Blind Quality Prediction of Stereoscopic 3D Images

Jiheng Wang, Qingbo Wu, Abdul Rehman, Shiqi Wang, and Zhou Wang
Dept. of Electrical & Computer Engineering, University of Waterloo; Waterloo, ON, Canada

Abstract

Blind image quality assessment (BIQA) of distorted stereoscopic pairs without referring to the undistorted source is a challenging problem, especially when the distortions in the left- and right-views are asymmetric. Existing studies suggest that simply averaging the quality of the left- and right-views well predicts the quality of symmetrically distorted stereoscopic images, but generates substantial prediction bias when applied to asymmetrically distorted stereoscopic images. In this study, we propose a binocular rivalry inspired multi-scale model to predict the quality of stereoscopic images from that of the single-view images without referring to the original left- and right-view images. We apply this blind 2D-to-3D quality prediction model on top of ten state-of-the-art base 2D-BIQA algorithms for 3D-BIQA. Experimental results show that the proposed 3D-BIQA model, without explicitly identifying image distortion types, successfully eliminates the prediction bias, leading to significantly improved quality prediction performance. Among all the base 2D-BIQA algorithms, BRISQUE and M_3 archive excellent tradeoffs between accuracy and complexity.

Introduction

Objective quality assessment of stereoscopic images/videos is a challenging problem [1], especially when the distortions are asymmetric, i.e., when there are significant variations between the types and/or degrees of distortions occurred in the left- and right-views. Recent subjective studies suggested that in the case of symmetric distortions of both views, simply averaging state-of-the-art 2D image quality assessment (IQA) measures of both views provides reasonably accurate quality predictions of stereoscopic images [2]. Compared with the case of symmetric distortions, quality assessment of asymmetrically distorted stereoscopic images is much more difficult. It was reported that there is a large drop in the performance of both Full-reference (FR) 2D-IQA and 3D-IQA models from quality predictions of symmetrically to asymmetrically distorted stereoscopic images [3].

No-reference (NR) or blind image quality assessment (BIQA) approaches predict perceived quality of a test image without referring to an original image that is assumed to have pristine quality [4]. BIQA is highly challenging not only because of the difficulty in accurately estimating human behaviors in evaluating image quality across different visual content, distortion types and distortion levels, but also because real-world applications such as online quality monitoring often require the image and video streams to be evaluated at high speed, ideally in real-time. Therefore, 3D-BIQA is an even more challenging problem, especially when the distortions in the left- and right-views are asymmetric.

Table 1 reports Pearsons linear correlation coefficient (PLCC) and Spearmans rank-order correlation coefficient (SRCC) between 3D image quality mean opinion scores (3DIQ-MOS) and averaging some state-of-the-art 2D-BIQA estimations of both views on LIVE 3D Image Database Phase II [3], Waterloo-IVC 3D Image Database Phase I [5] and Phase II [12]. The ten tested state-of-the-art 2D-BIQA algorithms include Blind Image Quality Index (BIQI) [13], BLind Image Integrity Notator using DCT-Statistics II (BLIINDS-II) [14], Blind/Referenceless Image Spatial QUality Evaluator (BRISQUE) [15], Codebook Representation for No-Reference Image Assessment (CORNIA) [16], Distortion Identification-based Image Verity and INtegrity Evalutation (DIIVINE) [17], Local Pattern Statistics Index (LPSI) [18], M_3 [19], Naturalness Image Quality Evaluator (NIQE) [20], Quality-Aware Clustering (QAC) [21] and Distortion Type Classification and Label Transfer (TCLT) [22]. Among them, BIQI, BLIINDS-II, BRISQUE, CORNIA, DIIVINE, M_3, TCLT are opinion-aware BIQA methods that require subject-rated images for training, and are trained using all images from LIVE Image Quality Assessment Database Release 2 [26]. LPSI, NIQE and QACS are opinion-free BIQA methods, and are tested directly with their default parameters.

