A NEW NO-REFERENCE STEREOSCOPIC IMAGE QUALITY ASSESSMENT BASED ON OCULAR DOMINANCE THEORY AND DEGREE OF PARALLAX

Ke Gu, *Student Member, IEEE,* Guangtao Zhai, *Member, IEEE,*
Xiaokang Yang, *Senior Member, IEEE* and Wenjun Zhang, *Fellow, IEEE*

Institute of Image Communication and Information Processing, Shanghai Jiao Tong University, Shanghai, China
Shanghai Key Laboratory of Digital Media Processing and Transmissions

ABSTRACT

In recent years, while stereoscopic images were becoming widely applied and the corresponding technologies were substantially developed, very few stereoscopic image quality assessment metrics were proposed, especially under the condition that there is no reference image available. This paper proposes a new no-reference stereoscopic image quality assessment algorithm based on the ocular dominance theory and degree of parallax. All of our tests using the Toyama database draw two valuable conclusions. First, the performances of stereoscopic image quality assessment methods are significantly affected by difference of image qualities between the left and right images. Second, to offset the discriminations of different degrees of parallax caused by various 3D image contents is required indeed. Experiments and comparative studies are provided to confirm the effectiveness of our proposed new stereoscopic image quality metric.

***Index Terms*—**Stereoscopic image, image quality assessment (IQA), no-reference, ocular dominance, parallax

1. INTRODUCTION

Three-dimensional (3D) imaging is an extensive research area, which ranges from entertainment, including videos and games, to specialized applications, such as the education and medicine. As more and more image processing operations have been specifically designed for stereoscopic images, the necessity of perceived stereoscopic image quality metrics shows gradually evident, especially under the condition that there is no reference image available.

Although the human subjective feelings between two-dimensional (2D) and 3D images are considerably different on watching, they have a very close contact with each other. It is stated in [1] that cells in the retina of each eye individually encode its received information, and then the information from both eyes is merged in Lateral Geniculate Nucleus (LGN). Thus, the final stereoscopic image is formed in the brain. Illuminated by this phenomenon, it suggests us to design a no-reference stereoscopic IQA method combining monoscopic image quality metrics and the relationship between 2D and 3D image qualities.

It is found that some stereoscopic IQA metrics were recently proposed [2]-[4], but few has been studied on the impact of distinct qualities between the left and right images on perceptual quality of the stereoscopic image. In our research, it is noticed that both 2D image qualities and 3D image quality show one type of nonlinear relationship. This coincides with a demonstration in [5] that the amplitude tends to be reduced with anticorrelated stimuli due to the responses of V1 complex cells. Further research, referring to the ocular dominance theory [6]-[7], points out that this nonlinear relationship is likely to be introduced by the difference of qualities between the left and right images, and in the meantime, it has significant influences on the perceptual quality of stereoscopic images.

In addition, different degrees of parallax caused by various 3D image contents also highly affect the prediction performance of image quality metrics, just coincided with the demonstration in [8] that the low subjective evaluation is introduced from high degree of parallax. Thus, except for applications of the ocular dominance theory, a valid offset on different degrees of parallax is also employed here.

Consequently, based on a 2D no-reference image quality metric [9], the ocular dominance theory and degree of parallax, we propose a novel Ocular dominance theory and Degree of parallax based Distortion Metric (ODDM) for the no-reference quality assessment of JPEG compressed stereoscopic images. The ODDM algorithm mainly includes two parts: sum of image quality predictions from the left and right images, and two compensations on the "distress" caused by different qualities between the left and right images, as well as different degrees of parallax caused by different 3D image contents.

The remainder of this paper is organized as follows. Section 2 first reviews the ocular dominance theory and defines "distress" induced by different qualities between the left and right images. The importance of degree of parallax is introduced, and then a corresponding offset method is described in detail in section 3. Section 4 proposes our ODDM approach. In Section 5, experimental results using the Toyama database [4] are reported and analyzed. Finally, concluding remarks are given in Section 6.

This work was supported in part by NSERC, NSFC (61025005, 60932006, 61001145), SRFDP (20090073110022), postdoctoral foundation of China 20100480603, 201104276, postdoctoral foundation of Shanghai 11R21414200 and the 111 Project (B07022).

