https://doi.org/10.2352/ISSN.2470-1173.2019.10.IQSP-309

The Image Quality Evaluation of HDR OLED Display

Dalin Tian[1], Lihao Xu[1], Ming Ronnier Luo*[1,2]

[1]State Key Laboratory of Modern Optical Instrumentation, Zhejiang University, Hangzhou, China
[2]School of Design, Leeds University, Leeds, UK
*m.r.luo@zju.edu.cn

Abstract

This study describes an experiment for image quality evaluation on a new OLED TV under high dynamic range (HDR) and wide colour gamut (WCG) condition. Eight attributes including peak brightness, blackness, colourfulness, contrast, reality, sharpness, texture details and overall image quality, were investigated. The results were used to understand the relationship between image quality and those attributes. An image quality model for OLED TV was derived using principle component analysis method. The model includes three main factors, spatial and colour quality, black level and brightness.

Introduction

The organic light emitting diode (OLED) displays have superior qualities than the conventional LCD displays, such as wide colour gamut (WCG), high dynamic range (HDR), and thinner panel thickness. It has been utilized in several commercial products such as smart phones, and televisions. One of the most desirable properties of an OLED display is its high rendering performance to provide a life-like appearance in terms of colour and image quality.

However, most of previous studies were concentrated on the image quality of traditional CRT and LCD [1, 2], and little work has been performed for OLED displays. Baek *et. al.* [3] compared a 2.2-inch OLED display a LCD display in terms of spectral power distribution, tone reproduction curve (TRC), luminance, contrast, correlated color temperature, 2D color gamut and spatial uniformity. It was found that the OLED gave better performance in the former four aspects but similar quality for the remaining attributes. Sun *et. al.* [4] proposed two characterization models and compared them with some classic display models. Although both models were reported to work well, they used some empirical formulas based on some assumptions. This implies they may not work well in practice. Ye *et. al.* [5] evaluated an OLED and two LCD TVs in terms of their performance using some rendering images. Seven types of image quality attributes as well as overall preference were evaluated and OLED TV was found to has a superiority over LCD TVs on the most of attributes studied, especially for peak brightness and contrast. Although the results confirmed that OLED could perform better than CRT and LCD, which attributes affected image quality most was unclear.

In the present study, a psychological experiment was performed to visually evaluate its performance in rendering real images and an analysis was performed to develop an image quality (IQ) model to correlate the overall IQ and each individual image attribute on the OLED TV.

Objective

A psychophysical experiment was conducted so that the visual perceptual data of OLED display was firstly collected using processed images. Eight attributes were assessed in this experiment, including peak brightness, blackness, colourfulness, contrast, reality, sharpness, texture details and overall image quality. Correlation coefficient matrix were calculated and principle component analysis (PCA) was further implemented to find the most influential attribute for the overall image quality. An image quality model was built for HDR OLED display.

Method

Experimental setup

A psychophysical experiment was carried out to investigate the image quality of an HDR OLED TV. A 65" Sony A1 OLED TV was adopted in this study. Its physical size was 1451 mm by 834 mm by 86 mm with a diagonal length of 1639 mm and its spatial resolution was 3840 pixels by 2160 pixels. It could achieve a luminance up to over 700 cd/m^2 and a quite low black level. Default settings were adopted when taken measurement under Cinema Pro mode. The TV settings for contrast enhancement was turned off. HDR and WCG functions were turned on and the OLED TV to approximate the BT.2020 standard [6] at this condition. Note BT.2020 is the most recent standard for digital cinema.

Twenty observers consisting of 11 females and 9 males participated in this experiment. Their ages ranged from 19 to 28 with an average of 23. They had experience to take part in psychophysical experiments and were trained before conducting the experiments. All observers undertook the Ishihara Test for colour defect to ensure that they had normal colour vision before carrying out this experiment. They sit in 2.5 meters in front of the center of OLED TV in a dark room (reflectance of the wall was 8%), which had no other light sources in the room except the display. Fig. 1 shows the experimental situation.

