© 2019 Society for Imaging Science and Technology

Assessing Colour Differences under a Wide Range of Luminance Levels Using Surface and Display Colours

Qiang Xu[1], Guihua Cui[2], Muhammad Safdar[3], Lihao Xu[4], Ming Ronnier Luo[1, 5] and Baiyue Zhao[1]*

[1]State Key Laboratory of Modern Optical Instrumentation, Zhejiang University, Hangzhou, China
[2]School of Physics & Electronics Information Engineering, Wenzhou University, Wenzhou, China
[3]The Norwegian Colour and Visual Computing Laboratory, NTNU - Norwegian University of Science and Technology, Gjøvik, Norway
[4]School of Digital Media and Art Design, Hangzhou Dianzi University, Hangzhou, China
[5]School of Design, Leeds University, Leeds, UK

*Corresponding Author: zhaobaiyue@zju.edu.cn

Abstract

Two experimental data sets were accumulated for evaluating colour differences under a wide range of luminance levels using the printed colours in a spectrum tuneable viewing cabinet and self-luminous colours on a display respectively. For the surface mode experiment, pairs of samples were assessed at 9 phases ranged from 0.25 to 1128 cd/m^2. For the luminous mode experiment, it was conducted in 6 phases ranged from 0.25 to 279 cd/m^2. There were 140 and 42 pairs of samples judged by 20 observers using a six-category scales for each phase. The results were used to establish the just noticeable difference (JND) at each luminance level and showed a great agreement between two modes of colours. These were used to test the performance of 5 uniform colour spaces and colour difference equations. Also, the spaces or equations were extended to improve the fit to the present data sets.

Introduction

More recently, in the display industry, high dynamic range (HDR) and wide colour gamut (WCG) displays have become commonplace. For human vision system, the luminance range is from 10-6 to 108 cd/m^2. The luminance of traditional standard dynamic range display is from 0.1 to several hundred cd/m^2. But for HDR display, it can be from 0.001 to several thousand cd/m^2. HDR means a greater range of luminance levels. The technology includes OLED, quantum dot LCD. They can have a peak white around 700-1000 cd/m^2 and a black level of 0.01 to 0.0001 cd/m^2. DCI-P3 (Digital Cinema Initiatives) is a popular wide gamut used in cinema projection and mobile display, larger than sRGB. And Reco. 2020 is a wider gamut used as the standard of ultra-high definition television. WCG means more colourful.

Due to the limitation of the camera sensor, a digital camera can capture the scene of a limited luminance range for each exposure. Details are lost from the over-exposure or under-exposure areas. Therefore, HDR has been developed as a solution. An HDR image is generated by capturing multi-exposure images. For HDR applications, it gives greater luminance contrast, the dark is darker, the white is whiter. The details of the pictures can be expressed more clearly.

For the evaluation of colour reproduction in high dynamic and wide gamut range, the conventional uniform colour spaces and colour-difference equations like CIELAB [1, 2], CIEDE2000 [3] and CIECAM02 [4-7], CAM16 [8-11] cannot be used. The new ones specially derived for HDR/WCG applications such as ICtCp [12] and $J_za_zb_z$ [13] need data to verify their performance. The goals of the research are to provide data to extent the conventional metrics and to test the models' performance in HDR applications.

Experiment

The experiment was divided into two. Experiments 1 and 2 studied the surface and luminous mode colours, respectively.

Experiment 1: Surface Mode

Experiment 1 was conducted in a viewing cabinet placed in a dark room. L^* of the background was 65. A spectrum tunable LED viewing cabinet was used. The light in the cabinet was set to CIE D65 and 1931 standard colorimetric observers. The experiment was divided into nine phases to investigate the colour difference thresholds at different luminance levels from very dark to very bright (0.25, 0.51, 0.9, 1.6, 2.8, 32, 111, 407 and 1128 cd/m^2). Neutral density filters were employed to obtain dark luminance levels lower than 2.8 cd/m^2. Table 1 shows lighting conditions. 140 pairs of printed samples were selected from our previous study [14]. They were distributed to surround seven colour centres. Colour pairs in each colour centre included two colour difference magnitudes (2 and 4 CIELAB units). For each magnitude of each centre, there were 2, 3 and 5 pairs in L^*b^*, L^*a^* and a^*b^* planes respectively. These were printed in the colour of seven centres and corresponding samples with no hair-line or gap between them.

