Quality Assessment for Dual-View Display System

Yuanchun Chen, Ning Liu, Guangtao Zhai, Ke Gu, Jia Wang, Zhongpai Gao and Yucheng Zhu
Insti. of Image Commu. & Infor. Proce., Shanghai Jiao Tong University, Shanghai, China
Shanghai Key Laboratory of Digital Media Processing and Transmissions
{chenyuanchun, ningliu, zhaiguangtao, gukesjtuee, jiawang, gaozhongpai, zyc420}@sjtu.edu.cn

***Abstract*—Spatial psychovisual modulation (SPVM) is a new information display technology, which aims to generate multiple visual percepts for different viewers on a single display simultaneously. After the proposal of SPVM, lots of efforts have been made and several applications (i.e., dual-view display system) have been implemented based on this technology. The dual-view display (DVD) system is considered as an effective digital image hiding system based on SPVM theory, but little work has been dedicated to the perceptual quality assessment of DVD system. Up to now, there is no clear and standard method to evaluate the performance of the dual-view display system. It is important for the viewers to see a clear and non-aliasing image when they are front of the screen. Therefore, in this paper, we will build a DVD database and carry out a subjective experiment to evaluate the performance of the DVD system, and then we investigate and analyze the performance of prevailing no-reference (NR) image quality metrics on the particular DVD system. We have a sufficient belief that this paper can supply the guideline for the performance on the DVD system and serve as a good testing bed for future research of SPVM technology.**

***Index Terms*—Spatial psychovisual modulation (SPVM), image quality assessment (IQA), dual-view display (DVD), subjective experiment, no-reference (NR)**

Fig. 1: The prototype system of the dual-view display. We denote the view through this polarization glasses or polarizing film as the personal views and the one watched directly from the screen as the shared views.

I. Introduction

As a new display technology, Temporal Psychovisual Modulation (TPVM) was lately proposed by *Wu* and *Zhai* in [1], [2], [3] and [4], which aims to generate multiple percepts for different viewers on the same medium. In TPVM system, the viewers wearing different active liquid crystal (LC) glasses with varying transparency levels can see different images (called *personal views*). The viewers without LC glasses can also see a semantically meaningful image (called *shared views*). The idea of TPVM can be extended to spatial domain, called spatial psychovisual modulation (SPVM). The implementation of dual-view display (DVD) system is based on the SPVM. It is possible to devise a type of dual-view system on a single display using the mismatch between resolutions of modern display devices and the HVS. Modern display devices now support very high pixel density. Meanwhile the human visual system (HVS) cannot distinguish image signals with spatial frequency above a threshold, as predicted by the contrast sensitivity function (CSF). Therefore, a stereoscopic display device can be utilized in DVD system, odd and even scan lines of the device are polarized in different directions. The PPI of this type of screen is usually very high that HVS cannot discriminate the mismatch between the left and right eye image. It is possible for this type of display device to broadcast a pair of images simultaneously on the screen in odd and even lines. We can only see the odd line or even line image when using a pair of glasses with both glasses in the same polarization direction. We denote the view through this polarization glasses as personal views and the one watched directly from the screen as the shared views. The prototype system of the DVD system is shown in Fig. 1. Supposing a pair of images X_1 and X_2 are displayed in the odd and even lines simultaneously. Using a pair of glasses matched to the odd lines, we can see personal views X_1. And with naked eyes, we see shared views $Y = X_1 + X_2$. The mechanism behind the DVD system is shown in Fig. 2

After the emergence of TPVM and SPVM, so far many researches have been conducted and some applications have been implemented. In order to prevent camcorder privacy, a new video projection technology was proposed in [5] and then extended via DLP in [6]. Another application of TPVM is information security, it was present in [7], meanwhile, in [8] information security can also be implemented in spatial domain by SPVM, which is an extension of the idea of TPVM. The technology of SPVM has also been widely employed in some other applications.

In the abovementioned applications, however, no standard and clear image quality assessment (IQA) method is used to evaluate the performance of DVD system. There exists a prob-

978-1-5090-5316-2/16/$31.00 © 2016 IEEE
VCIP 2016, Nov. 27 – 30, 2016, Chengdu, China

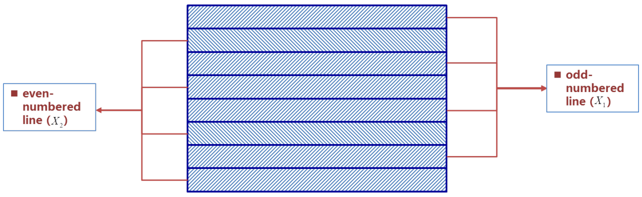


Fig. 2: The mechanism of the dual-view system on a single display.

