Stereoscopic Image Quality Assessment with The Dual-weight Model

Yucheng Zhu, Guangtao Zhai, Ke Gu, Zhaohui Che and Duo Li
Insti. of Image Commu. & Infor. Proce., Shanghai Jiao Tong University, Shanghai, China
Shanghai Key Laboratory of Digital Media Processing and Transmissions
{zyc420, zhaiguangtao, gukesjtuee,chezhaohui}@sjtu.edu.cn, liduoee@gmail.com

***Abstract*—Three-dimensional (3D) imaging technology has been growingly prevalent in today's world. But objective quality assessment of 3D images is a challenging task. In this paper, we try to investigate and develop the 2D image quality metrics into stereoscopic image quality assessment using a dual-weight model within the concept of free energy principle and human visual system. On the basis of a psychological measure, the free energy is a principle telling where supervises more and attracts human attention. We believe that the "surprise" can account for the binocular rivalry and introduce the inter-weight. Also, the intra-weight is computed according to the physiology of human visual system. We first do the intra-weighted 2D image quality assessment. With inter-weights, we then calculate the combining score of stereopairs. Experiments are performed on the symmetric LIVE3D-I and asymmetric LIVE3D-II databases. Results confirm that the proposed 3D IQA technique, appropriately computing and combining the inter- and intra-weights, is able to faithfully predict the visual quality of stereopairs.**

***Keywords*—Performance evaluation, objective evaluation techniques, image quality assessment (IQA), no-reference (NR), three-dimensional, stereoscopic, dual-weighted, binocular rivalry.**

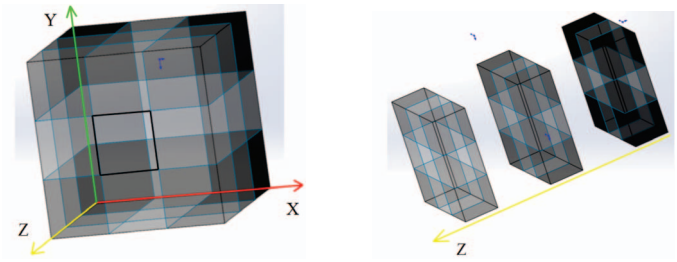
I. Introduction

Three-dimensional (3D) media services are becoming increasingly popular in today's world. The emerging technic is altering the way in which we entertain, communicate, and store information. People's love for the new display mode can be demonstrated by the number of 3D silver screens worldwide from 2006 to 2014. The number of screens in 2014 reaches 64905, which is 22.3% higher than that of 2013 and is 251.6 times as much as that of 2006. Apart from 3D movies, there is a glut of non-cinematic 3D content that is making its way to consumers, especially over wireless networks such as 3D on mobile devices, 3D TVs and some emerging wearable equipment [1]. However, the wireless transmission mode and the spotty displays imply the presence of types of distortions, which will be received by consumers and degrade the viewing experience. So we need systems to monitor, control and improve the visual quality of stereoscopic presentations. Image quality assessment (IQA), due to its capability of simulating human visual perception to image quality, is usually used to solve this problem.

Numerous approaches for 2D IQA have been proposed over the last several decades. In the current research of IQA, image quality metrics are usually classified into full-reference (FR) [2], [3], [4], [5], reduced-reference (RR) [6], [7], and no-reference (NR) [8], [9] methods depending on the accessibility to the original reference points [10]. Some of them perform quite well in predicting the subjective ratings on popular image quality databases [11], [12], [13].

Following the research of monoscopic image quality metrics, more and more objective 3D metrics have been developed. We can classify the existing 3D IQA approaches into three categories: 1) [14], [15], [16], [17], etc evaluate stereoscopic images using 2D IQA metrics without the use of depth information; 2) [18], [19], etc measure stereoscopic images considering the depth information with 2D-IQA models; 3) [20], [21], [22], etc focus on building 3D quality models directly without relying on existing 2D-IQA algorithms. In addition, several databases of 3D image sets as important tools in the research of 3D IQA have been proposed, such as LIVE 3D IQA Database including the symmetric Phase I dataset [1] and the asymmetric Phase II dataset [23], which make us convenient to analyze the performance of a variety of 3D quality metrics.

