https://doi.org/10.2352/issn.2169-2629.2021.29.381

Testing Colour Appearance Model Based UCS Using HDR, WCG and COMBVD Datasets

Qiang Xu[1], Muhammad Safdar[2] and Ming Ronnier Luo[1, 3]*

[1]State Key Laboratory of Modern Optical Instrumentation, Zhejiang University, Hangzhou, China
[2] Avient Corporation, Avient Colorants Singapore Pte. Ltd., 618684 Singapore, Singapore
[3]School of Design, Leeds University, Leeds, UK

Corresponding Author: m.r.luo@zju.edu.cn

Abstract

Two colour appearance models based UCSs, CAM16-UCS and ZCAM-QMh, were tested using HDR, WCG and COMBVD datasets. As a comparison, two widely used UCSs, CIELAB and IC_TC_P, were tested. Metrics of the STRESS and correlation coefficient between predicted colour differences and visual differences, together with local and global uniformity based on their chromatic discrimination ellipses, were applied to test models' performance. The two UCSs give similar performance. The luminance parametric factor k_L, and power factor γ, were introduced to optimize colour-difference models. Factors k_L and γ of 0.75 and 0.5, gave marked improvement to predict the HDR dataset. Factor k_L of 0.3 gave significant improvement in the test of WCG dataset. In the test of COMBVD dataset, optimization provide very limited improvement.

Introduction

The current trend of imaging devices such as displays, TVs, and cameras, is to include high dynamic range (HDR) and wide colour gamut (WCG) technologies. Human vision system has a luminance range from 10^{-6} to 10^{8} cd/m^2. The traditional standard dynamic range (SDR) display only covers a luminance range from 0.1 to hundreds cd/m^2. For HDR displays, the luminance range can be from 0.001 to several thousand cd/m^2. HDR displays provide larger contrast to show more details in images. DCI-P3 displays, a wider colour gamut than sRGB, have been widely used in cinema projection and mobile displays. New display primaries are being studied to obtain wider colour gamut, and ensure the accuracy of colour reproduction meanwhile. WCG provides more colourful and brighter images.

Colour appearance attributes including lightness, brightness, colourfulness, chroma, saturation, hue angle, are widely used to describe a colour. Many colour appearance models (CAM) have been developed to predict colour appearance under different viewing conditions, *e.g.*, CIECAM97s [1, 2], CIECAM02 [3], CAM16 [4], and ZCAM [5]. More recently, the two-dimensional colour appearance attributes including vividness, blackness, depth, whiteness, and clarity were also integrated with CIECAM02 [6] and ZCAM [5].

In 1997, the Commission Internationale de l'Eclairage (CIE) published CIECAM97s [1, 2]. Since then, it has been tested in imaging industry.

In 2002, CIECAM02 was recommended by the CIE [3] to supersede CIECAM97s to overcome some problems in imaging applications. CAM02-UCS [7] an extension of CIECAM02 to give accurate prediction of colour differences. The COMBined Visual Dataset (COMBVD) was used to derive CAM02-UCS. COMBVD consists of four different datasets, including RIT-DuPont [8], Witt [9], Leeds [10], and BFD [11, 12].

However, CIECAM02 had some shortcomings in the image processing of cross-media colour reproduction. Li *et al.* developed a revision of the CIECAM02 in 2016, named CAM16 [4], which overcomes the previous problem. The CAM16 has as good performance as the CIECAM02 in predicting the visual results, or even better. And its associated UCS, CAM16-UCS [4], was proposed to replace CAM02-UCS for colour difference evaluation. Lightness J, colourfulness M, and hue angle h, computed using the CAM16 model, are adjusted to obtain J', M' and h' to form the polar coordinate space of CAM16-UCS. The chromatic axis a' and b' in the cartesian coordinate space of CAM16-UCS are transformed from M' and h' in the polar coordinate.

CAM16-UCS predict change of colour appearance at different luminance levels through the F_L (luminance level adaptation) factor. Eq. 1 is nonlinear cone response transformation formular to complete luminance adaptation. In Eq. 1, R_a is the post-adaptation cone response, and similarly for the computations of G_a and B_a.

$$\{R_a,\ G_a,\ B_a\} = p_1 \cdot \frac{\left(\frac{F_L \cdot \{R_c, G_c, B_c\}}{100}\right)^{p_2}}{\left(\frac{F_L \cdot \{R_c, G_c, B_c\}}{100}\right)^{p_2} + p_3} + p_4\ , \quad (1)$$

where F_L is the luminance adaptation factor and R_c (G_c, B_c) is cone response. Parameters p_1 to p_4 are 400, 0.42, 27.13 and 0.1, respectively.

The luminance range in CAM16-UCS is not HDR. CAM16-UCS should be tested using experimental data under HDR conditions. In the previous study [13], CAM02-UCS performs very good for evaluating colour differences under different luminance levels from 0.25 to 1128 cd/m^2. CAM16-UCS is expected to have similar performance to CAM02-UCS or even better.

