The LLAB Model for Colour Appearance and Colour Difference Evaluation

M. Ronnier Luo
University of Derby, Design Research Centre
Mackworth Road, Derby DE22 3BL U. K.

Abstract

An ideal system of colorimetry should provide measures agreeing to what we see in three respects: colour specification, difference and appearance. A successful method to quantify these measures depends upon the reliability of psychophysical experimental data. These data sets have been accumulated and were used to derive the LLAB model. The model includes two parts: a chromatic adaptation transform and a uniform colour space. Tristimulus values under a particular set of illuminant/observer conditions are transformed to those of D65/2° conditions via a chromatic adaptation transform. A modified version of CIELAB is then used to calculate six perceived attributes: lightness (L_L), redness-greenness (A_L), yellowness-blueness (B_L), colourfulness (C_L), hue angle (h_L) and hue composition (H_L). The model gives similar degree of prediction in comparison with the other state of the art models using the accumulated data sets. The LLAB model demonstrates that it is possible to achieve a system, which provides precise measures to quantify colour match, difference and appearance.

Key Words: Colour Appearance, Colour Difference, Colour Appearance Model, Uniform Colour Space, Colour Difference Formula, Chromatic Adaptation Transform.

1. Introduction

To achieve successful cross-media colour reproduction requires technologies such as device calibration and characterisation, colour appearance modelling and colour gamut mapping. All these apply colorimetry.

The development of colorimetry can be divided into three phases: colour matching, colour difference and colour appearance. The initial stage intends to quantify whether two colours match each other and the second stage aims to precisely evaluate colour difference between a pair of colours. A number of measures were developed to suit these purposes such as CIELAB,[1] CIELUV,[1] CMC,[2] CIE94[3] and BFD.[4,5] However, these measures are limited to being used under fixed viewing conditions, i.e. the two stimuli in question should be presented using identical backgrounds and surrounds, and be viewed using identical illuminants and observers defined by the CIE. Obviously, they cannot effectively achieve crossmedia colour reproduction, in which the media involved include vastly different viewing conditions such as white point chromaticity, luminance level, and surround. This leads to the third stage of the development of colorimetry: to derive models quantifying colour appearance such as Hunt,[6] Nayatani[7] and RLAB.[8] These colour appearance models are capable of predicting changes of colour appearance by taking into account the above mentioned viewing parameters.

Various CIE technical committees have been formed and intend to encourage researchers and industrialists to gather as much as possible experimental data in order to evaluate various colour models and to provide new standards in these areas. However, this often leads to some controversies such as the competition between the CMC (ISO 105-JO3) and CIE94 formulae. It is likely that similar situations will occur in the colour appearance area. Ideally, a comprehensive model capable of quantifying colour matching, colour difference and colour appearance would provide the optimum solution for these problems.

This article describes a new colour model, named LLAB. The model was derived to fit a number of experimental data sets in the fields of colour appearance[9–15] and colour difference.[4,16,17] The results indicate that the LLAB model gave equal or better performance than the other state of the art models. This strongly implies that it is possible to adopt a universal model for all possible applications.

2. The LLAB Model

Three goals were initially established to develop a colour model. Firstly, the model should give a precise prediction of colour appearance between a pair of single stimuli or complex images under different viewing conditions. Hence, the model can be used not only for cross-media colour reproduction but also for colour rendering, metamerism and colour constancy applications. Secondly, the model should provide a uniform colour space for colour gamut mapping and colour difference evaluation. Finally, it should include simple mathematical equations and it should be easy to derive its reverse model. This is necessary to minimise the computer power and time required for processing a complex image which normally includes a couple of hundred thousand pixels.

With this in mind, the LLAB model was derived to fit the LUTCHI colour appearance data[9–15] and various colour difference data sets.[4,16,17] The form of the model and part of modelling techniques follow those of the Hunt[6] and RLAB[8] models.