2. OCULAR DOMINANCE THEORY

Unlike monoscopic image quality assessment, which basically depends on distortion or difference between the reference and distorted images, it is easy to notice that the perceptual stereoscopic image quality cannot completely be determined by distortion or difference between the reference and distorted images alone.

2.1. "Distress" of Stereoscopic Image Pairs

A large number of groups of stereoscopic image pairs are chosen from the Toyama database [4]. Every group includes three stereoscopic image pairs, which meet the requirements presented in Table I. And one exemplary group is shown in Fig. 1.

If we observe the group of images as shown in Fig. 1 with 3D shutter glasses, at a glance, we can immediately tell that the first stereoscopic image has the noticeable distortion. However, we can barely find any difference between the reference and distorted images for the second and third stereoscopic images. This phenomenon can be explained by the fact that there is one image of high quality in the image pairs and the visual transience effect makes observers feel good because of the rapid switch of 3D glasses shutters between the open and close states. Contrarily, since the qualities of both images in the first stereoscopic image pairs are not good, they are of much lower perceptual quality.

However, if viewing the second and third image pairs for sufficiently long time, we can feel uncomfortable or even dazzle. The MOS (Mean Opinion Score) from the Toyama database verifies this observation: The first image quality is always the highest of the three while the MOS value of the third image is usually a little higher than that of the second image. We believe that this phenomenon is caused by the "distress" feeling for the condition that large difference exists between qualities of the left and right images.

TABLE I. DEMANDS FOR THE CHOSEN IMAGE GROUPS FOR TESTING THE "DISTRESS" FEELING.

Image Index	MOS of left image	MOS of right image
1	Fair	Fair
2	High	Low
3	Low	High

2.2. Ocular Dominance Theory

When observing a stereoscopic image, if the left and right images are almost of the same quality, our eyes can function equally, and the images are merged smoothly in LGN [1]. However, when the left and right image qualities are quite different, e.g. the left image is clear but the right one is heavily contaminated as shown in Fig. 1 (c)-(d), the imbalance causes the inequality of the two eyes [10] and LGN will have difficulty in merging left and right images. The brain must tense the muscles around the outside of the eye that is receiving the low quality image so as to steady

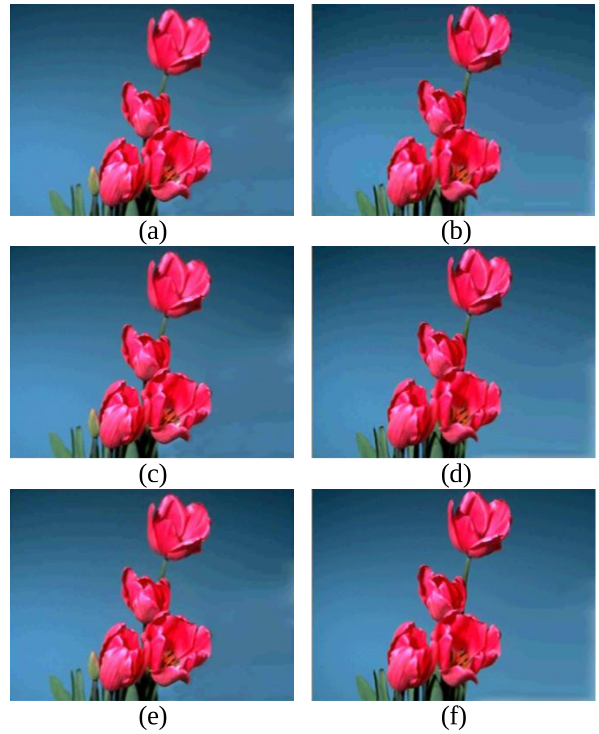


Figure 1. A group of three stereoscopic image pairs that meet the demands in Table I: (a)-(b), (c)-(d) and (e)-(f) are the left and right images of the first, second and third stereoscopic images, respectively

the view. And this makes the brain suffer. After a while, the muscles start to ache and the nerves begin to pain. And this unbalanced image pair may even result in amblyopia for a very long term viewing [10].