In 2017, Safdar *et al.* [14] proposed a perceptually uniform colour space, $J_za_zb_z$, for HDR and WCG applications. The Perceptual Quantizer (PQ) curve, used to encode a luminance range of 0.001 to 10,000 cd/m^2, is used for HDR conditions. Eq. 2 gives the nonlinear cone response transformation as PQ curve.

$$\{R', G', B'\} = \left(\frac{c_1 + c_2\left(\frac{\{R,G,B\}}{10{,}000}\right)^{\eta}}{1 + c_3\left(\frac{\{R,G,B\}}{10{,}000}\right)^{\eta}}\right)^{\rho}, \quad (2)$$

where R (G, B) is cone response. Parameters $c_1 = 3424/2^{12}$, $c_2 = 2413/2^7$, $c_3 = c_1 + c_2 - 1 = 2392/2^{12}$, $\eta = 2610/2^{14}$, and $\rho = 1.7 \times 2523/2^5$.

In $J_za_zb_z$, lightness axis J_z, and chromatic axis a_z and b_z, form the cartesian coordinate space. Chroma C_z, and hue angle h_z are transformed from a_z and b_z. COMBVD was one of the training datasets to derive $J_za_zb_z$.

A new colour appearance model based on $J_za_zb_z$, named ZCAM [5], was proposed to predict colour appearance attributes including brightness, lightness, colourfulness, chroma, hue angle, hue composition, saturation. In addition, it can predict the two-dimensional attributes, vividness, blackness, and whiteness. ZCAM performs similar to that of CAM16 in predicting colour appearance tested with a range of experimental datasets. ZCAM should give good performance for HDR and WCG applications, due to $J_za_zb_z$ deriving for these applications.

The goals of this study were to investigate how to apply colour appearance models in predicting colour differences under HDR conditions and to understand its performance in predicting COMBVD (SDR) and WCG data.

CIELAB [15] and IC_TC_P [16] were also tested together with the above models. CIELAB has been widely used for industrial colour-difference evaluation. IC_TC_P was specially designed to be used for HDR and WCG applications.

Experimental Data

COMBVD Dataset

The COMBined Visual Dataset (COMBVD), containing 3813 pairs of samples with an average colour difference of 2.6 in CIELAB units, was used to derive CIEDE2000, CAM02-UCS. COMBVD was applied as a training dataset by other UCSs, like CAM16-UCS, $J_za_zb_z$, etc. COMBVD consists of four different datasets, including RIT-DuPont [8], Witt [9], Leeds [10], and BFD [11, 12]. COMBVD samples are surface colours with hairline division between the two colours in a pair.

HDR Dataset

In the earlier study [13], an experiment was conducted covering a very large range of luminance levels in a spectrum tunable viewing cabinet in a darkened room. The surround factor in the calculation of colour appearance model was set as 'dim'. The luminance factor, Y_b, of the grey background in the cabinet is 34. The reference white was set at a correlated colour temperature (CCT) of 6500 *K* under nine luminance levels, *i.e.*, 0.25, 0.6, 1.1, 1.9, 3.3, 32, 111, 406 and 1128 cd/m^2. The adapting luminance, L_a, was calculated using the corresponding luminance multiplied by Y_b, and divided by 100. One hundred and forty printed sample pairs were prepared around seven colour centres, *i.e.*, red, yellow, yellow-green, blue-green, blue, purple, and black. Each centre contained twenty sample pairs, including two colour difference magnitudes (2 and 4 CIELAB units). Each magnitude consisted of 2, 3, and 5 pairs in $\Delta L^*\Delta b^*$, $\Delta L^*\Delta a^*$ and $\Delta a^*\Delta b^*$ planes, respectively. Sample pairs were printed in the colour of colour centres and corresponding samples with no hairline or gap between them.

The HDR dataset were obtained using a six-categories including '1' for 'no difference', '2' for 'just noticeable difference', '3' for 'small difference', '4' for 'acceptable difference', '5' for 'large difference' and '6' for 'extremely large difference' for assessing colour differences. Twenty normal colour vision observers (ten males and ten females) took part in the experiment. They had a mean age of 22 ranged from 18 to 25. Observers were asked to assess the colour difference of sample pairs. The mean category for each pair was calculated to represent the visual data (ΔV).

In total, 1,260 pairs (9 luminance levels ×140 pairs) data were accumulated. This set of data is named HDR dataset.

WCG Dataset

In the earlier study [17, 18], two experiments were carried out in a darkened room on an NEC PA302W display. The display peak white was set at a CCT of 6500 *K* and a luminance of 310 cd/m^2. The Gain-Offset-Gamma (GOG) model [19] was built to characterize the display, and the predictive accuracy of samples was 0.42 ΔE_{00}. The display is a WCG display with colour gamut larger than sRGB, and close to that of DCI-P3.