Table 1: The lighting conditions used in the experiment

Luminance (cd/m^2)	CCT	Duv	Ra
0.25	6592	-0.001	92
0.51	6487	0.003	98
0.9	6508	0.001	96
1.6	6490	0.002	96
2.8	6505	0.001	96
32	6548	-0.001	97
111	6508	0.003	92
407	6445	-0.004	95
1128	6505	-0.002	92

A Konica Minolta CS2000A spectroradiometer was employed to measure the tri-stimulus coordinates XYZ of sample pairs at different luminance levels. It can measure a wide luminance range from 0.0005 to 5000 cd/m^2. Luminance accuracy was ±2%. Chromaticity accuracy were x, y: ±0.002 and ±0.0015 for colours larger and smaller than 0.05 cd/m^2.

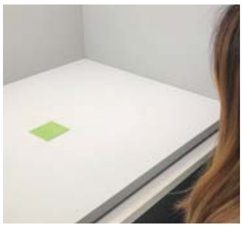
Six categories including '1' for 'no difference', '2' for 'just noticeable difference', '3' for 'small difference', '4' for 'acceptable difference', '5' for 'large difference' and '6' for 'extremely large difference' were employed for visual assessment of colour difference. Twenty normal colour vision observers (ten males and ten females) took part in the experiment. Their ages ranged from 18 to 25 years.

Experiment 1 was divided into nine phases. Each observer came 3 times and each time lasted about 1 hour. Each phase contained one luminance level. In total, 3 hours can finish all the experiments. These nine luminance levels were arranged in a random order for each observer. Figure 1(a) shows experiment setting. Observers sat 60 cm away from the sample pair. The sample pairs had a field of view of 3.5°. The illumination: observation geometry was 0°: 45°. Observers adapted to the viewing conditions for one minute in each phase. In each session, observers viewed the sample pairs following a random order. The mean category for each pair was calculated to represent the visual data (*ΔV*). Twenty sample pairs of grey colour centre were repeated in the formal experiment to test the intra-observer variability. In total, 28,800 observations were accumulated, i.e., (140 + 20) pairs × 9 luminance levels × 20 observers, where the 20 pairs were the repeated stimuli for quantifying intra-observer variation later.

Experiment 2: Display (Luminous Mode)

Experiment 2 was conducted on a 30-inch calibrated NEC PA302W display placed in a dark room with a peak white luminance set at 279 cd/m^2 and CCT at approximate 6500K. The display was characterised using the GOG model. The experiment was divided into six phases to investigate the colour difference thresholds at different luminance levels from very dark to bright (0.25, 2.9, 28, 58, 116 and 279 cd/m^2). Neutral density filters were also employed to obtain dark luminance levels. 42 pairs of samples were selected from the surface data set, again surrounding seven colour centres. Colour pairs in each colour centre included two colour difference magnitudes (2 and 4 CIELAB units). For each centre, there were 1, 1 and 2 pairs (2 CIELAB units) in L^*b^*, L^*a^* and a^*b^* planes respectively, 1 pair (4 CIELAB units) in L^*a^* plane and 1 pair (4 CIELAB units) in a^*b^* plane. These were displayed in the colour of seven centres and corresponding samples with no hair-line or gap between them. The entire screen of the display was covered by a piece black cardboard, for which a square window with 5 centimetres side length was shown in the middle of the screen. Only 42 pairs were used for the luminous mode. This was verified using Experiment 1 results, i.e. the results from the 42 and 140 pairs gave very similar results for each colour centre.