The LLAB model can be divided into two parts: the BFD[18] chromatic adaptation transform developed at the University of Bradford by Lam and Rigg, and a modified

CIELAB uniform colour space. The former is used to transform corresponding colours from a source illuminant to the reference illuminant: D65, and the latter calculates perceived attributes similar to those of CIELAB. These attributes vary under different luminances, surrounds, and achromatic backgrounds. This uniform colour space is only valid under the reference illuminant. The mathematical expression of the LLAB is given below:

Data Preparation

X_s Y_s Z_s	Tristimulus values of test sample under source adapting field
X_{os} Y_{os} Z_{os}	Tristimulus values of reference white under source adapting field
X_{or} Y_{or} Z_{or}	Tristimulus values of reference white under reference adapting field, which is fixed for the CIE D65 illuminant and the 1931 standard colorimetric observer (i.e. 95.05, 100.0, 108.88 for X_{or}, Y_{or} and Z_{or}, respectively).
L	Luminance (in cd/m^2) of the reference white under source adapting field
Y_b	Y value of achromatic background under source adapting field

Three parameters are required according to different viewing conditions used, i.e. F_S (Surround induction factor), F_L (Lightness induction factor), F_C (Colourfulness induction factor). These values correspond to each set of viewing conditions and are given in Table I.

Table I. The F_S, F_L and F_C parameters used in LLAB model.

Surface samples and images in average surround	F_S	F_L	F_C
Subtending 10°	3.0	0.0	1.00
Subtending 2°	3.0	1.0	1.00
CRT displays in dim surround	3.5	1.0	1.15
Transparency in dark surround	4.2	1.0	0.95

Stage 1. Compute the tristirnulus values (X_r Y_r Z_r) under the reference adapting field via the BFD chromatic adaptation transform.

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = M \begin{bmatrix} X/Y \\ Y/Y \\ Z/Y \end{bmatrix} \tag{1}$$

where

$$M = \begin{bmatrix} 0.8951 & 0.2664 & -0.1614 \\ -0.7502 & 1.7135 & 0.0367 \\ 0.0389 & -0.0685 & 1.0296 \end{bmatrix}$$

The RGB cone responses for the reference whites under reference and source adapting fields, and test colour under source adapting field are calculated using eq. (1) and are designated as R_{or}, G_{or}, B_{or}; R_{os}, G_{os}, B_{os} and R_s, G_s, B_s, respectively.

$$\begin{aligned} R_r &= R_{or}\, R_s / R_{os} \\ G_r &= G_{or}\, G_s / G_{os} \\ B_r &= B_{or}\, (B_s / B_{os})^{\beta}, \end{aligned} \tag{2}$$

where $\beta = (B_{os} / B_{or})^{0.0834}$

$$\begin{bmatrix} X_r \\ Y_r \\ Z_r \end{bmatrix} = M^{-1} \begin{bmatrix} R_r Y_s \\ G_r Y_s \\ B_r Y_s \end{bmatrix} \tag{3}$$

Stage 2. Calculate the appearance attributes: lightness (L_L), redness-greenness (A_L), yellowness-blueness (B_L), colourfulness (C_L), hue angle (h_L) and hue composition (H_L) under reference illuminant, and L, Y_b, Fs, FL and Fc determined under source adapting field.

$$L_L = 116\, f(Y)^z - 16 \tag{4}$$

$$A = 500\, [f(X) - f(Y)] \tag{5}$$

$$B = 200\, [f(Y) - f(Z)] \tag{6}$$

where $z = 1 + FL\, (Y_b / 100)^{1/2}$ (7)

If X/95.05, Y/100 or Z/108.88 > 0.008856

$$\begin{aligned} f(X) &= (X_r / 95.05)^{1/Fs} \\ f(Y) &= (Y_r / 100.0)^{1/Fs} \\ f(Z) &= (Z_r / 108.88)^{1/Fs} \end{aligned} \tag{8}$$

