

# A New Objective Video Quality Metric for Stereoscopic Video

Yan Zheng\*, Jungdong Seo\*\*, Kwanghoon Sohn\*

\*School of Electrical & Electronic, Yonsei University

\*\* School of Electrical & Electronic, Yonsei University

## Abstract

Although quality metrics for 2D video quality assessment have been proposed, the quality models on stereoscopic video have not been widely studied. In this paper, a new objective video quality metric for stereoscopic video is proposed. The proposed algorithm consider three factors to evaluate stereoscopic video quality: blocking artifact, blurring artifact, and the difference between left and right view of stereoscopic video. The results show that the proposed algorithm has a higher correlation with DMOS than the others.

## 1. Introduction

Stereoscopic video which contains two viewpoint video sequences is the simplest 3D video format. It is currently the most commercialized 3D video format. The 3D movie theater and glass-type 3DTV are based on the stereoscopic video. Quality assessment of stereo video could be obtained either through subjective tests or through objective assessment [1]. The best way to assess image and video quality would be to use subjective tests, which are defined in order to obtain correct, universal, and reliable quality evaluations. However, subjective tests are expensive and time-consuming, since a number of people and test conditions are required to obtain subjective scores and the analysis of the obtained results is not straightforward. Furthermore, the subjective metrics may not be feasible for most applications. To overcome this, researchers have made great efforts to develop objective quality metrics that can be used in real-world applications. The definition of objective assessment reliably predicting the perceived quality of images would be a great improvement in the quality assessment field.

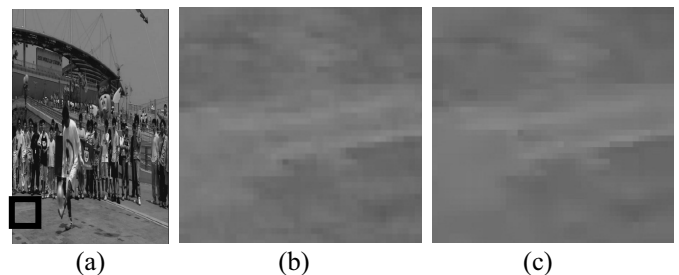
Although many quality metrics for 2D image quality assessment have been proposed, the quality models on stereoscopic images have not been widely studied. Hewage et al. [2] tested the performance of three quality metrics, including (PSNR), video quality model (VQM) proposed by NTIA [3], and structural similarity model (SSIM) [4], with respect to a subjective quality experiment on a series of coded stereoscopic images. The experimental results demonstrated that VQM is better than other two metrics while its performance is still not promising. Similar work has been done in [5]. Four metrics, as well as three approaches, called average approach, main eye approach, and visual acuity approach, were tested for measuring the perceptual quality of stereoscopic images. Further, disparity information was integrated into two metrics for the quality assessment [6]. **It was found** the disparity information has a significant impact on stereoscopic quality assessment, while its capability has not been studied adequately. In [7], only absolute disparity was used. It was found that added noise on the relatively large absolute disparity has greater influence than on other disparity. Subsequently, a metric called stereo sense assessment (SSA) based on the disparity distribution has been proposed.

In this paper, a new objective video quality metric for stereoscopic video is proposed. The proposed quality metric measure spatial impairments including blocking artifact and blurring artifact that can occur in compression and transmission and also measuring the difference between left and right views of stereoscopic video. Since the subjective quality of stereo video sequences was successfully measured by the proposed metric, the quality of the multimedia service can be guaranteed.

The rest of this paper is organized as follows: Section 2 presents research background of stereo video quality metric, section 3 describes the proposed algorithm in detail. Section 4 provides performance comparisons for proposed objective video quality metric. Section 5 provides our research conclusions.

