Measuring Depth Perception for Stereoscopic Images Using a Phase-shift Model

Qiuping Jiang, Shanshan Wang, Kemeng Li, Feng Shao*
Faculty of Information Science and Engineering, Ningbo University, 315211, Ningbo, China
Email: shaofeng@nbu.edu.cn

***Abstract*—We present an effective depth perception assessment index for stereoscopic images using a phase-shift model. To be more specific, we use Gabor filter to compute the responses of left and right images respectively, and proposed a phase-shift model for computing disparity maps based on phase gradient and phase difference information. Then, quality score for depth perception is obtained by measuring the similarity between the estimated disparities of the original and distorted stereoscopic images. Experimental results on two publicly 3D image quality assessment databases demonstrate that, in comparison with the most related existing methods, the devised algorithm achieves high consistency alignment with subjective assessment.**

***Index Terms*—Depth perception, phase-shift, quality assessment, binocular energy.**

I. INTRODUCTION

In recent years, there has been great progress in developing objective image quality assessment (IQA) metrics [1]. However, the development of three-dimensional (3D) image/video quality index is still in its early stage. Assessing the 3D image quality is a challenging issue because it is affected by image quality, depth perception and visual comfort [2,3]. It is particularly challenging how to evaluate the perceived quality of depth perception and visual comfort when the stereoscopic image consists of two views with different quality.

For measuring the perceived quality of stereoscopic images, several metrics have been proposed. Hwang *et al.* [4] devised a visual attention and depth assisted stereo image quality assessment by fusing the impact of stereo attention predictor, depth variation and stereo distortion predictor. Bensalma *et al.* [5] devised a Binocular Energy Quality Metric (BEQM) by modeling the complex cells responsible for the construction of the binocular energy, and evaluated the similarity between the binocular energy maps of the original and the distorted stereo-pairs. Chen *et al.* [6] constructed a “Cyclopean” view from the stereo-pair, and evaluated the ‘Cyclopean’ view by 2D quality metrics. De Silva *et al.* [7] proposed a quality metric for compressed stereoscopic video by extracting features that quantify the compression artifacts. Other relevant works can be found in [8-9].

Some studies have been conducted on visual comfort evaluation for stereoscopic images [10]. Wopking found that disparity magnitude is the main factor for visual fatigue by subjective experiments [11], and oversized horizontal parallax may exceed the Panum’s binocular fusion limit and then cause blurred vision and diplopia phenomenon. Seuntiens *et al.* found that the crosstalk between left and right images may also induce visual discomfort [12]. Speranza *et al.* found that motion is another factor of visual comfort for 3D videos [13]. Lambooij *et al.* proposed a visual comfort prediction model by using average disparity magnitude and global disparity as features [14]. Nojiri *et al.* designed an effective objective method by considering parallax distribution and depth motion for 3D videos [15]. Perceptual visual attention models are taken into account in the visual comfort prediction methods [16,17].

Several works in the literatures have been proposed in evaluating depth perception. Lebreton *et al.* characterized 3D materials on different depth perception scales [18]. Faria *et al.* proposed a stereoscopic depth perception approach inspired by the primary visual cortex using the stimulus response of the receptive field [19]. Currently, some metrics evaluated the two views of the stereoscopic images, disparity/depth images separately by 2D-IQA metrics, and then combined them into an overall score. Boev *et al.* [20] combined monoscopic and stereoscopic quality components from the ‘Cyclopean’ image and disparity map respectively for stereo-video evaluation. Benoit *et al.* [21] computed quality scores of both stereo-pair and the disparity map by 2D quality metrics, and then combined them to produce a final score. You *et al.* [22] investigated ten common 2D quality evaluators on a stereo-pair and on its disparity map, and found the optimal combination means which can yield the best performance.

In this paper, we proposed a depth perception assessment index for stereoscopic images using a phase-shift model. The main contributions of the paper are as follows: 1) we focus on phase-shift receptive field mechanism for depth perception computation; 2) we calculate the phase difference similarity between the estimated disparity shifts as the final quality index; 3) we demonstrate that depth perception is not the main visual cue in evaluating 3D visual quality under poor quality stereoscopic images. The rest of the paper is organized as follows. Section II presents the proposed depth perception assessment index. The experimental results are given and discussed in Section III, and finally conclusions are drawn in Section IV.

doi:10.4304/jsw.9.10.2665-2671

II. Proposed Depth Perception Assessment Index

We give a short overview on energy neurons modeling disparity-tuned cells in the visual cortex [23]. The output of a simple receptive field is formulated mathematically as a convolution of image I with the receptive field function f_v,

$$C_v(x,y,\varphi)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_v(x-\xi,y-\eta,\varphi)I(\xi,\eta)d\xi d\eta \quad (1)$$

Binocular complex cells combine the output of the receptive fields of left and right images as [24]

$$\begin{aligned} E_v &= \left\| C_{lv}+C_{rv} \right\|^2 \\ &= \left(\mathrm{Re}[C_{lv}]+\mathrm{Re}[C_{rv}]\right)^2+\left(\mathrm{Im}[C_{lv}]+\mathrm{Im}[C_{rv}]\right)^2 \end{aligned} \quad (2)$$

where Re[] and Im[] are real and imaginary parts of the response.

