GLOBAL QUALITY OF ASSESSMENT AND OPTIMIZATION FOR THE BACKWARD-COMPATIBLE STEREOSCOPIC DISPLAY SYSTEM

Yuanchun Chen[†], *Guangtao Zhai*[†], *Jiantao Zhou*[‡], *Zhaolin Wan*[♯], *and Lu Tang*[§]

[†]Institue of Image Commu. and Infor. Proce., Shanghai Jiao Tong University, Shanghai, China
[‡]Department of Computer and Information Science, University of Macau, China
[♯]School of Computer Science and Technology, Harbin Institute of Technology, China
[§]School of Medical Imaging, Xuzhou Medical University, Xuzhou, China

ABSTRACT

The backward-compatible stereoscopic display is a technology that a stereoscopic view is perceived with 3D glasses while a 2D version of the 3D image is concurrently available for naked-eye viewers on the same physical display medium. This unique functionality is achieved by an information display technology Temporal Psychovisual Modulation (TPVM), an interesting interplay between high refresh rate optoelectronic display, signal processing and psychophysics. However, the current performance of the system is not satisfactory, and it is a trade-off keeping simultaneously the best performance of the 3D view and the 2D view. The global quality of 3D scene and 2D scene plays a great important impact on user's quality-of-experience. In this paper, we are the first to put forward the concept of global quality, including two components: the quality of 3D view and 2D view. Then we have constructed the display system prototype and first proposed quality assessment criteria to evaluate the quality of both 3D scene and 2D scene, towards as guidance to improve their performances. We conduct subjective experiments to figure out when the system runs the best under the criteria. Experimental results demonstrate that we can effectively calculate the optimal quality of this system using the quality assessment criteria.

***Index Terms*—** Temporal psychovisual modulation, 3-dimensional, display technique, quality assessment criteria, backward-compatible stereoscopic display, global quality.

1. INTRODUCTION

With the rapid development of information display technology, the emerging three-dimensional (3D) stereoscopic displays have enriched and benefited our daily life. Especially in recent years stereoscopic displays have earned mass market acceptance with applications in movie industry, entertainment and engineering visualization. In general, stereoscopic displays are usually divided into two types: the ones requiring the use of 3D glasses [1] and the others free of glasses [2]. The naked-eye stereoscopic displays are also called autostereoscopic displays, and it can be divided into two subcategories namely those coupled with passive polarization glasses and those coupled with active shutter glasses. Currently the most popular 3D display paradigm shows the stereoscopic image pairs for the viewers' left and right eyes on the same display medium, whose working mechanism is temporal multiplexing. In other words, the stereo frames in the time-sequential stereoscopic displays are perceived by human eyes using a high refresh rate display and a pair of synchronized 3D shutter glasses. The 3D shutter glasses can pass images to the correct eye when synchronized with the display at 120Hz refresh rate. Viewers who are not wearing the shutter glasses will see both images superimposed, creating a "ghosted" double view where two copies of objects appear overlayed. At the same time, the autostereoscopic displays implemented by the spatial multiplexing suffers from the reduced spatial resolution. However, compared with the autostereoscopic displays, the time-sequential stereoscopic display has much less restrictive of viewing angles and distances [3] and lower inter-channel crosstalk since the left-eye and right-eye signals are separated.

It is not always desirable to require all the viewers to wear stereo glasses when they view 3D videos. For some viewers long duration of 3D viewing may cause visual fatigue or other side effects [4]. It would be preferable to allow those not wearing stereo glasses to see a single and unghosted image. An early attempt to address this problem is [5]: Didyk et al. referred to the problem as "backward compatible stereo", where a clean 2D image is derived by reducing the disparity between objects in the left and right images to a minimal threshold at the cost of a quality degradation of the 3D view. In this way, the 2D composite image which has smaller disparities is more acceptable to viewers nor wearing stereo glasses, but a ghost image remains. Recently the novel backward-compatible stereoscopic display technology based on TPVM proposed by Wu and Zhai [6] [7] has made it a reality, which aims to exhibit a 3D image and a 2D version of the 3D image simultaneously on the same physical display medium. Moreover, viewers with the 3D active shutter glasses can perceive the 3D image and with naked-eye concurrently see a clear and unghosted 2D image instead of a double and ghosting image. It provides the compatibility with the traditional 3D video signal as well as an optional 2D choice for viewers who feel uncomfortable with 3D experiences.