It can be observed from Table 1 that for most of the tested 2D-BIQA methods, simply averaging 2D-BIQA measures of both views provides reasonably accurate image quality predictions of symmetrically distorted stereoscopic images but there is a significant drop in the performance for asymmetrically distorted stereoscopic images on all tested 3D databases, which is consistent with the trend we have observed with the FR 2D-IQA methods [12].

A more straightforward way to examine the relationship between the perceptual quality of stereoscopic images and that of its single-view images is to perform subjective test on both 2D and 3D images [5]. It was found that for symmetrically distorted stereoscopic images, directly averaging the 2D image quality mean opinion scores (2DIQ-MOS) of both views provides excellent 3D image quality predictions, while for asymmetrically distorted stereoscopic images, a similar performance drop as of those in objective methods is observed [5]. The performance drop is largely due to the significant prediction bias that could lean towards opposite directions (either overestimate or underestimate image quality), depending on the distortion types and levels [5]. In [12], a binocular rivalry inspired multi-scale model to predict the quality of stereoscopic images from that of the single-view images was applied to 2DIQ-MOS scores and different base 2D-IQA measures. The experimental results showed that the quality prediction performance is significantly improved for most base 2D-IQA methods as well as with 2DIQ-MOS scores. Unfortunately, the model proposed in [12] requires access to the pristine reference stereopairs, which are not available in the case of 3D-BIQA.

In this work, we aim to develop an objective 3D-BIQA predictor. We take advantage of the previous findings on the relationship between the perceptual quality of stereoscopic images

Table 1 Performance comparison of 2D-BIQA models on LIVE and Waterloo-IVC 3D image database

	LIVE 3D Phase II				Waterloo-IVC Phase I				Waterloo-IVC Phase II			
	PLCC		SRCC		PLCC		SRCC		PLCC		SRCC	
2D-BIQA	Symm.	Asym.	Symm.	Asym.	Symm.	Asym.	Symm.	Asym.	Symm.	Asym.	Symm.	Asym.
Average 2DIQ-MOS	N/A	N/A	N/A	N/A	0.9801	0.8572	0.9657	0.8471	0.9799	0.8418	0.9696	0.8501
Proposed	N/A	N/A	N/A	N/A	0.9801	0.9460	0.9657	0.9324	0.9799	0.9433	0.9696	0.9302
Average IDW-SSIM	0.9368	0.7365	0.9229	0.6874	0.9638	0.7607	0.9480	0.7214	0.9377	0.7509	0.9056	0.7454
Proposed	0.9369	0.8593	0.9232	0.8517	0.9638	0.9283	0.9480	0.9042	0.9377	0.8859	0.9054	0.8660
Average BIQI	0.8350	0.6794	0.8278	0.6247	0.8674	0.7820	0.7965	0.7582	0.7343	0.6089	0.6239	0.5795
Proposed	0.8351	0.7455	0.8276	0.6971	0.8671	0.8864	0.7957	0.8528	0.7344	0.7695	0.6239	0.7071
Average BLIINDS-II	0.6649	0.5520	0.6243	0.5334	0.7583	0.5545	0.6766	0.5850	0.7774	0.4747	0.7586	0.5962
Proposed	0.6649	0.6146	0.6237	0.6011	0.8061	0.7925	0.6913	0.5854	0.8223	0.7389	0.7829	0.5951
Average BRISQUE	0.8688	0.6993	0.8491	0.6670	0.9371	0.8012	0.9102	0.7800	0.9293	0.6577	0.9022	0.7957
Proposed	0.8575	0.7513	0.8493	0.7365	**0.9371**	**0.8873**	**0.9096**	**0.8339**	**0.9293**	**0.8780**	**0.9023**	**0.8627**
Average CORNIA	0.8748	0.7026	0.8748	0.6778	0.9097	0.8252	0.8638	0.8150	0.9370	0.6490	0.9160	0.7877
Proposed	**0.8749**	**0.7567**	**0.8752**	**0.7398**	0.9096	0.8847	0.8635	0.8470	0.9369	0.8762	0.9164	0.8535
Average DIIVINE	0.7835	0.6035	0.7808	0.5536	0.5872	0.3974	0.5753	0.3261	0.6512	0.4449	0.6234	0.3920
Proposed	0.7697	0.6764	0.7806	0.6196	0.5872	0.4937	0.5753	0.4269	0.6513	0.5905	0.6238	0.4910
Average LPSI	0.7611	0.4898	0.7483	0.4051	0.7642	0.5512	0.7165	0.5891	0.7526	0.5859	0.7809	0.6076
Proposed	0.7611	0.6056	0.7488	0.4979	0.7642	0.6118	0.7165	0.6353	0.7526	0.7092	0.7809	0.7207
Average M_3	0.8512	0.7136	0.8480	0.7126	0.9418	0.8265	0.9147	0.8114	0.9114	0.7618	0.8713	0.7553
Proposed	0.8329	0.7695	0.8476	0.7668	**0.9418**	**0.9154**	**0.9147**	**0.8942**	**0.9114**	**0.8644**	**0.8717**	**0.8399**
Average NIQE	0.7801	0.7383	0.7592	0.7120	0.8276	0.7871	0.6395	0.6617	0.7429	0.6468	0.5282	0.5758
Proposed	0.7801	0.7751	0.7585	0.7557	0.8276	0.8311	0.6395	0.6596	0.7429	0.7656	0.5284	0.6103
Average QACS	0.8821	0.7388	0.8727	0.6960	0.8077	0.7627	0.5620	0.6448	0.7715	0.7125	0.5305	0.6567
Proposed	**0.8822**	**0.8169**	**0.8728**	**0.7858**	0.8077	0.8291	0.5620	0.6456	0.7716	0.7994	0.5303	0.6757
Average TCLT	0.8252	0.6961	0.7972	0.6564	0.8609	0.6624	0.7632	0.6245	0.8135	0.6075	0.7099	0.5735
Proposed	0.8253	0.7093	0.7981	0.6845	0.8609	0.8438	0.7620	0.7612	0.8136	0.7558	0.7098	0.7177