Currently, two different models have been proposed for complex cells tuned by disparity [25]

1) Phase-shift: corresponding left and right receptive fields have different phases, e.g.,

$$E_{Dv}(x,y,\varphi)=\left\| C_{lv}(x,y,\varphi+2\pi v\frac{D}{2})+C_{rv}(x,y,\varphi-2\pi v\frac{D}{2}) \right\|^2 \quad (3)$$

2) Position-shift: corresponding left and right receptive fields are centered at shifted positions, e.g.,

$$E_{Dv}(x,y,\varphi)=\left\| C_{lv}(x+\frac{D}{2},y,\varphi)+C_{rv}(x-\frac{D}{2},y,\varphi) \right\|^2 \quad (4)$$

In the phase-shift model, the preferred disparity can be estimated by $D=\Delta\phi_{lr}/2\pi v$, where $\Delta\phi_{lr}=\phi_l-\phi_r$ being the phase difference between the left and right images, and ν is the central frequency of the cell. Also, the preferred disparity can be given by $D=\hat{d}$, where $\hat{d}$ is the position shift that maximizes the E_{Dv}.

However, the above binocular energy model is not directly suitable for the distorted stereoscopic images, because depth perception (with stimuli from the position shift) may not be able to correctly characterized in the case. The analysis results in our previous work [8] have revealed that phase shift between two views provides the main cue for binocular disparity; then, if the disparity of the input is not matched to the preferred disparity, the phase shift will tend not to be zero, and the binocular energy response will change. In order to measure depth perception, the existing technologies [20-22] directly evaluate the quality of the estimated disparity maps. However, position-shift mechanism is not suited for disparity computation for the distorted stereoscopic images.

In a mathematical model of depth perception, the neurons can be treated by 2D filter tuned to different scales and orientations. In this work, we use Gabor filter to compute the responses of left and right images, respectively. Firstly, we define the phase gradient extended from 2D case as

$$\begin{aligned} \mathbf{k}_{org}(x,y;\omega_m,\theta) &= \left[k_x^{org}(x,y;\omega_m,\theta), k_y^{org}(x,y;\omega_m,\theta)\right] \\ &= \left[\frac{\partial\phi_{lr}^{org}(x,y;\omega_m,\theta)}{\partial x}, \frac{\partial\phi_{lr}^{org}(x,y;\omega_m,\theta)}{\partial y}\right] \end{aligned} \quad (5)$$

where the phase derivatives are taken over the mean phase of the complex filter responses in the left and right images

$$\phi_{lr}^{org}(x,y;\omega_m,\theta)=\frac{\phi_l^{org}(x,y;\omega_m,\theta)+\phi_r^{org}(x,y;\omega_m,\theta)}{2} \quad (6)$$

where $\phi_l^{org}=\arctan\left(\frac{\mathrm{Im}(C_{lv}^{org})}{\mathrm{Re}(C_{lv}^{org})}\right)$ and $\phi_r^{org}=\arctan\left(\frac{\mathrm{Im}(C_{rv}^{org})}{\mathrm{Re}(C_{rv}^{org})}\right)$ being the phases of left and right images, respectively.

Also, the phase difference between the left and right images is calculated as

$$\Delta\varphi_{LR}^{org}(x,y;\omega_m,\theta)=\varphi_L^{org}(x,y;\omega_m,\theta)-\varphi_R^{org}(x,y;\omega_m,\theta) \quad (7)$$

Taking into account the above equations, the estimated disparity vector is computed as the following Eq.(8).

Since the estimated disparity vectors $\mathbf{d}_{org}(x,y;\omega_m)$ and $\mathbf{d}_{dis}(x,y;\omega_m)$ are detected as phase shifts in the spectrum, we measure the similarity between the disparity vectors. Specifically, we define the phase difference between them as in Eq.(9).

Here, $\langle\ \rangle$ represents the inner product of their gradient values in x and y direction, and T_1 is a positive constant to increase the stability of the phase difference. In this paper, T_1=0.85.