However, the most fatal flaw of the system is that the 2D view suffers from ghosting artifacts because of the simple blending of the time-sequential displayed frames. What is more, as the left-eye and right-eye images are of equal brightness in this system, it results in low contrast for naked-eye viewers. In order to improve the performance of 2D view, the system has allowed left-eye image more brightness than right-eye image, leading to reducing the 3D depth perception to some extent. Therefore, it is a trade-off between the quality of 2D view and 3D view. Since the performance of 3D scene and 2D scene has a great important impact on user's quality-of-experience, we aim to establish an assessment criterion to evaluate the quality of backward-compatible stereoscopic display system and for optimization like [8, 9].

The rest of this paper is organized as follows: In Section 2, we design a prototype for the backward-compatible stereoscopic display system. Section 3 describes the through user study where we investigate the performance of 3D view and 2D view. The proposed quality assessment criteria is elaborated in the Section 4. In Section 5, we present the associated results and analyses on experiments. At last, the conclusion is given in Section 6.

2. THE DISPLAY SYSTEM PROTOTYPE

In the backward-compatible stereoscopic display system, we first give some definitions here. Y^L and Y^R denote the left-eye view and right-eye view of a 3D scene. Meanwhile let Y_l, Y_r and Y_i be vectors of image pixels, which is containing all the possible brightness. And let the functions MAX($\cdot$) find the maximum brightness in the vector. We denote the maximum possible brightness of Y_L as the max_L = *MAX*(Y_L), and similarly define max_R. Let α_R equal to $max_R/max_L \leq 100\%$, which means the brightness of the image Y_R relative to Y_L.

As shown in Fig. 1, the prototype for the backward-compatible stereoscopic display system is made up of three 3D (120Hz) DLP projectors of the same type, a pair of active shutter glasses, two polarizers, an IR emitter and a polarization-preserving screen. In our system, the NVIDIA® 3D Vision™ [10] on a single graphic card have been adopted to sync and drive three displays as a "single large surface" with the resolution of 3840×720. To realize the "single large surface" , we have adopted the NVIDIA GeForce GTX 690 graphics card in our system, which can provide three synchronical DVI signals at any time. When the graphics card connects the three projectors, we can adjust the position of the displayed images projected by the three projectors to make them overlap together completely.

In our concurrent 3D/2D display system, to generate the 3D view for the viewers wearing glasses, the first unpolarized 3D projector synced to LCD active shutter glasses is used to display the left- and right-eye images at 120Hz refresh rate. The second projector displays the third image Y_I, the inverse of the right-eye image and is linearly polarized, to cancel out the left-eye image. The LCD glasses contain an orthogonal linear polarizer that blocks the image Y_I from the second projector. The third projector is applied as a brightness compensation to improve the contrast ratio and quality of the synthesized 2D scene. In this concurrent 3D/2D display system, on one hand, when the viewers view the stereo images with the 3D glasses, the perceived 3D visual depth varies with the relative brightness between Y_R and Y_L,that is, α_R. On the other hand, when the viewers view the 2D scene not wearing glasses, the quality of the composite 2D view also varies with the α_R because it would lead to the degradation of the contrast and brightness. Generally, the smaller of the α_R, the better of the 2D view's quality.

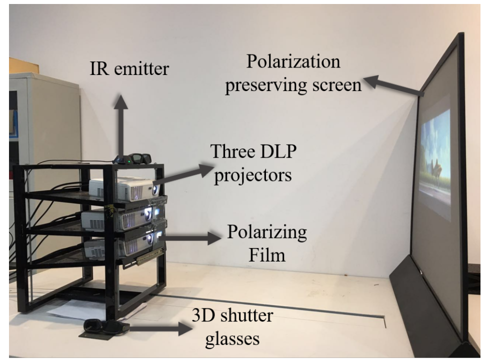

Fig. 1. Backward-compatible stereoscopic display system.

3. STUDY OF USER'S EXPERIENCE

Compared with the traditional 3D display system, the concurrent 3D/2D display system provides an additional clean and unghosted 2D view, and meanwhile the brightness disparity of left-eye and right-eye images reduces the 3D visual perception. What is more, since only the α_R is adjustable in this system, the quality of the 3D and 2D views increases or decreases depending on the choice of α_R. In order to estimate our system's viability, we have conducted subjective quality assessment experiments, where 16 inexperienced subjects were invited to see the test images at a viewing distance of 1.5 meter to the screen [11].