The additional dimension of 3D content brings about many important issues. When human eyes focus on one point, the areas outside the cone angle form a blurry image because the areas outside the macular central fovea have few cones. And the areas remote from the focus point show blurry ghosting because of the binocular disparity between two eyes and the focusing mechanism of human eye. However, when presented with the stereograph, eyes are focusing on the fixed screen without refractive accommodation. It means that objects in the stereograph of different distances can be seen clearly at the same time. Also, two eyes perceive different contents. When dissimilar monocular stimuli are presented to corresponding retinal locations of the two eyes, the binocular rivalry occurs. In this paper, we take these factors into account. These above-mentioned factors may influence the perceived 3D quality. We resort to free energy principle to emulate the result of binocular rivalry and compute the inter-weight. It is worth mentioning that Zhai *et al*. first apply the brain theory into IQA research in [6] and introduce a new approach based on free-energy principle. Some brain theories have been unified within the free-energy framework. In [24] Frison indicates that the uncertainty will be removed by human to infer the meaningful part from visual stimuli during the inference process of human brain. It is natural that there exists a gap between the real scene and the brain's prediction due to the fact that the internal generative model cannot be universal. It is the gap that makes human "surprise", and thus attracts attention. In other words, when images presented to two eyes are different, the image causing more surprise draws more attention. And we believe that the surprise can account for the binocular rivalry. Besides, we do the spatial division with the depth information and



(a) (b)

Fig. 1: (a) and (b) show the extension of stereograph. Z-axis represents the depth of the stereograph. The area with larger z-value is closer to viewers. The grey-scale level of nine cubes represents the influence of the cube space on the quality assessment.

calculate the intra-weight. With inter- and intra-weights, we calculate the score of stereopairs.

The rest of this paper is organized as follows. Section II first presents the proposed IQA model. In Section III, the effectiveness of the dual-weighted algorithms is proved by comparison of the experimental results with those obtained by existing relevant models. Finally, concluding remarks are given in Section IV.

II. Dual-weighted Stereoscopic Image Quality Assessment

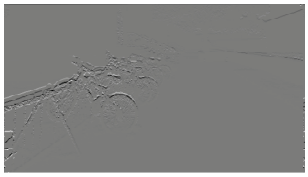
When we observe one object, two eyes will form a certain binocular convergence angle to see the object clearly. The areas outside the cone angle form a blurry image because the areas outside the macular central fovea have few cones. And the areas remote from the focus point show blurry ghosting because of the binocular disparity between two eyes and the focusing mechanism. However, when presented with stereograph, eyes are focusing on the fixed screen without refractive accommodation. It means that objects of different depth imformation in the screen can be seen clearly at the same time without blurry ghosting. So when we watch a stereograph, we can see clearly the objects of different distances. In this situation, we try to find those areas that are dominant in the quality assessment.

We think the quality of important areas have relatively larger influence on the final quality assessment. For the reason that the stereograph displays the artificial three-dimensional scene, we can see clearly the objects of different depth. So we think those areas that are closer to viewers and of better three-dimensional visual effects will play a more important role in the quality assessment. We use the depth map and take the spatial information into account:

$$w_{intra}^{(i)} = \sum_{j} \cos[\frac{\pi}{2} \cdot (1 - \frac{d_{i,j}}{d_{max} - d_{min}}) \cdot w_{xy}^{i}] + k \cdot g_i, \quad (1)$$

where the $d_{i,j}$ is the jth pixel of the ith block of the depth map. The first part of this equation can be illustrated by using Fig .1. The stereograph is expanded into three dimensions by adding the z-axis which represents the depth of stereograph. The area with larger z-value is closer to viewers. The grey-scale level of nine cubes represents the influence of the cube space on the quality assessment. We think the influence will change slowly if the object is close enough or far away. So the cosine function is used for calculation and the values of d_{max} and d_{min} are fixed. The w_{xy}^{i} is the weight of areas in x-y plane. We think the center regions of one image have greater influence considering the environment of subjective assessment that the human eye position is right in front of the center of screen. So the cube circled by the black rect in the Fig .1 (a) is given the largest weight.

