97s2 was then used to transform those *J, C* and *h* values to tristimulus values. Finally, they were converted to RGB values via the GOG model for displaying on the CRT. There were 9 lightness scales and 54 chroma and 27 light series scales investigated. The experimental conditions are summarised in Table 3. For each scale, there were 9 models which resulted in 36 pairs (9 x 8/2) for pair comparison. In total, 7200 estimations were made.

Table 2. Anchor Colours for Each Scale

			Start			End		
Scales	Scales	Symbol	L*	C*	h	L*	C*	h
Lightness	Neutral	L	5	0	-	95	0	-
	Red	CR	50	0	-	50	90	42
	Green	CG	50	0	-	50	70	139
	Blue	CB	50	0	-	50	90	305
Chroma	Yellow	CY	50	0	-	50	50	102
	Magenta	CM	50	0	-	50	90	330
	Cyan	CC	50	0	-	50	30	197
	Red	SLR	95	0	-	50	90	42
Light Series	Green	SLG	95	0	-	50	70	139
	Blue	SLB	95	0	-	50	90	305

Table 3. Experimental Condition

Scale	Parameter	No. of Scales	No. of Pairs	Repetition	No. of Observer	Total Judgement
Lightness	L*(5-95),C*=0	1	36	2	10	720
Chroma	L*=50, C*(0-max.)	6	216	2	10	4320
Light Series	L*(50-95),C*(0-max)	3	108	2	10	2160

Observer Precision and Accuracy

A measure of wrong decision (WD%) developed by McLaren[16] was used to indicate the observer repeatability and observer accuracy. For investigating observer repeatability, wrong decision represents the number of disagreements between the observer's first judgement and second judgement. For observer accuracy, wrong judgement represents the number of disagreements between an individual observer's judgement and the overall panel judgement. The WD% for observer repeatability in this study is 19% for the hue constancy study and 40% for the uniformity study. The WD% for observer accuracy is 19% for hue constancy and 38% for the uniformity study. The WD% is twice higher for the uniformity results than for the hue constancy results. This is caused by the fact that large perceived differences in hue were observed, while the space uniformity was similar for the different spaces. Hence, observers found it is relatively easy to judge hue constancy.

Testing Colour Spaces' Uniformity and Hue Constancy

The results for space uniformity were also analysed based upon Thurstone's law of comparative judgement[17] and reported in terms of z-score as shown in Table 4, together with ranking values. The ranking was determined not only by z-sore but also its 95% confidence interval. The bold and underlined, and bold italic figures indicate the best and worst performed models respectively. The higher the z-score, the better the model performed.

Comparing the results for the lightness scale, the VISUAL and DIN99d spaces performed the best, followed by CIELAB, GLAB and LCDLAB, with the IPT series spaces and CAM97s2 in particular, the worst. Comparing all the light series scales, MUNIPT and IPT performed the best and DIN99d, the worst. Comparing the chroma scales, LCDIPT, MUNIPT and VISUAL spaces performed the best and again DIN99d, the worst. Overall, comparing the uniformity of all scales, the VISUAL results, and all spaces based upon IPT gave the best performance, followed by LCDLAB, GLAB, CIELAB, then CAM97s2 and DIN99d, the worst. The IPT series spaces outperformed the other models in uniformity.

The results for hue constancy were also analysed and reported in terms of z-score as shown in Table 5 together with ranking values. Comparing the performance of light series scales, the IPT based scales and the average visual results performed the best. DIN99d performed the worst. The others are in between without significant difference between them. For the performance of all scales, IPT series scales again performed the best and DIN99d performed the worst. The others gave intermediate results without significant difference. The results clearly showed that there is a lack of hue constancy in the blue direction for CIELAB space4. This problem occurs for all CIELAB related spaces: GLAB, LCDLAB. Overall, the spaces based on IPT are the most hue constant spaces. They also provide the most uniform steps and are simple to use. Hence, they should be confidently used in practical applications.