In Experiment 1 [17], 12 colour centres were selected and each colour centre consisted of 16 sample pairs. Three of 12 colour centres were grey, red and blue, recommended by the CIE and widely investigated for colour difference research. Other centres located close to the boundary of the display colour gamut. The pairs had two levels of colour difference magnitudes, 3 or 6 CIELAB units. For each magnitude of each colour centre, 5, 1, 1, and 1 sample colours distributed in $\Delta a^*\Delta b^*$ plane, $\Delta L^*\Delta a^*$ plane, $\Delta L^*\Delta b^*$ plane, and on ΔL^* axis, respectively.

In Experiment 2 [18], 16 colour centres were selected and each colour centre consisted of 14 sample pairs. Colour centres in Experiment 2 were selected to fill in the gap between the most saturated colour regions in Experiment 1 and less saturated colours in COMBVD. All sample pairs had only one colour difference magnitude, 3 CIELAB units. For each colour centre, 11, 1, 1, and 1 sample colours distributed in $\Delta a^*\Delta b^*$ plane, $\Delta L^*\Delta a^*$ plane, $\Delta L^*\Delta b^*$ plane, and on ΔL^* axis, respectively.

Sample pairs had colours of colour centres and corresponding samples with no hairline or gap between them. The colour difference of the sample pair displayed in the centre of the screen was assessed against the grey scale pairs shown at the top of the display. The grey-scale method has been widely used for assessing colour differences [20]. The grey scale consisted of five grey-scale pairs, *i.e.*, *GS-1* to *GS-5*, with measured colour difference of 0.0, 1.4, 3.0, 6.0, and 11.7, in CIELAB unit, respectively.

The background was set to a neutral grey with a luminance factor, Y_b, of 13.56. The adapting luminance, L_a, was 42. The surround factor in the calculation of colour appearance model was 'dim'. In Experiments 1 and 2, 18 and 20 normal colour vision observers (half males and half females) with a mean age of 23 from 22 to 25 years old took part respectively. Observers assessed the colour difference of sample pairs using a value from 1 to 5, with one decimal place. The grey scale values (*GS*) judged by observers were converted to visual colour difference values (ΔV) through Eq. 3.

$$\Delta V = 0.7999e^{0.5567GS} - 1.2359 \quad (3)$$

In total, 416 pairs (192 + 224) data were accumulated, named WCG dataset.

Metrics to test models' performance

The CAM based UCSs, CAM16-UCS, ZCAM-QMh and ZCAM-JCh were tested using the three datasets, HDR, WCG and COMBVD.

The standard residual sum of square (*STRESS*) metric [21, 22] calculated from Eq. 4 was used to indicate the disagreement between two sets of data compared.

$$STRESS = \left(\frac{\sum_{i=1}^{n}(A_i - FB_i)^2}{\sum_{i=1}^{n}F^2B_i^2}\right)^{1/2} \times 100, \tag{4}$$

with $F = \sum_{i=1}^{n} A_i^2 / \sum_{i=1}^{n} A_i B_i$, where n is the number of sample pairs and F is a scaling factor to adjust A and B data sets on to the same scale. The percent *STRESS* values are always between 0 and 100. Values of *STRESS* near to zero indicate better agreement between two sets of data.

The *STRESS* value between predicted colour differences (ΔE) and visual differences (ΔV) was calculated to indicate the performance of colour difference models. The correlation coefficient (r) was also reported.

Another testing method is to compare the global and local uniformity based on chromatic discrimination ellipses. A colour-difference ellipse is given by Eq. 5.

$$\Delta E^2 = g_{11}\Delta a^2 + g_{12}\Delta a\Delta b + g_{22}\Delta b^2, \tag{5}$$

where coefficients g_{11} to g_{22} are optimized to achieve the lowest *STRESS* between the calculated colour difference using the ellipse model and the visual difference, a and b are cartesian coordinates in colour spaces. In Eq. 5 ΔL is ignored due to lack of sample pairs with lightness difference from most colour centres in the above datasets.

The local uniformity [23] is measured by calculating root mean square error (*RMSE*) between the ratios of semi axes (A/B) and that of a circle (A/B = 1), in Eq. (6). The global uniformity [23] is measured by calculating the coefficient of variation (CV) between the size (S) of each ellipse and the average of all ellipses ($\bar{S}$), in Eq. (7).

$$Local = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{A_i}{B_i} - 1\right)^2} \times 100\% \tag{6}$$

$$Global = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(S_i - \bar{S})^2}}{\bar{S}} \times 100\% \tag{7}$$

The local uniformity represents the shape of chromatic discrimination ellipses. Small local uniformity value corresponds to ellipses close to circles, indicating that sample colours around the colour centre have equally perceived colour difference. The global uniformity represents the size of ellipses. Small global uniformity value corresponds to ellipses having similar sizes, indicating that different colour regions have uniform proportion.

