Objective No-reference Stereo Video Quality Assessment Research

Yan Li[1] **Yu Yang**[2]

Abstract. With the development of the 3d industry, 3d video quality assessment method has become one of the hot spots. But now the main research directions are concentrated in reference stereoscopic video evaluation method and subjective evaluation method, which also have made some progress. Objective no-reference stereoscopic video evaluation method based on its extensive applicability has been getting more and more attention. In this paper, we introduce the three-dimensional video evaluation technology research status, and then refer to 3d video quality measurement and a common evaluation model Finally, we do a brief overview about its developing tendency.

Keywords: 3d; quality; assessment; developing

1.1 Introduction

Nowadays, three-dimensional (3D) stereo media is becoming immersive media to increase visual experience as natural in various applications ranging from entertainment [1] to more specialized applications such as remote education, industrial, medical, robot navigation, and many other fields. But the birth of a 3d video is typically through filming, production processing, compression coding, signal transmission, decoding and stereo display .Every link of processing may cause damage of the video quality, giving access to information or video post-processing difficulties. Although subjective assessment is the most accurate method for perceived image quality, it is time consuming, and expensive. Therefore, there need to be a more perfect and standard stereo video evaluation method to promote the development of 3d industry.

1.2 The Research Status

Although, several conventional 2D objective quality assessment methods [2] have been proposed based on peak signal-to-noise ratio (PSNR), mean squared error (MSE), or human visual system (HVS) characteristics in the last two decades for images/videos. However, still now no comparable effort has been devoted to the quality assessment for 3D/stereoscopic images. 3d video evaluation started late, the first method is to extend the 2d image evaluation method to 3d image evaluation. In 2008, a method based on improved PSNR of the stereo image quality evaluation was proposed [3]. Due to the different signals for the visual impact of the human visual system is different, so the method is inconsistent with the human eye vision system..In [4], the quality of depth map is used as a 3d video quality measurement.. A full reference (FR) quality metric for the assessment of stereoscopic image pairs using the fusion of 2D quality metrics and of the depth information is proposed in[5].This method is not fully considering human visual characteristics and does not have a wide range of applicability. With the deepening of the research on human visual system, some of the human visual characteristics is introduced into the evaluation method. In [6], the quality of 3D videos stored as monoscopic color videos that augmented by pixel depth map and finally this pixel information used for color coding and depth data. In recent years Tianjin University, Ningbo University and other schools have puts forward some reference

---

[1] Yan Li(18811557875)
Communication University of China ,Beijig,Chiana
e-mail: liyanchuanmei0124@126.com

[2] YuYang (13641221931)
Communication University of China ,Beijig,Chiana
e-mail: young_rain@bbi.edu.cn

stereoscopic video evaluation methods. In [7] a three-dimensional objective evaluation model combined of human visual perception characteristics of visual sensitivity, sensitivity function and direction in the wavelet domain is proposed. The model respectively from the left and right viewpoint and the depth perception evaluation analysis, concluded that the overall image quality evaluation and the subjective sense of consistency, but respectively in the left and right viewpoint and depth perception evaluation model is not very good. The model based on the viewpoint about image quality and depth perception quality, strongly depends on the statistical characteristic values, such as edge, texture, flat, etc. Parallax is obtained by a range of search and parallax matching need a lot of time. The model parameters are based on the statistical characteristics, so the application range of the model was limited. In [8], a no reference stereo video evaluation model is proposed. The method is based on image segmentation, through calculating the extent of the blockiness, and then combines disparity estimation to the assessment of the quality of video. The results show that the model performs quite well over wide range of stereo image content and distortion levels. But this method is mainly for JPEG image coding, application scope is limited. There are, therefore, to establish a set of widely applied three-dimensional image/video quality objective evaluation model is of great practical significance.

1.3 Proposed Objective Method

Objective quality evaluation method is that can be established from video quality and Stereo image comfort.

1.3.1 Video quality

The video quality evaluation method mainly is the level of distortion in the image. At present, in the process of video decoding, it generally adopts the peak signal-to-noise ratio (peak signal noise ratio, PSNR) and mean square error (mean square error, MSE) measure the distortion degree of video sequences. In[10], a method based on PSDR is proposed for stereoscopic color images. They use average PSDR of the left and right images as the basis of image quality assessment. Author thinks that people use a single eye to see flat images and stereoscopic images is no difference. The study is evaluated that the FR metric of 2D quality assessment can be used for an extension to 3D.Though the PSDR is a simple and widely used image quality objective evaluation indicators but the relevance to visual quality is very poor.