(a) Original image



(b) Edge map

(c) Intra-weight map

Fig. 2: (a) is the left-view image of one stereopair, (b) is the edge map of the depth map and (c) is the intra-weight map blurred with Gaussian filter.

The second term of Eq. 1 helps to modify the weight. The g_i is derived from:

$$E(n,m) = \frac{1}{2\pi}(\frac{1}{\sigma_1^2} e^{-(x^2+y^2)/2\sigma_1^2} - \frac{1}{\sigma_2^2} e^{-(x^2+y^2)/2\sigma_2^2}) \quad (2)$$
$$* d(x,y),$$

$$g_i = \sum_{n,m} E_i'(n,m), \quad (3)$$

where the $E(n,m)$ is the edge map of the depth map $d(x,y)$ and the $g(i)$ is the sum of the ith block of the edge map. The edge of the depth map imply that the z-value of the object in the stereograph alters rapidly. In other words, the object is approximately perpendicular to the x-y plane or the object is isolated from the background. Those areas should have larger weights.

Those weights, obtained theoretically by calculation, are called intra-weights in this paper. Because they reflect the importance between the blocks in the stereograph. In Fig .2, (a) is the left-view image of one stereopair, (b) is the edge map of the depth map and (c) is the intra-weight map of blocks blurred with Gaussian filter. We can see that the areas around the bicycles are of higher importance for its higher grey level.

What is worth mentioning, when we conduct the experiments in this paper, the depth information is provided by the database. However, there is no need for us to use the very precise depth information to do the computation and those common depth-extraction algorithm will be enough if the depth is not provided.

A stereopair contains two images and the existing binocular disparity between them contributes to the three-dimensional visual effects. When dissimilar monocular stimuli are presented to corresponding retinal locations of the two eyes, the binocular rivalry occurs. And we want to figure out the weights between the two images to simulate the binocular rivalry.

Given an image signal I, the free-energy principle [24] suggests that the cognitive process is governed by an internal generative model $\mathcal{G}$ in the brain. Given different scenes or images, the model $\mathcal{G}$ will adapt itself through varying a parameter vector θ [6]. Then the "surprise" caused by image I can be computed by integrating the joint distribution $P(I, \theta|\mathcal{G})$ over the space of model parameter θ

$$-\log P(I|\mathcal{G}) \quad = \quad -\log \int P(I, \theta|\mathcal{G}) d\theta. \tag{4}$$

We can use an auxiliary posterior distribution $Q(\theta|I, \mathcal{G})$ to calculate the surprise of I in (4). Referring to [6], we can drop the latent model assumption $\mathcal{G}$ in our analysis for simplicity, since the behavior of the model can be characterized by parameter θ. By letting the auxiliary term into (4) and using Jensen's inequality we have

$$-\log P(I) \quad \leq \quad -\int Q(\theta|I) \log \frac{P(I, \theta)}{Q(\theta|I)} d\theta. \tag{5}$$

The right hand side of (5) is defined as the free energy:

$$F(I, \theta) \quad = \quad -\int Q(\theta|I) \log \frac{P(I, \theta)}{Q(\theta|I)} d\theta. \tag{6}$$

By noticing that $P(I, \theta) = P(\theta|I)P(I)$, we can write (6) into

$$\begin{aligned} F(I, \theta) &= \int Q(\theta|I) \log \frac{Q(\theta|I)}{P(\theta|I)P(I)} d\theta \\ &= -\log P(I) + \int Q(\theta|I) \log \frac{Q(\theta|I)}{P(\theta|I)} d\theta \\ &= E_Q[-\log P(I|\theta)] + KL\left(Q(\theta|I) \| P(\theta)\right). \end{aligned} \tag{7}$$

Here the term $KL\left(Q(\theta|I) \| P(\theta)\right)$ measures the distance between the recognition densities and the true prior of the model parameters. The term $E_Q[-\log P(I|\theta)]$ is the averaged entropy of predicting I.