and that of its single-view images. We assume that existing successful 2D-BIQA methods are reliable for evaluating single-view images, and what is missing is an effective blind 2D-to-3D prediction model to combine single-vew quality scores, so as to eliminate the prediction bias for asymmetric distortions. Therefore, we opt to use a two-stage approach. The first stage builds a binocular rivalry inspired multi-scale 2D-to-3D quality prediction model without referring to the original images. In the second stage, this quality prediction model is applied to combine state-of-the-art 2D-BIQA estimations of both views, resulting in a 3D image quality estimation.

Blind 2D-to-3D Quality Prediction

Motivated by existing vision studies on binocular rivalry [6, 7], where it was found that for simple ideal stimuli, an increase in contrast enhances the predominance of one view against the other [8, 9, 10, 11], we showed that the strength of view dominance in binocular rivalry of stereoscopic images is related to the relative energy of the two views [12]. However, the computation of the relative energy involves the original left- and right-view images [12], which are not available in the case of 3D-BIQA. In this work, to overcome the problem, we apply a divisive normalization transform (DNT) [23, 24, 25] to the distorted left- and right-view images, and then estimate the strength of view dominance from DNT domain representations.

A DNT is typically built upon a linear image decomposition, followed by a divisive normalization stage [27]. The linear transformations may be discrete cosine transform or wavelet-type of transforms. Here, we assume a wavelet image decomposition, which provides a convenient framework for localized representation of images simultaneously in space, frequency (scale) and orientation. The DNT image representation of the image is then calculated by dividing each wavelet coefficient by a local energy measure based on its neighboring coefficients. The DNT image representation is not only an effective way to reduce the statistical redundancies between wavelet coefficients [27], but is also highly relevant to biological vision [28]. In [29], a divisive normalization framework was applied to develop a computational model that trades off between binocular rivalry and suppression, and the predictions were confirmed with psychophysical tests. This binocular perceptual relevance of divisive normalization representation leads us to design a multi-scale and multi-orientation DNT domain 2D-to-3D prediction model without referring to the original left- and right-view images.