If X/95.05, Y/100 or Z/108.88 ≤ 0.008856

$$\begin{aligned} f(X) &= [(0.008856^{1/Fs}\ 16/116)/0.008856]\ X_r / 95.05 + 16/116 \\ f(Y) &= [(0.008856^{1/Fs}\ 16/116)/0.008856]\ Y_r 1\, 100.00 + 16/116 \\ f(Z) &= [(0.0088S6^{1/Fs} - 16/116)/0.008856]\ Z_r 1\, 108.88 + 16/116 \end{aligned} \tag{9}$$

$$C_L = [4.907 + 0.162\, C + 10.92 \ln (0.638 + 0.07216\, C)]\, F_C S_C \tag{10}$$

where $C = (A^2 + B^2)^{1/2}$ (11)

and $S_C = 1.0 + 0.47 \log (L) - 0.057\, [\log(L)]^2$ (12)

$$h_L = \tan^{-1} (B/A) \tag{13}$$

$$A_L = C_L \cos (h_L) \tag{14}$$

$$B_L = C_L \sin (h_L) \tag{15}$$

$$H_L = H_{L1} + (H_{L2} - H_{L1})(h_L - h_{L1})/(h_{L2} - h_{L1}) \tag{16}$$

where H_{L1} and H_{L2} are either 0, 50, 100, 150, 200, 250, 300, or 350 according to whether R, Y50R, Y, G50Y, G, B50G, B or R50B, respectively, and are the hue compositions having nearest lower and higher values of h_L. The h_{L1} and h_{L2} values are the hue angles having nearest lower and higher values of h_L respectively. The values of h_L, H_L and NCS expression are given in Table II.

Table II. Values for converting hue angle to hue composition.

h_L	H_L	R	Y	G	B	NCS expression
25	0	100	0	0	0	R
62	50	50	50	0	0	R50Y
93	100	0	100	0	0	Y
118	150	0	50	50	0	Y50G
165	200	0	0	100	0	G
202	250	0	0	50	50	G50B
254	300	0	0	100	0	B
322	350	50	0	0	50	B50R

Stage 3. Calculate the LLAB colour difference

$$\Delta E_L = [(\Delta L_L / l)^2 + (\Delta C_L / c)^2 + \Delta H_L^2)^{1/2} \qquad (17)$$

where

$$\Delta L_L = L_{L,bat} - L_{L,std}$$
$$\Delta C_L = C_{L,bat} - C_{L,std}$$
$$\Delta H_L = 2 (C_{L,bat} C_{L,std})^{1/2} \sin (\Delta h/2)$$
$$\Delta h_L = h_{L,bat} - h_{L.std}$$

where the LLAB coordinates with subscripts of 'std' and 'bat' represents those for standard and batch samples. The lightness weight, l, equals to 1.0, 1.5 and 0.67 for perceptibility, acceptability and large colour difference pairs respectively. The chroma weight, c, should always equal to one.

3. Features of the LLAB Model

The main features in the LLAB model are summarised below.

The LLAB model adopts the BFD chromatic adaptation transform. The transform only operates the complete chromatic adaptation. This results in a 'white point' remaining 'white' regardless of which illuminant is used. No attempt was made to account for incomplete adaptation due to not enough evidence found in the previous experimental results.[9-15]

The model includes a modified version of the CIELAB space to predict six appearance attributes: lightness (L_L), redness-greenness (A_L), yellowness-blueness (B_L), colourfulness (C_L), hue angle (h_L) and hue composition (H_L). The lightness, colourfulness and hue composition attributes have been consistently used in the previous magnitude estimation experiments[9-14] and have proven to be the most efficient, precise and easy to understand attributes in scaling colour appearance.

It is well known that the luminance of the surround condition is a dominant factor affecting colour appearance, i.e. the image contrast is smaller under the dim and dark surround in comparison with that under average surround. This effect is taken into account by the Fs parameter, i.e. 3.0, 3.5 and 4.2 for the average, dim and dark surrounds respectively.

The LLAB model's colourfulness scale (C_L) is based upon that of CMC. It corrects the problem of non-uniformity occurring in the CIELAB space,[2-5] which predicts much smaller differences around the neutral axis than in the other areas. It also includes a function (Sc) for modelling the effect on increase of colourfulness due to increase of luminance level, known as the Hunt effect. In addition, the colourfulness induction factor (Fc) is introduced allowing for the medium/surround difference.

