Prediction of Chinese color system appearance scales using various color appearance models

Yusheng Lian (廉玉生), Xiuze Wang (王秀泽), and Wei Meng (孟 伟)

Institute of Science, Information Engineering University, Zhengzhou 450001

Received January 4, 2007

The chromaticities of the Chinese color system dataset are applied to eight color appearance models (CAMs). Models used are: CIELAB, Hunt, Nayatani, RLAB, LLAB, ZLAB, CIECAM97s, CIECAM02. Three color appearance attributes (lightness, chroma, and hue) are discussed for their uniformity, in terms of the constant perceptual nature of the Chinese color system dataset. The results show that no particular model can excel at all metrics. Comparison can lead to the conclusion that Chinese color system appearance scales can be predicted only slightly poorer than Munsell appearance scales using the eight CAMs.

OCIS codes: 330.1690, 330.1720, 330.1730.

Color appearance model (CAM) is an important tool which solves the problem of the color fidelity display or communication under complex illumination and viewing conditions and between cross-media[1]. Hence, there is a strong need by color imaging engineers to integrate a CAM with color management systems. Several CAMs have been developed and refined in recent years, each derived with a different approach and stressing the various aspects of perception to a greater or less degree. So, these models, most widely known and used, should be tested using some available data groups. And Chinese color system data can be used to test these models because of its uniformity. This study compares several modern CAMs with respect to their abilities to predict uniformly appearance scales of Chinese color system data, hereafter, compares the results of Chinese color system with Munsell color system. Models used were: CIELAB, Hunt, Nayatani, RLAB, LLAB, ZLAB, CIECAM97s, CIECAM02[2].

Chinese color system has a structure similar to Munsell system. It is built along the three perceptual quantities: hue, chroma (describing saturation), and value (describing lightness). In the album, adjacent color samples represent equal intervals of visual perception. It consists of 1338 samples arranged on 40 pages of constant hue[3,4]. The number of samples in Chinese color system is about a half of that in Munsell system[5].

Input data of these models are the chromaticity coordinates of the Chinese color system data and the model-specific parameters for viewing conditions. The concrete parameters used to calculate model coordinates for Chinese color system data sets are: CIELAB: $X_n = 95.02$, $Y_n = 100$, $Z_n = 108.81$; Hunt: $X_w = 95.02$, $Y_w = 100$, $Z_w = 108.81$, $L_A = 63.6619$, $T = 6504$, $N_c = 1$, $N_b = 75$, $Y_b = 20$; Nayatani: $X_n = 95.02$, $Y_n = 100$, $Z_n = 108.81$, $Y_o = 20$, $E_o = 5000$, $E_{or} = 1000$; RLAB: $X_n = 95.02$, $Y_n = 100$, $Z_n = 108.81$, $\sigma = 1/2.3$, $D = 1$; LLAB: $X_n = 95.02$, $Y_n = 100$, $Z_n = 108.81$, $F_S = 3$, $F_L = 1$, $F_C = 1$, $Y_b = 20$, $D = 1$; ZLAB: $X_n = 95.02$, $Y_n = 100$, $Z_n = 108.81$, exps $= 0.345$, $L = 63.6619$, $F = 1$; CIECAM97s: $X_w = 95.02$, $Y_w = 100$, $Z_w = 108.81$, $L_A = 63.6619$, $C = 0.69$, $N_c = 1$, $F = 1$, $Y_b = 20$, $F_{LL} = 1$, $D = 1$; CIECAM02: $X_w = 95.02$, $Y_w = 100$, $Z_w = 108.81$, $L_A = 63.6619$, $C = 0.69$, $N_c = 1$, $F = 1$, $Y_b = 20$. Parameters were chosen to consistently and appropriately represent the viewing conditions recommended for Chinese samples: daylight (Illuminate D_{65}) and average surrounding[6].

All these models predicted the perceptual color attributes of lightness, chroma, and hue. These three color appearance attributes are divided into three dimensions. The discussion focuses on each of these separately. These three dimensions in any model cannot be appropriately combined. This is because that, in the original scaling experiments of Chinese samples, observers adjust each dimension of color separately.