Although, the impact of coding distortions on the perceived stereoscopic image quality of an asymmetric image pair depends on the visual appearance of the artifact, where blockiness appears to be much more disturbing than other factors [12].That is to say, the extent of blockiness can be used as a main evaluation indicator of image distortion. In [7], firstly a differencing signal along each horizontal line is calculated. They estimate blockiness in spatial domain based on segmented local features. Then, combined with zero-crossing estimates the fuzzy degree of the image.

1.3.2 3D depth perception

3D depth perception has always been one of the main factors of image quality assessment and is a complex process. We believe it has a strong correlation with objects/structural information of a scene content that is near to the viewers. In [11] the next generation of 3d content production environment has carried on the investigation and analysis. They discuss parallax effects on stereo feeling by five groups' subjective experiments. They concluded Parallax is the key factors of stereoscopic image have stereo feeling and will directly affect the quantity of the three-dimensional images. In [12], the quality of 3D videos stored as monoscopic color videos that augmented by pixel depth map and finally this pixel information used for color coding and depth data. The relationship between the perceived overall image quality and the perceived depth are discussed in [13].

Human is the final receiver of stereoscopic images and judge the comfort of stereoscopic images. Comfort of stereo image effect will eventually perceive through people's eyes. The primary function of the human visual system (HVS) is to extract structural or edge information from the viewing field [14]. Therefore, Human visual perception is very sensitive to edge detection, and consequently, perceive distortions should be strongly dependent on local features such as edge, and non edge.

In [15], the factors that affect the comfort of binocular stereo images are studied by experiments, such as brightness colour, contrast, crosstalk, parallax ,etc. Experimental results show that the two view as long as the stereo image does not match the degree above a certain threshold, the visual comfort will seriously reduce. Therefore, we can be according to around image matching degree to evaluation of image quality.

Stereo vision can be divided into two parts: physical stereo vision and mental stereo vision. Theory and experiment have proved that the physiological characteristic of stereo vision has made greatest contribution to the stereoscopic images, meanwhile it is

also the main means by which people have deep feelings, reflecting the depth of the objective things. The body's physiological factors have important impacts on people's stereo visual effects, among these factors which impact on the stereoscopic vision, parallax between two eyes is the most important one, so eye-parallax is also considered as the main factor of binocular stereo imaging. Paper[16] points out that through the experiment study the binocular parallax increases to more than a certain threshold, it will significantly reduce the stereoscopic image visual comfort .Excessive parallax will reduce the three-dimensional visual comfort is discussed in[17].So the stereoscopic image evaluation method based on disparity estimation is an important way to evaluate visual comfort.

1.4 An Common Objective Evaluation Model

In this section we introduce a no-reference (NR) quality assessment method for stereoscopic images based on segmented local features of distortions and disparity. In order to combine artifacts and disparity features to build a stereo quality assessment model. The specific process as shown in figure1.
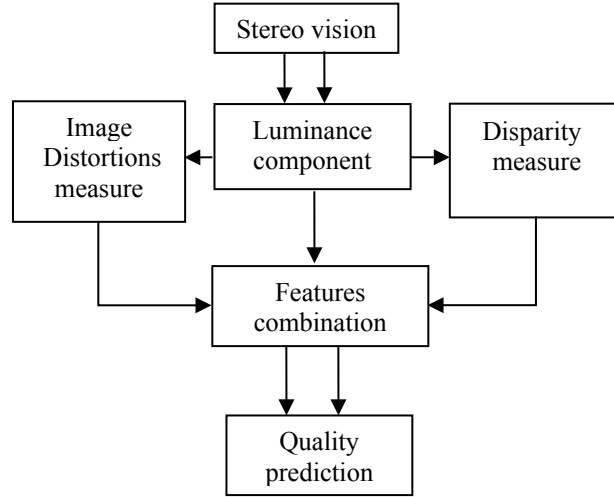

Fig1.1: NR quality evaluation method.