The ocular dominance theory can well justify the condition of viewing stereoscopic image pair with the great discrepancy of image quality. So in this research, we define the "distress" as

$$G_1(X,Y) = F(X,Y) \bullet \|X - Y\|_\infty \quad (1)$$

where X and Y are the left and right eye responses respectively, and $F(X,Y)$ denotes a "distress" degree parameter. Setting $F(X,Y)$ equal to 1, we approximately get:

$$G_2(X,Y) = \|X - Y\|_\infty . \quad (2)$$

Inspired by the definition of ODI in [6]-[7]:

$$ODI = \frac{[left\_eye\_repsonse]}{[left\_eye\_repsonse]+[right\_eye\_repsonse]}, \quad (3)$$

we believe that ODI is a good characterization to constitute a link between both 2D image qualities and the final 3D image quality. Therefore, we can rewrite the Eq. (1) by taking ODI into $F(X,Y)$ as:

$$H(X,Y) = ODI \bullet \|X - Y\|_\infty = \frac{X}{X+Y} \bullet \|X - Y\|_\infty . \quad (4)$$

3. PARALLAX

Except that the ocular dominance theory exists in the combination of both 2D image qualities, we found another category of additive effect between different groups of stereoscopic images. We test three monoscopic image pairs of the whole thirteen 3D image groups, and set every group of three image pairs to meet the demands presented in Table II, which lists the chosen three kinds of different degrees of JPEG distortion for complete testing.

The results of experiments illustrated in Fig. 3 support our assumption, and meanwhile, two abnormal findings are observed. Firstly, we noticed that blue "×" and green "*" in Fig. 3(a) with high qualities of both 2D images have the unusual low qualities of their corresponding 3D images, which violates our common sense that the higher qualities of left and right images the higher quality of their forming stereoscopic image. And moreover, the similar phenomenon also exists in Fig. 3(b)-(c). Secondly, it is not difficult to find that the abnormal quantities of every group of stereoscopic images are extremely consistent in Fig. 3(a)-(c). Testified by all the other stereoscopic images, these two facts stated above exist widely.

TABLE II. REQUIREMENTS FOR THE CHOSEN THREE IMAGE PAIRS FOR TESTING DEGREE OF PARALLAX.

Image Pairs Index	JPEG Compressive Quality of Left Image	JPEG Compressive Quality of Right Image
1	27	79
2	79	27
3	55	55

This pair of phenomena cannot be produced by the relationship between qualities of left and right images, or else it should not appear in Fig. 3(c), in which each pair of images have the same JPEG compressive qualities. It is easy to observe a strange phenomenon that the quantities above are immensely consistent. Enlightened by an important result in [8] that the low subjective evaluation appeared for high degree of parallax, and furthermore, it is demonstrated that this degree is independent of degree of distortion, we have a reason to believe that the visual discomfort is introduced from different degrees of parallax, which are caused by different stereoscopic image contents. So, we define the degree of parallax by:

$$d^{\theta} = cos^{-1}\left(\frac{L(M)\bullet R(M)}{\|L(M)\|_2 \bullet \|R(M)\|_2}\right) \tag{5}$$

where L and R represent left and right images respectively, and M indicates the middle region of the image.

4. THE PROPOSED QUALITY METRIC

Understandably, the quality of each monoscopic image in the stereoscopic image pair affects the final quality. So it is possible for us to devise a 3D image quality metric based on a no-reference 2D image quality metric ([9] is chosen here), the ocular dominance theory and degree of parallax.