The LLAB model's lightness scale includes a z function for predicting the change of lightness due to different Y values of the background (lightness contrast effect). The F_L factor was used to switch this effect on and off ($F_L = 1$ and 0) for the small and large sized stimuli respectively.

The model's hue angle is identical to that of the CIELAB and its hue composition is based upon the work of Derefeldt and Sahlin.[19] They established a look-up-table between the NCS's hue triangles and CIELAB's hue angles.

The LLAB model includes the lightness and colourfuluess weighting values for calculating colour difference. The lightness (l) and colourfulness (c) weights are only applied for calculating colour difference, not colour appearance attributes.

4. The LLAB Model's Performance

The LLAB model's performance is compared with the state of the art models and formulae using various colour appearance and colour difference experimental data sets.

4.1 Cross-Media Image Reproduction

An experiment[15] was carried out to assess the colour reproduction quality of CRT (softcopy) images processed via different colour models against a hardcopy (original) image illuminated in a viewing cabinet. Six images were used and each was assessed by a panel of nine observers. The experiment was divided into seven phases according to whether the same or different chromaticities for the hardcopy's and softcopy's reference white were used. The results described, using a seven-point image quality scale, were analysed using paired comparison and category judgement methods. Four Phases (A to D) which had different chromaticities for the hardcopy's and softcopy's reference white are used here. These phases were chosen to reflect the colour appearance models' performance in predicting the chromatic adaptation effect, the binocular memory matching[20] technique was used. The paired comparison results (z scores) for the Hunt, Nayatani, RLAB and LLAB models are summarised in Fig. 1. The z score is linearly correlated with the visual response and is proportional to the quality of the image, i.e. the higher the z score, the better the performance of the model. For each model, a vertical line was also drawn to indicate the range of ±2 units standard deviation (95% confidence limit). If the z score of one model is within another model's 95% confidence limit, the two models are considered not to be signifiicantly different.

Figure 1 shows that the Nayatani and Hunt models gave the best prediction when the colour temperatures between two viewing fields were close (Phases A and B), but much poorer performance for Phases C and D with large colour temperature difference between the two fields. The LLAB model predicted the best in Phases C and D, especially for Phase D (matching A to D65). It had a similar performance as the RLAB model in Phases A and B, and gave an average performance in the other seven phases. The Hunt and RLAB models gave the average level of prediction for all Phases. In Phase D, all models except LLAB model predicted the results badly.

4.2 The LUTCHI Colour Appearance Data Sets

A set of experimental data was accumulated over a period of nine years at Loughborough University of Technology Computer Human Interface (LUTCHI) Research Centre. The detailed description can be found in references 9 to 14. A brief summary is given below.

Many colours were assessed by a panel of six to seven observers using the magnitude estimation method. Each colour was described in terms of lightness, colourfulness and hue. The experiment was conducted under fifty sets of viewing conditions such as variations of light source, luminance level, background, sample size and medium/surround. This data set includes over 100,000 visual estimates. The data is divided into three subsets according to the three media used: surface, CRT and transparency.

Again, the Hunt, Nayatani, RLAB and LLAB models were tested using three data sets. The results in terms of coefficient of variation (CV) value are summarised in Figs.