The performance of models' lightness linearity is illustrated in Fig. 1, in which the model lightness is plotted against Chinese value[7]. A good lightness scale should be linear with Chinese value. A linear fit is done on each set of model lightness data. For comparison, correlation coefficients for regression lines are shown in Table 1. From it one can observe that the level of correlation is high as shown by the correlation coefficient R^2 of 0.9944 or higher. Moreover, the differences of correlation coefficient among these models are very small. It can be confirmed by linearity of each tie-line in Fig. 1. The larger the value of correlation coefficient, the better

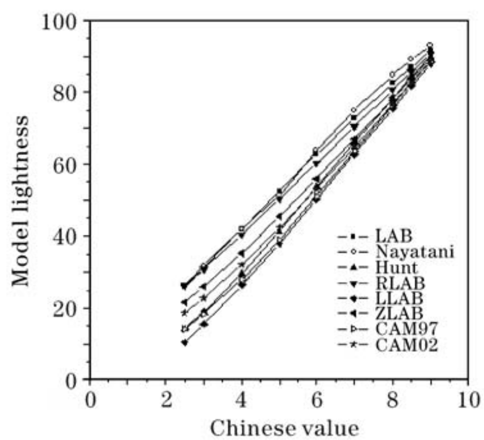

Fig. 1. Lightness linearity of the eight CAMs.

1671-7694/2007/100617-04 © 2007 Chinese Optics Letters

Table 1. Results from Regression of Model Lightness versus Chinese Value

Model	R^2 of Linear Fit
CIELAB	0.9994
Nayatani	0.9955
Hunt	0.9972
RLAB	0.9998
LLAB	0.9994
ZLAB	0.9985
CAM97	0.9944
CAM02	0.9976

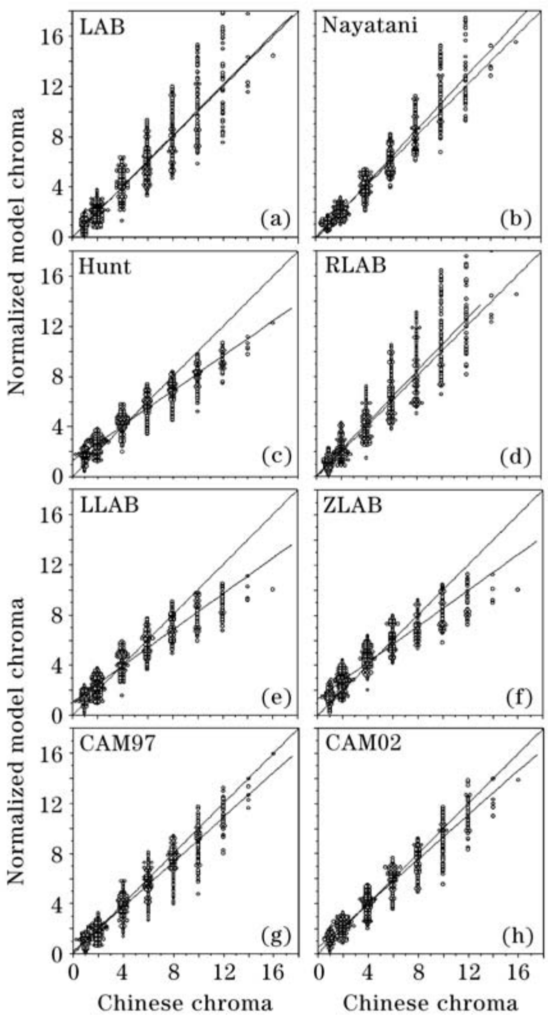

Fig. 2. Model chroma performance.

the lightness linearity. So, lightness linearity of RLAB is better than the others. R^2 values for all models are distributed at approximate 1.0. In a word, all models perform quite well in lightness linearity.