1.4.1 Image Distortions Measure

We believe that 3D depth perception is strongly dependent on objects, structures, or textures edges of stereo image content. Blockiness and edge distortions are taken as the image distortion degree index of this model. In[18] blockiness is one of the most annoying artifacts for human eyes and the zero -crossing values show a decreasing trend with the increasing edge distortion level. Therefore, we take into account the maximum blockiness and edge distortion measures between the left and right views. Block effect is mainly due to the discontinuity between image blocks and blocks, rather than all the pixels in the image distortion. According to this, there are two main methods for measuring block effect can be used. Edge of the test need the comparison of the original image, so its not suitable for this topic. We choose to block between adjacent pixel difference methods, Because this method is simple and .has good correlation with subjective evaluation results. Specific process is as follows.(1) Block based on image segmentation, we calculate blockiness and zero-crossing of each (8×8)block of the stereo image pair separately.(2)We use a simple block based segmentation algorithm[19] to classify edge and non-edge blocks of an image. and then average each value of blockiness and zero-crossing separately for edge, and non edge blocks of each image of the stereo pair.(3)we update the total blockiness and zero-crossing values by some weighting factors that are optimized by an optimization algorithm. Finally, the overall blockiness, and zero-crossing of each stereo image pair are calculated by the following equation. The method parameters and weighting factors (w1 to w4)are must be estimated by an optimization algorithm with the subjective test data.

$$B = B_e^{w1} \bullet B_n^{w2}$$

(1.1)

$$Z = ZC_e^{w3} \bullet ZC_n^{w4}$$

(1.2)

1.4.2 Relative Disparity Measure

To measure disparity, we use a fixed block based difference zero-crossing (DZC) approach which is discussed in [17].Firstly, we divide the left view into nonoverlapping blocks (8x8) and distinguish between edge and non edge blocks. Secondly, block-based DZC is estimated in the two corresponding blocks between the left and right images. Thirdly, we average the DZC rate values separately for edge and non edge blocks and values are update values with some weighting factors. The relative parallax is estimated by the following formula.

$$DC = AC_e^{w5} \bullet AC_n^{w6}$$

(1.3)

1.4.3 Features Combination.

In order to combine artifacts and disparity features to develop a stereo quality assessment model. We consider several different equations for validation, finally an optimal equation is obtained as follows, α , β (subjective experiment data).

$$S = \alpha + \beta B \bullet Z$$

(1.4)

In theory, the objective quality evaluation model and the subjective perceptual quality evaluation result should be a linear relationship, but in the process of subjective perception experiment design will bring in some nonlinear factors. If we compare the objective model with subjective perceptual quality assessment models directly, it is difficult to effectively persuade the validity of the model. Therefore, We consider a logistic function as the nonlinearity property between the human perception and the physical features. Finally, the obtained MOS prediction Mops, is obtained by the following equation [20]. Here, Particle Swarm Optimization (PSO) algorithm is used for optimization [21].

$$MOS_p = \frac{4}{1+\exp[-1.02179(S-3)]} + 1$$

(1.5)

1.4.4 Result

This model uses stereo images from IRCCyN/IVC testing library in French Nantes university and MICT testing library in university of Toyama in Japan to verify the result of the experiment. In order to get quantitative indicators, the standard performance evaluation process is hired, which was advocated by the video quality expert group (VQEG) by calculating mean absolute error (AAE), mean square error (RMSE), outlier ratio (OR), linear correlation coefficient (CC) to compare with the subjective ratings. The result shows that the performance of our method is superior compared to the published method both for the training and testing databases. Finally, it is compared with the recently raised evaluative method which was based on a kind of no reference stereo video. A 2d image evaluative method and a full reference (FR) stereo image evaluative method were compared. The result shows that the method is superior to other methods, and it is more practical. The experimental results are given in paper [8]. In this paper, these results illustrate the accuracy of current evaluation model.

