https://doi.org/10.2352/issn.2169-2629.2019.27.14
© 2019 Society for Imaging Science and Technology

Developing a visual method to characterize displays

Yu Hu, Ming Ronnier Luo*, Baiyue Zhao and Mingkai Cao; State Key Laboratory of Modern Optical Instrumentation, Zhejiang University; Hangzhou, China
****m.r.luo@zju.edu.cn***

Abstract

The goal is to develop a display characterization model to include the personal vision characteristics. A two-step model for visually characterizing displays was developed. It was based on the concept of half-toning technique for obtaining gamma factor for each colour channel, and unique hue concept for achieving 3x3 matrix coefficients, respectively. The variation can be presented by the optimized RGB primaries for each observer. The typical difference between the individual and the measured ground truth is 2.2 in terms of CIEDE2000 units.

Introduction

Nowadays, 3C products are full of our lives, mobile phone, laptop, notebook computer, TV, etc. They all include a display, which is the window for us to contact to the outside world. The quality of the displays has been highly desired. They have different sizes from small to large and also have varieties of technology, such as LCD, LED, OLED, etc.

Each display was characterized before shipping to the market. This will allow the same image to be truthfully reproduced onto different displays. In the colour management terms, a colorimetric reproduction can be achieved. The current typical set up is standard RGB, or sRGB system [1]. The characterization procedure has been well established in the production line. The topic of characterization models has also been extensively studied [2, 3]. Different models were proposed such as GOG, PLCC or 1D-LUT, PLVC or 3D-LUT, etc. Each model can give a reasonable prediction to the measurement data typically measured by a tristimulus colorimeter or a spectroradiometer.

Two-stage model is widely used to characterize displays with well channel independence and chromaticity constancy. The first step involves the relationship between digital counts and relative luminance of each channel, which usually named display opto-electronic transfer function (OETF) [4]. The second stage uses a 3 × 3 matrix to obtain CIE XYZ tristimulus from luminance signal.

The gain-offset-gamma (GOG) model proposed by Berns [2] is the most common two-stage model to characterize cathode ray tube (CRT) displays. Although the originally typical OETF of liquid crystal display (LCD) is modeled by S-curve function [5, 6], GOG model can also perform pretty well on most LCD with a trend of using the power function OEFT [7]. First step of GOG model (gamma correction) which is given in equations (1) and (2)

$$L_{i,d_i} = \begin{cases} \left(k_{i,g}\left(\frac{d_i}{255}\right) + k_{i,o}\right)^{\gamma_i}, & \left(k_{i,g}\left(\frac{d_i}{255}\right) + k_{i,o}\right)^{\gamma_i} \geq 0 \\ 0 & , \left(k_{i,g}\left(\frac{d_i}{255}\right) + k_{i,o}\right)^{\gamma_i} < 0 \end{cases} \quad (1)$$

$$k_{i,o} = 1 - k_{i,g} \quad (2)$$

where L_{i,d_i} is relative luminance of each channel and d_i is digital counts. $k_{i,g}$, $k_{i,o}$ and γ_i represent gain, offset and gamma respectively.

The equation (3) calculates device-independent CIE XYZ tristimulus

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} X_{R,255} & X_{G,255} & X_{B,255} \\ Y_{R,255} & Y_{G,255} & Y_{B,255} \\ Z_{R,255} & Z_{G,255} & Z_{B,255} \end{bmatrix} \begin{bmatrix} L_{R,d_r} \\ L_{G,d_g} \\ L_{B,d_b} \end{bmatrix} + \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix} \quad (3)$$

where (X_0, Y_0, Z_0) is black point's tristimulus values. The 3 × 3 matrix contains the XYZ value at maximum level of each channel.

Piecewise Linear Chromaticity Constancy (PLCC) is an alternative method to the GOG model. It is based on a functional approximation by applying a linear interpolation between measurements [3]. The PLCC method will outperform the GOG method with enough linearization samples. Meanwhile, the PLCC model's error increases as the number of samples is N reduced. For N < 10 the GOG model achieves better performance than PLCC [7].