Let $(I_{d,l}, I_{d,r})$ be the left- and right-view image pairs of the distorted stereoscopic images. To compute the DNT representation of $I_{d,l}$ and $I_{d,r}$, we first apply a 3-scale, 2-orientation steerable pyramid wavelet transform [30] to decompose $I_{d,l}$ and $I_{d,r}$ into 6 oriented subbands (2 for each scale) and a highpass and a lowpass residual subbands, respectively.

At the i-th oriented subband, for each center coefficient y_c, we define a DNT neighboring vector Y that contains 11×11 coefficients from the same subband (including the center coefficient itself). As such, the corresponding DNT center coefficient $\tilde{y}_c$ at

the i-th normalized subband is computed as

$$\tilde{y}_c = \frac{y_c}{\sum |Y| + c}, \tag{1}$$

where $c = 1$.

For each divisive normalized subband, we estimate its energy by computing the local variances at each spatial location, i.e., the variances of local patches extracted around each spatial location from the DNT coefficients are computed, for which an 11×11 circular-symmetric Gaussian weighting function $\mathbf{w} = \{w_i | i = 1, 2, \cdots, N\}$ with standard deviation of 1.5 samples, normalized to unit sum ($\sum_{i=1}^{N} w_i = 1$), is employed. The resulting mean energies for the i-th normalized subband in the $I_{d,l}$ and $I_{d,r}$ are denoted as $E_{i,l}$ and $E_{i,r}$, respectively. The overall energy estimations in both views are computed as the sum of the energies of all divisive normalized subbands

$$g_l = \sum_{i=1}^{6} E_{i,l} \qquad \text{and} \qquad g_r = \sum_{i=1}^{6} E_{i,r}, \tag{2}$$

Here g_l and g_r are estimations of the level of dominance of the left- and right-views, respectively. Given the values of g_l and g_r, the weights assigned to the left- and right-view images are given by

$$w_l = \frac{g_l^2}{g_l^2 + g_r^2} \qquad \text{and} \qquad w_r = \frac{g_r^2}{g_l^2 + g_r^2}, \tag{3}$$

respectively.

Finally, the overall prediction of 3D image quality is calculated by a weighted average of the left- and right-view image quality:

$$Q^{3D} = w_l Q_l^{2D} + w_r Q_r^{2D}, \tag{4}$$

where Q_l^{2D} and Q_r^{2D} denote the 2D image quality of the left- and right-views, respectively.

Validation

We use three 3D image quality databases to test the proposed algorithm, which are the Waterloo-IVC 3D Image Databases Phase I and Phase II and the LIVE 3D Image Database Phase II. All these databases contain both symmetrically and asymmetrically distorted stereoscopic images. The parameters of the proposed blind 2D-to-3D quality prediction method are selected empirically when working with Waterloo-IVC database Phase I, but are completely independent of Waterloo-IVC database Phase II and the LIVE database Phase II.