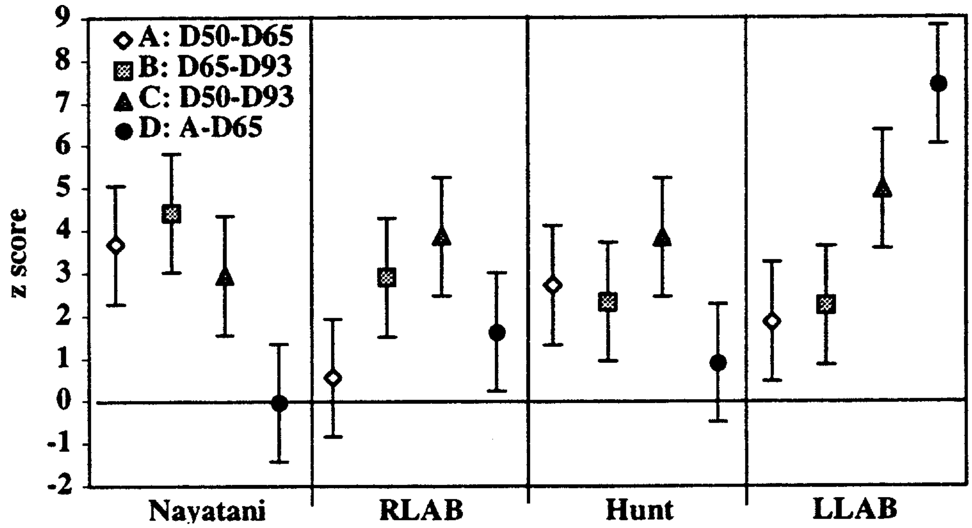

Figure 1. Comparison of the quality of colour reproduction (z score) between Nayatani, RLAB, Hunt and LLAB models.

2a to 2c for lightness, colourfulness and hue results respectively. For a perfect agreement between two sets of data, the CV should equal zero (i.e. 0% error). The observer precision representing the degree to which a single observer agrees with the average of a group of observers is also plotted. The precisions of 14, 17 and 8 CV units were found for the lightness, colourfulness and hue respectively. A reliable model should give error of prediction to the visual results as good as the observer precision.

For lightness comparison (Fig. 2a), the results show that the LLAB model performed the best for surface and CRT media and the Hunt model for transparency. This is caused by that the topmost part of the lightness results in 35 mm projected transparency data set, from 85 to 100 has a sharp transition towards reference white. This results in a separate Hunt lightness scale[12] (Jp) which was developed specially for predicting this effect. However, the LLAB's lightness scale still performed the second best amongst all models tested. Both LLAB and Hunt models gave a satisfactory prediction to the visual results, i.e. the prediction errors are less than the visual precision (CV = 14).

For predicting colourfulness results (Fig. 2b), the Hunt model performed the best and the LLAB ranked the second. The LLAB model predicted worse than Hunt model by about 3 CV units. Attempts were made to improve the fit allowing for the effect of background on the LLAB model's colourfulness predictions. This did produce better performance in predicting the colour appearance data sets, but worse prediction to the colour difference data sets (see later). Both Hunt and LLAB models' predictive errors were slightly larger than that of visual precision (CV = 17). For hue comparison (Fig. 2c), the LLAB and Hunt models gave the similar performance and outperformed the other models with a large margin and their errors of predictions are similar to the CV values of visual precision.

The Hunt colour appearance model has been refined over the years using this data set. However, the results from Section 4.1 show that the model did not perform well using complex images. In addition, another drawback associated with the Hunt model is its complexity. For image processing tasks such as cross-media reproduction, a reverse model is required to obtain the corresponding tristimulus values in order to reproduce colours with same colour appearance from the source device/viewing conditions to those of destination conditions. The reverse Hunt model is quite complicated and many parameters need to be predetermined. Additionally, the reverse model cannot be analytically expressed, so that a numerical approach is the only solution. For a typical complex image including a couple of hundred thousand pixels, the processing power and time required is considerable. This makes the practical application almost impossible. The LLAB model gave a similar or better degree of prediction to the LUTCHI data and is a much simpler than the Hunt model.

4.3 Colour Difference Data Sets

Three data sets were accumulated to test five colour-difference formulae: CIELAB,[1] CMC,[2] CIE94,[3] BFD[5] and LLAB. The latter four formulae are modified versions of the CIELAB. These data sets were accumulated by Luo and Rigg.[4,17] They collected a large number of the previously published colour discrimination data sets and carried out phychophysical experiment, in which pairs of textile samples corresponding to many of colour centres were assessed against a grey scale. The results were used to combine these data sets together to form two combined sets of perceptibility (CP) and acceptability (CA) data including 2776 and 1613 pairs respectively. The average colour differences for both data sets is about 3 CIELAB ΔE units. These were used to derive the BFD colour difference formula.[5] A similar method was used to combine another set of data[17] under illuminant A (CILA) including 1053 pairs of samples with a mean CIELAB ΔE of 3. These three sets were used to test the five formulae and the results are summarised in Fig. 3 in terms of PF/4, in which the Performance Factor (PF) was a measure devised by Luo and Rigg[4] for evaluating colour-difference formula's performance using colour difference data sets. It is a combination of four different statistical measures. Its magnitude is similar to that of CV, i.e. a PF/4 value of 30 indicates a 30% predictive error.