The GOG and PLCC both transfer device-dependent RGB space to device-independent XYZ space, which was based on standard observer colour matching function (CMF). However, two observers never perceive the same colour from the same stimulus because of different CMFs. Heckaman and Ho [8] studied the colour accuracy of variety display technologies among 1000 simulated observer CMFs. The mean colour difference across these technologies centers around 4 CIELAB units. It shows that the observer variations should be taken into consideration to achieve better colour accuracy for individual observer.

The methods to measure CMFs [9, 10] usually take a large amount of time, which is difficult to apply to everyday life. And each individual has different colour vision varied according to age, eye condition, personal preference, ambient lighting conditions. With the common place of mobile devices, we can argue why we cannot have the device considering our own eyes? This research was conducted to answer this question.

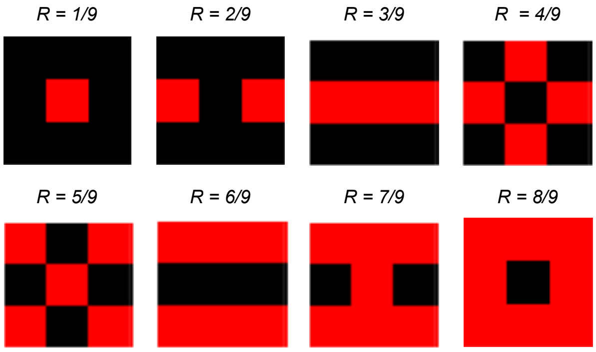


***Figure 1.** Distribution of peak color and black for a 3 × 3 pixels block [13]*

Table 1. Specification of tested displays and their characterization performance*

Brand	***Model***	***Panel***	***Size (")***	***White***		***Peak L cd/m²***	***GOG***	***Gamma***	***PLCC***
				x	***y***				
Eizo	CG243W	IPS	24.1	0.3194	0.3367	192.1	0.28	0.42	0.23
NEC	PA272W	AH-IPS	27	0.3145	0.3277	150.2	0.65	0.63	0.47
NEC	PA302W	IPS	29.8	0.3183	0.3229	280.9	0.57	0.63	0.54
Alienware	17R3	IPS	17.3	0.2955	0.2913	289.5	1.50	1.69	0.49
Apple	A1701	IPS	10.5	0.3084	0.3300	429.8	1.02	0.87	0.97
Sony	SVT112A2WP	IPS	11.6	0.3123	0.3375	350.1	1.16	1.64	0.49
Google	Pixel	AMOLED	5	0.3132	0.3418	486.5	1.07	1.59	1.09

*Note the last 3 columns reported the performance of 3 characterization models in terms of ΔE^*_{00} units.

The goal of this work is to develop a display characterization model to include both the personal vision and display characteristics. The hypothesis is that the model should represent this individual's vision and give a more comfort performance for the display. Some visual methods [11-14] have been developed to estimate OETF using luminance matching task. While different relative luminance levels of half tone patterns is designed as the target, observer controls the digital count of another patch to have equal bright as the target. Then OETF is derived from those data points. Xiao *et al.* [13] designed eight half-tone images (Figure 1) with relative luminance varied from 1/9 to 8/9. The result reveals that this method can reliably estimate OETF. The methods above only characterize OTEF visually, therefore the 3 × 3 matrix contained ground truth data is still needed to complete the characterization model.

For developing the 3x3 matrix, Karatzas and Wuerger [15] developed a colour calibration method based on unique hue judgement. Unique hues were first introduced by Hering [16] as the hues of four fundamental chromatic percepts regardless of lightness and saturation: unique yellow (UY) and unique blue (UB), unique red (UR) and unique green (UG). Unique hue judgements are not significantly influenced by language or age. For Karatzas's method, users have to do unique hue judgement on two displays, one is reference display and another one is test display. The transfer matrix was established using two unique hue plane based on device-dependent RGB space. It is not applicable to normal users who do not have a reference display around.