Blind 2D-to-3D quality prediction with 2DIQ-MOS

We first test the proposed blind 2D-to-3D quality prediction model on all 3D images in Waterloo-IVC database by applying it to the ground truth 2DIQ-MOS scores. The PLCC and SRCC values between 3DIQ-MOS and the predicted Q^{3D} value for all stereoscopic images and for each test image group are given in Table 2 (refer to Table 3 for categories of Waterloo-IVC database). The comparison results with our FR 2D-to-3D quality prediction model [31] are also given in Table 2. The corresponding scatter plots are shown in Figure 1. From Table 2, it can be observed that the proposed blind model performs as well as the FR 2D-to-3D quality prediction model [31]. The proposed model outperforms the direct averaging method in almost all cases, and the improvement is most pronounced in the case of strong asymmetric distortions (Group 3D.2) or when all test images are put together (All 3D image case). By comparing different columns of Figure 1, we observe the impact of the proposed blind 2D-to-3D prediction model on each image distortion type. For different distortion types, although the direct averaging method produces different distortions and levels of quality prediction biases, the proposed method, which does not attempt to recognize the distortion types or give any specific treatment for any specific distortion type, removes or significantly reduces the prediction biases for all distortion types. Moreover, for the mixed distortion case that provides the strongest test on the generalization ability of the model, the proposed method maintains consistent performance.

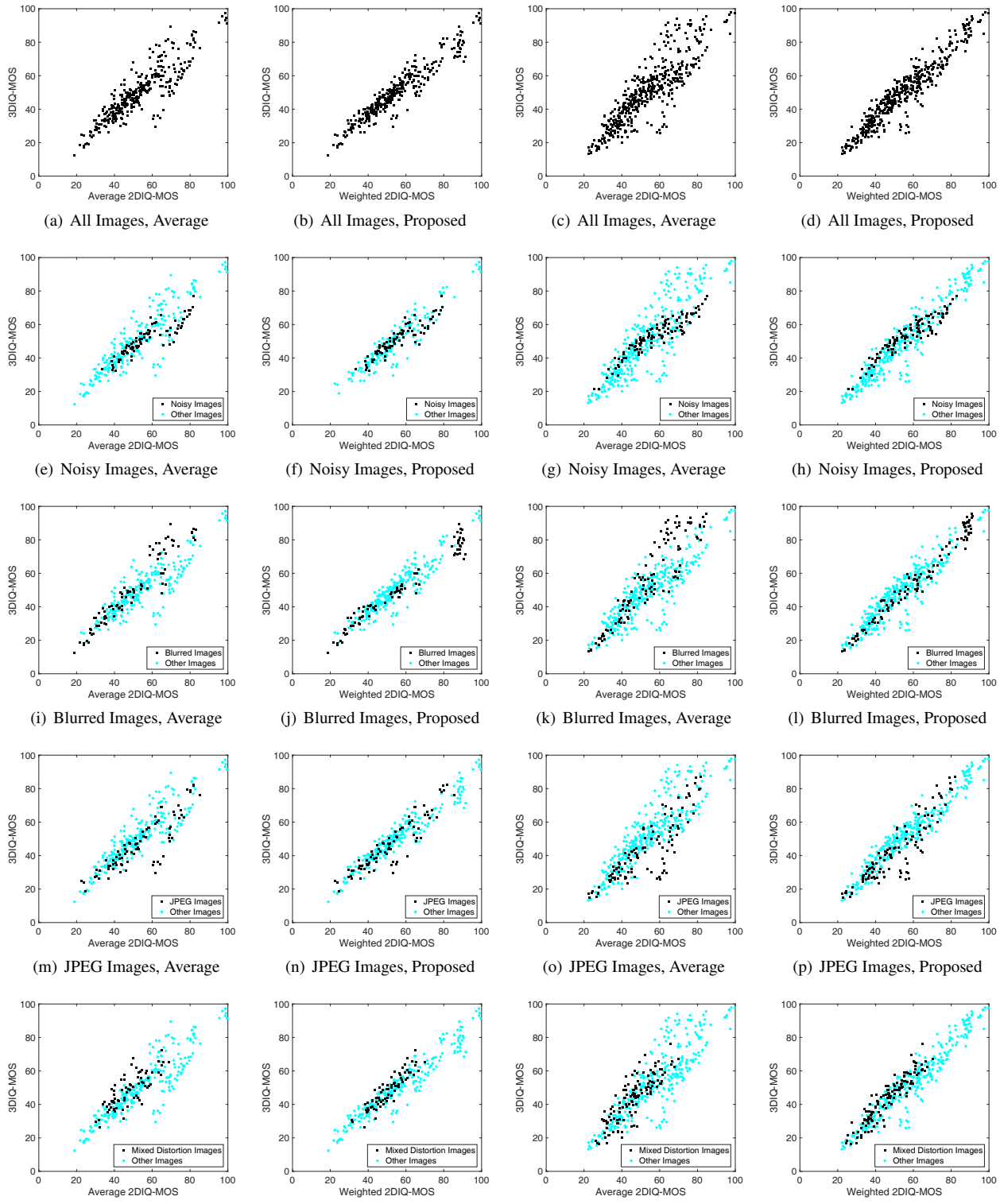
Table 2 Performance comparison of 2D-to-3D quality prediction models (direct average, FR [31], and the proposed blind prediction model), where the single-view quality is given by ground truth 2DIQ-MOS