The same 6-point category method was used as the surface mode experiment. Experiment 2 was divided into six phases. Each observer came for only once and finished all the six phases. The whole experiment took about one hour. Each phase contained one luminance level. These six luminance levels arranged in a random order. Each observer sat on a chair and kept their eyes 60 cm in front of the display. The sample pairs had a field of view of 2.5°. The illumination: observation geometry was 0°: 0°. Figure 1(b) shows experiment setting. Each phase started with dark adaptation for one minute. Observers viewed the sample pairs following a random order. After each pair a black image was presented for two seconds to eliminate the after-image effect caused by the visual persistence. The mean category for each pair was calculated to represent the visual data (*ΔV*). Six sample pairs of grey colour centre were repeated in the formal experiment to test the intra-observer variability. In total, 5,760 observations were accumulated, i.e., (42 + 6) pairs × 6 luminance levels × 20 observers, where 6 pairs were repeated.



(a)

(b)

Figure 1. *Experimental situation for a) Experiment 1 and b) Experiment 2.*

Results & Discussion

Inter and Intra Observer Variability

The STRESS value [15] calculated from equation (1) was used to indicate the disagreement between two sets of data compared.

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

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. In colour-difference studies, a STRESS value exceeding 35 is typically an indicator of the poor performance of the colour-difference formula [15].

Table 2: Inter observer variability expressed as STRESS value

Experiment 1 Surface Colour		Experiment 2 Display	
Luminance (cd/m^2)	STRESS	Luminance (cd/m^2)	STRESS
1128	15	---	---
407	15	279	19
111	15	116	19
---	---	58	20
32	15	28	19
2.8	20	2.9	23
1.6	21	---	---
0.90	25	---	---
0.51	23	---	---
0.25	25	0.25	26
mean	19	mean	21

Inter-observer variability was first investigated. Table 2 shows inter-observer variability as STRESS. The STRESS value was calculated between mean of all the observers and each individual observer's results. In Experiment 1, the average STRESS value from all observers represent inter-observer variability. It was found the values to be ranged from 15 (at the brightest level) to 25 (at the darkest level) with a mean of 19. In Experiment 2, inter observer variability ranged from 19 (at the brightest level) to 26 (at the darkest level) with a mean of 21. Mean intra observer variability for Experiments 1 and 2 were 12 and 10, respectively. A clear trend can be found in Table 2 that observers are less consistent for dark luminance levels.

Colour Difference Thresholds

For each sample pair, the results from the two categories ('1' (no difference) and '2' (just noticeable difference)) were judged as 'not perceptible' pair. The number of these pairs divided by the total number of pairs is called not perceivable percentage (NP%). The NP% of 50% was regarded as perception threshold, which represents half of the observers can perceive the colour difference of the sample pair but the other half cannot. The NP% values were plotted against colour differences calculated from one of the five colour models (CIELAB, CIEDE2000, CAM02-UCS, $J_za_zb_z$ and ICtCp). Probability distribution curves were then fitted to the NP% data. Figure 2 gives an example for CAM02-UCS formula to fit Experiment 1 visual data in NP%. Each dot in the figure represents a sample pair. The abscissa is the colour difference, and the ordinate is the probability. The colour difference threshold (ΔE_t) was defined to correspond to 50 NP% at each luminance level. The data having NP% below 5% or above 95% were removed from the calculation due to large experimental noise for very large and small colour differences.

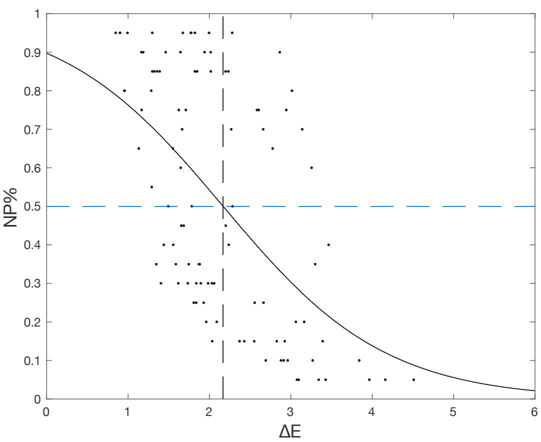


Figure 2. Probability distribution curves. The abscissa and ordinate are ΔE and NP% respectively.