Result and Discussion

Selecting a UCS for ZCAM

ZCAM-JCh applies lightness and chroma, which are relative colour appearance attributes. ZCAM-QMh applies brightness and colourfulness, which are absolute colour appearance attributes. Absolute colour appearance attributes would vary according to the luminance level of light source. They should respond to the change of colour difference under different levels in HDR applications.

The *STRESS* and correlation coefficient r between the predicted colour differences (ΔE) calculated using ZCAM-JCh and ZCAM-QMh, and the visual differences (ΔV), were computed using HDR, WCG and COMBVD datasets. Table 1 lists the testing results in terms of *STRESS* and r (in brackets) values. Fig. 1 shows the scatter plots of ΔV against ΔE for the three datasets. From Table 1 and Fig. 1, ZCAM-JCh and ZCAM-QMh performs almost the same for WCG and COMBVD datasets. For testing using HDR dataset using *STRESS* measure, ZCAM-JCh gave the same performance as ZCAM-QMh, but a much smaller r value for ZCAM-JCh. This can be shown clearly by comparing Fig. 1(a) and 1(d), for which the latter is marked less scatter than the former, for ZCAM-JCh and ZCAM-QMh respectively. Therefore, ZCAM-QMh was selected as UCS for ZCAM.

Table 1: *STRESS* and *r* (in brackets) values between *ΔE*, predicted by ZCAM-JCh and ZCAM-QMh, and *ΔV* of HDR, WCG and COMBVD datasets.

	HDR	WCG	COMBVD
ZCAM-JCh	35 (0.72)	38 (0.89)	39 (0.79)
ZCAM-QMh	35 (0.81)	38 (0.89)	39 (0.79)

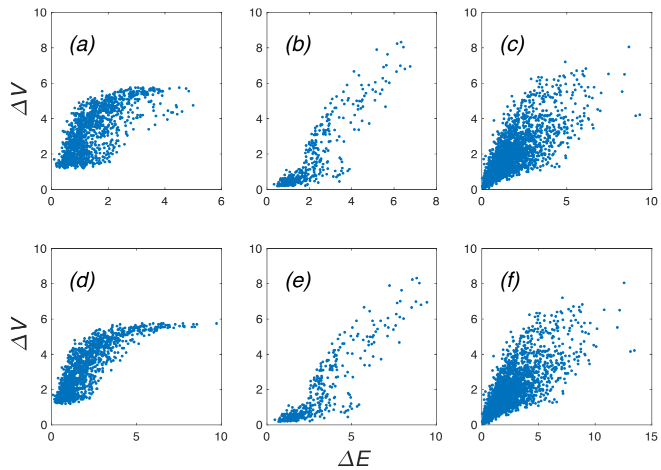

Figure 1. Scatter plots of ΔV against ΔE: (a) HDR, (b) WCG, (c) COMBVD for ZCAM-JCh; and (d) HDR, (e) WCG, (f) COMBVD for ZCAM-QMh

STRESS Testing and Factors Optimization

The *STRESS* values between the predicted colour differences (ΔE) calculated using CIELAB, IC_TC_P, CAM16-UCS and ZCAM-QMh, and the visual differences (ΔV) from the HDR, WCG and COMBVD datasets.

For different datasets, with or without gap between the two colours on a sample pair, using surface or display colours, the luminance parametric factor, k_L, was employed to adjust lightness difference. For HDR dataset without gap using surface colours, the optimized k_L of different colour difference models has an average value about 0.75, ranged from 0.44 (CIELAB) to 0.99 (IC_TC_P) with a standard deviation of 0.22, for CIELAB, CIEDE2000, CAM02-UCS, IC_TC_P and $J_zA_zB_z$ in [13]. For WCG dataset without gap using display colours, the optimized k_L of different colour difference models has an average value about 0.3 [18]. COMBVD dataset with gap using surface colours, as a training dataset, was applied to derive many colour difference euqtions and uniform colour spaces [4, 5, 7, 24]. So, k_L was set as 1 for COMBVD dataset.

Factor k_L was fixed at 0.75, 0.3 and 1 for HDR, WCG and COMBVD respectively. Then, the power factor γ, were introduced to optimize colour difference model, Eq. 8.

$$\Delta E = \sqrt{\left(\frac{\Delta L}{k_L}\right)^2 + \Delta C^2 + \Delta H^2}^{\,\gamma}, \tag{8}$$

where $\Delta C = 2\sqrt{C_1 C_2}\sin\left(\frac{h_2 - h_1}{2}\right)$. Parameter γ was optimized to minimize the *STRESS* between ΔE and ΔV.

Table 2 lists the optimization results in terms of *STRESS* values. From Table 2, parameter k_L affected HDR dataset slightly, but improves the performance of predicting WCG dataset markedly. Parameter γ gives significant improvement to HDR dataset, but little improvement to COMBVD and WCG datasets. For HDR dataset, the optimized γ values are close to 0.5, values given in the brackets. Therefore, factor γ was fixed at 0.5 for HDR dataset. For WCG and COMBVD datasets, factor γ was set at 1.