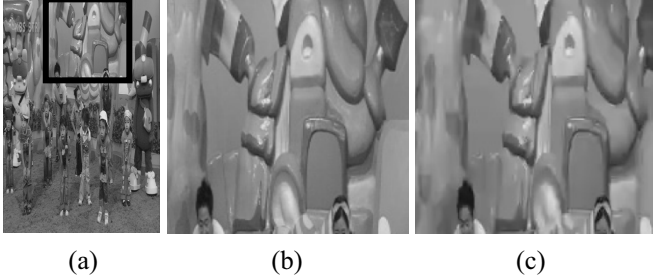
## 2. Background

Compressed video sequences often contain spatial impairments such as blocking and blurring artifacts. Since the JPEG image compression standard uses block-based DCT (discrete cosine transform), as the coding ratio increases, the compressed image show more blocking artifacts. This phenomenon also occurs in video compression. Moreover, block-based motion-compensation/estimation in video compression can propagate the blocking artifact to adjacent frame. Fig. 1(c) shows the blocking artifacts that can occur in compressed video. Therefore, blocking metrics can be used to measure the perceptual quality of compressed images and video sequences.



**Fig. 1.** Example of blocking artifacts on 'Soccer' sequence. (a) The original sequence. (b) The sub image of original frame. (c) The sub image of compressed frame.

Besides the blocking artifacts, a modern image compression standard such as JPEG2000 and modern video compression standards such as H.264 and HEVC may also produce blurring artifacts. These blurring artifacts damage the sharpness of scenes and the perceptual quality can be degraded. Therefore, the blurring metric should also be considered to measure the perceptual quality. Fig. 2 shows an example of when the sharpness of the scene is degraded.

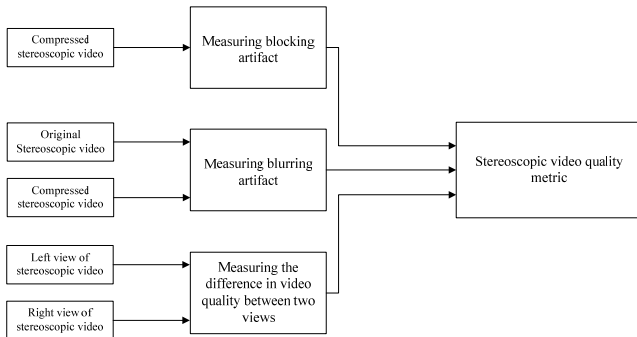


**Fig. 2.** Example of blurring artifacts on ‘Dance’ sequence. (a) The original sequence. (b) The sub image of original frame. (c) The sub image of compressed frame.

While besides blocking and blurring artifacts, stereoscopic video offers a sense of presence and reality because it delivers depth information to users, unlike 2D video. When left and right views are shown to user’s left and right eyes, the user perceives depth information. Video quality of each view has a significant effect on the stereoscopic video quality, as well as the depth information, because users perceive stereoscopic video based on each view. Therefore, stereoscopic video quality metric should be able to consider 3D effect quality and video quality of each view as 2D video.

### 3. An objective video quality metric for Stereoscopic Video

Based on the above analysis, we assume video quality is affected by three factors, blocking artifact, blurring artifact and difference between left and right views. The proposed metric predicts the quality of stereoscopic video by combining values of three items as shown Fig. 3.



**Fig. 3.** Block diagram of the proposed algorithm

#### 3.1 Measuring the blocking artifact

F. Pan introduced a new method that estimated blocking artifact of images by using edge direction information to compute the orientation changes of the edge pixels and their relationship with the blocking artifacts [8]. This method does not need the exact location of the block boundary thus is in-

variant to the displacement, rotation and scaling of the images. In this paper, we use this method to measure blocking artifact since its good performance and efficient in terms of computational complexity and memory usage.