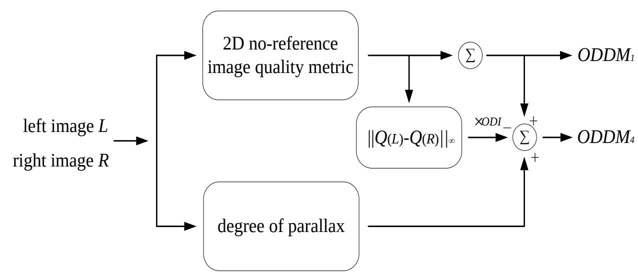
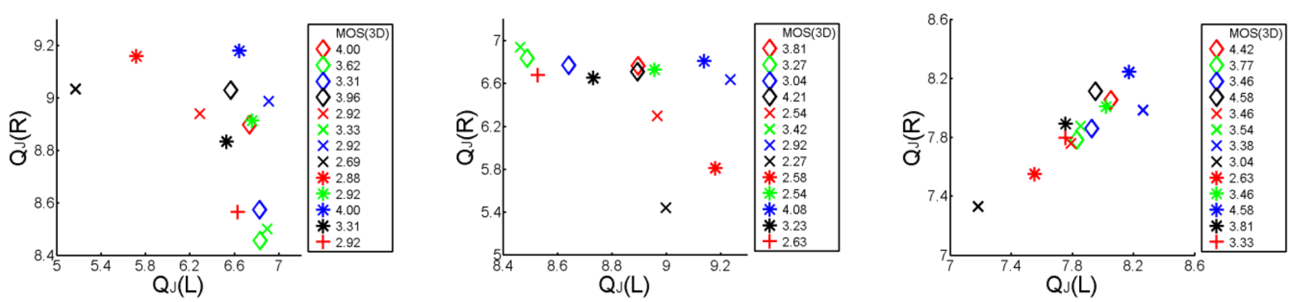


Figure 2. Four steps of ODDM framework.

The proposed ODDM algorithm includes four steps (only $ODDM_1$ and $ODDM_4$ are displayed in Fig. 2, and the omitted $ODDM_2$ and $ODDM_3$ have similar steps): First, calculate the left and right image qualities; Second, estimate ocular dominance difference between both 2D image qualities; Third, evaluate the compensating quantity introduced from different degrees of parallax; Fourth, predict the quality score by combining the 2D quality scores as well as the compensating quantity, and reducing ocular dominance difference. Roughly using Q_J (the abbreviation of Q_{JPEG} defined in [9]) as 2D IQA prediction, some middle products ($ODDM_1$ to $ODDM_3$) and our final $ODDM_4$ are given as follows:



(a) (b) (c)

Figure 3. Illustration of the (a): first index; (b): second index; (c): third index of image pairs in Table II. $Q_J(L)$ and $Q_J(R)$ (defined in Section 4) are preditions of left and right image qualities by [9], and MOS(3D) indicates MOS values of stereoscopic images.

$$ODDM_1 = X + Y = Q_J(L) + Q_J(R) \tag{6}$$

$$ODDM_2 = ODDM_1 - G_2(X,Y) \tag{7}$$

$$ODDM_3 = ODDM_1 - H(X,Y) \tag{8}$$

$$\begin{aligned} ODDM_4 &= ODDM_1 - H(X,Y) + d^{\theta} \\ &= Q_J(L) + Q_J(R) + \cos^{-1}\left(\frac{L(M)\bullet R(M)}{\|L(M)\|_2 \bullet \|R(M)\|_2}\right) \\ &\quad - \frac{Q_J(L)}{Q_J(L)+Q_J(R)} \bullet \|Q_J(L) - Q_J(R)\|_{\infty} \end{aligned} \tag{9}$$

where $Q_J(L)$ and $Q_J(R)$, as the responses of X and Y, indicate prediction values of left and right image qualities.

5. EXPERIMENTAL RESULTS

Mappings of these four metrics values to subjective scores are obtained using nonlinear regression with a 4-parameter logistic function as suggested by VQEG [11]:

$$q(x) = \frac{\beta_1 - \beta_2}{1 + exp\left(-(x-\beta_3)/\beta_4\right)} + \beta_2 \tag{10}$$

with x being the input score and $q(x)$ the mapped score and β_1 to β_4 are free parameters to be determined during the curve fitting process.

Four generally used performance metrics, Pearson Linear Correlation Coefficient (PLCC), Spearman Rank-order Correlation Coefficient (SRCC), Average Absolute prediction Error (AAE) and Root Mean Square prediction Error (RMSE) as suggested by VQEG [11], are applied to further evaluate the competing ODDM based IQA metrics. Their values and [4] are illustrated in Table III and scatter plots (after nonlinear regression) of $ODDM_3$ and $ODDM_4$ are shown in Fig. 4.