		Waterloo-IVC Phase I		Waterloo-IVC Phase II	
Group	Method	PLCC	SRCC	PLCC	SRCC
All	Average	0.8835	0.8765	0.8763	0.8820
	FR [31]	0.9561	0.9522	0.9568	0.9477
	Proposed	0.9509	0.9413	0.9507	0.9382
Symm.	Average	0.9801	0.9657	0.9799	0.9696
	FR [31]	0.9801	0.9657	0.9799	0.9696
	Proposed	0.9801	0.9657	0.9799	0.9696
Asym.	Average	0.8572	0.8471	0.8418	0.8501
	FR [31]	0.9522	0.9452	0.9511	0.9424
	Proposed	0.9460	0.9324	0.9433	0.9302
3D.1	Average	0.9801	0.9657	0.9799	0.9696
	FR [31]	0.9801	0.9657	0.9799	0.9696
	Proposed	0.9801	0.9657	0.9799	0.9696
3D.2	Average	0.6613	0.5433	0.6121	0.5874
	FR [31]	0.9286	0.9160	0.9414	0.9497
	Proposed	0.9370	0.9141	0.9459	0.9482
3D.3	Average	0.9666	0.9164	0.9471	0.8898
	FR [31]	0.9714	0.9307	0.9602	0.9318
	Proposed	0.9772	0.9427	0.9665	0.9380
3D.4	Average	0.9223	0.8271	0.9225	0.8798
	FR [31]	0.9656	0.9357	0.9549	0.9320
	Proposed	0.9664	0.9366	0.9652	0.9496

Blind 2D-to-3D quality prediction with 2D-BIQA

Before applying the proposed blind 2D-to-3D quality prediction model on the base 2D-BIQA methods, we first examine these 2D-BIQA methods' abilities in predicting 2DIQ-MOS scores of single-view images in Waterloo-IVC 3D Image Database. Table 4 reports PLCC and SRCC values between 2DIQ-MOS scores and 2D-BIQA estimations. From Table 4, it can be seen that BRISQUE, CORNIA and M_3 archive the highest correlations with subjective data among all tested 2D-BIQA methods.

We then test the proposed blind 2D-to-3D quality prediction model by applying it to different base 2D-BIQA approaches on



(a) All Images, Average (b) All Images, Proposed (c) All Images, Average (d) All Images, Proposed

(e) Noisy Images, Average (f) Noisy Images, Proposed (g) Noisy Images, Average (h) Noisy Images, Proposed

(i) Blurred Images, Average (j) Blurred Images, Proposed (k) Blurred Images, Average (l) Blurred Images, Proposed

(m) JPEG Images, Average (n) JPEG Images, Proposed (o) JPEG Images, Average (p) JPEG Images, Proposed

(q) Mixed Distortion, Average (r) Mixed Distortion, Proposed (s) Mixed Distortion, Average (t) Mixed Distortion, Proposed