We hope to compute the surprise. For operational amenability we can hypothesize the generative model $\mathcal{G}$ to be a 2D linear autoregressive (AR) model for its high description capability for natural images. The AR model is defined as

$$x_n \quad = \quad \boldsymbol{\chi}^k(x_n)\boldsymbol{\alpha} + \varepsilon_n \tag{8}$$

where x_n is the n_{th} pixel, $\boldsymbol{\chi}^k(x_n)$ is a row-vector that consists of k nearest neighbors of x_n, $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3, ..., \alpha_k)^T$ is the vector of AR coefficients and the ε_n is the error term. Under the large sample condition, the free energy equals the total description length of image I. So we estimate the AR coefficients by minimizing the description length

$$\hat{\boldsymbol{\alpha}} \quad = \quad \arg\min_{\alpha}(-\log P(I|\boldsymbol{\alpha}) + \frac{k}{2}\log N) \tag{9}$$

where N is the data sample size. We fix the order and the training set size of the model and thus turn the comparison process into residual minimization

$$\hat{\boldsymbol{\alpha}} \quad = \quad \arg\min_{\alpha} ||x - \mathbf{X}\boldsymbol{\alpha}||_2 \tag{10}$$

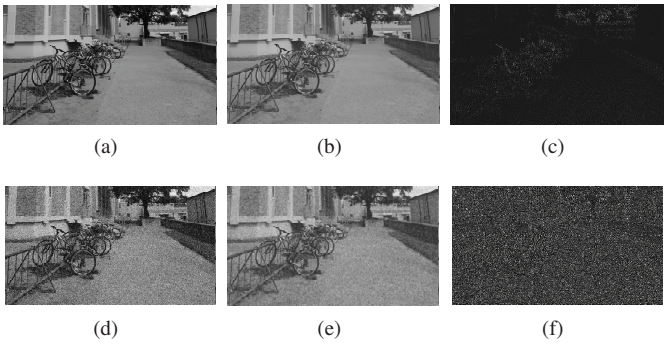


Fig. 3: Error map construction: (a) is of good quality, (d) is spoiled by white noise. (b) and (e) show the predicted images. (c) and (f) show the final computed error maps after scaling.

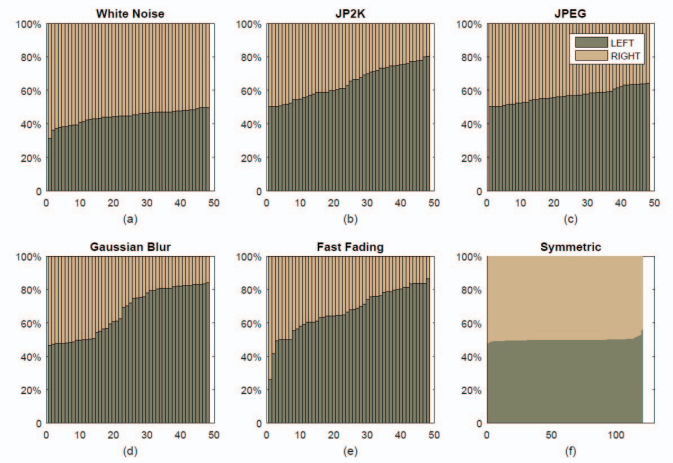


Fig. 4: Statistical analysis of the Phase II dataset: On Phase II dataset, the left-view images are of good quality or slightly spoiled, but the right-view images are almost heavily spoiled. (a) - (f) are the statistical results. Their x-axis represents the number of images and y-axis represents the ratio of $R(E)$ between stereopairs. Two colors of bar respectively represent left view and right view. (a) - (e) are results of asymmetric stereopairs with different kinds of distortions. (f) is the result of symmetric stereopairs.

where $x = (x_1, x_2, ..., x_N)^T$ and $\mathbf{X}(n, :) = \boldsymbol{\chi}^k(x_n)$. And the parameter can be solved as $\hat{\boldsymbol{\alpha}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T x$. In this case, the parameter θ of the model can be well described by $\hat{\boldsymbol{\alpha}}$.