Table 2: The original and optimized STRESS values between ΔE and ΔV of HDR, WCG and COMBVD datasets.

		CIELAB	ICTCP	CAM16-UCS	ZCAM-QMh
HDR	origin	43	47	33	35
	k_L	39	48	32	38
	k_L, γ	34 (0.44)	23 (0.38)	24 (0.56)	<u>19</u> (0.47)
WCG	origin	65	49	46	38
	k_L	35	28	23	28
	k_L, γ	34	28	23	27
COMBVD	origin	43	44	31	39
	γ	37	38	29	35

CIELAB is a colour space and can only apply to an environment defined by a fixed reference white (with input of relative *XYZ* tristimulus values) and cannot predict appearance at different luminance levels. It was modified to consider HDR viewing conditions. The first kind of input used normalized *XYZ* values for each individual luminance level. This is normally how is used. The second one used the absolute *XYZ* values with the reference white at luminance level of 1128 cd/m^2. They are named CIELAB1 and CIELAB2, respectively.

The above UCSs were tested using the fixed parameters. For HDR dataset, k_L and γ were 0.75 and 0.5, respectively. For WCG dataset, k_L and γ were 0.3 and 1, respectively. For COMBVD dataset, both k_L and γ were set to 1. Table 3 lists the testing results in terms of *STRESS* and *r* (in bracket) units. The *STRESS* and *r* values for the best performed original model and the best k_L, γ model were marked with underlines, and bold and underlined respectively. Note that for COMBVD dataset, the original and k_L, γ models were set to one. Fig. 2 shows the scatter plots in Table 3. For each dataset, the model performed the best is bold and underlined.

From Table 3, CAM16-UCS and ZCAM-QMh performed the best for the original and k_L, γ models, from the test of all the three datasets. This implies great advantages to use CAM based UCSs in predicting colour difference.

The results showed that CAM16-UCS performed the best for the original model from the test of COMBVD dataset, and the best for the k_L, γ model from the test of WCG and the mean datasets. Also, ZCAM-QMh performed the best for the original model from the test of WCG and the mean datasets, and the best for the k_L, γ model from the test of HDR dataset. ZCAM-QMh outperformed CAM16-UCS for the original model from the test of HDR and the mean datasets, in terms of *r* values. But in terms of *STRESS* values, CAM16-UCS outperforms ZCAM-QMh for the original model from the test of HDR dataset. The difference between *STRESS* and *r* tests indicate that, for the original colour-difference model of CAM16-UCS in the test of HDR dataset, a lower *STRESS* value indicates better agreement between quantities while a corresponding lower correlation value indicates that the quantities are more scattered than ZCAM-QMh. Overall, for the k_L, γ model, CAM16-UCS performed the best, followed by ZCAM-QMh, IC_TC_P, CIELAB1, and CIELAB2 the worst.

Table 3: Testing results in terms of *STRESS* and *r* (in brackets) values between ΔE and ΔV of HDR, WCG and COMBVD datasets, together with the mean values. Factors k_L = 0.75, γ = 0.5 for HDR; k_L = 0.3, γ = 1 for WCG; k_L = 1, γ = 1 for COMBVD, respectively.

		HDR	WCG	COMBVD	Mean
CIELAB1	origin	43 (0.28)	65 (0.48)	43 (0.76)	50 (0.50)
	k_L, γ	34 (0.45)	35 (0.89)		37 (0.70)
CIELAB2	origin	74 (0.46)	65 (0.48)	43 (0.76)	61 (0.56)
	k_L, γ	60 (0.53)	35 (0.89)		46 (0.73)
ICTCP	origin	47 (0.70)	49 (0.75)	44 (0.75)	46 (0.73)
	k_L, γ	25 (0.78)	28 (0.92)		32 (0.82)
CAM16-UCS	origin	<u>33</u> (0.69)	46 (0.81)	<u>31</u> (<u>0.86</u>)	<u>37</u> (0.79)
	k_L, γ	24 (0.77)	<u>**23**</u> (<u>**0.95**</u>)		<u>**26**</u> (<u>**0.86**</u>)
ZCAM-QMh	origin	35 (<u>0.81</u>)	<u>38</u> (<u>0.89</u>)	39 (0.79)	<u>37</u> (<u>0.83</u>)
	k_L, γ	<u>**19**</u> (<u>**0.87**</u>)	28 (0.93)		28 (<u>**0.86**</u>)

The *F*-test in Eq. 9 was conducted to test the differences between UCSs. There is significant difference between the two colour-difference models when $F < F_C$ or $F > 1/F_C$.

$$F = \frac{STRESS_{\Delta E1}^2}{STRESS_{\Delta E2}^2} \qquad (9)$$

In the test using HDR dataset (1260 pairs), F_C was 0.90 with a 95% confidence level. The *F* values between the k_L, γ model of ZCAM-QMh and those of other UCSs are less than 0.63, indicating ZCAM-QMh outperformed the other UCSs significantly.