Figure 1. *3DIQ-MOS versus predictions from 2DIQ-MOS of 2D left- and right-views. First column, predictions by direct averaging the 2DIQ-MOS scores of both views on Waterloo-IVC Phase I; Second column, predictions by the proposed model on Waterloo-IVC Phase I; Third column, predictions by direct averaging 2DIQ-MOS scores of both views on Waterloo-IVC Phase II; Fourth column, predictions by the proposed model on Waterloo-IVC Phase II.*

Table 3 Categories of test images on Waterloo-IVC 3D Image Database

Group	Description
2D.0	Pristine single-view images
2D.1	Distorted single-view images
3D.0	Pristine stereopairs
3D.1	Symmetrically distorted stereopairs with the same distortion type and distortion level
3D.2	Asymmetrically distorted stereopairs with distortion on one view only
3D.3	Asymmetrically distorted stereopairs with the same distortion type but different levels
3D.4	Asymmetrically distorted stereopairs with mixed distortion types and levels

Table 4 Performance comparison of 2D-BIQA models on Waterloo-IVC 3D image database (Single-view Images)

	Waterloo-IVC Phase I		Waterloo-IVC Phase II	
2D-BIQA	PLCC	SRCC	PLCC	SRCC
BIQI	0.8694	0.7790	0.7098	0.5798
BLIINDS-II	0.8169	0.5977	0.5436	0.6752
BRISQUE	**0.9223**	**0.8912**	**0.9238**	**0.8892**
CORNIA	**0.9260**	**0.8861**	**0.9306**	**0.8934**
DIIVINE	0.5943	0.5804	0.5660	0.5631
LPSI	0.7537	0.6679	0.7478	0.6962
M_3	**0.9353**	**0.9040**	**0.9038**	**0.8534**
NIQE	0.8423	0.6685	0.7384	0.5616
QACS	0.7920	0.5543	0.7692	0.5452
TCLT	0.8471	0.6977	0.7900	0.6345

all three databases. Note that exactly the same blind 2D-to-3D quality prediction model obtained from 2DIQ-MOS and 3DIQ-MOS scores with Waterloo-IVC 3D image database Phase I is used and thus this blind 2D-to-3D prediction model is completely independent of any base objective 2D-BIQA approaches. Table 1 reports PLCC and SRCC values between 3DIQ-MOS and the predicted Q^{3D} value with the direct averaging method and the proposed blind 2D-to-3D quality prediction model. Note that the cases of using 2DIQ-MOS and Information content and Distortion Weighted Structural SIMilarity (IDW-SSIM) [31], a highly competitive FR 2D-IQA method, are also included for comparison.

From Table 1, it can be seen that the proposed method significantly improves most base 2D-BIQA methods on both databases. On the Waterloo-IVC 3D Image Database, BRISQUE, CORNIA and M_3 perform better than all competing 2D-BIQA methods with both the direct averaging and the proposed prediction model, which is consistent with their performance on single-view images. Interestingly, the performance of these 2D-BIQA methods approximates that of IDW-SSIM, which gives the most accurate prediction among FR 2D-IQA methods on both Waterloo-IVC 3D and LIVE 3D database [12]. This suggests that a good 2D-BIQA method can predict symmetrically distorted stereoscopic images with good accuracy, and when properly combined with a 2D-to-3D quality prediction model, can also well predict asymmetrically distorted stereoscopic images.

On the LIVE 3D Image Database Phase II, the proposed method achieves the best performance in the case of using QAC and also pronounces competitive performance with BRISQUE, CORNIA and M_3. However, there is a large gap when compared with the FR IDW-SSIM's prediction performance. This suggests that there is still potential to further improve 2D-BIQA methods in terms of robustness and generalizability.

We have also compared the proposed method with state-of-the-art 3D-BIQA approaches [32, 3, 33, 34] on the LIVE 3D image database Phase II. The PLCC and SRCC values are reported in Table 5. From Table 1 and Table 5, it can be seen that the proposed method, when combined with QAC, performs better than [32, 33, 34] but not as good as [3], which is a training-based method and the results reported here are median performance of 1000 trails, each using 80% of the data in LIVE 3D Phase II database for training and the remaining 20% for testing.