$$\begin{aligned} \mathbf{d}_{org}(x,y;\omega_m) &= \left[\Delta d_x^{org}(x,y;\omega_m), \Delta d_y^{org}(x,y;\omega_m)\right] \\ &= \begin{bmatrix} \sum_{\theta=1}^{4} k_x^{org}(x,y;\omega_m,\theta)\cdot k_x^{org}(x,y;\omega_m,\theta) & \sum_{\theta=1}^{4} k_x^{org}(x,y;\omega_m,\theta)\cdot k_y^{org}(x,y;\omega_m,\theta) \\ \sum_{\theta=1}^{4} k_y^{org}(x,y;\omega_m,\theta)\cdot k_x^{org}(x,y;\omega_m,\theta) & \sum_{\theta=1}^{4} k_y^{org}(x,y;\omega_m,\theta)\cdot k_y^{org}(x,y;\omega_m,\theta) \end{bmatrix}^{-1} \\ &\times \begin{bmatrix} \sum_{\theta=1}^{4} k_x^{org}(x,y;\omega_m,\theta)\cdot \Delta\varphi_{LR}^{org}(x,y;\omega_m,\theta) \\ \sum_{\theta=1}^{4} k_y^{org}(x,y;\omega_m,\theta)\cdot \Delta\varphi_{LR}^{org}(x,y;\omega_m,\theta) \end{bmatrix} \end{aligned} \quad (8)$$

$$
\begin{aligned}
Q_{DP}(x,y) &= \arccos\left(\frac{\left\langle \mathbf{d}_{org}(x,y;\omega_m),\mathbf{d}_{dis}(x,y;\omega_m)\right\rangle + T_1}{\left\|\mathbf{d}_{org}(x,y;\omega_m)\right\|\cdot\left\|\mathbf{d}_{dis}(x,y;\omega_m)\right\| + T_1}\right) \\
&= \arccos\left(\frac{\left(\Delta d_x^{org}(x,y;\omega_m)\cdot\Delta d_x^{dis}(x,y;\omega_m) + \Delta d_y^{org}(x,y;\omega_m)\cdot\Delta d_y^{dis}(x,y;\omega_m)\right) + T_1}{\sqrt{\left(\Delta d_x^{org}(x,y;\omega_m)\right)^2 + \left(\Delta d_y^{org}(x,y;\omega_m)\right)^2}\cdot\sqrt{\left(\Delta d_x^{dis}(x,y;\omega_m)\right)^2 + \left(\Delta d_y^{dis}(x,y;\omega_m)\right)^2} + T_1}\right)
\end{aligned}
\tag{9}
$$

Finally, quality score of the depth perception is decided via averaging individual quality scores of pixels

$$
Q = \frac{\sum_{(x,y)\in\Omega} Q_{DP}(x,y)}{N} \tag{10}
$$

where N is the number of pixels of the image.

The phase difference between the estimated disparity vectors reflects the range of distortion degrees in an image. The higher estimated value, the larger distortion rang, and thus the lower depth perception quality. Here we present one example to illustrate this point above. The first row of Fig.1 shows: (a) JPEG compressed version, (b) JPEG2000 compressed version, (c) Gaussian blurred version, and (d) White Noise distorted version of left image of ‘Balloons’ test sequences from NBU IQA database. The second row of Fig.1 shows the estimated disparity maps from the stereoscopic images in the first row (used the stereo matching algorithm in [26]). The third row of Fig.1 shows the phase difference maps from the stereoscopic images in the first row by the proposed method. The difference mean opinion scores (DMOS) values for the Gaussian blurred, JPEG compressed, JPEG2000 compressed and White Noise distorted stereoscopic images are 30.6087, 31.4783, 29.4348 and 30.1304, respectively, that is, the subjective measures for these distorted stereoscopic images are similar. It is clearly demonstrated that the evaluated quality of the estimated disparity by stereo matching have weak correlations with the subjective perceived quality, while the proposed quality scores are more consistent with the DMOS values.

III. Experimental Results And Analyses

A. *Databases And Performance Measures*

In the experiment, two publicly available 3D IQA databases: NBU 3D IQA Database [8], and LIVE 3D IQA Phase I Database [6], are used to verify the performance of the proposed metric for stereoscopic images. The NBU 3D IQA Database consists of 312 distorted stereoscopic pairs generated from 12 reference stereoscopic images. Five types of distortions, JPEG,