Table 1.1 Evaluation results for training and testing

Method	Training			
	CC	AAE	RMSE	OR
Model	0.964	0.276	0.336	0.061
Method , FR	0.945	0.310	0.381	0.065
Using 2D mean	0.916	0.406	0.513	0.074
	Testing			
Model	0.940	0.339	0.393	0.037
Method , FR	0.929	0.370	0.441	0.082
Using 2D mean	0.883	0.438	0.559	0.061

The MOS versus MOSp of this proposed method for training and testing images are respectively shown in Figures1. The symbols "*" and "+" respectively, indicate MOSp points for the databases of training and testing. The subjective experiments results indicate this model has sufficient prediction performance.And, of course, the images from the video library in this model are based on the JPEG image encoding format, so the application is not wide, the future research direction is to apply it among images which are encoded to various formats.
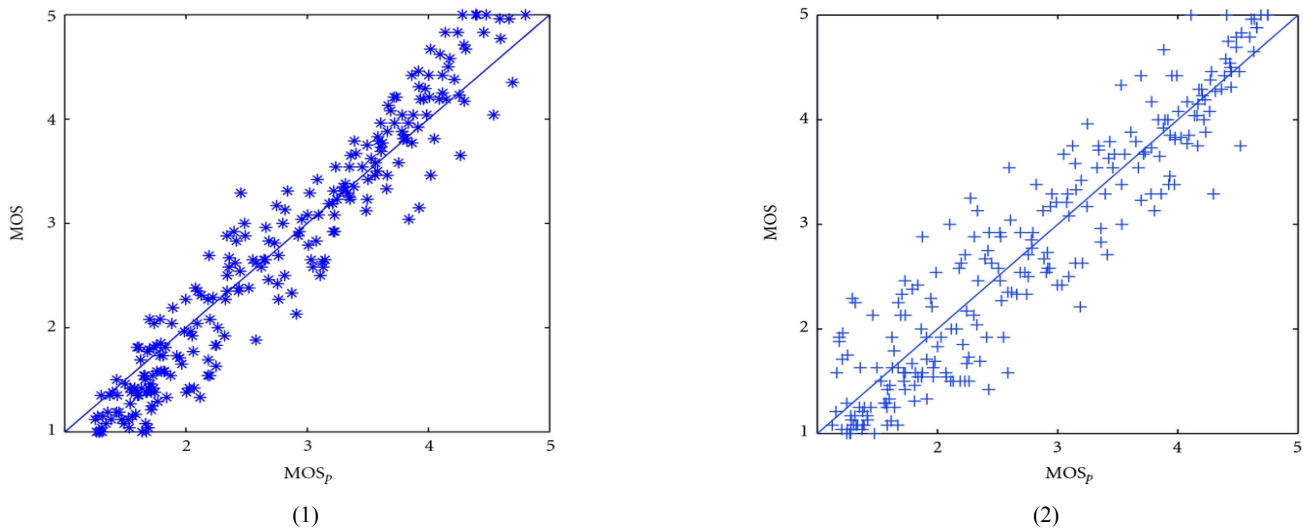

(1)   (2)

Fig1.2: Proposed method's MOS versus MOSp. (1) Training, (2) Testing.

1.5 Conclusion

Overall, no reference stereo video evaluation method is still a relatively new field of study, having broad application prospects. The research achievements are modelling and optimization. As researchers continue to the human visual system in-depth study of this problem, it will be grad still very limited, which has not yet formed a set of standard methods. The advantage is that the original image does not need to be embedded into the quality control system or the image processing system thus greatly improving the practical applications. At present, this kind of model in adaptability and accuracy has a larger gap to the actual demand. Future research needs from the perspective of simulation of HVS, using of multichannel characteristics of human visual perception system, such as brightness features, coequally resolved.