(a) (b)

Fig. 3: The setup of subjective experiment. (a) is on subjective quality assessment for personal views. (b) performs subjective quality assessment for shared views.

lem of DVD system that it will appear aliasing phenomenon between two kinds of views, leading to a significant impact on the viewers under normal viewing conditions. Without the assessment criteria, we are unable to correctly tell whether the system achieves the optimum visual effect or not. So we first investigate and develop the image quality metrics for dual-view display system by most mainstream IQA algorithms.

Following the research line of IQA tasks, numerous existing approaches have been proposed over the last several decades. IQA can be basically divided into two categories: subjective and objective assessments. The first subjective assessment is often regarded as the ultimate quality criterion, but it is of expensive and time-consuming shortages and furthermore impractical for real-time processing systems. As a consequence, objective metrics have become an intensely research topic during the past years. Based on the availability of the original image, image quality metrics are further classified into three types, namely full-reference (FR) [9], [10], reduced-reference (RR) [11], [12], and no-reference (NR) [13], [14] methods depending on the accessibility to the original reference points [15]. Some of them perform substantially well in predicting the subjective ratings on popular image quality databases [16], [17]. The target of IQA is to design metrics for objective evaluation of quality in a way that is consistent with subjective human sensations. The largest number of objective image quality assessment metrics are FR-IQA algorithms [18], [19], which assume that the original image signal is completely known. Those quality measures are usually of high performance.

For the convenience of the research about the effects of different factors on image quality, we build the DVD image database. In the DVD database, we choose eleven images as the original image in the KODAK database. Then we can save the images on the display when the system is running under the same conditions. This database consists of 420 images generated from eleven pristine images in DVD system. It will be divided into two groups, i.e., the personal views and the shared views. Each group contains 210 images, which make us convenient to analyze the performance of a variety of DVD quality metrics.

The remainder of this paper is organized as follows: Section II details the process of constructing the image database and subjective experiment. Section III gives the objective quality metric for the DVD database by different NR algorithms and the associated results and analyses on experiments. At last, we conclude this paper in Section IV.

II. Subjective Quality Assessment

To facilitate our research on the quality assessment, we first establish a special image database which contains two groups of images with different dynamic range and the subjective quality evaluations in the form of MOS.

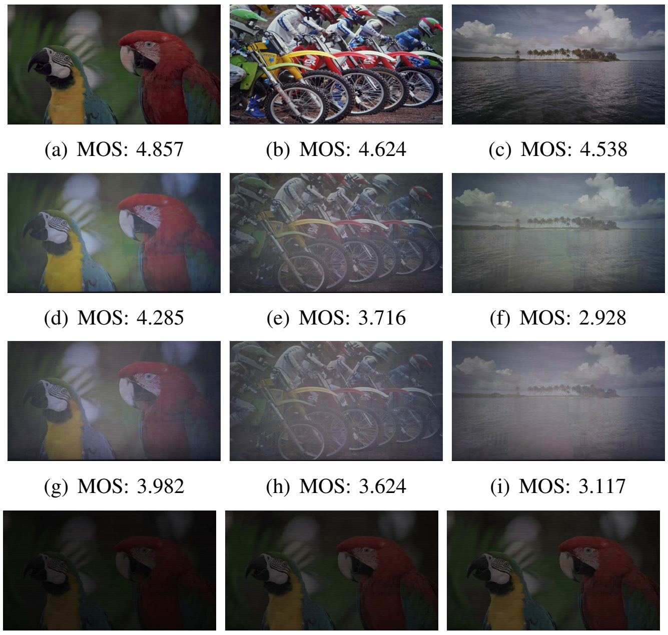
A. Database Construction

To make the research convenient, we construct a prototype system, which are composed of polarized glasses and a interlaced polarization 3D LCD display. When the DVD system is running, people can see the shared views without glasses and the personal views with glasses. The subjective experiment setup is shown in Fig. 3. In order to evaluate the performance of DVD system, we construct a new DVD database with two typical factors, that is, image complexity and image dynamic range. It is worth mentioning that both above two factors substantially affect human perceptions of image/video quality.