3.1. User Study on 2D View

The first experiment aims to investigate at what level of contrast the viewers would perceive the ghosting artifacts in the new synthesized 2D image and provide their subjective scores with the help of 10-category discrete scale, which can make the subjective values more distinguishable. In this experimen-

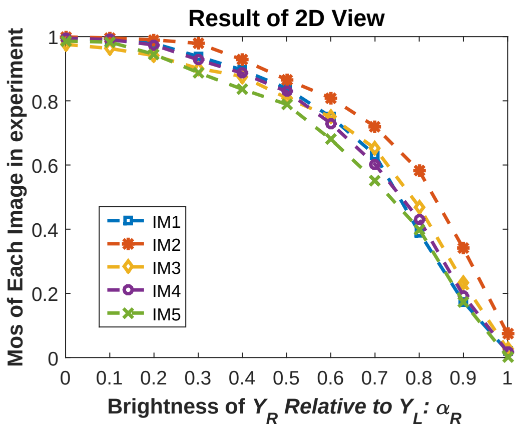

Fig. 2. In the system, the relationship between MOS and α_R in each kind of image are shown for naked-eye viewers in this figure. Each line in this figure stand for a kind of stereo image, which including 110 images with different α_R.

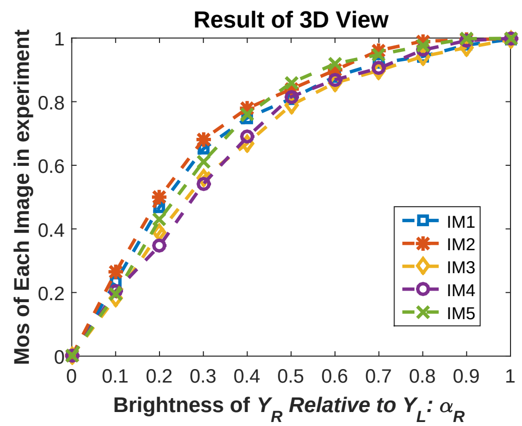

Fig. 3. The relationship between MOS and α_R in each kind of image are shown in this figure with 3D shutter glasses view in the backward-compatible stereoscopic display system.

t, the α_R, the brightness of the image Y_R relative to Y_L, was allowed to be increased progressively from 0% to 100% by 10 equilong degrees. The 15 original 3D images in the side by side format with different image complexity are chosen as the test materials, which each stereo image generates 11 images with 11 different values of α_R. Subjects were instructed to respond and mark only when they felt their eyes had perceived the interfering ghosting, typically answering immediately for α_R near 0% and waiting about a minute for α_R near 100%. As the partial adaptation to new light levels occurs immediately, with additional adaptation over approximately 20 minutes [12]. We presented subjects with 165 3D images totally at random in the concurrent 3D/2D display system. Each subject viewed 15 test images at eleven values of α_R (totally 165 images) in sequence, with n = 16 subjects participating in each experiment. By observing the Fig. 2, experimental results show at low values viewers uniformly perceive the nearly best quality of the synthesized 2D image without any ghosting. Only when the values of α_R at very high up to 0.7 do viewers gradually find the objectionable ghosting. Meanwhile the contrast ratio of the composed 2D image go through a severe reduction when the range of α_R is from 0.7 to 1.

3.2. User Study on 3D view

In our concurrent 3D/2D display system, we display a brighter image to 3D viewers' left eyes than to their right eyes. However, stereoscopic vision is degraded when the images seen by each eye become dissimilar [13]. Small brightness differences may be imperceptible, but an all-black right-eye image obviously precludes stereoscopic vision. This experiment is to quantify roughly depth perception when the α_R varies with the range from 100% to 0%. As similar to the previous experiment, the test images were displayed in sequence according to the value of α_R from the 100% to 0%. All the subjects need to wear a pair of 120Hz active shutter glasses while testing. The subjects were asked to respectively mark each image based on the depth of 3D vision they perceived. As shown in this Fig. 3, we can find the curve's variation trend of the 3D view is just opposite to that of the 2D view. For each group images, the quality of the 3D vision becomes better with the increasing of the α_R from 0 to 1.

4. QUALITY ASSESSMENT CRITERIA

By observing the results of above subjective experiments, it is surprising to find that these curvilinear shapes of 2D view and 3D view are both similar to the curve of sigmoid function. In order to get the unified law of this display system in mathematic, we can fit these experimental data with a standard sigmoid function.