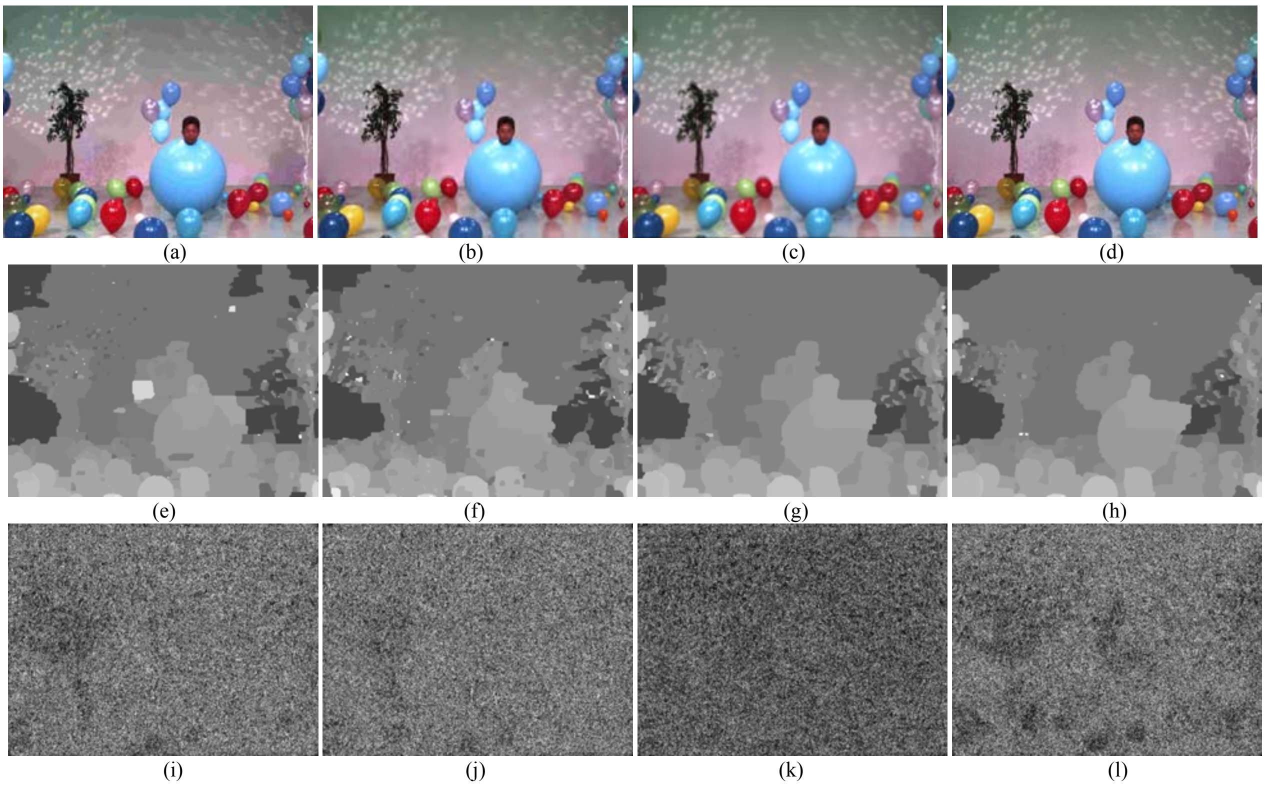
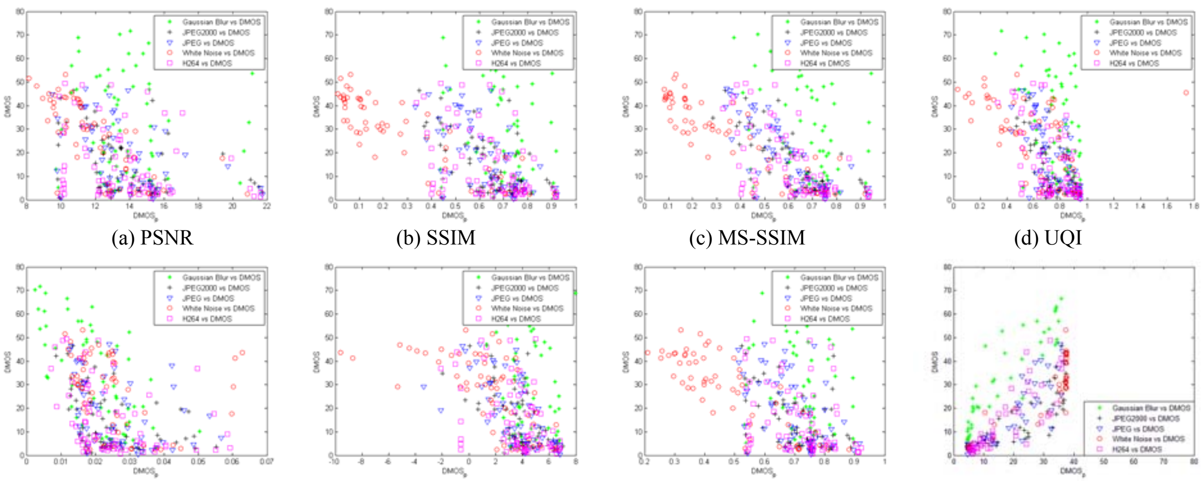


Figure 1. Examples of quality degraded left images of ‘Balloons’ test sequences from NBU 3D IQA database: (a) JPEG compressed version; (b) JPEG2000 compressed version; (c) Gaussian blurred version; (d) White Noise distorted version; (e) The estimated disparity map of (a); (f) The estimated disparity map of (b); (g) The estimated disparity map of (c); (h) The estimated disparity map of (d); (i) The calculated phase difference map of (a); (j) The calculated phase difference map of (b); (k) The calculated phase difference map of (c); (l) The calculated phase difference map of (d).

TABLE 1.
PERFORMANCE OF THE PROPOSED METHOD AND THE OTHER SCHEMES ON THE TWO DATABASES.