Next, we use the input image I in a point-wise manner to estimate the predicted version I_p via $\boldsymbol{\chi}^k(x_n) \cdot \hat{\boldsymbol{\alpha}}$. The point-wise error ε_n can then be pooled to get the error map $E(I)$ and entropy of the errors can also be computed as $R(E) = \sum -P(\varepsilon)\log P(\varepsilon)$ with $P(\varepsilon)$ being the probability distribution of the errors.

By above analysis, the linear AR model is used to approximate the generative model, thereby to predict an image that the HVS perceives from the outside stimulus. And we compute the $R(E)$ from the predicted image which contributes to the surprise in human brain. We choose two images for analysis. One is the reference image from 3D LIVE database and the other is the spoiled version with white noise. In Fig. 3, (a)-(f) show the predicted images and the scaled error maps. We

TABLE I: Comparison of 3D IQA Models: Pearson linear correlation coefficient (PLCC) against DMOS on the **Phase I** dataset. Italics indicates an NR (blind) algorithm.

Algorithm	Type	JP2K	JPEG	WN	BLUR	FF	All
Benoit [18]	FR	0.940	0.641	0.925	0.949	0.747	0.903
Hewage [16]	RR	0.904	0.531	0.896	0.798	0.670	0.830
You [19]	FR	0.878	0.487	0.941	0.920	0.730	0.881
Gorley [26]	FR	0.485	0.312	0.796	0.853	0.365	0.451
Akhter [20]	NR	*0.906*	*0.729*	*0.905*	*0.618*	*0.660*	*0.427*
Dw-SSIM [2]	FR	0.875	0.495	0.928	0.911	0.723	0.872
Dw-FSIM [27]	FR	0.941	0.675	0.934	0.937	0.817	0.928
Dw-VIF [3]	FR	0.916	0.516	0.929	0.951	0.842	0.918
Dw-BRISQUE [8]	NR	*0.899*	*0.648*	*0.940*	*0.928*	*0.859*	*0.913*

TABLE II: Comparison of 3D IQA Models: Pearson linear correlation coefficient (PLCC) against DMOS on the **Phase II** dataset.

Algorithm	Type	JP2K	JPEG	WN	BLUR	FF	All
Benoit [18]	FR	0.784	0.853	0.926	0.535	0.807	0.748
Hewage [16]	RR	0.664	0.734	0.891	0.450	0.746	0.558
You [19]	FR	0.905	0.830	0.912	0.784	0.915	0.800
Gorley [26]	FR	0.372	0.322	0.874	0.934	0.706	0.515
Akhter [20]	NR	*0.776*	*0.786*	*0.722*	*0.795*	*0.674*	*0.568*
Dw-SSIM [2]	FR	0.766	0.666	0.939	0.911	0.880	0.817
Dw-FSIM [27]	FR	0.821	0.844	0.958	0.958	0.927	0.846
Dw-VIF [3]	FR	0.792	0.785	0.836	0.983	0.901	0.848
Dw-BRISQUE [8]	NR	*0.747*	*0.815*	*0.854*	*0.953*	*0.931*	*0.813*

TABLE III: Break down of performance on symmetrically and asymmetrically distorted stimuli in the **Phase II** dataset. Spearman rank-order correlation coefficient (SROCC) numbers are reported.

Algorithm	Type	Symmetric	Asymmetric
Benoit [18]	FR	0.860	0.671
Hewage [16]	RR	0.656	0.496
You [19]	FR	0.914	0.701
Gorley [26]	FR	0.383	0.056
Akhter [20]	NR	*0.420*	*0.517*
Dw-SSIM [2]	FR	0.852	0.792
Dw-FSIM [27]	FR	0.925	0.787
Dw-VIF [3]	FR	0.934	0.823
Dw-BRISQUE [8]	NR	*0.888*	*0.756*

intuitively deduce that the $R(E)$ of the image spoiled by white noise is larger because its error map seems of higher grey scale. Experimental results prove our deduction. Fig. 4 gives the statistical results on Phase II dataset. The figure shows that when an asymmetric stereopair is spoiled by white noise, the worse image will cause more surprise. However, other listed kinds of distortion lead to opposite results. Besides, the ratios of symmetrically distorted and asymmetrically white-noise stereopairs almost go halves, which implies that the convenient average-weighted method can perform well on the two datasets so as to reduce the computing time. And we find that there are some similarities between the entropy of error map and the standard deviation of the image.