This study aims to develop a visual method to characterize display, including gamma correction and device-independent transformation. A visual experiment carried out on a display to produce a visual display model for individual observer. The performance of visual display model was inter-compared between different observers and also contrast with the ground truth data (measured by spectrophotometer).

Experimental

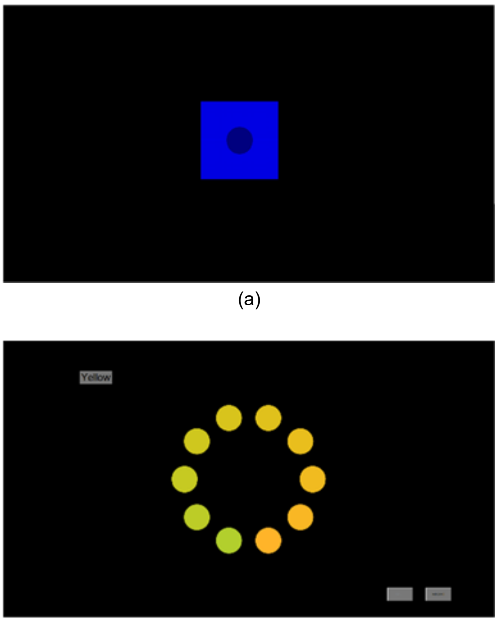
Characterize displays using instrument

Before developing the visual method, we used spectrophotometer Konica Minolta CS2000A measured 7 different displays including PC monitor, laptop, tablet and mobile. Most of them are LCD except one mobile has OLED panel. Table 1 shows the specification of 7 displays. The linearization sample set was the set of three colour ramps. The RGB values is starting from 0 to 255 with 15-unit interval for each channel. The testing sample set was the Macbeth ColorChecker chart. It's impossible to obtain black level and white point data using visual method. So the basic scheme of visual method simplifies GOG to power function (Gamma), which can be described by equation (4) and (5).

$$L_{i,d_i} = \left(\frac{d_i}{255}\right)^{\gamma_i} \tag{4}$$

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} X_R & X_G & X_B \\ Y_R & Y_G & Y_B \\ Z_R & Z_G & Z_B \end{bmatrix} \begin{bmatrix} L_{R,d_r} \\ L_{G,d_g} \\ L_{B,d_b} \end{bmatrix} \tag{5}$$

Three display models were tested including GOG, PLCC and Gamma. The performance was evaluated by the CIEDE2000 (ΔE^*_{00}) [17] colour difference between the measured 24 XYZ values and those predicted by the display model. Table 1 also shows the models' performance. With 18 steps sample set, the PLCC model generally out-performed the GOG model and the Gamma model, except for Apple iPad Pro 2017. Additionally, the Gamma model was slightly worse than the GOG model, but still had very good performance for all the displays tested. It was decided to adopt Gamma model in the visual method.



(a)

(b)

Figure 2. *Experiment interface for (a) visual gamma correction; (b) unique hue selection.*

Apparatus

The NEC PA302W display, with resolution of 2560 ×1600 pixels, driven by a Dell computer was used to stimulus presentation in the main experiment. The display had a correlated colour temperature (CCT) of about 6500K with peak white luminance of 280 cd/m^2. The visual experiment was conducted in a completely darkened room to avoid the effect of ambient light.

Observers

Thirty-five observers (23 males and 12 females) ranging from 19 to 28 years of age (mean = 23, std. dev. = 1.45) participated in the experiment. All the observers had normal color vision, as tested using the 24 Plate Ishihara Color Vision Test. They were naïve in regard to the goal of the experiment. Ten observers of them had basic colour science knowledge.

Procedure

Upon arrival, the observer filled in the personnel data sheet and performed the Ishihara Color Vision Test. The experimenter then provided an instruction and answered the questions raised by each observer. They were then escorted to the experimental room and being seated in front of display with distance of 50 cm. Each observer took 1 minute of adaptation. In general, each observer took around 30 min in total to complete the experiment.