IQA model	NBU (312 images)				LIVE I (365 images)				average			
	PLCC	SRCC	KRCC	RMSE	PLCC	SRCC	KRCC	RMSE	PLCC	SRCC	KRCC	RMSE
PSNR	0.3935	0.3926	0.2810	15.7935	0.2574	0.2488	0.1680	15.8448	0.3255	0.3207	0.2245	15.8192
SSIM	0.4354	0.4510	0.3201	15.4659	0.2921	0.2786	0.1895	15.6825	0.3638	0.3648	0.2548	15.5742
MS-SSIM	0.5210	0.5482	0.4010	14.6641	0.2950	0.2948	0.2005	15.6678	0.4080	0.4215	0.3008	15.1660
UQI	0.4889	0.4847	0.3412	14.9859	0.3293	0.2741	0.1888	15.4827	0.4091	0.3794	0.2650	15.2343
VIF	0.5681	0.4859	0.3374	14.1381	0.3423	0.3094	0.2106	15.4071	0.4552	0.3977	0.2740	14.7726
VSNR	0.4746	0.4776	0.3386	15.1215	0.3183	0.2696	0.1863	15.5452	0.3965	0.3736	0.2625	15.3334
FSIM	0.4043	0.4148	0.2949	15.7125	0.2090	0.1711	0.1108	16.0354	0.3067	0.2930	0.2029	15.8740
Proposed	**0.7559**	**0.7969**	**0.5824**	**10.6624**	**0.8373**	**0.8396**	**0.6327**	**8.9654**	**0.7966**	**0.8183**	**0.6076**	**9.8139**



(a) PSNR (b) SSIM (c) MS-SSIM (d) UQI

(e) VIF (f) VSNR (g) FSIM (h) Proposed

Figure 2. Scatter plots of predicted quality scores against the subjective scores (DMOS) of the eight methods on the NBU 3D IQA databases.

JP2K, Gblur, WN and H.264, are symmetrically applied to the left and right reference stereoscopic images at various levels. The LIVE 3D IQA Phase I Database consists of 365 distorted stereoscopic pairs generated from 20 reference stereoscopic images. Five types of distortions, JPEG, JP2K, Gblur, WN and FF, are symmetrically applied to the left and right reference stereoscopic images at various levels for the LIVE 3D IQA Phase I Database.

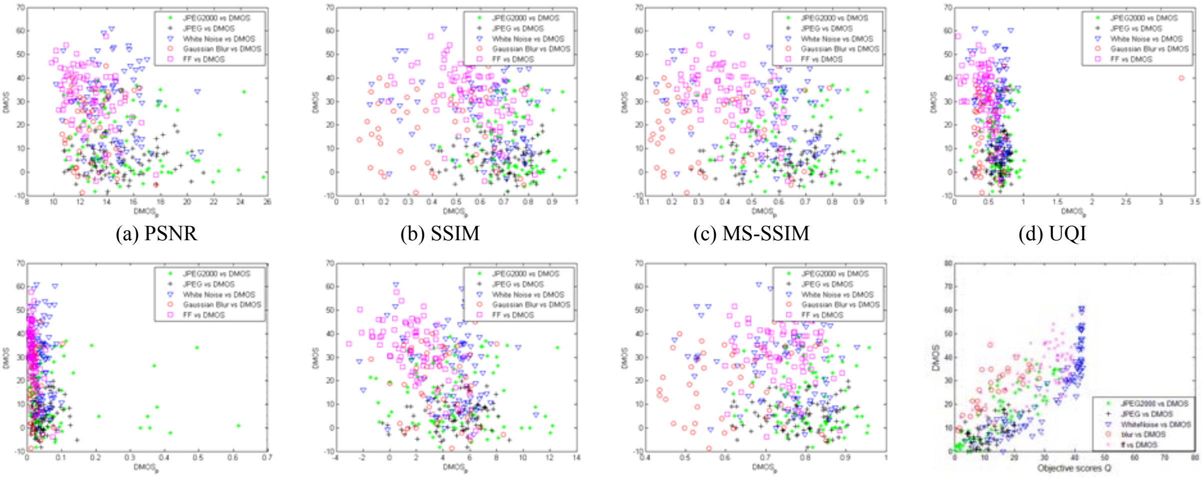
In the paper, four commonly-used performance indicators are used to benchmark the proposed metric against the relevant state-of-the-art techniques: Pearson linear correlation coefficient (PLCC), Spearman rank order correlation coefficient (SRCC), Kendall rank-order correlation coefficient (KRCC), and root mean squared error (RMSE), between the objective and subjective scores. For a perfect match between the objective and subjective scores, PLCC=SRCC=KRCC=1, and RMSE=0. For the nonlinear regression, we use the following five-parameter logistic function [27]:

$$DMOS_p = \beta_1 \cdot \left(\frac{1}{2} - \frac{1}{1+\exp(\beta_2 \cdot (x-\beta_3))} \right) + \beta_4 \cdot x + \beta_5 \quad (11)$$

where β_1, β_2, β_3, β_4 and β_5 are determined by using the subjective scores and the objective scores.