We give some definitions here. The Q_{3d}, Q_{2d} and Q_{glo} denote the quality of 3D view, the quality of 2D view and the global quality of the display system respectively. Therefore Q_{glo} contains two components: Q_{3d} and Q_{2d}. We would calculate Q_{3d} and Q_{2d} respectively to get Q_{glo}.

4.1. Formulation of the Display System

In the system, besides the α_R plays the great impact on the performance of the 3D view and 2D view, the number of 3D viewers and 2D viewers also makes a difference to the global performance of the 3D/2D sharing viewing system. We let M and N be the number of 3D viewers and 2D viewers. It should be noting that the weight of Q_{3d} and Q_{2d} are $M/(M+N)$ and $N/(M+N)$ respectively. Since the shape and variation trend of Q_{2d} and Q_{3d} are like the sigmoid function curve, we can let Q_{2d} equal to $2(\frac{1}{1+e^{-k\alpha_R}} - 0.5)$ and Q_{3d} equal to $-2(\frac{1}{1+e^{-p(\alpha_R-1)}} - 0.5)$. Then it can be formulated as the following equations:

$$Q_{glo} = \frac{M}{M+N} Q_{3d} + \frac{N}{M+N} Q_{2d} \qquad (1)$$
$$= \frac{2M}{M+N}\left(\frac{1}{1+e^{-k\alpha_R}} - 0.5\right) + \frac{-2N}{M+N}\left(\frac{1}{1+e^{-p(\alpha_R - 1)}} - 0.5\right)$$

where k and p are the parameters to be confirmed according to the fitting results, M and N are respectively the number of the 3D viewers and the 2D viewers. By observing the formula, we can find the quality of the 3D and 2D views increases or decreases only depending on the choice of α_R. To get the optimal performance of this system, we can calculate the solution of equation (1). First we have proved the equation (1) does not exist analytical solutions, the detailed verification process is not mentioned here for the limited space. Here we will estimate the α_R by simulation in the following sections.

4.2. Parameter Estimation and Calculating the α_R

Given the above quality assessment criteria model, we can estimate the parameters k and p based on the subjective experimental data. To estimate parameters k and p in the equations Q_{2d} and Q_{3d}, we need to solve the following problem:

$$\min_{k} \sum_{i=1}^{N} \|\mathbf{S_i} - \mathbf{Q_{2d}}\|_F^2 \qquad (2)$$

$$\min_{p} \sum_{i=1}^{N} \|\mathbf{S_i} - \mathbf{Q_{3d}}\|_F^2 \qquad (3)$$

where the S_i is the Mean of Scores (MOS) of the subjective experiment. i is the number of the subjects participating in the test and here N equals to 16. In the equation (2), there is only one parameter k to be confirmed since the S_i has already been known. Therefore we can easily to estimate the range of k which is from 4.405 to 5.025 with 95% confidence bounds. As similar to the equation (2), the parameter p also can be confirmed to the range from 4.626 to 5.125 with 95% confidence bounds in equation (3). Then we can substitute the concrete values of k and p into the equation (1), and get the global quality of the display system Q_{glo}:

$$Q_{glo} = \frac{2M}{M+N}\left(\frac{1}{1+e^{-4.715\alpha_R}} - 0.5\right) + \frac{-2N}{M+N}\left(\frac{1}{1+e^{-4.875(\alpha_R - 1)}} - 0.5\right) \qquad (4)$$

5. EXPERIMENTS AND DISCUSSIONS

In this session, we will make the simulation experiment to verify our proposed quality assessment criteria. Without loss

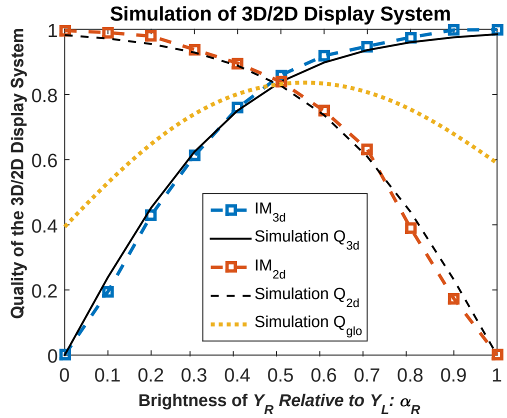

Fig. 4. The simulation of Q_{3D}, Q_{2D} and Q_{glo} in the system.