Figure 2 shows the chroma performance for all models. The data points are color coded to approximate the input Chinese color sample. In the figure, the normalized model chroma is plotted against Chinese chroma. All model chroma need to be adjusted to the same scale. It is calculated using

$$C_{\mathrm{model,norm}} = (C_{\mathrm{C}} C_{\mathrm{model}})/C_{\mathrm{model,ave}}, \quad (1)$$

where C_{C} (Chinese chroma) is always 6, C_{model} is model chroma, and $C_{\mathrm{model,ave}}$ is the average model chroma of the 40 colors at chroma = 6, value = 5.

A linear regression was done to these normalized chroma data, so that we can know whether the interception is significantly different from zero, and whether the values lie inside the 95% confidence intervals for the slope. The results of regression analysis were listed in Table 2. The interceptions on the vertical axes of the best-fit lines in Fig. 2 are a measure of how well the models predict the chroma of near neutral colors; ideally they should always be zero. The p-value is used to make the judgment as to the significance of the coefficients. It can be considered as the maximum choice of α for which the null hypothesis can be rejected. Since the goal is to make the interception equal 0, we desire the p-value to be high.

The chroma performance for all models can be judged by the scattering of the points. For the perfect model, the points should be coincident with the 45° lines. It can be calculated by the CV values given by[8,9]

$$\mathrm{CV} = \frac{\sqrt{\frac{1}{n}\sum (V_i - fP_i)^2}}{\overline{V}} \times 100 \quad (2)$$

with

$$f = \frac{\sum V_i P_i}{\sum P_i^2}, \quad (3)$$

where V is the chroma for Chinese sample i, P is its corresponding model chroma, and n is the number of samples used. The lower the CV value is, the better the performance will be. The CV values are summarized in Table 2. From Table 2, we can see that CIECAM02 has a CV value of 19.680 outperforming the other models, followed by CIECAM97s. The CIECAM97s and Nayatani models give similar performance within a CV value of 1.201. The worst models are CIELAB and RLAB. But note that their interceptions are the smallest among these models. This implies that there is not a model which has the satisfying values of both interception and CV.

According to the results listed in Table 2, for the interception, all models gave dissimilar performance, in which CIELAB and RLAB performed the best, followed by CIECAM02 and Nayatani. For these four models, interception values are very small. So, these four models can predict the chroma of near neutral colors more accurately. What's more, the same four models also give better performance of slope except CIECAM02. They had slopes of unity within their 95% confidence intervals. Both of these are the goal for this linear regression. From the statistical values of slope shown in Table 2, we can see that the other five models show some chroma compression. According to results of regression analysis and CV values, CIECAM02 gave reasonable performance compared with other models in predicting the chroma of Chinese color system.

The Chinese color system was designed to be perceptually uniform in terms of hue, which means that a set of straight lines radiating from the center representing constant hue and the angle between two adjacent hue

Table 2. Results from Linear Regression of Normalized Model Chroma versus Chinese Chroma

Model	Interception		Slope				CV
	Value	p-Value	Value	p-Value	Lower 95%	Upper 95%	Value
CIELAB	0.024	0.732	1.010	0	0.985	1.035	29.476
Nayatani	−0.192	0	1.079	0	1.059	1.098	22.852
Hunt	1.302	0.659	0.695	0	0.682	0.708	24.649
RLAB	0.033	0	1.045	0	1.019	1.072	30.237
LLAB	1.086	0	0.715	0	0.700	0.731	25.816
ZLAB	1.346	0	0.719	0	0.702	0.736	27.815
CAM97	0.198	0	0.892	0	0.876	0.909	21.651
CAM02	0.177	0	0.884	0	0.870	0.906	19.680