It can be seen that our ODDM have achieved better results and moreover, as expected, the performance gain of $ODDM_4$ which makes full use of the ocular dominance theory and degree of parallax achieves the best accuracy.

6. CONCLUSION

This paper proposes a novel no-reference stereoscopic image quality assessment approach combining 2D no-reference IQA method, the ocular dominance theory and degree of parallax. It is worth pointing out here that stereoscopic image quality is considerably influenced by difference of qualities between the left and right images. And the reliable offset of different degrees of parallax caused by different 3D image contents can make the stereoscopic image quality prediction more accurate. Results of experiments verify that our proposed ODDM methods have superior accuracy for stereoscopic image quality assessment.

TABLE III. PLCC, SRCC, AAE AND RMSE VALUES (AFTER NONLINEAR REGRESSION) OF [4] AND $ODDM_{1\text{-}4}$ ON TOYAMA DATABASE.

METRICS	PLCC	SRCC	AAE	RMSE
$ODDM_1$	0.8934	0.8897	0.4136	0.5234
$ODDM_2$	0.9207	0.9076	0.3594	0.4548
$ODDM_3$	0.9318	0.9266	0.3358	0.4228
[4]	0.9350	—	0.3500	0.4210
$ODDM_4$	0.9578	0.9500	0.2594	0.3354

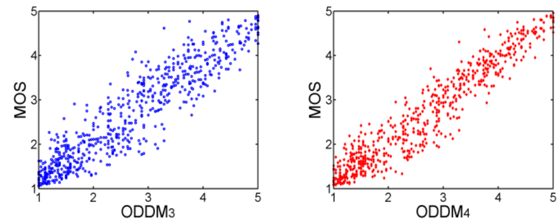


Figure 4. Scatter plots after nonlinear regression of MOS vs. $ODDM_3$ and $ODDM_4$ on Toyama database [4].

7. REFERENCES

[1] Zhou Wang and Alan C. Bovik, Modern Image Quality Assessment. Synthesis Lectures on Image, Video & Multimedia Processing, Morgan & Claypool Publishers, 2006.

[2] Benoit, P. Le. Callet, P. Campisi, and R. Cousseau, “Quality assessment of stereoscopic images,” EURASIP Journal on Image and Video Processing, vol. 2008, pp.1–13, 2008.

[3] Hewage, C.T.E.R., Martini and M.G, “Reduced-reference quality evaluation for compressed depth maps associated with colour plus depth 3D video,” Proceedings IEEE International Conference on Image Processing, September 2010.

[4] Roushain Akhter, Z. M. Parvez Sazzad, and Y. Horita, and J. Baltes, “No reference stereoscopic image quality assessment,” in Proceedings SPIE, vol. 7524, San Jose, CA, USA, January 2010.

[5] Read, J. C. A., Parker, A. J. & Cumming, B. G., “A simple model accounts for the response of disparity-tuned V1 neurons to anticorrelated images,” Visual Neuroscience, pp. 735–753, 2002.

[6] Macy A., Ohzawa I. and Freeman R. D., “A quantitative study of the classification and stability of ocular dominance in the cat's visual cortex,” Exp Brain Res 48: 401–408, 1982.

[7] M. Nomura, G. Matsumoto and S. Fujiwara, “A binocular model for the simple cell,” Biological Cybernetics, vol. 63, no. 3, pp. 237–242, 1990.

[8] S. Yano, S. Ide, T. Mitsuhashi, and H. Thwaites, “A study of visual fatigue and visual comfort for 3D HDTV/HDTV images,” Displays, vol. 23, pp. 191–201, 2002.

[9] Z. Wang, H. R. Sheikh and A. C. Bovik, “No-reference perceptual quality assessment of JPEG compressed images,” Proceedings IEEE International Conference on Image Processing, September 2002.

[10] J. Hoeve, “Amblyopia and squint,” Documenta Ophthalmologica, vol. 7–8, no. 1, pp. 392–421, 1954.

[11] VQEG, “Final report from the video quality experts group on the validation of objective models of video quality assessment”, March 2000, http://www.vqeg.org/.