In the test using WCG dataset (416 pairs), F_C was 0.82 with a 95% confidence level. The *F* values between the k_L, γ model of CAM16-UCS and those of other UCSs are less than 0.67, indicating CAM16-UCS outperformed the other UCSs significantly.

In the test using COMBVD dataset (3813 pairs), F_C was 0.94 with a 95% confidence level. The *F* values between the k_L, γ model of CAM16-UCS and those of other UCSs are less than 0.63, indicating CAM16-UCS outperformed the other UCSs significantly.

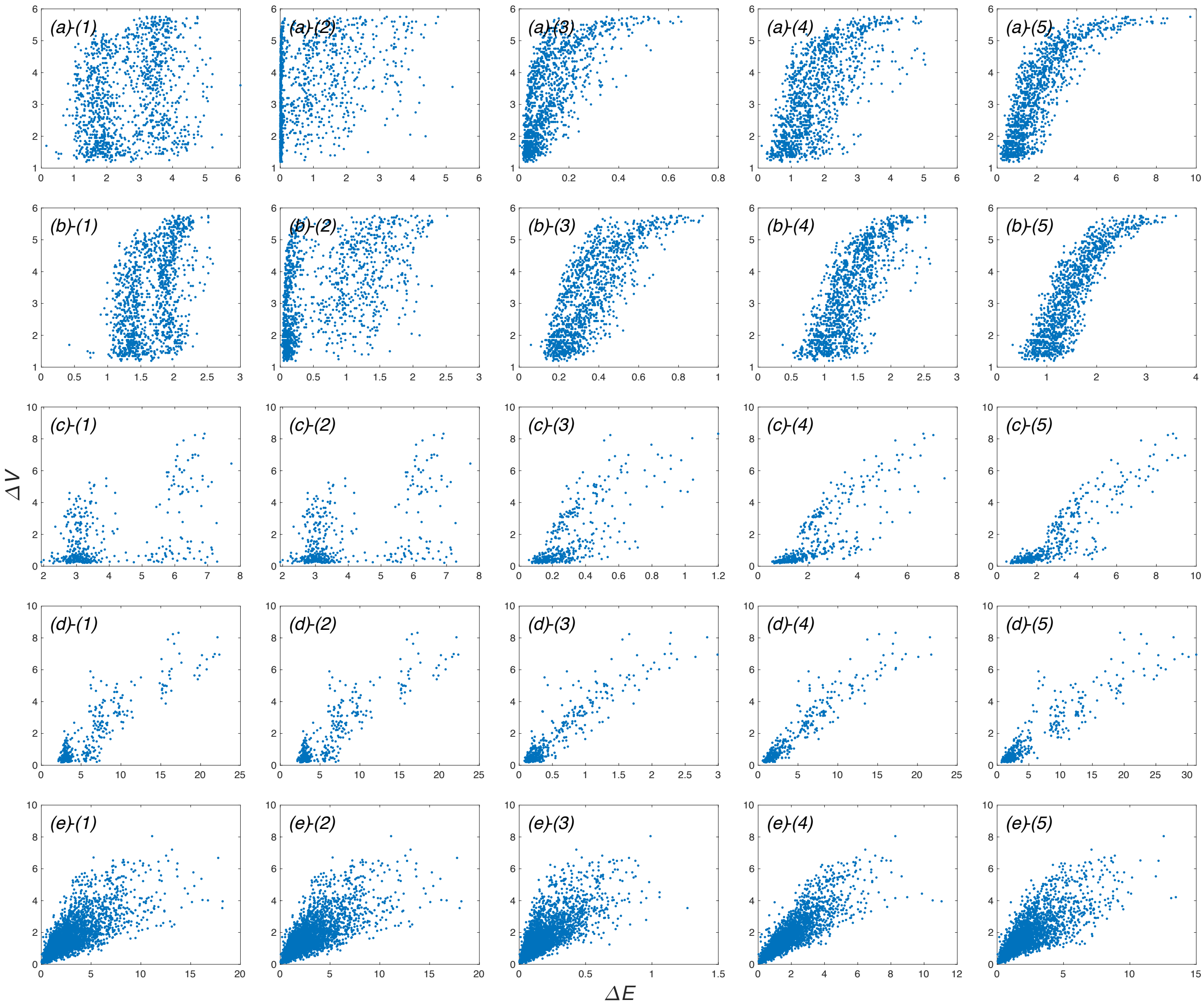

Figure 2. *Scatter plots of ΔV against ΔE: (a) HDR, (b) the k_L, γ models for HDR, (c) WCG, (d) the k_L, γ models for WCG, (e) COMBVD; -(1) CIELAB1, -(2) CIELAB2, -(3) IC_TC_P, -(4) CAM16-UCS, -(5) ZCAM-QMh*

Testing uniformity using Chromatic Discrimination Ellipses

Although parameters k_L and γ improve colour-difference models, they may give little influence to chromatic discrimination ellipses. The ellipses from HDR dataset were fitted using Eq. 5, and then local and global uniformity in Eqs. 6 and 7 were computed.