Figure 3 plots ΔE_t values at all luminance levels for the two experiments. It can be found that the trends for CIELAB, CIEDE2000 and CAM02-UCS are quite similar, a decrease of ΔE_t with luminance levels. The curves of ICtCp and $J_za_zb_z$ increase of ΔE_t as luminance increases. This could be due to the Perceptual Quantizer (PQ) curve already imbedded in both models for luminance adaptation. Another trend can be found that the ΔE_t values from the two experiments are very similar. The ΔE_t values were used later to optimize colour difference formulae as HDR correction factors (c) given below.

$$c_i = \frac{\Delta Et_1}{\Delta Et_i} \tag{2}$$

where i ranges from 1 to 9 in Experiment 1, and from 1 to 6 in Experiment 2.

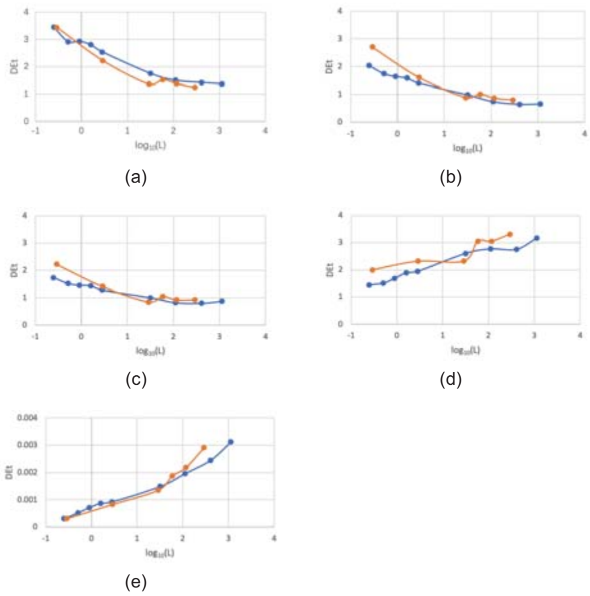


Figure 3. Colour difference threshold curves for a) CIELAB; b) CIEDE2000; c) CAM02-UCS; d) ICtCp; e) $J_za_zb_z$ respectively. The surface and luminous media are plotted in blue and orange colours respectively

Testing the Performance of Colour Models

The STRSS was again calculated between the predicted ΔE values and ΔV values to indicate the five colour models' performance. All luminance levels' data were combined for the surface and luminous model respectively in the calculations. Each colour model had five versions, designated as ΔE_1 to ΔE_5. Their generic structures are given below. For all models except CIEDE2000, the R_T term was set at zero.

$$\Delta E_1 = \sqrt{(\Delta L)^2 + (\Delta C)^2 + (\Delta H)^2 + R_T(\Delta C)(\Delta H)}$$

$$\Delta E_2 = c\sqrt{(\Delta L)^2 + (\Delta C)^2 + (\Delta H)^2 + R_T(\Delta C)(\Delta H)}$$

$$\Delta E_3 = c\sqrt{\left(\frac{\Delta L}{K_L}\right)^2 + (\Delta C)^2 + (\Delta H)^2 + R_T(\Delta C)(\Delta H)}$$

$$\Delta E_4 = c\left(\sqrt{(\Delta L)^2 + (\Delta C)^2 + (\Delta H)^2 + R_T(\Delta C)(\Delta H)}\right)^\gamma$$

$$\Delta E_5 = c\left(\sqrt{\left(\frac{\Delta L}{K_L}\right)^2 + (\Delta C)^2 + (\Delta H)^2 + R_T(\Delta C)(\Delta H)}\right)^\gamma$$

where c is the HDR correction factor as shown in Figure 3; K_L is the lightness parametric factor; γ is a power factor. Tables 3 and 4 list these factors.

The original colour model is ΔE_1. ΔE_2 is the HDR model which is based on the HDR correction factors derived from the present study. ΔE_3 model is an extension of ΔE_2 by including the lightness parametric factor (K_L). ΔE_4 is the ΔE_3 model with a power factor (γ) as introduced by Huang et al [16]. ΔE_5 is a full model to include all the 3 corrections, c, K_L and γ. Figures 4(a) and 4(b) summarize all models' performance for Experiment 1 and Experiment 2 results respectively.