Step I: visual gamma correction

The same half-tone images were used in the visual gamma correction as Xiao *et al.*, which shows in Figure 1. A Matlab GUI experiment interface was designed to perform the visual gamma correction as shown in Figure 2(a). A uniform disk with a 2° diameter was displayed on a half-tone image (6° × 6°) with the total black surrounding. Observer used arrow keys to adjust the digital count of uniform disk. “Up arrow” and “down arrow” keys were coarse adjustment with 8-unit adjustment, while “left arrow” and “right arrow” keys performed fine adjustment with 1 unit changing.

If the lightness of uniform disk matched with the half-tone image, observer clicked “enter” key to change to the next sample and digital count will be recorded for OETF estimation. The experiment sequence of channel and relative luminance level were randomized. To obtain intra-observer variation, the same experiment was repeated within several min interval. Each observer finished 48 times (3 channels × 8 luminance levels × 2 repeats) matching, which took around 15 mins.

Table 2. Lightness and chroma of test sample for each unique hue.

	L^*	C^*_{ab}
Red	35	65
	55	99
Yellow	70	70
	80	90
Green	40	50
	60	70
Blue	55	35
	65	45

Step II: Unique hue selection

As shown in Figure 2(b), an annulus of circle patches with the same lightness and chroma but different hue angles presented on the display. Each patch had diameter of 3° of visual angle and was arranged along a 7° annulus. Observers were asked to select unique colour from those 10 patches. Top-left shown which unique colour was the target. UY and UB were obtained by selecting the patches which were neither red nor green. Likewise, UR and UG were the patches which do not contain yellow or blue. And observers used the mouse to click the patch to choose unique hue colour.

The test colours were chosen from another experiment. That experiment was aimed to establish unique hue loci to test the hue linearity in different uniform colour spaces. There were 10 colour centres selected for each unique hue page. The 8 colours selected were within the colour gamut of the display and also had high chroma to be identified as a pure red, yellow, green and blue. Table 2 lists each unique colour defined by L^* and C^*_{ab} values. For each test sample, ten patches were arranged in the order of hue angle, as verified by Xiao *et al.* [18]. They found that similar results would be obtained between the randomized or sequential order but lower observer variability came from the sequential order.

The transform function between XYZ and sRGB was used to calculate RGB values for ten patches due to absent of measured display model during the characterization process. Starting from RGB values obtained from previous experiment, they were transformed to XYZ values via sRGB formula, then to CIELAB in term of L^*, C^*_{ab}, and h. Only h changed to 10 different values dependent on preset hue angle range. Finally, reverse transform of sRGB was used to calculate 10 RGB values which displayed on screen at one trial.

At each trail, the initial hue angle range was set at ±18° with 4° intervals. Then the smaller hue range derived from the previous response. The smaller range was set at ±9° with 2° intervals. After two rounds of selection, the RGB values were saved for succeeding optimization process. To obtain intra-observer variation, the same experiment was also repeated within several min interval. Each observer finished 36 times (8 samples × 2 hue ranges × 2 repeats) selection. This took around 10 mins.