Figure 1. *experimental environment*

Table 1. Definition of each attributes

Attributes	Definition
Peak brightness	The amount of the light from the brightest region in the image
Blackness	The amount of the light from the darkest region in the image
Colourfulness	The amount of colour in an image appears to be more or less chromatic
Contrast	The range of luminance and colour to make the image to be distinguishable
Reality	The closeness of the image to the scene in the real world (memory colour, size, object, environment)
Sharpness	The discrimination of the boundaries between zones of different tones or colours.
Texture details	The ability to see the details in an image, especially in the shadow area.
Overall image quality	The quality of the image by comparing it with an imaginary image that have perfect quality.

A categorical judgement method was adopted in this study. For each phase, observers were asked to view and rate a displayed image in terms of 8 image attributes including peak brightness, blackness, colourfulness, contrast, reality, sharpness, texture details and overall image quality. Table 1 describes all 8 attributes. Each attribute had six categories, ranging from extremely bad, bad, a little bad to a little good, good and extremely good. For blackness, a higher score represented more blackness. Six original images with ten different rendering having different contents and features were included. All of them had the same resolution as the display and thus occupied the full field of view. All images/scales were arranged in a random order for each observer.

Image rendering

Six original images with ten different rendering having different contents and features were included. All of them had the same resolution as the display and thus occupied the full field of view. All images/scales were arranged in a random order for each observer.

All these images were processed using manipulation methods to generate a larger test database. Three image attributes were manipulated including lightness, chroma and sharpness. All the rendering methods are explained below.

Lightness Manipulation

To modify the lightness level of an image, the RGB values of an image was first transformed into XYZ value using The Society of Motion Picture and Television Engineers (SMPTE) inverse function, then to Jab value using the $J_z a_z b_z$ inverse function where J value represented the lightness ($J_z a_z b_z$ is a uniform color space for HDR and WCG application) [7]. And the J value should be normalized into the range of 0 to 100. After that, wo linear and two sigmoid functions were used on J channel. The linear functions were expressed as below:

$$J_{output} = \mathrm{S} * J_{input}$$

where S value were 0.7 and 1.2 respectively. And as for two sigmoid functions, they could be expressed as:

$$J_{output} = \frac{T}{\left[\frac{1}{1+\mathrm{M}^{\mathrm{E}}}\right] * \left\{1 + \left[\frac{M}{J_{input}}\right]^{\mathrm{E}}\right\}}$$

where M=1.23 and E=1.45, and T value were 1 and 1.2 respectively.

After these four different manipulations, the output J value should be de-normalized and then transferred to XYZ using $J_z a_z b_z$ reverse function and then to RGB using SMPTE reverse function. The sigmoid function reduced the lightness in the dark areas of the original image but increased the lightness in light areas.

Chroma Manipulation

Similarly, the chroma manipulation also contained two linear modification and two sigmoid modification. After transferred into $J_z a_z b_z$, the chroma value of each pixel was calculated using az and bz with function:

$$\mathrm{C} = \left({a_z}^2 + {\mathrm{b}_z}^2\right)^{\frac{1}{2}}$$

And similarly the linear and sigmoid function were:

$$C_{output} = \mathrm{S} * C_{input}$$

$$C_{output} = \frac{T}{\left[\frac{1}{1+\mathrm{M}^{\mathrm{E}}}\right] * \left\{1 + \left[\frac{M}{C_{input}}\right]^{\mathrm{E}}\right\}}$$

For linear function, S were 0.7 and 1.2, and for sigmoid function T were 1 and 1.2, respectively.

Sharpness manipulation

A high frequency emphasis filter was used and is given below. Since the cut-off frequency parameter (d) was smaller, the image became sharper due to a reduction in low-frequency information.

$$\mathrm{Filter} = 1 + 1.5 * \left\{1 - \exp\left[-\frac{\mathrm{x}^2}{2 * \mathrm{d}^2}\right]\right\}$$

where x is spatial frequency and d = 1024×P (cut-off frequency parameter). P equals 1/11 and 1/19 for small and large sharpness manipulation respectively.

Summary of image rendering

After each image was rendered with an algorithm, it was also checked to ensure over 95% pixels were within the colour gamut. As a result, each source image generated 10 reproductions. In total, 66 images were obtained, i.e. (10 reproductions + 1 original image) * 6. Each image should be judged in terms of eight image attributes as introduced above. Besides, 6 out of these 66 image were repeated. These 6 repeat images were only assessed by three attributes: overall image quality, colourfulness, and contrast.