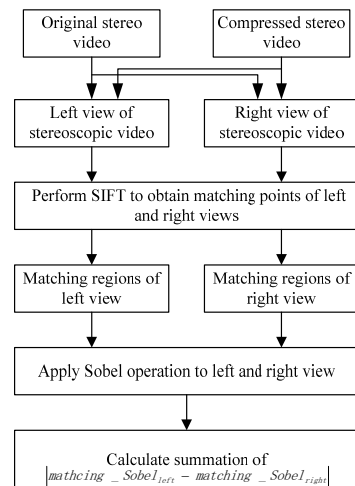
#### 3.2 Measuring the blurring artifact

Blurring in a video sequence is due to the attenuation of the high spatial frequencies, which commonly occurs during filtering or visual data compression. One effective way to examine blurring artifact is to check the distribution of a high-frequency component between the original frame and the compressed frame. Therefore, first we apply an edge detector to the luminance component of the image. Sobel filter is used because of its simple process and good performance. Noise and insignificant edges are removed by applying a threshold (Thr1) to the gradient image. In our experiment, an average value of the edge magnitude in the original image is chosen as the threshold value.

Then to detect blurring artifact we compare edge magnitude for the original image with the compressed image. In this procedure, we also assign a threshold (Thr2). If the difference is larger than Thr2, we assume that blurring is occurring in the region. In this paper, half of standard deviation of the edge magnitude is chosen as Thr2. Then the difference between the edge magnitude values of the original and the compressed image accumulates in the blurred area. The accumulated value indicates the degree of blurring artifact.

#### 3.3 Measuring the difference between two views

It is totally different when we view a 2D image and a stereo image. Viewing a 2D image, two eyes can only see a same image. While as for a stereo image, two eyes see two images with the standard disparity. With the combinative effect of two eyes, stereo image distinguishes significantly from 2D image. When two eyes viewing stereo image, human brain can generate stereo sense from the standard disparity information, which is the distinction of the stereo image. Applying coding, distortions can lead to modify this difference between left and right views. Therefore, by measuring the sense quality of stereo vision, we can predict the quality of stereoscopic video. Fig. 4 presents a detection process for difference between stereoscopic views.



**Fig. 4** Detection the difference between two views

First, we perform SIFT (Scale Invariant Feature Transform) [9] to obtain matching points of left and right views of stereoscopic video. SIFT is a well-known algorithm for extracting distinctive invariant features from images that can be used to perform reliable matching between different views of an object or scene. The algorithm was published by David Lowe in 1999. After we obtained matched regions of left and right views, Sobel operator is used in matched regions of left video and right video to detect edge regions. Last, we calculate the difference in video quality between two views of the original video pair and the processed video pair. The difference is obtained by accumulating the difference between left and right images in the corresponding regions. Then, the accumulated difference value of the processed video pair is divided by the accumulated difference in the original video pair for normalization.

$$|matching\_Sobel_{left} - matching\_Sobel_{right}|$$

The more similar between the absolute matching image of the original and the coded stereo image is, the better the two eyes disparity quality shows, and the better the stereo sense presents.

After calculating the score of the three items, we obtain weights for each item for the stereoscopic video quality metric. The weights are calculated using a linear regression, assuming that the relationship between the response variable and the vector of input variable is approximately linear.

#### 4. Experiment result

To confirm performance of the proposed algorithm, we carried subjective evaluation for stereoscopic videos. The experimental environment for the subjective evaluation is summarized in Table 1.

Table 1. Experimental environment

Test video	48 stereoscopic videos (8 original video + 40 compressed video)
Resolution	640×480
Frame rate	30 fps
subjects	20
Display	SAMSUNG 3DTV with shutter glasses
Viewing distance	3 times of display height
Screen size	24 inch / 16:9

Test sequences for subjective evaluation are eight VGA (640×480) sequences. Processed video is compressed using Joint Multi-view Video Model 8.0 (JMVM 8.0) software which was adopted as an international standard for multi-view video coding in JVT (Joint Video Team) [11]. Fixed QP (Quantization Parameter) values were used. The values were 25, 30, 35, 40 and 45. The results of subjective evaluation are

represented in Fig. 5.