B. *Overall Assessment Performance*

For evaluating the proposed algorithm, the relevant existing 2D-IQA schemes, e.g., PSNR, SSIM [28], MS-SSIM [29], UQI [30], VIF [31], VSNR [32] and FSIM [33], have been compared. In order to characterize depth perception, the 2D-IQA schemes directly predict the quality of the estimated disparity maps (stereo matching algorithm [26] is used in this paper). The PLCC, SRCC, KRCC and RMSE of the eight schemes on the two databases are given in Table.1, where the three best metrics have been highlighted in boldface. It is clearly shown that directly evaluating disparity quality does not improve the performance, because the quality of the estimated disparity highly dependent on the stereo matching algorithm, and the 2D quality metric for disparity maps does not align with the human perception of disparity. Fig.2 and Fig.3 show the scatter plots of predicted quality scores against subjective quality scores (in terms of DMOS) for the eight schemes on the two databases, respectively. From the figures we find that, in the case of low quality stereoscopic images (e.g., DMOS is larger than 40), the predicted depth perception quality is limited within a certain range (the predicted DMOS is lower than 40) by the proposed scheme. The phenomenon is acceptable because depth perception is not the main visual cue in 3D perception for the low quality case; that is, image quality and visual comfort will dominate the 3D



(a) PSNR (b) SSIM (c) MS-SSIM (d) UQI

(e) VIF (f) VSNR (g) FSIM (h) Proposed

Figure 3. Scatter plots of predicted quality scores against the subjective scores (DMOS) of the eight methods on the LIVE 3D IQA databases.

TABLE 2
PERFORMANCE COMPARISON OF THE EIGHT SCHEMES ON EACH INDIVIDUAL DISTORTION TYPE IN TERMS OF PLCC.

	Criteria	PSNR	SSIM	MS-SSIM	UQI	VIF	VSNR	FSIM	Proposed
NBU	JPEG	0.5174	0.6688	0.8289	0.6388	0.5257	0.6800	0.5824	**0.8746**
	JP2K	0.5338	0.6055	0.7461	0.7233	0.6200	0.6526	0.5534	**0.8213**
	Gblur	0.1712	0.2616	0.2754	0.3983	0.6422	0.4281	0.1462	**0.9384**
	WN	0.7942	0.8809	0.8916	0.6249	0.4564	0.7373	0.8550	**0.9343**
	H246	0.2806	0.4238	0.6139	0.5163	0.5313	0.4427	0.3143	**0.8374**
LIVE I	JPEG	0.7091	0.0226	0.3636	0.5496	0.4023	0.4550	0.1220	**0.8994**
	JP2K	0.6815	0.6887	0.8492	0.8647	0.8675	0.8004	0.5493	**0.9257**
	Gblur	0.7091	0.7472	0.8301	0.8636	0.9206	0.8318	0.7473	**0.9278**
	WN	0.7656	0.8888	**0.9027**	0.7571	0.7685	0.7475	0.8407	0.6269
	FF	0.4226	0.5419	0.5837	0.6396	0.6451	0.6769	0.2082	**0.7074**

TABLE 3
PERFORMANCE COMPARISON OF THE EIGHT SCHEMES ON EACH INDIVIDUAL DISTORTION TYPE IN TERMS OF SRCC.

	Criteria	PSNR	SSIM	MS-SSIM	UQI	VIF	VSNR	FSIM	Proposed
NBU	JPEG	0.4977	0.6520	0.8021	0.6376	0.4798	0.6900	0.5425	**0.8875**
	JP2K	0.4097	0.5350	0.7208	0.6086	0.5240	0.5610	0.4639	**0.8976**
	Gblur	0.0951	0.1413	0.2282	0.3295	0.6048	0.2003	0.0812	**0.9403**
	WN	0.7761	0.8504	0.8760	0.6298	0.4326	0.7123	0.8056	**0.8876**
	H246	0.2718	0.4195	0.5997	0.4765	0.4469	0.4117	0.3512	**0.8829**
LIVE I	JPEG	0.1082	0.0288	0.3259	0.5155	0.3672	0.3014	0.0724	**0.8780**
	JP2K	0.6658	0.6667	0.8222	0.8285	0.8273	0.7720	0.5075	**0.8892**
	Gblur	0.6177	0.5794	0.7775	0.8076	0.8810	0.7677	0.4705	**0.9289**
	WN	0.7632	0.8852	**0.8979**	0.6784	0.7348	0.7432	0.8425	0.5253
	FF	0.3779	**0.5327**	0.5102	0.5815	0.5164	0.5668	0.0471	0.5196

perception quality in assessment (in agreement with the subjective observations in [34]).

C. *Performance On Individual Distortion Types*

To more comprehensively evaluate the prediction performance of the proposed method, we compare the eight schemes on each type of distortion. The PLCC and SRCC results are listed in Tables.2-3, where the best metrics have been highlighted in boldface. One can see that the proposed scheme obtains the best performance on the most of the distortion types. Even though some 2D metrics may perform well on some specific types of distortions, e.g., MS-SSIM is an effective measure for noisy images, but it is not able to faithfully measure the quality of images impaired by other types of distortions.