In order to reduce the complexity of computation and motivated by the fact that the receptive fields of the human visual cortex are localized in space [25], we further partition the distorted image into $M \times N$ patches. We calculate the standard deviation of the divided error map first. For those blocks with larger values we replace the standard deviation with local entropy:

$$R^{'} = \Sigma_i \, \eta_i, \quad \eta_i \in \{\alpha\sigma_i + \beta, \; R(E_i)\} \tag{11}$$

In (11) σ_i is the standard deviation of i_{th} block and $R(E_i)$ is entropy of local error map. We use the term $\alpha\sigma_i + \beta$ is because that there is a roughly linear relation between σ_i and $R(E_i)$ according to observation of experimental results. So we can get $R^{'}$ by partitioning distorted image, local computation to obtain $R(E_i)$ and σ_i and doing the appropriate substitution. And the inter-weights between the left view and right view can be figured out by:

$$\begin{aligned} w^l_{inter} &= R^{'}_l/(R^{'}_l + R^{'}_r), \\ w^r_{inter} &= R^{'}_r/(R^{'}_l + R^{'}_r). \end{aligned} \tag{12}$$

The final assessment will be done by:

$$S_{3D} = \Sigma_{l,r}(\Sigma_i w^i_{intra} \cdot S^i_{2D}) \cdot w^{l,r}_{inter}, \tag{13}$$

where the S^i_{2D} is the score predicted by 2D quality assessment methods in the ith block. And after the dual-weighted process, we can get S_{3D}. The diagram of proposed model is shown in Fig. 5.

III. Experiments and Analysis

We use LIVE 3D Image Quality Database to validate the effectiveness of our method. This database was constructed in two phases. Phase I contains symmetrically distorted stimuli while Phase II has both symmetrically and asymmetrically

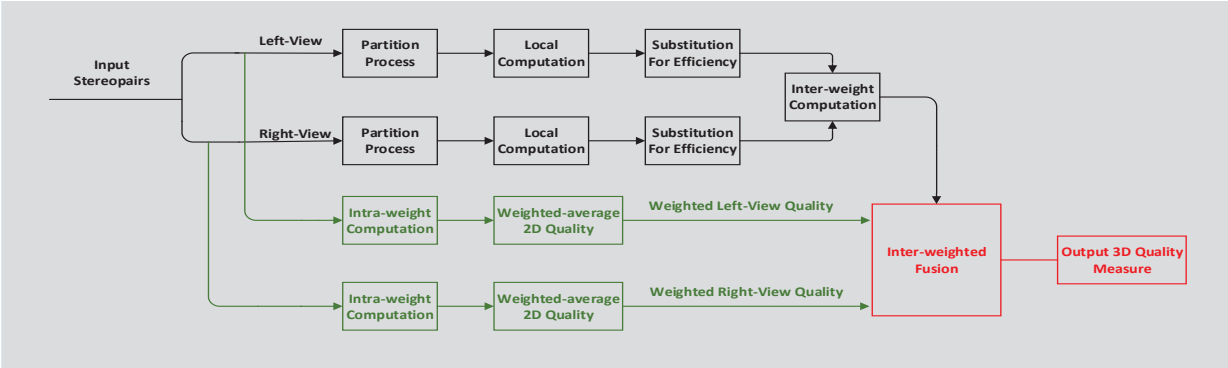


Fig. 5: Diagram of our proposed dual-weighted model: The green portion calculates the intra-weight and then compute the weighted 2D quality. The black portion shows the process to get the inter-weight between two views. The red portion calculates averaged scores with inter-weights and intra-weighted 2D scores and finally gets the result.

distorted stimuli. Both phases used five types of distortions: compression using the JPEG and JPEG2000 compression standards, additive white Gaussian noise, Gaussian blur and a fast-fading model. Phase I has 20 pristine stereopairs and 365 distorted stereopairs, while Phase II has 8 pristine stereopairs and 360 distorted stereopairs.