Fig. 3 shows chromatic discrimination ellipses from HDR dataset in the above UCSs tested. Since ellipses from WCG and COMBVD were presented in the previous work [18], they were not shown here. Table 4 lists local and global uniformity of ellipses from HDR dataset in above UCSs.

Fig. 3 clearly showed a Hunt effect [25] as colour centres become more colourful with an increase of luminance level except CIELAB1. This is caused by the fact that the space is not a colour appearance model, which does not take luminance levels into account. From Fig. 3 and Table 4, ZCAM-QMh has the best local uniformity, indicating that chromatic discrimination ellipses are close to circles (Fig. 3(e)). CIELAB1 has the best global uniformity, and very similar sizes of ellipses under difference luminance levels (Fig. 3(a)), as the reason explained above. Figs 3(b), CIELAB2, and 3(c), IC_TC_P, show a large variation in ellipse size, *i.e.*, smaller ellipses are close to achromatic colour and larger ellipses for high chroma colours. ZCAM-QMh and CAM16-UCS performed well for both local and global uniformity, and ZCAM-QMh slightly better. For CAM16-UCS, the ellipse sizes close to the achromatic colour seem to be a lot larger than those of chromatic colours.

Table 4: Local and global uniformity based on chromatic discrimination ellipses

	CIELAB1	CIELAB2	ICTCP	CAM16-UCS	ZCAM-QMh
Local	71	100	112	87	68
Global	54	174	66	66	63

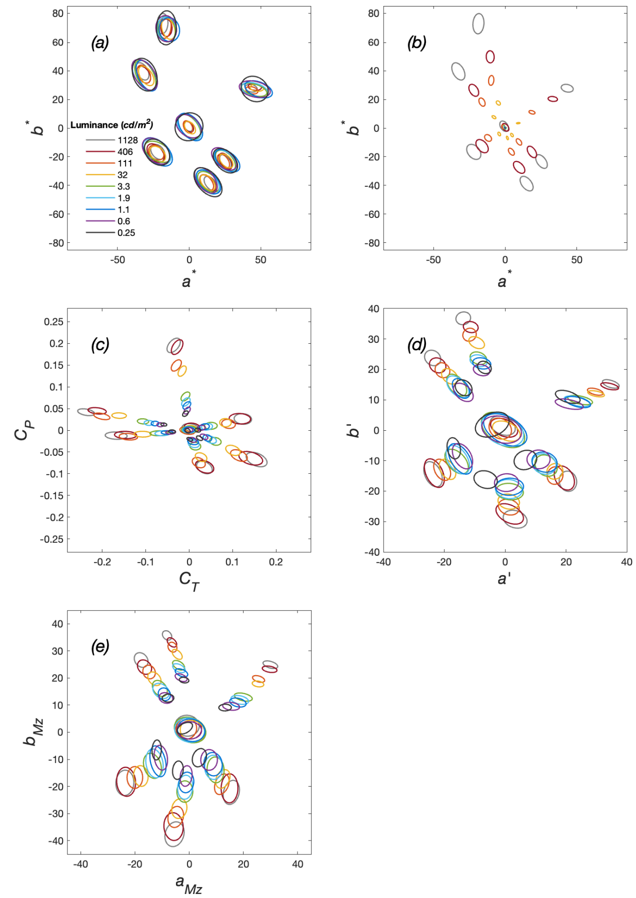

Figure 3. *Chromatic discrimination ellipses from HDR dataset in (a) CIELAB1, (b) CIELAB2, (c) IC_TC_P, (d) CAM16-UCS, (e) ZCAM-QMh*

Conclusion

Two UCSs based on two colour appearance models, CAM16-UCS and ZCAM-QMh, were tested using three datasets, HDR, WCG and COMBVD. As a comparison, CIELAB and IC_TC_P were tested. Two metrics, *i.e.*, *STRESS*, local and global uniformity, were calculated to indicate the performance of predicting colour difference. The two CAM based UCSs outperformed the other UCSs using the three datasets. The luminance parametric factor, k_L, and power factor, γ, were introduced to improve the prediction of colour difference. The power factor improved HDR dataset efficiently. The lightness parametric factor gave significant improvement to WCG dataset. COMBVD dataset was improved limitedly by all parameters. For HDR dataset without gap using surface colours, k_L was 0.75. For WCG dataset without gap using display colours, k_L was 0.3. For COMBVD dataset with gap using surface colours, k_L was 1. The γ was set to 0.5 for HDR dataset. The γ was set to 0.5 for both WCG and COMBVD datasets. The testing results indicated that CAM16-UCS and ZCAM-QMh performed the best considering all the three datasets. ZCAM-QMh significantly outperformed the other UCSs for HDR dataset. CAM16-UCS significantly outperformed the other UCSs for WCG and COMBVD datasets. In the test of chromatic discrimination ellipses, ZCAM-QMh and CAM16-UCS performed well for both local and global uniformity, and ZCAM-QMh gave slightly better performance.