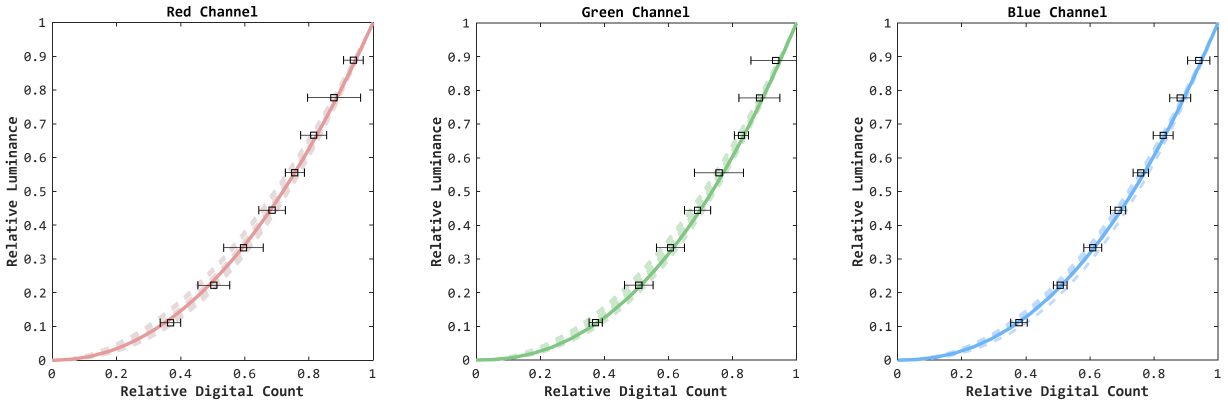
Table 3. Intra- and inter-observer variability in terms of MCDM

			MCDM		
			Mean	***Max***	***STDEV***
Visual Gamma Correction	Intra	R	0.43	1.63	0.36
		G	0.49	2.20	0.43
		B	0.19	0.58	0.11
	Inter	R	0.68	2.10	0.44
		G	0.88	3.20	0.67
		B	0.27	0.77	0.14
Unique Hue Selection	Intra	UR	1.46	4.05	0.95
		UY	1.47	3.20	0.73
		UG	1.11	4.43	0.82
		UB	1.20	3.76	0.93
		All	1.31	4.43	0.87
	Inter	UR	2.43	5.59	1.27
		UY	2.71	5.89	1.41
		UG	2.23	5.65	1.27
		UB	3.29	6.85	1.55
		All	2.66	6.85	1.42

Results and Discussion

Observer variations

Firstly, the inter- and intra-observer variability were calculated to evaluate how reliable the experiment data were. The



(a) (b) (c)

Figure 3. *The estimated OETF from measured data (solid line), the mean results (square) with 95% confidence interval and the predicted OETF (dashed line) for 35 observers, for (a) red channel, (b) green channel, and (c) green channel.*

mean colour difference to the mean value (MCDM) using CIEDE2000 colour difference formula was calculated. The mean value for inter-observer MCDM was calculated by averaging the data of 35 observers; the mean value to calculate he intra-observer MCDM was the average of 2 repeated data for each observer. Table 3 lists the inter- and intra-observer variability results. For visual gamma correction results, the intra-observer variations were ranged from 0.02 to 2.2 and inter-observer variation were from 0.06 to 3.2. The largest MCDM of inter- observer variability occurred at green channel (3.20 ΔE^*_{00}), whereas the maximum intra-observer MCDM was 2.2 ΔE^*_{00}. For unique hue selection task, the intra-observer values were ranged from 0.19 to 4.43 and inter-observer were from 0.34 to 6.85 for all observer in terms of MCDM. The above MCDM values was similar to those reported by Xiao *et al.* [13, 18]. The results showed that the observer variation of visual gamma correction is much smaller than that the observer variability of unique hue selection task.

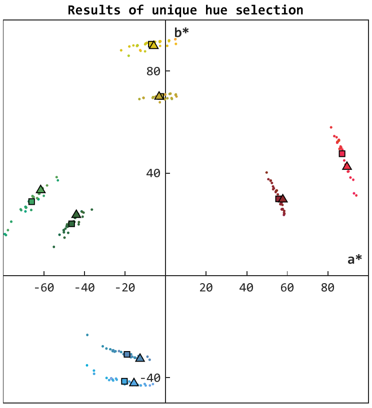


Figure 4. *The unique hue selection results of all observers plotted on a*b* diagram (dot); The mean results for each test sample (square), and the mean results from previous study which will be used as optimization target (triangle).*

Visual gamma correction

In Figure 3 the relative luminance was plotted as a function of input digital count of the matching uniform disk. Figure 3 shown the average matching results of all observers with the 95% confidence interval, and ground truth data based on measured results of 18 steps colour-ramp samples (solid line). The power function (dashed line) was used to fit OETF for each observer, as shown in Figure 3. Each observer's predicted OETF is within a small range of ground truth data, therefore, there is little bias introduced by visual matching which agrees well with PLCC and the 3th order polynomial in Xiao's study [13].