1.6 References

1. A . Kubota, A. Smolic, M. Mag n or, M . Tanimoto, T.Chen, and C . Zhang , "Multiview imaging and 3DTV," IEEE Signal Processing Magazine , vol. 24, no. 6, pp. 10－21, 2007
2. Z. M. P. Sazzad, and Y. Horita, "Local region-based image quality assessment independent of JPEG and JPEG2000 coded color images," J. of Electronic Imaging, vol. 17(3), pp. 033002-1-17, Jul-Sep 2008
3. Yang J.C., Hou C.P. Shen L.L., and Zhang, Z.Y.. Objective Evaluation Method for Stereo Image Quality Based on PSNR. Journal of [4]A. Tikanmaki, and A. Gotchev, "Quality assessment of 3D video in rate allocation experiments," in Proc. IEEE ISCE, Algarve, Portugal, Apr., 14-16, 2008
4. Leon G, Kalva H, Furht B,3D video quality evaluation with depth quality variations,3DTVConference 2008:301－304
5. Mao X.Y., Yu M., Zhou J.M., et al.. A Stereoscopic Image Objective Assessment Model Based on Weighted Structure Similarity. 2010 AsiaPacific.Conference on Information Theory, Xian, November 12-14, 2010.
6. A. Tikanmaki, and A. Gotchev, "Quality assessment of 3D video in rate allocation experiments," in Proc.IEEE ISCE, Algarve, Portugal, Apr., 14-16, 2008.Tianjin University Science and Technology, 2008,
7. Akhter R., Parvez Sazzad Z.M., Horita Y., and Baltes J.. No-Reference Stereoscopic Image Quality Assessment. Neil A. Dodgson, Proc. of SPIE-IS&T Electronic Imaging, SPIE. 2010.
8. Sazzad Z M, Akhter R, Baltes J, et al. Objective no-reference stereoscopic image quality prediction based on 2d image features and relative disparity[J]. Advances in Multimedia, 2012, 2012: 8
9. Wang Z, Sheikh H R, Bovik A C. No-reference perceptual quality assessment of JPEG compressed images[C]//Image Processing. 2002. Proceedings. 2002 International Conference on. IEEE, 2002, 1: I-477-I-480 vol. 1.
10. Tikanmaki A, Gotchev A, Smolic A, et al. Quality assessment of 3D video in rate allocation experiments[C]//Consumer Electronics, 2008. ISCE 2008. IEEE International Symposium on. IEEE, 2008: 1-4.

11. You J, Xing L, Perkis A, et al. Perceptual quality assessment for stereoscopic images based on 2D image quality metrics and disparity analysis[C]Proc. Int. Workshop Video Processing and Quality Metrics, Scottsdale, Arizona, USA. 2010.

12. Hewage C, Worrall S T, Dogan S, et al. Prediction of stereoscopic video quality using objective quality models of 2-D video[J]. Electronics letters, 2008, 44(16): 963-965.

13. Z. Wang , Rate scalable foveated image and video communications Department of ECE,The University of Texas at Austin, 2003

14. FrankL.Kooi,Alexander Toet.Visual comfort of binocular and 3D displays[J],Display,2004,25(2-3):99-108.

15. Seuntiens P. Visual experience of 3D TV[J]. doctor doctoral thesis, Eindhoven University of Technology, 2006.

16. D.V.Meegan,L.B.Stelmach, and W.J.Tam, "Unequal weighting of monocular inputs in binocular combination: implications for t he compression of stereoscopic imager y," Journal of Experimental Psychology : Applied ,vol.7,no.2,pp.143–153, 2001

18. B. P. McKinnon, Point, line segment, and region-based stereo matching for mobile robotics [M.S. thesis],Department of Computer Science, University of Manitoba, 2009

19. R. Akhter, Perceptual image quality for stereoscopic vision[M.S. thesis] , Department of Computer Science , University of Manitoba, 2011

20. J. Kennedy and R . E berhart, "Particles warm optimization," in Proceeding s of the IEEE International Conference on Neural Networks, pp. 1942–1948, Per t h, Australia, December 1995.

21. Campisi P., Callet P.L., Marini E.. Stereoscopic Images Quality Assessment.15th European Signal Processing Conference, Poznan, Poland, 2007:2109-2114

22. Sazzad Z.M.P., Yamanaka S., Kawayoke Y., Horita Y.. Stereoscopic Image Quality Prediction [J]. In Proc. IEEE QoMEX, San Diego, CA, USA, July 29-31, 2009.

23. Z. M. P. Sazzad, S. Yamanaka, Y. Kawayoke, and Y. Horita, "Stereoscopic image quality prediction," in Proceedings of the International Workshop on Quality of Multimedia Experience (QoMEx 09), pp. 180–185, San Diego, C A, USA, July 2009.

24. Z. Wang , A . C . B ov ik, H. R . Sheikh, and E . P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," IEEE Transact ions on Image Processing, vol. 13, no.4, pp. 600–612, 2004