References

[1] M. R. R. Luo, "The structure of the CIE 1997 Colour Appearance Model (CIECAM97s)," Color Res. Appl. 23, 138-146 (1998).
[2] P. J. Alessi, M. Fairchild, K. Hashimoto, R. Hunt, M. R. Luo, L. Mori, Y. Nayatani, T. Seim, H. Sobagaki, and K. Richter, "The CIE 1997 interim colour appearance model (simple version), CIECAM97s," in Wiley Subscription Services, Inc., A Wiley Company(1998).
[3] "A colour appearance model for colour management systems: CIECAM02," CIE 159: 2004.
[4] C. Li, Z. Li, Z. Wang, Y. Xu, M. R. Luo, G. Cui, M. Melgosa, M. H. Brill, and M. Pointer, "Comprehensive color solutions: CAM16, CAT16, and CAM16-UCS," Color Res. Appl. 42, 703-718 (2017).
[5] M. Safdar, J. Y. Hardeberg, and M. Ronnier Luo, "ZCAM, a colour appearance model based on a high dynamic range uniform colour space," Opt. Express 29, 6036-6052 (2021).
[6] Y. J. Cho, L.-C. Ou, G. Cui, and R. Luo, "New colour appearance scales for describing saturation, vividness, blackness, and whiteness," Color Res. Appl. 42, 552-563 (2017).
[7] M. R. Luo, G. Cui, and C. Li, "Uniform colour spaces based on CIECAM02 colour appearance model," Color Res. Appl. 31, 320-330 (2006).
[8] R. S. Berns, D. H. Alman, L. Reniff, G. D. Snyder, and M. R. Balonon‐Rosen, "Visual determination of suprathreshold color‐difference tolerances using probit analysis," Color Res. Appl. 16, 297-316 (1991).
[9] K. Witt, "Geometric relations between scales of small colour differences," Color Res. Appl. 24, 78-92 (1999).
[10] D. H. Kim, and J. H. Nobbs, "New weighting functions for the weighted CIELAB colour difference formula," in Proceedings of the AIC(1997), pp. 446-449.
[11] M. R. Luo, and B. Rigg, "BFD (l:c) colour-difference formula Part 1-Development of the formula," J. Soc. Dyers Colour. 103, 86-94 (1987).
[12] M. R. Luo, and B. Rigg, "BFD (l:c) colour-difference formula Part 2-Performance of the formula," J. Soc. Dyers Colour. 103, 126-132 (1987).
[13] Q. Xu, G. Cui, M. Safdar, L. Xu, and B. Zhao, "Assessing Colour Differences under a Wide Range of Luminance Levels Using Surface and Display Colours," Color and Imaging Conference 2019, 355-359 (2019).
[14] M. Safdar, G. Cui, Y. J. Kim, and M. R. Luo, "Perceptually uniform color space for image signals including high dynamic range and wide gamut," Opt. Express 25, 15131-15151 (2017).
[15] "Colourimetry," CIE 015: 2018.
[16] "ICtCp White Paper," Dolby.
[17] B. Zhao, Q. Xu, and M. R. Luo, "Color difference evaluation for wide-color-gamut displays," J. Opt. Soc. Am. A 37, 1257-1265 (2020).
[18] Q. Xu, B. Zhao, G. Cui, and M. R. Luo, "Testing uniform colour spaces using colour differences of wide colour gamut," Opt. Express 29 (2021).
[19] R. S. Berns, "Methods for characterizing CRT displays," Displays 16, 173-182 (1996).
[20] "Textiles - Tests for colour fastness - Part A02: Grey scale for assessing change in colour," ISO 105-A02.
[21] "Recommended Method for Evaluating the Performance of Colour-Difference Formulae," CIE 217:2016.
[22] P. A. Garcia, R. Huertas, M. Melgosa, and G. Cui, "Measurement of the relationship between perceived and computed color differences," J. Opt. Soc. Am. A 24, 1823-1829 (2007).
[23] M. Huang, H. Liu, G. Cui, and M. R. Luo, "Testing uniform colour spaces and colour-difference formulae using printed samples," Color Res. Appl. 37, 326-335 (2012).
[24] "Colorimetry -- Part 6: CIEDE2000 colour-difference formula," ISO/CIE 11664-6:2014(E).
[25] HUNT, and W. G. R., "Light and Dark Adaptation and the Perception of Color," J Opt Soc Am 42, 190-199 (1952).