(1) Image dynamic range: The image dynamic range will make a significant impact on perceived image quality in DVD system. At the same time, it is a trade-off coefficient to balance the visual effect between the personal views and the shared views. If we set the coefficient of the image dynamic range equal to 0.5, the dynamic range of the personal views is from 128 to 255, and the dynamic range of the shared views is from 0 to 127. It means that, we make the shared views turn whiter and the personal views darker than the original images.

(2) Image complexity: The complexity of an image tells many aspects of the image content and is an important factor in the selection of source material for the DVD database. The higher of the image complexity, the more spatial information in this image. The image complexity will make a remarkable effect on the performance of the DVD system.

Image materials of our database includes 420 images which can be divided into 2 groups, group one represents the shared views, which means people see the images without the polarized glasses. The other one shows the personal views, which stems from the personal views by simulating the effects of polarized glasses. Scenario contains image with seven dynamic range degrees ranging from the lightest to the worst. Fig. 4 gives some example images in the database. As can be clearly seen, the looming contour of the image caused by different degrees of image dynamic range and image



(a) MOS: 4.857 (b) MOS: 4.624 (c) MOS: 4.538

(d) MOS: 4.285 (e) MOS: 3.716 (f) MOS: 2.928

(g) MOS: 3.982 (h) MOS: 3.624 (i) MOS: 3.117

(j) MOS: 2.931 (k) MOS: 3.514 (l) MOS: 4.235

Fig. 4: Twelve images with resolutions of 1920 $\times$ 1080 used in the DVD database. First row, the shared views with the full dynamic range in DVD system. Second row and third row, the shared views with different dynamic range degrees. Fourth row, the personal views with different dynamic range degrees in DVD system. Three columns are of different image complexity except the fourth row.

complexity distributes irregularly or unequally in the images, e.g. Fig. 4 (d),(e),(f),(g),(h) and (i). It should be pointed out that all the images in DVD database are generated when DVD system is running and the image resolution is 1920$\times$1080.

B. Subjective Experiment

Next, we performed subjective experiment on our image database to gather subjective quality evaluations. Single-stimulus (SS) method was applied in our experiment. Twenty-five inexperienced subjects were invited for the assessment task. For easy operation, we designed an interactive system to automatically display the test images and collect the subjective quality scores using graphical user interface (GUI) in MATLAB, similar to that used in [20]. Most of viewers were college students with various kinds of majors to display the images and collect subjective ratings. The subjective experiment includes two stages: the training stage and the rating stage. In the training stage, subjects previewed some example images in DVD database. The training stage would prevent subjects from rating arbitrarily. Moreover, training images will not appear in the rating procedure. The experimental environment was arranged according to the recommendations specified by ITU-R BT.500-12 [21]. To be illustrated, the illuminance of the testing room keeps low. The viewing distance is fixed at four times the image height.

The subjects were asked to provide their overall perception of quality on a continuous scale from 0 to 5. The presentation order of the images was randomized for each subject. After

TABLE I: Exprimental results on the DVD database.

Methods	PLCC	SRCC	KRCC
S_3 [22]	0.1946	0.1648	0.1071
NFERM [14]	0.3007	0.3070	0.2121
NIQE [23]	0.5935	0.4841	0.3467
FISBLIM [24]	**0.7260**	**0.6808**	**0.5096**
SISBLIM [25]	0.4046	0.3392	0.2377
FISH [26]	0.3271	0.3180	0.2188
ARISM [27]	0.3325	0.2873	0.2088

subjective experiment, we collected the opinion scores from all the subjects and excluded those unreliable scores in the light of the guidelines in ITU-R BT.500-12 [21]. One of the twenty-five subjects was removed and the remaining scores were averaged to obtain the final mean opinion scores (MOSs).

III. Experimental Results and Discussions

In this section, we will investigate the performance of existing mainstream NR-IQA objective methods to evaluate the visual quality of the images in the DVD database. The DVD database is used as testing bed for performance evaluation and comparison. The images in the DVD database do not include so-called reference images since they were taken by real cameras, which makes it impossible to use FR- and RR-IQA methods. Those testing IQA metrics are summarized as follows. The first group includes six general-purpose methods, S_3 [22], NFERM [14], NIQE [23], FISBLIM [24] and SISBLIM [25]. The second group are composed of four distortion-specific methods, FISH [26] and ARISM [27].