Development of 3x3 matrix

Figure 4 shows the unique hue selection results measured by spectrophotometer. And the mean results for each test sample and the mean results from our pilot study both are also plotted in Figure 4. Table 4 shows the comparison between two sets of unique hue data. In spite of not using display model to find unique hue, this study still surprisingly produced very similar result to our previous study. It clearly proved that unique hue was a stable colour indicator. There were the smallest colour difference on the yellow region of 0.73 ΔE^*_{00} and the largest colour difference on blue region of 4.02 ΔE^*_{00}.

Table 4. Comparison between two sets of unique hue results. (this study ■; previous study ▲)

	h (■)	*h* (▲)	|Δ*h*|	|Δ*H*|	Δ*E*₀₀
Red	28.3	27.3	1.1	1.2	0.87
	28.7	25.4	3.3	5.7	2.80
Yellow	92.0	92.6	0.6	0.8	0.79
	94.4	93.7	0.7	1.1	0.73
Green	156.4	151.7	4.7	4.1	2.27
	156.4	151.6	4.8	6.0	2.82
Blue	238.3	248.9	10.6	6.5	4.02
	243.9	249.7	5.8	4.6	2.69

It can be seen that difference between each observer's selection result and mean observer result could represent the difference between their CMF and standard observer CMF in some degree. In other words, if standard observer CMF was replaced to individual's CMF, the unique hue result would be more close to mean results of unique hue.

So an optimization method was proposed to obtain the 3 × 3 colour transfer matrix. The input data included visual gamma values and unique data set, while the variables was the matrix. And there were two constraints which related to white point. The

Y value of white was set around 100, and the chromaticity was close to D65 (sRGB assumption). And objective function was to minimize sum of $|\Delta H|$ which compared to target unique hue angle. Figure 5 shows the optimal matrix for each observer in form of display primaries. It also shows the typical individual characteristics of the 35 individual observers participated in the current study.

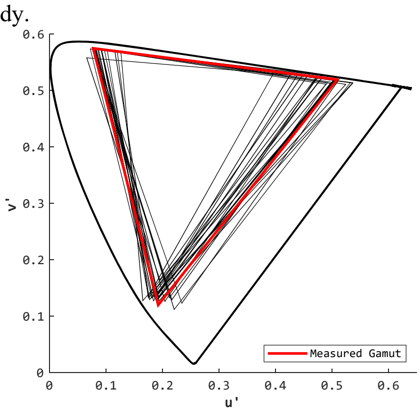


***Figure 5.** The optimal matrix for each observer in form of display gamut (black line); The measured gamut (bold red line).*

Performance evaluation

To evaluate the performance of the visual method, the measured XYZ values of 24 MCCC samples were compared with those ground truth data with the XYZ values predicted from visual model. Because this visual model is two-stage model including visual gamma and visual matrix, there are three combinations for the visual model. First one is visual gamma combined with ground truth matrix (VG). Second one is ground truth gamma with visual colour matrix optimized from unique hue data (UH). And the last one is including both visual data (VGUH).

Table 5. Performance of three visual model.

ΔE_{00} (35 obs.)	*Mean*	*Min*	*Max*	*STDEV*
VG	1.80	0.65	4.47	0.97
UH	1.61	0.94	4.73	0.89
VGUH	2.21	0.65	5.01	1.08

Table 5 shows performance of three visual models. It shows that method only applied one stage's visual data will have better performance than the method applied both stages' visual data when compared with measuring data. But all of them have reasonably good performance. However, VGUH model should represent individual observer's characteristics. It includes two parts, VG and UH, for which the VGUH can be approximated using sqrt(VG^2+UH^2).