Table 3. RMS Errors of Ten Hues for Various Models

	5B	5BG	5G	5GY	5Y	5YR	5R	5RP	5P	5PB
LAB	1.157	0.268	0.344	3.386	2.292	6.893	2.648	1.457	0.938	1.986
Nayatani	1.616	0.297	1.388	0.606	10.222	15.158	2.685	1.084	1.237	2.726
Hunt	0.94	0.265	0.383	0.917	2.381	4.483	0.897	1.978	0.959	1.316
RLAB	0.967	0.26	0.356	2.95	2.768	5.586	1.94	1.461	0.885	2.085
LLAB	1.272	0.29	0.466	3.878	3.206	5.928	2.305	2.069	1.162	2.683
ZLAB	1.355	0.638	0.223	4.254	3.823	6.341	2.492	1.573	1.79	2.844
CAM97	0.88	0.42	0.738	1.163	1.98	4.693	0.877	1.002	1.236	1.284
CAM02	1.062	0.653	1.407	1.661	3.086	7.435	1.264	0.839	2.123	1.538

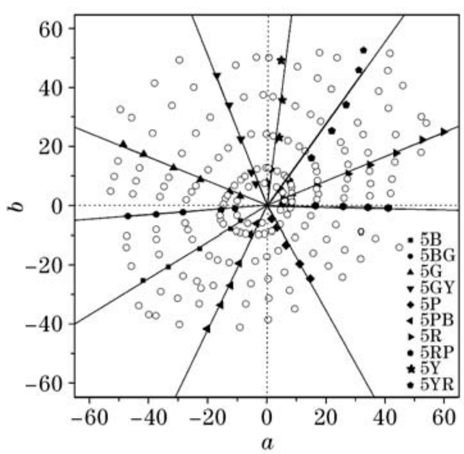

Fig. 3. Coordinates of Chinese color samples in CIECAM02 chromaticity diagram.

lines should be equal to each other in the a-b chromaticity diagram[10,11]. Such a plot for CIECAM02 is illustrated in Fig. 3. Figure 3 shows Chinese color data set with value of 6 in the CIECAM02 chromaticity diagram. Actually, these lines are curved in the chromaticity diagram, particularly for yellow and yellow-red hues. Take these hues in mass, their linearities are well.

In order to evaluate the performance of each constant-hue line, we try to define root mean square (RMS) error to express the linearity for each hue line. RMS errors are given as

$$\mathrm{RMS} = \sqrt{\sum \frac{(Y_i - BX_i)^2}{(1+B^2)N}} \bigg/ \bar{C}_6, \quad (4)$$

where Y and X are coordinates in chromaticity diagram based on various models for Chinese sample i, N is the number of samples used, and B is the inclination for regression lines. The smaller the value of RMS error is, the better the performance of prediction will be. For both 5 principal hues and 5 intermediate hues, the results of $V = 6$ in terms of RMS are given in Table 3. Clearly, for ten hues, 5Y has the greatest RMS error, followed by 5YR and 5PB. So the linearities of constant-hue lines for 5Y, 5YR, and 5PB are worse than the others. For 5BG, the RMS value is the smallest; it significantly outperforms the other hues by a very large margin. It means that most of these models have good performance of hue lines through the blue-green region and have bad performance in the red-yellow region. Comparing the data in each row of the table, we can conclude that CIECAM97 and Hunt perform the best and second best for the linearity of each line, respectively. Also, CIECAM02 performs very close to both of them and gives a satisfactory prediction to hues.

The hue circle in the Chinese color system is divided into 40 hues and is designed to divide the complete hue circle into equal perceptual intervals. Therefore, good performance means that predicted principal hue lines and intermediate hue lines should be 36° apart[12]. In Table 4, values of the angles between two adjacent hue lines and the standard deviations are the indications to evaluate whether the distribution of ten hue lines is uniform on the chromaticity plane or not. The results show that most of these models give similar performance to the hue spacing, however, Nayatani performs slightly more poorly than the others.