When calculating performance, we firstly mapped the results of the objective quality metric to subjective ratings through nonlinear regression of a five-parameter logistic function as suggested by VQEG [28]. Then, we employed five commonly used metrics which are Pearson Linear Correlation Coefficient (PLCC), Spearman Rank ordered Correlation Coefficient (SRCC), Kendall Rank Correlation Coefficient (KRCC), Average absolute prediction error (AAE) and Root Mean-Squared Error (RMSE) respectively. Among them, PLCC evaluates IQA method's prediction accuracy, SRCC evaluates the prediction monotonicity, KRCC is another metric used to evaluate the prediction monotonicity, AAE predict the average absolute error and RMSE points out the prediction consistency. A good quality measure is expected to achieve high values in PLCC and SRCC, while low values in AAE and RMSE. We list the performance results of objective quality metrics in Table I.

By observing Table I, a comparison of these state-of-the-art general-purpose NR-IQA methods and popular blind measures is conducted on DVD database. We can get that the general-purpose IQA methods can assess the quality of the DVD images due to their general QA ability for distorted images. While most image sharpness assessment methods perform better than the general-purpose IQA methods to the DVD images. Surprisingly, it is found that the BRSIQUE cannot discriminate the images' differences while the FISBLIM performs the best among all the computing IQA models. Also, the experiment results have proved that the two factors play a remarkable impact on the performance of the DVD system. As shown in Fig. 4, the trade-off coefficient in second row is set to 0.2, with the image complexity vary from the first column to the third column, we can find that the performance of the DVD image turns worse. The image have appeared the aliasing phenomenon. While the coefficient is set to 0.4 in the third row, the shared views turn whiter than the second row. At the same time, the images in fourth row do not appear any artificial aliasing. They all turn darker as the trade-off coefficient of dynamic range varies. Besides, we take the three columns for the comparison together, we will find that the image of higher dynamic range gets a better visual effects in the DVD system when the trade-off coefficient is set to a relatively low value, and the image with higher image complexity is more suitable for placing in the shared views instead of the personal views. However, image sharpness assessment is not enough for quality assessment. In the future work, we will focus on proposing a specific objective quality metric for the DVD database based on the reasonable modification of the good-performing FISBLIM.

IV. Conclusion

In this paper, we have investigated into the problem of quality assessment of dual-view system on a single display, which caused by two factors, image complexity and image dynamic range. Firstly, we introduced a specific DVD database to facilitate this research. The image database was constructed through subjective experiment. Afterwards, we conducted more than ten NR-IQA objective assessment methods to evaluate the performance of the DVD system. Experimental results demonstrated that the state-of-the-art blind IQA methods and image sharpness assessment methods for the DVD database have some discrepancy with subjective evaluations.

V. Acknowledgment

This work was supported in part by the National Science Foundation of China under Grants 61422112, 61371146, 61521062, 61331014 and National High-tech R&D Program of China under Grant 2015AA01590.

References

[1] X. Wu and G. Zhai, Image/information display system and method based on temporal psycho-visual modulation, Jun. 21 2012, US Patent App. 14/128,509.

[2] X. Wu and G. Zhai, Temporal psychovisual modulation: A new paradigm of information display [Exploratory DSP], Signal Processing Magazine, IEEE, vol. 30, no. 1, pp. 136141, 2013.

[3] G. Zhai and X. Wu, Multiuser collaborative viewport via temporal psychovisual modulation [Applications Corner], Signal Processing Magazine, IEEE, vol. 31, no. 5, pp. 144149, 2014.

[4] Gao Z, Zhai G, Zhou J. Factorization Algorithms for Temporal Psychovisual Modulation Display[J]. IEEE Transactions on Multimedia, 2016, 18(4): 614-626.

[5] X. Wu and G. Zhai, Backward compatible stereoscopic displays via temporal psychovisual modulation, in SIGGRAPH Asia 2012 Emerging Technologies. ACM, 2012, p. 4.

[6] G. Zhai and X. Wu, Defeating camcorder piracy by temporal psychovisual modulation, Journal of Display Technology, vol. 10, no. 9, pp. 754757, 2014.

[7] Z. Gao, G. Zhai, X. Wu, X. Min, and C. Zhi, DLP based antipiracy display system, in Visual Communications and Image Processing Conference, 2014 IEEE. IEEE, 2014, pp. 145148.