In total, 10,920 judgements were made, i.e. 20 (observers) * [66 (images) * 8 (attributes) +6 (repeats) * 3 (attributes)].

Data analysis and results

Torgerson's law of categorical judgment was applied to transform the raw data into z-score values, for which a higher z-

Table 2. Correlation coefficient matrix

	blackness	brightness	colorfulnes	contrast	reality	sharpness	texture
blackness	1.00	0.33	0.28	0.47	-0.03	0.25	0.04
brightness	0.33	1.00	0.10	0.38	0.18	0.13	0.00
colorfulnes	0.28	0.10	1.00	0.86	0.45	0.65	0.62
contrast	0.47	0.38	0.86	1.00	0.50	0.68	0.57
reality	-0.03	0.18	0.45	0.50	1.00	0.56	0.60
sharpness	0.25	0.13	0.65	0.68	0.56	1.00	0.82
texture	0.04	0.00	0.62	0.57	0.60	0.82	1.00

score means a more positive perception such as brigher, blacker, more colourful, etc. Inter-observer variability and intra-observer variability were evalued in terms of coefficient of variation (CV) to represent the consistency and repeatability of observers. For a perfect agreement, CV should be zero. A CV of 30, means 30% variation between the two sets of data. For the inter-observer variability, the best and worst observers gave 0.25 and 0.48 CV units, respectively, with a mean of 0.37. And as for the intra-observer variability, the best and worst observers gave 0.29 and 1.17 CV units, respectively, with a mean of 0.37. These values indicate this experimental results were quite reasonable, compared with previous study.

Correlation coefficient (r) was used to compare all possible combination of two attributes. Table 2 summarises all the r results. Larger r values stand for a larger relativity. Table 2 shows that colourful and contrast were correlated well with each other with an r value of 0.86, and sharpness and texture details were also correclted well (r= 0.82).

Principle component analysis (PCA) was also carried out to find the most influential components for the overall image quality. Seven image quality attributes (peak brightness, blackness, colourfulness, contrast, reality, sharpness, and texture details) were used here in order to determine what these components actually represent. Table 3 shows the factor loading between the components and seven independent variables were computed.

Table 3. Principle component analysis of seven attributes

	components		
	1	2	3
texture	0.91	-0.04	-0.07
sharpness	0.87	0.18	0.03
colorfulness	0.79	0.41	-0.04
reality	0.77	-0.29	0.34
contrast	0.74	0.51	0.27
blackness	0.06	0.90	0.20
brightness	0.03	0.22	0.95
distribution(%)	48.10	20.12	16.12

The 'distribution (%)' row explains the percentage of variance accounted for by each component. For example, the first component account for nearly 48% of the variability in the seven independent variables used for determining image quality. The summary of the distribution of these three components were nearly 84%, which meant they could describe the space.

Practically most of the colour related attributes were integrated in Component one (**Q_1**), including spatial and colour quality. Component two (**Q_2**) and three (**Q_3**) represent blackness and peak brightness respectively.

Afterwards a linear regression was conducted between the three components and image quality. Thus, the following formula is obtained.

$$\text{Image quality}=0.59*Q_1+0.12*Q_2+0.02*Q_3+\text{constant}$$

The correlation coefficient between the measured image quality and predicted image quality using this formula was 0.89, indicating that this formula is accurate to predict the visual image quality evaluation.

This formula implies that spatial and colour quality is an important property to influence the image quality, which agreed with previous study to be a commonly known vital attribute, related with gamut volume and spatial resolution. Black level turns out to be also important in the overall image quality. In addition, the brightness of an image did not show very high contribution with image quality. These three components implies that to enhance the image quality we should focus on two aspects: HDR property (blackness and brightness), like higher brightness and lower black level; and spatial and colour quality, i.e. higher resolution and wider colour gamut.

Conclusion

A study on the image quality evaluation of HDR OLED display was conducted. A psychophysical experiment was carried out to correlate the overall image quality with seven individual attributes. The results showed that besides the colour quality, the blackness of an image could also influence the visual image quality, which is more important than that of brightness. An image quality model for HDR display was proposed.