Pearson linear correlation coefficient (PLCC) and Spearman rank-order correlation coefficient (SROCC) are used to evaluate performance of our approach. PLCC can be considered as a measure of prediction accuracy, while SROCC measures the monotonicity by ignoring the relative distance between the data. The higher SROCC and PLCC values indicate better performance in terms of correlation with human opinion.

We experimented on the Phase I dataset to compare the performance of dealing with symmetrically distorted stereopairs. We apply our dual-weighted method to the 2D IQA models: SSIM[2], FSIM[27], VIF[3] and BRISQUE[8]. We compared the performance of those dual-weighted algorithms with several 3D IQA models: Benoit[18], Hewage[16], You[19], Gorley[26] and Akther[20]. For the symmetrically distorted stereopairs, the inter-weights almost go halves because there is little or no binocular rivalry in the symmetrically distorted stimuli. Our algorithms achieve high performance on the Phase I dataset as shown in Table I. Among individual distortion types, the method is effective except for the JPEG. We research the reason why the proposed method is not effective for the JPEG. In [28] subjective tests show that for JPEG compression distorted stimuli, range, disparity, and luminance activities of distorted patches in the low DMOS group are significantly higher than those in the high DMOS group. It means that binocular suppression, luminance and range masking also have effects on the viewing of stereoscopic images which are not considered in our method.

The Phase II dataset has both symmetrically and asymmetrically distorted stimuli. Tables II-III show the results against the mixed dataset. The inter-weight plays a role here. The dual-weighted algorithms deliver good performance compared against other algorithms as shown in Tables II. Besides, we further compare our method with other algorithms with the completely asymmetrically distorted stimuli. And the results in Tables III show that the dual-weighted methods are effective when predicting the symmetrically or asymmetrically distorted stereopairs.

What is worth mentioning, the proposed dual-weight method is a universal model since the component within our framework can be replaced with other more effective dedicated metrics to improve the prediction accuracy of algorithm.

IV. Conclusion

In this paper, we try to investigate and develop the 2D image quality metrics into stereoscopic image quality assessment using the dual-weight model. The resulting method focuses on the acquiring of intra-weight based on the properties of HVS and the inter-weight caused by binocular rivalry based on the brain theory. Under both symmetrically and asymmetrically distorted conditions, our algorithms deliver competitive performance on the 3D LIVE Database. And what we propose is a universal model since the component within our framework can be replaced with other more effective dedicated metrics.

Acknowledgment

This work was supported in part by the National Science Foundation of China under Grants 61422112, 61371146, 61521062, 61527804, the Foundation for the Author of National Excellent Doctoral Dissertation of China under Grant 201339 and National High-tech R&D Program of China under Grant 2015AA015905.

References

[1] A. K. Moorthy, C.-C. Su, A. Mittal, and A. C. Bovik, "Subjective evaluation of stereoscopic image quality," *Signal Processing: Image Communication*, vol. 28, no. 8, pp. 870–883, 2013.

[2] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *Image Processing, IEEE Transactions on*, vol. 13, no. 4, pp. 600–612, 2004.

[3] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *Image Processing, IEEE Transactions on*, vol. 15, no. 2, pp. 430–444, 2006.

[4] G. Zhai, W. Zhang, X. Yang, S. Yao, and Y. Xu, "Ges: a new image quality assessment metric based on energy features in gabor transform domain," in *Circuits and Systems, 2006. ISCAS 2006. Proceedings. 2006 IEEE International Symposium on*. IEEE, 2006, pp. 4–pp.

[5] K. Gu, G. Zhai, X. Yang, and W. Zhang, "An efficient color image quality metric with local-tuned-global model," in *Image Processing (ICIP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 506–510.

[6] G. Zhai, X. Wu, X. Yang, W. Lin, and W. Zhang, "A psychovisual quality metric in free-energy principle," *Image Processing, IEEE Transactions on*, vol. 21, no. 1, pp. 41–52, 2012.