This validates that depth quality does not directly equal to the quality of estimated disparity maps.

V. CONCLUSIONS

In this paper, we devised a satisfactory computational model to provide an effective depth perception assessment index for stereoscopic images. More specifically, we use Gabor filter at different scales and orientations to compute the responses of left and right images. Then, we proposed a phase-shift model for computing disparity maps that is based on phase gradient and phase difference information. Finally, we calculate the similarity between the estimated disparity maps of the original and distorted stereoscopic images as the final quality index. Compared with state-of-the-art 2D image quality assessment (2D-IQA), the proposed metric performs better in terms of both accuracy and efficiency on two publicly available 3D IQA databases. In the future work, we will further explore how to combine image quality, depth perception and visual comfort in modeling 3D visual perception.

ACKNOWLEDGMENT

This work was supported by the Natural Science Foundation of China (grant 61271021), and the Natural Science Foundation of Ningbo (grant 2012A610039).

REFERENCES

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

[2] A. K. Moorthy, A. C. Bovik , "A survey on 3D quality of experience and 3D quality assessment", in *Proc. of SPIE: Human Vision and Electronic Imaging XVIII*, vol. 8651, no. 86510M, San Jose, CA, USA, 2013.

[3] R. Vlad, P. Ladret, A. Guérin, "Three factors that influence the overall quality of the stereoscopic 3D content: image quality, comfort and realism", in *Proc. of SPIE: Human Vision and Electronic Imaging XVIII*, vol. 8651, San Jose, CA, USA, March 2013.

[4] J. J. Hwang, and H. R. Wu, "Stereo image quality assessment using visual attention and distortion predictors", *KSII Transactions on Internet and Information Systems*, vol. 5, no. 9, pp. 1613-1631, 2011.

[5] R. Bensalma, and M. C. Larabi, "A perceptual metric for stereoscopic image quality assessment based on the binocular energy", *Multidimensional Systems and Signal Processing*, vol.24, no.2, pp. 281-316, June 2013.

[6] M.-J. Chen, C.-C. Su, D.-K. Kwon, L. K. Cormack, A. C. Bovik, "Full-reference quality assessment of stereopairs accounting for binocular rivalry", *Signal Processing: Image Communication*, vol. 28, no. 9, pp. 1143-1155, Oct. 2013.

[7] V. De Silva, H. K. Arachchi, E. Ekmekcioglu, A. Kondoz, "Towards an impairment metric for stereoscopic video: A full-reference video quality metric to assess compressed stereoscopic video", *IEEE Transactions on Image Processing*, vol. 22, no. 9, pp. 3392-3404, Sept. 2013.

[8] C. Hu, F. Shao, G. Jiang, M. Yu, F. Li, and Z. Peng, "Quality assessment for stereoscopic images by distortion separation," *Journal of Software*, vol. 9, n. 1, pp. 37-43, Jan. 2014.

[9] S. Ryu, D. H. Kim, K. Sohn, "Stereoscopic image quality metric based on binocular perception model", in *Proc. of IEEE Intentional conference on Image Processing*, pp. 609-612, Orlando, Florida, Sept. 2012.

[10] M. Lambooij, W. Ijsselsteijn, "Visual discomfort and visual fatigue of stereoscopic displays: A review", *Journal of Imaging Science and Technology*, vol. 53, no. 3, pp. 030201-030201-14, 2009.

[11] M. Wopking, "Viewing comfort with stereoscopic pictures: An experimental study on the subjective effects of disparity magnitude and depth of focus", *Journal of the Society for Information Display*, vol. 3, no. 3, pp. 101-103, 1995.

[12] P. Seuntiens, L. Meesters, W. Ijsselsteijn, "Perceptual attributes of crosstalk in 3D images", *Displays*, vol. 26, no. 4-5, pp. 177-183, 2005.

[13] F. Speranza, W. J. Tam, R. Renaud, and N. Hur, "Effect of disparity and motion on visual comfort of stereoscopic images", in *Proc. of the SPIE*, vol. 6055, pp. 94-103, 2006.

[14] M. Lambooij, W. Ijsselsteijn, I. Heynderickx, "Visual discomfort of 3-D TV: Assessment methods and modeling", *Displays*, vol. 32, no. 4, pp. 209-218, 2011.

[15] Y. Nojiri, H. Yamanoue, S. Ide, F. Okano, "Parallax distribution and visual comfort on stereoscopic HDTV", in *Proc. of IBC*, pp. 373-380, 2010.

[16] H. Sohn, Y. J. Jung, S. Lee, H. W. Park, and Y. M. Ro, "Attention model-based visual comfort assessment for stereoscopic depth perception", in *Proc. of International Conference on Digital Signal Processing*, Corfu Island, Greece, pp. 1-6, 2011.