Table 4. The Performance of Uniform Colour Spaces in Uniformity Test

Uniformity	**ALL**		**All light Series**		**All Chroma**		**All Lightness**	
Colour Model	**Z-score**	**Ranking**	**Z-score**	**Ranking**	**Z-score**	**Ranking**	**Z-score**	**Ranking**
CIELAB	-0.053	2	0.156	2	-0.178	2	0.270	2
GLAB	-0.026	2	-0.213	3	0.031	2	0.253	2
LCDLAB	0.030	2	-0.045	3	0.061	2	0.185	2
DIN99d	***-0.328***	4	***-1.061***	4	***-0.237***	3	0.857	1
CAM97s2	-0.118	3	0.116	2	-0.061	2	***-1.996***	4
IPT	0.027	2	0.334	1	-0.024	2	-0.468	3
LCDIPT	0.142	1	0.170	2	**0.197**	1	-0.070	3
MUNIPT	**0.166**	1	**0.436**	1	0.114	1	-0.111	3
VISUAL	0.160	1	0.107	2	0.096	1	**1.079**	1

Table 5. The Performance of Uniform Colour Spaces in Hue Constancy Test

Hue constancy	**ALL**		**All light Series**		**All Chroma**	
Colour Model	**Z-score**	**Ranking**	**Z-score**	**Ranking**	**Z-score**	**Ranking**
CIELAB	-0.037	2	-0.157	3	-0.044	2
GLAB	-0.098	2	-0.288	3	-0.084	2
LCDLAB	-0.079	2	-0.247	3	-0.069	2
DIN99d	***-0.19***	3	***-0.624***	3	-0.104	2
CAM97s2	-0.084	2	0.064	2	-0.071	2
IPT	0.123	1	**0.371**	1	0.099	1
LCDIPT	0.104	1	0.286	1	0.071	1
MUNIPT	**0.130**	1	0.306	1	0.096	1
VISUAL	**0.131**	1	0.289	1	**0.107**	1

Conclusion

In our earlier study, it was found that the colour discrimination data can be divided into three categories: large difference, small difference and the Munsell data. Three colour spaces based upon CIELAB and IPT were developed to fit each category. Three of these representative spaces together with some popularly used spaces were evaluated by performing psychophysical experiments. Ten scales (one lightness, six chroma and three light series) of nine spaces were studied in terms of the space uniformity and hue constancy. Pair comparison method was used to compare all 90 combinations of the scales. The results showed that the IPT based spaces performed well and outperformed the other spaces. They should be confidently used in practical applications.

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*, Rochester, USA 626-629, 2001.
2. M. R. Luo, S. Y. Zhu and G. H. Cui, Testing the uniformity of colour spaces, *CIE Expert Symposium*, Scottsdale, Arizona, 7-10 November, 2001.
3. 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.
4. 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*, 8-13, 1998.
5. S. S. Guan and M. R. Luo, A Colour-Difference Formula for Assessing Large Colour Differences, *Col. Res. Appl.,* **24**, 344-355, 1999.
6. S. Y. Zhu, M. R. Luo, G. H. Cui and B. Rigg, Comparing different colour discrimination data sets, ibid.
7. Commission Internationale de l'Éclairage: Colorimetry, *CIE* 15.2-1986.
8. G. H. Cui, M. R. Luo, B. Rigg, G. Roesler and K. Witt, Uniform colour spaces based on the DIN99 colour difference formula, *Color Res. Appl.*. **25**, 282-290, 2002.
9. C.J. Li, M.R. Luo and R.W.G. Hunt, A revision of the CIECAM97s model, *Color Res. Appl.*. **25** 260-266, 2000.
10. M. R. Luo, G. Cui and B. Rigg, The development of the CIE 2000 colour difference formula, *Color Res. Appl.* **26** 340-350, 2001.
11. D. L. MacAdam, Uniform Color Scales, *J. Opt. Soc. Am.*, **55**, 1619-1702,1974.
12. S. Badu, Large Colour Differences between Surface Colours, Ph.D. thesis, University of Bradford, 1986.
13. M. R. Pointer and G. G. Attridge, Some Aspects of the Visual Scaling of Large Colour Differences, *Col. Res. Appl.,* **22**, 298-307, 1997.
14. R. S. Berns, J. J. Motta and M. E. Grozynski, CRT Colorimetry. Part I: Theory and Practice, *Col. Res. Appl.,* **18**, 299-314, 1993.
15. R. S. Berns, J. J. Motta and M. E. Grozynski, CRT Colorimetry. Part II: Metrology, *Col. Res. Appl.,* **18**, 315 325, 1993.
16. K. McLaren, The precision of textiloe colour matches in relation to colour difference measurements, *Proc. 1st AIC Congr.*, Color 69, Muster-Schmidt, Gottingen, 688-707, 1970.
17. L. L. Thurstone, A law of comparative judgement. *Psycholog. Rev.*, **34**, 273-286, 1927.

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.