of generality, we just take one kind of stereo image as an example. Since we have got the MOS of each image in the subjective test, we first fit the MOS data with the above equations Q_{3d} and Q_{2d}. Then we give a concrete value to the number of the 2D viewers and 3D viewers. Here we let M and N be the 60 and 40 (We make the weight of M bigger than N.). Therefore we can plot the Q_{glo} in the above figure window. As shown in Fig. 4, IM_{3d} and IM_{2d} describe the relationship between the MOS and different α_R in 3D view with 3D shutter glasses and in 2D view without glasses. Simulation Q_{3D} and Q_{2D} are respectively the fitted curves of the IM_{3d} and IM_{2d}. Since the Q_{glo} consists of Q_{3D} and Q_{2D} with different weights, we can achieve the optimal global quality of this system where the two curves of Q_{3D} and Q_{2D} intersect. Here when the α_R reaches about 0.55, the global quality of the display system achieves the optimum. When substituting all the known parameters into the solution of equation (4), we can get the accurate α_R to optimize the system.

6. CONCLUSION

This paper has constructed a prototype for the backward-compatible stereoscopic display system. We are the first to propose the concept of global quality aiming at this system in the quality assessment field. The quality assessment criteria is proposed to evaluate the global quality of both 3D scene and 2D scene, towards as guidance to improve their performances. Sufficient study on user's quality-of-experience has been done to assess the performance of the system. Experimental results confirm that we can effectively calculate the optimal quality of this system with the quality assessment criteria.

7. ACKNOWLEDGE

This work was supported by the National Science Foundation of China (61422112, 61371146, 61521062, 61527804), National High-tech R&D Program of China (2015AA015905), and Science and Technology Commission of Shanghai Municipality (15DZ0500200).

8. REFERENCES

[1] J. Su, M. Johnson, D. Miller, G. Wernig, and D. Burckhard, “3-d stereo glasses,” Sep. 22 2009, US Patent D600,738.

[2] J. Harrold and G. Woodgate, “Lcd spatial light modulator as electronic parallax barrier,” Jun. 23 2004, Patent 0,833,183.

[3] A. J. Woods, “Crosstalk in stereoscopic displays: a review,” *JEI*, vol. 21, no. 4, pp. 040 902-040 902, 2012.

[4] Lambooij M, IJsselsteijn W A, Heynderickx I. Visual discomfort of 3D TV: Assessment methods and modeling[J]. Displays, 2011, 32(4): 209-218.

[5] P. Didyk, T. Ritschel, E. Eisemann, K. Myszkowski, and H.-P. Seidel, “A perceptual model for disparity,” *ACM TOG*, vol. 30, no. 4, pp. 96:1-96:10, Jul. 2011.

[6] X. Wu and G. Zhai, Backward compatible stereoscopic displays via temporal psychovisual modulation. *ACM SIGGRAPH Asia*, 2012.

[7] X. Wu and G. Zhai, Temporal psychovisual modulation: A new paradigm of information display, *IEEE SPM*, 2013.

[8] K. Gu, D. Tao, J.-F. Qiao, and W. Lin, “Learning a no-reference quality assessment model of enhanced images with big data,” *IEEE Trans. Neural Netw. Learning Syst.*, 2017.

[9] K. Gu, J. Zhou, J.-F. Qiao, G. Zhai, W. Lin, and A. C. Bovik, “No-reference quality assessment of screen content pictures,” *IEEE Trans. Image Process.*, 2017.

[10] N.Corporation, “NVIDIA 3D Vision,” [Online]. Available: *http://www.nvidia.com/object/3d-vision-main.html*, 2016.

[11] K. Gu, M. Liu, G. Zhai, X. Yang, and W. Zhang, “Quality assessment considering viewing distance and image resolution,” *IEEE Trans. Broadcasting*, vol. 61, no. 3, pp. 520-531, Sep. 2015

[12] Standing L G, Dodwell P C, Lang D. Dark adaptation and the Pulfrich effect[J]. Attention, Perception and Psychophysics, 1968, 4(2): 118-120.

[13] Cormack L K, Stevenson S B, Schor C M. Interocular correlation, luminance contrast and cyclopean processing[J]. Vision research, 1991, 31(12): 2195-2207.