For hue performance of these models, it is easy to draw

Table 4. Values of the Angle Included between Two Adjacent Hue Lines and the Standard Deviation (STD)

	YR-R	Y-YR	GY-Y	G-GY	BG-G	B-BG	PB-B	P-PB	RP-P	R-RP	STD
CAM02	31.986	28.109	28.768	46.453	27.088	26.985	32.899	54.952	59.429	23.330	12.819
CAM97	28.283	29.059	31.220	47.481	26.816	27.780	33.608	54.357	59.115	22.279	12.822
LAB	33.327	25.400	27.111	45.064	25.069	33.323	48.150	45.680	48.291	28.584	9.745
LLAB	32.739	26.163	27.832	44.326	25.098	33.396	47.460	46.139	48.788	28.059	9.602
Nayatani	38.384	20.330	27.565	45.031	28.810	25.414	31.151	61.391	56.845	25.078	14.102
RLAB	32.326	27.236	28.651	44.213	24.812	33.821	47.779	45.512	48.382	27.268	9.431
ZLAB	31.850	24.670	26.895	45.700	26.760	33.297	43.862	44.829	52.652	29.485	9.859
Hunt	27.533	29.541	31.677	47.388	26.713	28.082	34.101	55.474	58.200	21.291	12.922

the conclusion that curvature of hue lines are satisfactory, however, the distribution of hue lines on the chromaticity plane is not ideal. Various models gave a similar performance for predicting hues of Chinese color system.

Prediction of Munsell appearance scales using various color appearance models has been studied by Wyble[7]. Comparing the above analysis for Chinese color system with Wyble's study for Munsell color system, we can find that various models gave excellent performance in predicting both data sets. This is expected because the structure of two systems is similar.

In predicting both Chinese color system and Munsell system, all models perform quite well in lightness linearity and overall chroma performance. The results of lightness linearity and chroma performance for two systems are similar. However, in predicting hue linearity and hue spacing of two systems using various models, the performances are different. Hue linearity of R and Y is better than the other hues for Munsell system, however, hue linearity of Y is the worst among these hues. The distribution of hue lines on the chromaticity plane for Munsell color samples is better than that for Chinese color samples. Using various models, Chinese color system appearance scales can be predicted only slightly worse than Munsell appearance scales.

Another important issue to understand is that the number of Chinese color samples is less than the number of Munsell samples. So, the statistical results for Munsell system are more reliable than that for the Chinese color system. These models were not derived from the Chinese color system data sets. Therefore, it is not surprising that several of these models give slightly worse performance in predicting Chinese color system data sets.

These modern CAMs can predict uniformly the appearance scales of Chinese color system data sets very well. None of these models excell at all metrics. Eight models can predict the appearance scales of two color systems reasonably well, and the hue results for Chinese color system are worse than those for Musell system.

Y. Lian's e-mail address is lianyusheng@126.com.

References

1. B. Chai, N. Liao, and D. Zhao, Chin. Opt. Lett. **2,** 637 (2004).
2. M. D. Fairchild, *Color Appearance Models* (John Wiley and Sons, New York, 2005) p.183.
3. The Chinese Color System (in Chinese) Chinese National Bureau of Technical Supervision (1995).
4. Y. Li and Y. Zhao, China Illuminating Engineering Journal (in Chinese) **6,** (2) 1 (1995).
5. G. Wyszecki and W. S. Stiles, *Color Science* (John Wiley and Sons, New York, 1982) p.507.
6. D. Wang, Q. Jing, X. Sun, Z. Lin, and J. Zhang, Acta Psychologica Sin. (in Chinese) **29,** 225 (1997).
7. D. R. Wyble and M. D. Fairchild, Color Res. Appl. **25,** 132 (2000).
8. S.-S. Guan and M. R. Luo, Color Res. Appl. **24,** 331 (1999).
9. S. Y. Zhu, M. R. Luo, G. Cui, and B. Rigg, in *Proceeding of the 10th Color Imaging Conference* (2002).
10. F. Ebner and M. D. Fairchild, in *Proceeding of the 6th Color Imaging Conference* (1998).
11. H. Huan, K. Xiao, and D. Zhao, Progress in Natural Science **11,** 536 (2001).
12. M. R. Luo, G. Cui, and C. Li, Color Res. Appl. **31,** 320 (2006).