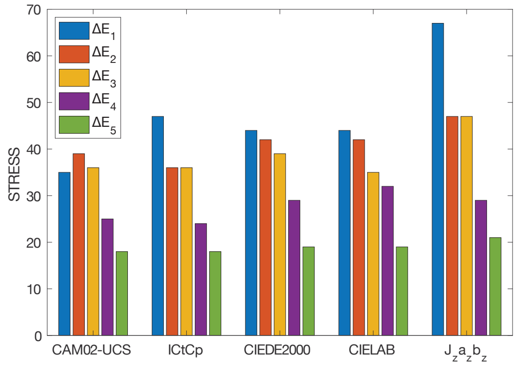
Table 3: The optimized factors for each space or formula in Experiment 1 (Surface mode)

		ΔE_3	ΔE_4	ΔE_5
CAM02-UCS	K_L value	0.66	1.00	0.31
	γ value	1.00	0.45	0.42
ICtCp	K_L value	0.91	1.00	0.33
	γ value	1.00	0.40	0.42
CIEDE2000	K_L value	0.66	1.00	0.22
	γ value	1.00	0.42	0.39
CIELAB	K_L value	0.50	1.00	0.15
	γ value	1.00	0.43	0.40
$J_za_zb_z$	K_L value	0.81	1.00	0.14
	γ value	1.00	0.35	0.30

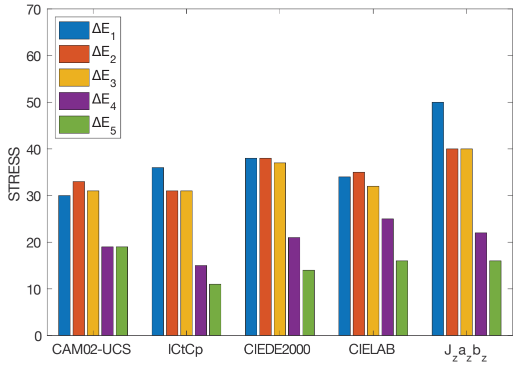
Table 4: The optimized factors for each space or formula in Experiment 2 (Luminous mode)

		ΔE_3	ΔE_4	ΔE_5
CAM02-UCS	K_L value	0.79	1.00	1.05
	γ value	1.00	0.72	0.72
ICtCp	K_L value	1.10	1.00	0.53
	γ value	1.00	0.49	0.43
CIEDE2000	K_L value	0.81	1.00	0.36
	γ value	1.00	0.41	0.38
CIELAB	K_L value	0.71	1.00	0.30
	γ value	1.00	0.43	0.40
$J_za_zb_z$	K_L value	0.85	1.00	0.29
	γ value	1.00	0.36	0.33

From Figure 4, comparing all ΔE_1 models for both modes, $J_za_zb_z$ performed the worst (STRESS of 67, 50), followed by ICtCp (47, 36), CIELAB (44, 34), CIEDE2000 (44, 38), and CAM02-UCS (35, 30). However, it can also be found that all models markedly improved from the original (ΔE_1) with the largest improvement for ΔE_5 model for each space or formula, except that CAM02-UCS did not improve by applying the HDR correction factors. Comparing the improvement at each stage, the individual correction for HDR correction or K_L factor made smaller improvement than that of γ and all 3 combined.



(a)



(b)

Figure 4. *Testing the performance of ΔE1 to ΔE5 for CAM02-UCS, ICtCp, CIEDE2000, CIELAB and $J_za_zb_z$ using the results from a) Experiment 1 and b) Experiment 2, respectively.*

Conclusion

Two psychophysical experiments were carried out to investigate the change of colour difference from very bright to very dark luminance level for the surface and display colours. The results were used to establish look up table including JND tolerance at different luminance levels. It can be used to extend uniform colour spaces and colour difference equations for HDR application. Five colour models, CIELAB, CIEDE2000, CAM02-UCS, $J_za_zb_z$ and ICtCp were tested. The performance of colour models was tested and the results showed that CAM02-UCS performed the best, CIELAB, CIEDE2000, and ICtCp performed mediocre, and $J_za_zb_z$ performed slightly the worst. By introducing the HDR colour correction factor (c), the power exponent factor (γ) and lightness parametric factor (K_L), all models' performance greatly improved and they all gave similar performance.

Acknowledgement

The authors like to thank the support from the Chinese Government's National Science Foundation (Project Number on 61775190).