References

[1] N. Katoh, T. Deguchi and R. S. Berns, "An Accurate Characterization of CRT Monitor (I) Verifications of Past Studies and Clarifications of Gamma," *Optical Review,* vol. 8, pp. 305-314, 2001.

[2] Y. Kwak and L. Macdonald, "Characterisation of a desktop LCD projector," *Displays,* vol. 21, pp. 179-194, 2000.

[3] Y. Baek, H. Kim, S. Park, and Y. Kim, "Colorimetric Characteristics Evaluation of OLED and LCD," in *International Meeting on Information Display*, 2008, pp. 509-512.

[4] P. L. Sun and R. M. Luo, "Color Characterization Models for OLED Displays," in *SID Symposium Digest of Technical Papers*, 2014, pp. 1453-1456.

[5] Z. Ye, J. Qiu, H. Xu, M. R. Luo, and S. Westland, "Image Quality Evaluation of HDR Displays," in *SID International Symposium, Seminar & Exhibition*, 2017.

[6] ITU, "Parameter values for ultra-high definition television systems for production and international programme exchange," in *ITU-R Recommendation BT.2020*, 2012.

[7] M. Safdar, G. Cui, Y. J. Kim, and M. R. Luo, "Perceptually Uniform Color Space for Image Signals Including High Fynamic Range and Wide Gamut.," Optics Express, vol. 25, p. 15131, 2017.

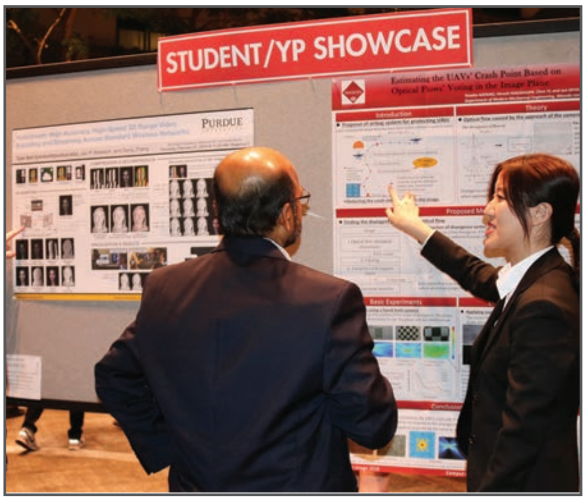
JOIN US AT THE NEXT EI!

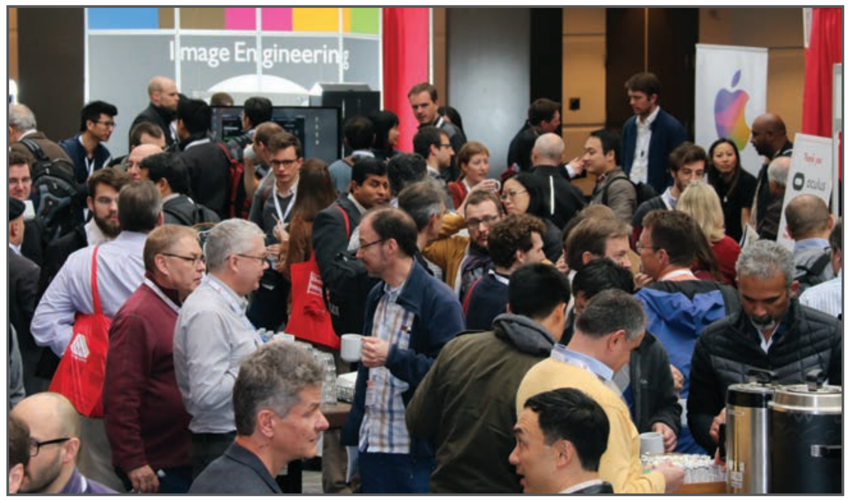
IS&T International Symposium on

Electronic Imaging

SCIENCE AND TECHNOLOGY

Imaging across applications . . . Where industry and academia meet!





• SHORT COURSES • EXHIBITS • DEMONSTRATION SESSION • PLENARY TALKS •
• INTERACTIVE PAPER SESSION • SPECIAL EVENTS • TECHNICAL SESSIONS •

www.electronicimaging.org

IS&T
imaging.org