[7] 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.

[8] 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.

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

[10] 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.

[11] 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

[12] N. Ponomarenko, O. Ieremeiev, V. Lukin, K. Egiazarian, L. Jin, J. Astola, B. Vozel, K. Chehdi, M. Carli, F. Battisti *et al.*, "Color image database tid2013: Peculiarities and preliminary results," in *Visual Information Processing (EUVIP), 2013 4th European Workshop on*. IEEE, 2013, pp. 106–111.

[13] K. Gu, G. Zhai, W. Lin, and M. Liu, "The analysis of image contrast: From quality assessment to automatic enhancement," *IEEE Trans. Cybernetics*, vol. 45, 2015.

[14] J. Wang, A. Rehman, K. Zeng, S. Wang and Z. Wang, "Quality prediction of asymmetrically distorted stereoscopic 3D images," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 3400-3414, Nov. 2015.

[15] Jiheng Wang, Kai Zeng, and Zhou Wang, "Quality prediction of asymmetrically distorted stereoscopic images from single views," in *Multimedia and Expo (ICME), 2014 IEEE International Conference on*. IEEE, 2014, pp. 1–6.

[16] C. Hewage, S. T. Worrall, S. Dogan, and A. Kondoz, "Prediction of stereoscopic video quality using objective quality models of 2-d video," *Electronics letters*, vol. 44, no. 16, pp. 963–965, 2008.

[17] Y. Zhu, G. Zhai, K. Gu, and M. Liu, "Blindly Evaluating Stereoscopic Image Quality with Free-Energy Principle," in *Circuits and Systems (ISCAS), 2016 IEEE International Symposium on*. [accepted].

[18] A. Benoit, P. Le Callet, P. Campisi, and R. Cousseau, "Quality assessment of stereoscopic images," *EURASIP journal on image and video processing*, vol. 2008, pp. Article–ID, 2008.

[19] J. You, L. Xing, A. Perkis, and X. Wang, "Perceptual quality assessment for stereoscopic images based on 2d image quality metrics and disparity analysis," in *Proc. of International Workshop on Video Processing and Quality Metrics for Consumer Electronics, Scottsdale, AZ, USA*, 2010.

[20] R. Akhter, Z. P. Sazzad, Y. Horita, and J. Baltes, "No-reference stereoscopic image quality assessment," in *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 2010, pp. 75 240T–75 240T.

[21] K. Gu, G. Zhai, X. Yang, and W. Zhang, "No-reference stereoscopic iqa approach: From nonlinear effect to parallax compensation," *Journal of Electrical and Computer Engineering*, vol. 2012, 2012.

[22] F. Shao, W. Lin, S. Wang, G. Jiang, and M. Yu, "Blind image quality assessment for stereoscopic images using binocular guided quality lookup and visual codebook."

[23] M.-J. Chen, L. K. Cormack, and A. C. Bovik, "No-reference quality assessment of natural stereopairs," *Image Processing, IEEE Transactions on*, vol. 22, no. 9, pp. 3379–3391, 2013.

[24] K. Friston, "The free-energy principle: a unified brain theory?" *Nature Reviews Neuroscience*, vol. 11, no. 2, pp. 127–138, 2010.

[25] D. H. Hubel and T. N. Wiesel, "Receptive fields and functional architecture of monkey striate cortex," *The Journal of physiology*, vol. 195, no. 1, pp. 215–243, 1968.

[26] P. Gorley and N. Holliman, "Stereoscopic image quality metrics and compression," in *Electronic Imaging 2008*. International Society for Optics and Photonics, 2008, pp. 680 305–680 305.

[27] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "Fsim: a feature similarity index for image quality assessment," *Image Processing, IEEE Transactions on*, vol. 20, no. 8, pp. 2378–2386, 2011.

[28] M.-J. Chen, L. K. Cormack, and A. C. Bovik, "Distortion conspicuity on stereoscopically viewed 3d images may correlate to scene content and distortion type," *Journal of the Society for Information Display*, vol. 21, no. 11, pp. 491–503, 2013.