Fig. 3 shows that CIELAB formula performs the worst and BFD formula gave slightly more accurate prediction

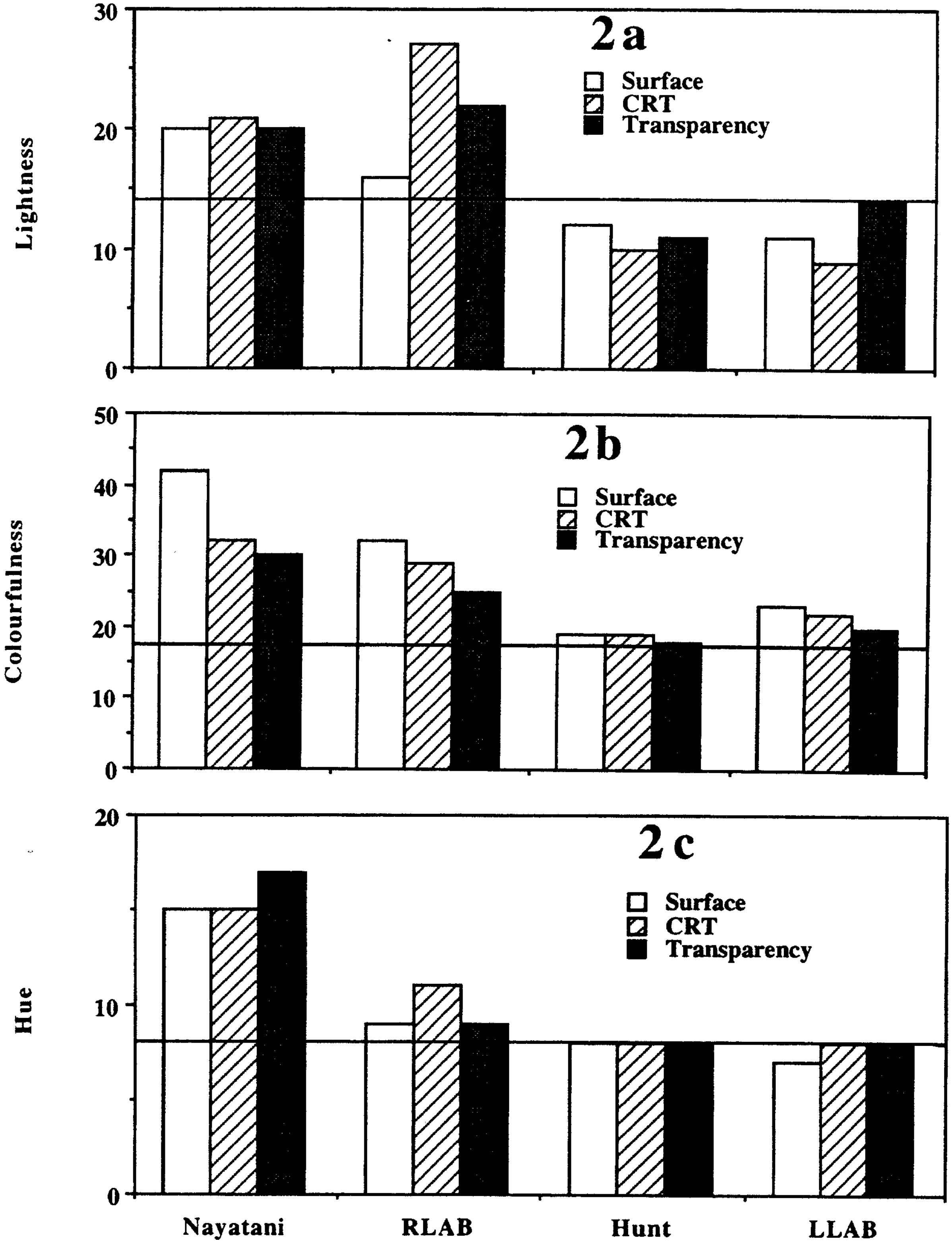

Figure 2. Comparison of the prediction errors (CV) between Nayatain, RLAB, Hunt and LLAB models; (a) for lightness, (b) for colourfulness and (c) for hue.