[8] Z. Gao, G. Zhai, and X. Min, Information security display system based on temporal psychovisual modulation, in Circuits and Systems (ISCAS), 2014 IEEE International Symposium on. IEEE, 2014, pp. 449452.

[9] K. Gu, S. Wang, G. Zhai, S. Ma, X. Yang, and W. Zhang, "Content-weighted mean-squared error for quality assessment of compressed images," *Signal, Image and Video Processing*, vol. 10, no. 5, pp. 803-810, Jul. 2016.

[10] K. Gu, S. Wang, H. Yang, W. Lin, G. Zhai, X. Yang, and W. Zhang, "Saliency-guided quality assessment of screen content images," *IEEE Trans. Multimedia*, vol. 18, no. 6, pp. 1-13, Jun. 2016.

[11] K. Gu, G. Zhai, X. Yang, and W. Zhang, "A new reduced-reference image quality assessment using structural degradation model," *in Proc. IEEE Int. Symp. Circuits and Syst.*, pp. 1095-1098, May 2013.

[12] M. Liu, K. Gu, G. Zhai, P. Le Callet, and W. Zhang, "Perceptual reduced-reference visual quality assessment for contrast alteration," *IEEE Trans. Broadcasting*, 2016, to appear.

[13] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Trans. Image Process.*, vol. 21, no. 12, pp. 4695-4708, Dec. 2012.

[14] K. Gu, G. Zhai, X. Yang, and W. Zhang, "Using free energy principle for blind image quality assessment," *IEEE Trans. Multimedia*, vol. 17, no. 1, pp. 50-63, Jan. 2015.

[15] W. Lin and C.-C. J. Kuo, "Perceptual visual quality metrics: A survey," *Journal of Visual Communication and Image Representation*, vol. 22, no. 4, pp. 297–312, 2011.

[16] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik, "LIVE image quality assessment Database Release 2," [Online]. Available: http://live.ece.utexas.edu/research/quality

[17] 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.

[18] A. Liu, W. Lin, and M. Narwaria, "Image quality assessment based on gradient similarity,"?*Image Processing, IEEE Transactions on*, vol. 21, no. 4, pp. 1500–1512, 2012.

[19] K. Gu, S. Wang, G. Zhai, W. Lin, X. Yang, and W. Zhang, "Analysis of distortion distribution for pooling in image quality prediction," *IEEE Trans. Broadcasting*, vol. 62, no. 2, pp. 446-456, Jun. 2016.

[20] K. Gu, G. Zhai, W. Lin, and M. Liu, "The analysis of image contrast: From quality assessment to automatic enhancement," *IEEE Trans. Cybernetics*, vol. 46, no. 1, pp. 284-297, Jan. 2016.

[21] I. Rec, "Bt. 500-11, methodology for the subjective assessment of the quality of television pictures," International Telecommunication Union, 2002

[22] C. T. Vu, T. D. Phan, and D. M. Chandler, S3: A spectral and spatial measure of local perceived sharpness in natural images, *IEEE Trans. Image Process., vol.* 21, no. 3, pp. 934-945, May. 2012.

[23] A. Mittal, R. Soundararajan, and A. C. Bovik, Making a completely blind image quality analyzer, *Signal Processing Letters, IEEE, vol.* 20, no. 3, pp. 209-212, 2013.

[24] K. Gu, G. Zhai, M. Liu, X. Yang, W. Zhang, X. Sun, W. Chen, and Y. Zuo, "FISBLIM: A five-step blind metric for quality assessment of multiply distorted images," *in Proc. IEEE Workshop on Signal Processing Systems*, pp. 241-246, Oct. 2013.

[25] K. Gu, G. Zhai, X. Yang, and W. Zhang, "Hybrid no-reference quality metric for singly and multiply distorted images," *IEEE Trans. Broadcasting*, vol. 60, no. 3, pp. 555-567, Sept. 2014.

[26] P. V. Vu and D. M. Chandler, A fast wavelet-based algorithm for global and local image sharpness estimation, *Signal Processing Letters, IEEE, vol.* 19, no. 7, pp. 423-426, 2012.

[27] K. Gu, G. Zhai, W. Lin, X. Yang, and W. Zhang, "No-reference image sharpness assessment in autoregressive parameter space," *IEEE Trans. Image Process.*, vol. 24, no. 10, pp. 3218-3231, Oct. 2015.

[28] VQEG, "Final report from the video quality experts group on the validation of objective models of video quality assessment," Mar. 2000, http://www.vqeg.org/.