References

[1] CIE 116-1995, "Industrial Colour-Difference Evaluation", Vienna: CIE Central Bureau, 1995.

[2] CIE142-2001, "Improvement to Industrial Colour Difference Evaluation", Vienna: CIE Central Bureau, 2001.

[3] M. R. Luo, G. Cui, B. Rigg, "The Development of The CIE 2000 Colour-Difference Formula: CIEDE2000", Colour Research & Application, 26(5):340-350, 2001

[4] N. Moroney, M. D. Fairchild, R. W. G. Hunt, C. Li, M. R. Luo, "The CIECAM02 Colour Appearance Models", in Proceeding of Tenth Colour Imaging Conference (CIC10), Scottsdale, AZ: IS&T. 23–27, 2002.

[5] C. Li, M.R. Luo, R.W.G. Hunt, N. Moroney, M.D. Fairchild, T. Newman, "The Performance of CIECAM02", in Proceeding of Tenth Colour Imaging Conference (CIC10), Scottsdale, Arizona. 28–32, 2002.

[6] CIE Publication 159:2004. "A Colour Appearance Model for Colour Man- Agement Systems: CIECAM02", Vienna: CIE Central Bureau, 2004.

[7] M.R. Luo, G. Cui, C. Li, "Uniform Colour Spaces Based on CIECAM02 Colour Appearance Model", Colour Res Appl, 31: 320– 330, 2006.

[8] C. Li, Z. Li, Z. Wang, Y. Xu, M.R. Luo, G. Cui, "Manuel Melgosa, Michael H. Brill and Michael Pointer, Comprehensive colour solutions: CAM16, CAT16 and CAM16-UCS", Colour Res Appl, 42:703–718, 2017.

[9] M. Withouck, K.A.G. Smet, W.R. Ryckaert, P. Hanselaer, "Experimental Driven Modelling of The Colour Appearance of Unrelated Self-Luminous Stimuli: CAM15u", Optics Express, 23: 12045–12064, 2015.

[10] W. J. Huang, Y. Yang and M.R. Luo, "Verification of the CAM15u Colour Appearance Model and the QUGR Glare Model", Lighting Res. Technol, doi.org/10.1177/1477153517734402, 2017.

[11] C. Li, X. Liu, K. Xiao, Y.J. Cho, and M.R. Luo, "An Extension of CAM16 for Predicting Size Effect and New Colour Appearance Perceptions", 26th Colour and Imaging Conference, November 12-16, Vancouver, Canada, pages 264-267, 2018.

[12] P. Boher, T. Leroux, P. Blanc, "New ICtCp and $J_za_zb_z$ Colour Spaces to Analyze the Colour Viewing-Angle Dependence of HDR and WCG Displays", SID Symposium Digest of Technical Papers, 49, 169-172. 10.1002/sdtp.12511, 2018.

[13] M. Safdar, G. Cui, Y.J. Kim, M.R. Luo, "Perceptually Uniform Colour Space for Image Signals Including High Dynamic Range and Wide Gamut", Optics Express, 25. 15131. 10.1364/OE.25.015131, 2017.

[14] F. Mirjalili, M.R. Luo, G. Cui, J. Morovic, "A Parametric Colour Difference Equation to Evaluate Colour Difference Magnitude Effect for Gapless Printed Stimuli", Colour and Imaging Conference, 123-127, 10.2352/ISSN.2169-2629.2018.26.123, 2018.

[15] P.A.R. García, H.M. Melgosa, G. Cui, "Measurement of the Relationship Between Perceived and Computed Colour Differences", J. Opt. Soc. Am. A, 24(7), 1823–1829, 2007.

[16] M. Huang, G. Cui, M. Melgosa, S. Manuel, C. Li, M.R. Luo and H. Liu, "Power functions improving the performance of Colour -difference formuls", Optical Express, 597-610, 2015.

Author Biography

Qiang Xu received his BS in Optical Engineering from Zhejiang University in June 2018 and he is now a Master student supervised by Professor Ming Ronnier Luo at Zhejiang University in September 2018. His research work is on colour difference under HDR luminance range and determination of contrast sensitivity function at different colour centres.