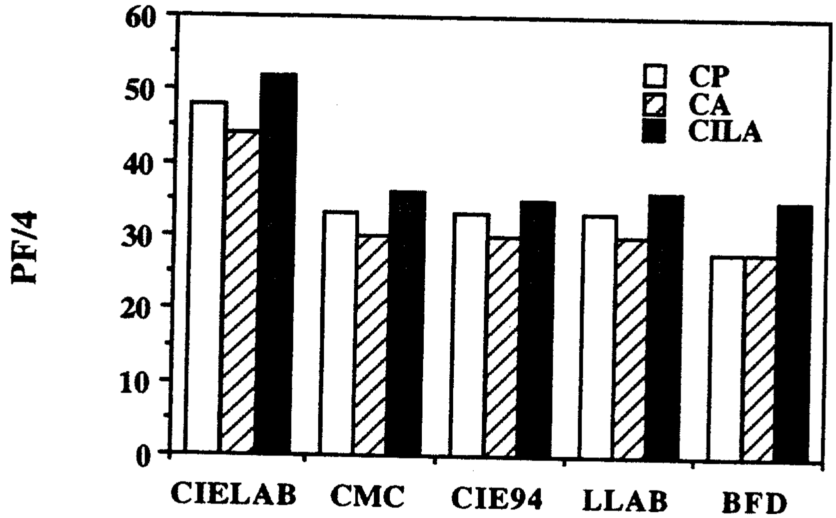

Figure 3. Comparison of the prediction errors (PF/4) between CIELAB, CMC, CIE94, BFD and LLAB colour difference formulae.

than the other formulae as expected. This formula was derived to fit these data sets. The results clearly indicate that there is hardly any difference between CMC, CIE94 and LLAB formulae. All formulae were derived to fit data sets under the daylight sources so that they gave a worse fit to the CILA data than the other two sets under daylight. For evaluating the perceptibility colour differences, the lightness weight of 1 is used for CMC, BFD and LLAB formulae except for the CIE94 ($l = 2$). The lightness weights of 2.0 and 1.5 are required for the acceptability data for CMC and CIE94, and BFD and LLAB respectively. The CIE94 formula applied $l = 2$ for both perceptibility and acceptability data sets (unlike the others requiring different weights). This indicates that there are some differences between these formulae, at least in calculating the lightness difference. The LLAB(1:1) model gives a satisfactory prediction of the three sets of visual results tested and is at least as good as those of CMC(1:1) and CIE94(2:1) formulae.

5. Conclusion

A new colour model, LLAB, was derived to quantify the degree of colour matching, colour difference and colour appearance. It includes a modified CIELAB uniform colour space for colour specification, gamut mapping and colour difference evaluation purposes. This space is further incorporated with the BFD chromatic transform for applications such as cross-image reproduction, assessment of metamerism and colour constancy, and quantification of colour rendering property.

The model's performance was also compared with the other spaces and models using different data sets. The results show that the LLAB model is capable of precisely quantifying the change of colour appearance under a wide range of viewing parameters such as light sources, surrounds/media, achromatic backgrounds, sizes of stimuli and luminance levels. The model gave a similar or better performance than the Hunt model's lightness and hue predictions, but not for colourfulness predictions. However, the LLAB model still outperformed the other models.

The model was also tested using various colour difference data sets. It gave the similar performance as the state of the art colour difference formulae: CMC, CIE94 and BFD. For evaluating small to medium colour differences, LLAB(1:1) is preferable for perceptibility data, while for acceptability data LLAB (1.5:1) gives the best results.