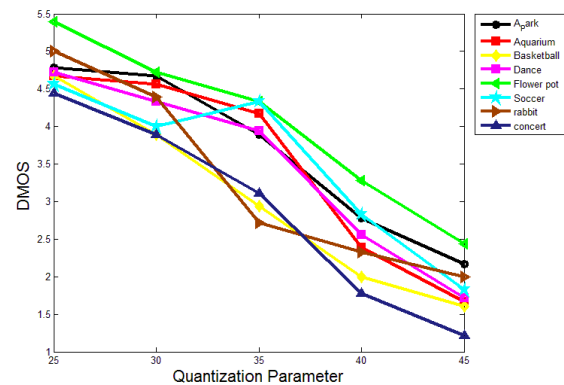


Fig. 5. Subjective evaluation results for test sequences

Fig. 6, Fig. 7 and 8 presents scores for each item on test sequence. In detection of blocking artifact, the result of 'basketball' sequence shows a high score of blocking artifact, and in detection of difference between left and right views, 'Aquarium' shows a lower result than other regular scores while 'rabbit' gives a larger result than other

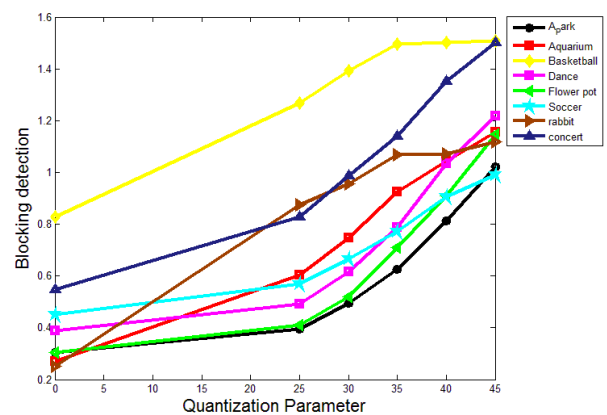


Fig. 6. Result of blocking artifact

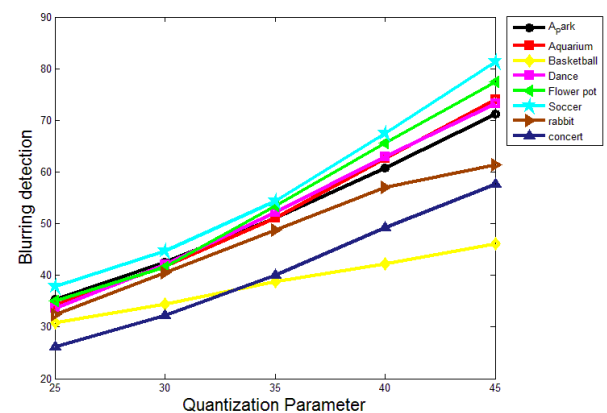


Fig. 7. Result blurring artifact

We combine three items using linear regression and compare the final results with conventional algorithms, which are PSNR, SSIM, and SSIM local disparity distortion (Ddl1)[10]. SSIM and PSNR are widely used for 2D video. We obtained

the objective scores by averaging the results of each vies. SSIM Ddl1 is proposed for stereoscopic video evaluation which is based on SSIM considering 3D effect.

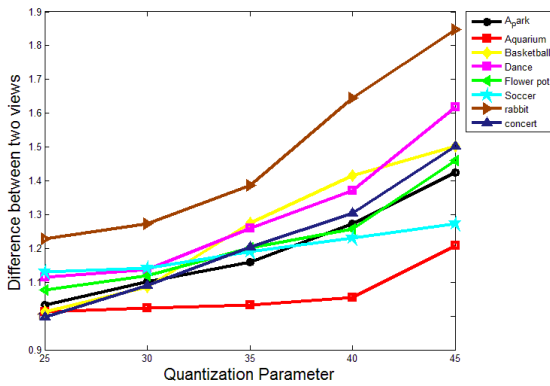


Fig. 8. Result of difference between two views

Figure 9 and Table 2 show scatter plots and correlation factors of the algorithms with Differential Mean Opinion Score (DMOS), respectively. In terms of Correlation Coefficient (CC), the result of the proposed algorithm is better than the others.