Further work is ongoing to investigate the visual characterization method based on the display having non-D65 peak white. The results are also quite promising.

Conclusion

This study proposed a visual method to characterize display. The goal is to include the personal vision characteristics and to give a more comfort performance. The visual method not only can estimate the OETF, but also can obtain colour transform matrix without measuring instruments. Visual experiments were conducted, which collected visual data including half-tone luminance match task and unique hue selection task. Thirty-five observers participated the experiment. A visual model was optimized using the data for each individual observer. The visual model had pretty good performance in agreement with the ground truth model based on the instrumental measuring data.

Acknowledgment

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

References

[1] IEC 61966-2-1: “Multimedia systems and equipment-Colour measurement and management - Part 2-1: Colour management - Default RGB colour space-sRGB.” Geneva, Switzerland, IEC (1999).

[2] R.S. Berns, R.J. Motta, M.E. Gorzynski, “CRT Colorimetry. Part I: theory and practice,” Color Res. Appl. 18, 299–314 (1993).

[3] D.L. Post, C.S. Calhoun, “An evaluation of methods for producing desired colours on CRT monitors,” Color Res. Appl. 14, 172–186 (1989).

[4] “Parameter values for the HDTV standards for production and international programme exchange”, ITU-R BT.709-6 (2015).

[5] Y. Kwak, L. MacDonald, “Characterisation of a desktop LCD projector.” Displays 21, 179–194 (2000).

[6] Y. Wang and H. Xu, “Colorimetric characterization of liquid crystal display using an improved two-stage model,” Chin. Opt. Lett. 4, 432-434 (2006).

[7] M. Vazirian, “Colour characterisation of LCD display systems." Diss. University of Leeds (2018).

[8] R.L. Heckaman, J. Ho, “Observer Metamerism and Its Effect on Color Accuracy in Display Media Technology,” SID DIGEST, 723–726 (2016).

[9] A.D. North, M.D. Fairchild, “Measuring color-matching functions. Part I.” Color Res. Appl. 18, 155–162 (1993).

[10] A.D. North, M.D. Fairchild, “Measuring color-matching functions. Part II. New data for assessing observer metamerism.” Color Res. Appl. 18, 163–170 (1993).

[11] E. Colombo, A. Derrington, “Visual calibration of CRT monitors.” Displays 22, 87–95 (2001).

[12] A. Neumann, A. Artusi, L. Neumann, G. Zotti, W. Purgathofer, “Accurate display gamma functions based on human observation. Color Res. Appl. 32, 310–319 (2007).

[13] K. Xiao, C. Fu, D. Karatzas, S. Wuerger, “Visual gamma correction for LCD displays.” Displays 32, 17–23 (2011).

[14] L. To, R.L. Woods, R.B. Goldstein, E. Peli, “Psychophysical contrast calibration”. Vision Res. 90, 15–24 (2013).

[15] D. Karatzas, S.M. Wuerger, “A hardware-independent colour calibration technique.” Annals of the BMVA 3, 1-11 (2007).

[16] E. Hering, “Outlines of a theory of the light sense.” (Harvard University Press, Cambridge, MA, 1964).

[17] M. R. Luo, G. Cui, B. Rigg, "The development of the CIE 2000 colour-difference formula: CIEDE2000," Color Res. Appl. 26, 340-350, 2001.

[18] K. Xiao, S. Wuerger, C. Fu, D. Karatzas, “Unique hue data for colour appearance models. Part I: Loci of unique hues and hue uniformity.” Color Res. Appl. 36, 316–323 (2011).

Author Biograph

Yu Hu received his BSc degree from College of Optical Science and Engineering, Zhejiang University, Hangzhou, China. He is currently a PhD student supervised by Professor Ming Ronnier Luo at Zhejiang University since 2015. His research interests include lighting quality, display colour perception, and observer metamerism. Email: daniel_hu@zju.edu.cn