6. Reference

1. CIE, Colorimetry (2nd Edition), CIE Publication No. 15.2 (1986).
2. F.J.J. Clarke. R. McDonald, and B. Rigg, Modification to the JPC79 colour-difference formula, *J. Soc. Dyers Col.* **100** pp. 128-132 (1984).
3. D. H. Alman, CIE TC 1-29 Industrial color-difference evaluation progress report, *Color Res. Appl.* **18,** pp. 137-139 (1993).
4. M.R. Luo and B. Rigg, BFD(l:c) colour-difference formula, Part I—Development of the formula, *J. Soc. Dyers. Col.*, **103** pp. 86-94 (1987).
5. M.R. Luo and B. Rigg, BFD(l:c) colour-difference formula, Part II—Performance of the formula, *J. Soc. Dyers. Col.*, **103** pp. 126-132 (1987).
6. R. W. G. Hunt, An improved predictor of colourfulness in a model of colour vision, *Color Res Appl.* **19**, pp. 23-33 (1994).
7. Y. Nayatani, Revision of chroma and hue scales of a nonlinear color-appearance model, *Color Res. Appl.*, **20**, pp. 143-155 (1995).
8. M. D. Fairchild, Visual evaluation and evolution of the RLAB color space, *The Second IS&T/SID Color Imaging Conference: Color Science, Systems and Applications*, pp. 9-13 (1994); (see page 85, this publication).
9. M. R. Luo, A. A. Clarke, P. A. Rhodes, A. Schappo, S. A. R. Scrivener, and C. J. Tait, Quantifying colour appearance. Part I. LUTCHI colour appearance data, *Color Res. Appl.* **16** pp. 166-180 (1991).
10. M. R. Luo, A. A. Clarke, P. A. Rhodes, A. Schappo, S. A. R. Scrivener, and C. J. Tait, Quantifying colour appearance. Part II. Testing colour models performance using LUTCHI colour appearance data, *Color Res. Appl.* **16** pp. 181-197 (1991).
11. M. R. Luo, X. W. Gao, P. A. Rhodes, H. J. Xin, A. A. Clarke,

S. A. R. Scrivener, Quantifying colour appearance. Part III. Supplementary LUTCHI colour appearance data, *Color Res. Appl*. **18** pp. 98-113 (1993).
12. M. R. Luo, X. W. Gao, P. A. Rhodes, H. J. Xin, A. A. Clarke, S. A. R. Scrivener, Quantifying colour appearance. Part IV. Transmissive media, *Color Res. Appl*. **18** pp. 191-209 (1993).
13. M. R. Luo, X. W. Gao, S. A. R. Scrivener, Quantifying colour appearance. Part V. Simultaneous contrast, *Color Res. Appl*. **20** pp. 18-28 (1995).
14. W. G. Kuo, M. R. Luo and H. E. Bez, Various chromatic-adaptation transformations tested using new colour appearance data in textiles, *Color Res. Appl*. **20** pp. 313-327 (1995).
15. M. C. Lo, M. R. Luo and P. A. Rhodes, Evaluating colour models' performance between monitor and print Images, *Color Res. Appl*. **21** pp. 000 000 (1996).
16. M.R. Luo and B. Rigg, Chromaticity-discrimination ellipses for surface colours, *Color Res. Appl*., **11** pp. 2542 (1986).
17. M. R. Luo and B. Rigg, A colour-difference formula for surface colours under illuminant A, *J. Soc. Dyers. Col.*, **103** pp. 161-167 (1987).
18. K. M. Lam, *Ph.D. thesis*, Metamerism and colour constancy, University of Bradford (1985).
19. G. Derefeldt and C. Sahlin, Transformation of NCS data into CIELAB colour space, *Color Res. Appl*. **11** pp. 146-152 (1986).
20. K. Braun, and M. D. Fairchild, Viewing environments for cross-media image comparison, *IS&T's 47th Annual Conference/ ICPS*, pp. 391-396 (1994); (see page 8, this publication).

published previously in SPIE, Vol. 2658, page 261