[17] Y. J. Jung, H. Sohn, S. Lee, H. W. Park, and Y. M. Ro, "Predicting visual discomfort of stereoscopic images using human attention model", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 23, no. 12, pp. 2077-2082, 2013.

[18] P. Lebreton, A. Raake, M. Barkowsky, P. Le Callet, "Evaluating depth perception of 3D stereoscopic videos", *IEEE Journal of Selected Topic in Signal Processing*, vol. 6, no. 6, pp. 710-720, Oct. 2012.

[19] F. Faria, J. batista, H. Araújo, "Stereoscopic depth perception using a model based on the primary visual cortex", *PloS One*, vol. 8, no.12, pp. e80745, 2013.

[20] A. Boev, A. Gotchev, K. Egiazarian, A. Aksay, G. B. Akar, "Towards compound stereo-video quality metric: a specific encoder-based framework", in *Proc. of IEEE Southwest Symposium on Image Analysis and Interpretation*, Denver, Colorado, pp. 218-222, 2006.

[21] A. Benoit, P. Le Callet, P. Campisi, and R. Cousseau, "Quality assessment of stereoscopic images", *EURASIP Journal on Image and Video Processing*, vol. 2008, Article ID 659024, 13 page, Jan. 2009.

[22] 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.

[23] Theimer, Wolfgang M., and Hanspeter A. Mallot, "Phase-based binocular vergence control and depth reconstruction using active vision", *CVGIP: Image Understanding*, vol. 60, no. 3, pp. 343-358, 1994.

[24] R. Blake, H. Wilson, "Binocular vision", *Vision Research*, vol. 51, no. 7, pp. 754-770, April 2011.

[25] D. J. Fleet, H. Wagner, and D. J. Heeger, "Neural encoding of binocular disparity: energy models, position shifts and phase shifts", *Vision Research*, vol. 36, no. 12, pp. 1839-1857, June 1996.

[26] V. Kolmogorov, and R. Zabih, “Computing visual correspondence with occlusions using graph cuts,” in *Proc. of IEEE International on Computer Vision*, Vancouver, Canada, vol. 2, pp. 508-515, Jul. 2001.
[27] P. G. Gottschalk, and J. R. Dunn, “The five-parameter logistic: a characterization and comparison with the four-parameter logistic”, *Analytical Biochemistry*, vol. 343, no. 1, pp. 54-65, Aug. 2005.
[28] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: from error visibility to structural similarity”, *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612, April 2004.
[29] Z. Wang, E. P. Simoncelli, and A. C. Bovik, “Multi-scale structural similarity for image quality assessment”, in *Proc. of IEEE Asilomar Conf. Signals, Systems, and Computers*, pp. 1398-1402, 2003.
[30] Z. Wang and A. C. Bovik, “A universal image quality index”, *IEEE Signal Processing Letters*, vol. 9, no. 3, pp. 81-84, March 2002.
[31] H. R. Sheikh, and A. C. Bovik, “Image information and visual quality”, *IEEE Transactions on Image Processing*, vol. 15, no. 2, pp. 430-444, Feb. 2006.
[32] D. M. Chandler, S. S. Hemami, “VSNR: A wavelet-based visual signal-to-noise ratio for natural images”, *IEEE Transactions on Image Processing*, vol. 16, no. 9, 2284-2298, 2007.
[33] L. Zhang, L. Zhang, X. Mou, and D. Zhang, “FSIM: A feature similarity index for image quality assessment”, *IEEE Transactions on Image Processing*, vol. 20, no. 8, pp. 2378-2386, Aug. 2011.
[34] M. Lambooij, W. IJsselelsteijn, D. G. Bouwhuis, I. Heynderickx, “Evaluation of stereoscopic images: Beyond 2D quality”, *IEEE Transactions on Broadcasting*, vol. 57, no. 2, pp. 432-444, 2011.

Qiuping Jiang received his B.S.degree from Department of Information and Enginerring, China Jiliang Univerisity, Hangzhou, China, in 2012. He is now working towards his M.S. degree in Ningbo University, Ningbo, China. His current research interests include image and video processing and stereoscopic image quality assessment.

Shanshan Wang received her B.S.degree from Department of Information and Enginerring, East China Institute of Technology, Nanchang, China, in 2011. She is now working towards her M.S. degree in Ningbo University, Ningbo, China. Her current research interests include image and video processing and stereoscopic image quality assessment.

Kemeng Li received his B.S.degree from Department of Information and Enginerring, China Jiliang Univerisity, Hangzhou, China, in 2013. She is now working towards his M.S. degree in Ningbo University, Ningbo, China. His current research interests include image and video processing and stereoscopic image quality assessment.

Feng Shao received his B.S. and Ph.D degrees from Zhejiang University, Hangzhou, China, in 2002 and 2007, respectively, all in Electronic Science and Technology. He is currently an associate professor in Faculty of Information Science and Engineering, Ningbo University, China. His research interests include video coding, image processing and perception, etc.