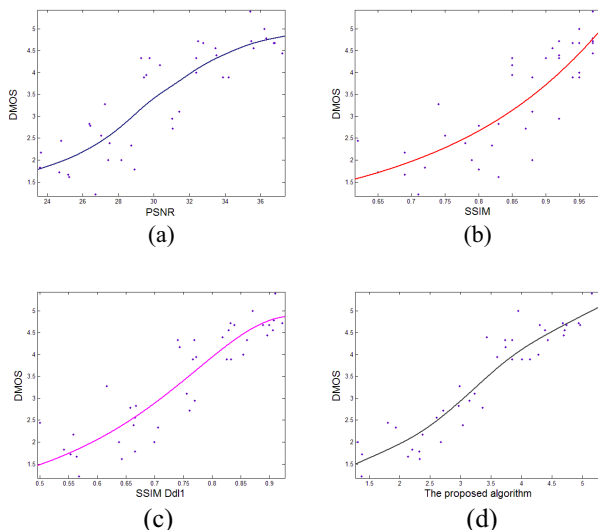


Fig. 9 Scatter plot of algorithm. (a) PSNR. (b) SSIM.(c) SSIM Ddl1. (d) The proposed algorithm.

Table 2. Performances comparison

	PSNR	SSIM	SSIM Ddl1	Proposed algorithm
cc	0.85	0.82	0.89	0.92
RMSE	0.607	0.432	0.411	0.377

## 5. Conclusion

In this paper, a new objective video quality metric for stereoscopic video is proposed. The proposed algorithm considers three factors: blocking artifact, blurring artifact, and difference between left and right view of stereoscopic video. The results show that the proposed algorithm has a higher

correlation with DMOS than the others. However, we need further research to achieve a more accurate objective quality metric considering eye fatigue in 3D video, because the distortion in compressed video depends on the characteristics of the codec and eye fatigue critically affects stereoscopic video quality, unlike 2D video.

## Reference

- [1] Meesters L.M.J., IJsselsteijn W.A., Seuntjens P.J.H., "A Survey of Perceptual Evaluations and Requirements of ThreeDimensional TV", IEEE Transactions on Circuits and Systems for Video Technology, Volume 14, Issue 3, March, 2004, pp:381 – 391
- [2] C.T.E.R.Hewage, S.T.Worrall, S. Dogan, and A.M. Kondoz, "Prediction of stereoscopic video quality using objective quality models of 2-D video," Electronics Letters, 44(16), pp. 963-965, Jul. 2008.
- [3] M. Pinson, and S. Wolf, "A new standardized method for objectively measuring video quality," IEEE Trans. Broadcasting, 50(33), pp. 312-322, Sep. 2004.
- [4] Z. Wang, A.C. Bovik, H.R. Sheikh, et al., "Image quality assessment: from error visibility to structural similarity," IEEE Trans. Image Processing, 13(4), pp. 600-612, Apr. 2004.
- [5] P. Campisi, P.L. Callet, and E. Marini, "Stereoscopic images quality assessment," in Proc. 15th European Signal Processing Conference (EUSIPCO), Poznan, Poland, 2007.
- [6] A. Benoit, P.L. Callet, P. Campisi, et al., "Quality assessment of stereoscopic images," EURASIP Journal Image and Video Processing, doi:10.1155/2008/659024, (2008)
- [7] Y. Jiachen, H. Chunping, Y. Zhou, et al., "Objective quality assessment method of stereo images," in Proc. IEEE 3DTV, Potsdam, Germany, 2009.
- [8] F. Pan, X. Lin, S. Rahardja, E. Ong, W. S. Lin, Using edge direction information for measuring blocking artifacts of images. Multidimensional Systems and Signal Processing, 2007, 18:297-308
- [9] Lowe, D.G. 1999. Object recognition from local scale-invariant features. In International Conference on Computer Vision, Corfu, Greece, pp. 1150–1157.
- [10] A. Benoit, P.L. Callet, P. Campisi, R. Cousseau, Quality assessment of stereoscopic images, EURASIP J. Image Video Process. doi:10.1155/2008/659024 (2008)
- [11] ISO/IEC JTC1 WG11 MPEG, Text of ISO/IEC 14496-10:200X/FDAM 1 Multi-view Video Coding, w9978 (2008)