Table 5 Performance comparison of 3D-BIQA models on LIVE 3D image database Phase II

	PLCC		SRCC	
3D-BIQA	Symm.	Asym.	Symm.	Asym.
Akhter [32]	N/A	N/A	0.4200	0.5170
Chen [3]	N/A	N/A	0.9180	0.8340
Gu [33]	0.0994	0.2271	0.1760	0.1141
Shao [34]	0.9119	0.5651	0.8966	0.5244

Computational complexity analysis

Speed is another important performance factor in evaluating a BIQA method. We use program running time in the test stage of all competing methods as an estimate of computational complexity. The average processing time for a single-view image and for a stereopair on Waterloo-IVC database Phase II and LIVE database Phase II is summarized in Table 6. The system platform is Intel(R) Core(TM) i7-3770 @3.40GHz, 16.0 GB RAM and Windows 7 64-bit version. All methods are tested with the MATLAB R2015a software. Note that the resolutions of the single-view images are 1920×1080 and 640×360 for Waterloo-IVC database Phase II and LIVE database Phase II, respectively. Also, the average running time for the proposed 2D-to-3D quality prediction model is 1.8079s for Waterloo-IVC database Phase II and 0.2175s for LIVE database Phase II. The total processing time for a stereopair should be computed as twice of the processing time for a single-view image plus the time for 2D-to-3D prediction. From Table 6, it can be seen that BLIINDS-II and DIIVINE are the slowest, while LPSI and M_3 are the fastest. Generally speaking, BRISQUE and M_3 achieve excellent tradeoffs between accuracy and complexity.

Conclusion and Discussion

We propose a binocular rivalry inspired multi-scale model to predict the quality of stereoscopic images from that of the single-view images without referring to the original left- and right-view images. We apply the proposed blind 2D-to-3D quality prediction model to ten state-of-the-art base 2D-BIQA measures for 3D-BIQA. Experimental results show that the proposed blind model, without explicitly identifying image distortion types, successfully eliminates the prediction bias observed in direct averag-

Table 6 Complexity comparison of 2D-BIQA models on 3D image databases based on average running time (seconds)

2D-BIQA	Waterloo 2D	Waterloo 3D	LIVE 2D	LIVE 3D
BIQI	0.7211	3.2501	0.4539	1.1253
BLIINDS-II	435.1705	872.1489	48.4902	97.1979
BRISQUE	**0.5219**	**2.8517**	**0.1510**	**0.5195**
CORNIA	3.9315	9.6709	2.4399	5.0973
DIIVINE	86.1006	174.0091	11.8495	23.9165
LPSI	**0.1220**	**2.0519**	**0.0130**	**0.2435**
M_3	**0.3095**	**2.4269**	**0.0347**	**0.2869**
NIQE	1.3471	4.5021	0.1181	0.4537
QACS	**0.4539**	**2.7157**	**0.0483**	**0.3141**
TCLT	8.7635	19.3349	1.1605	2.5385

ing method, leading to significantly improved quality prediction of stereoscopic images. Among all the base 2D-BIQA methods, BRISQUE and M_3 achieve excellent tradeoffs between accuracy and complexity.

References

[1] C.-C. Su, A. K. Moorthy, and A. C. Bovik, Visual quality assessment of stereoscopic image and video: Challenges, Advances, and Future trends, in Visual Signal Quality Assessment, 185212, (2015).

[2] A. K. Moorthy, C.-C. Su, A. Mittal, and A. Bovik, Subjective evaluation of stereoscopic image quality, Signal Processing: Image Communication, 28, 8, 870-883, (2013).

[3] M.-J. Chen, L. K. Cormack, and A. C. Bovik, No-reference quality assessment of natural stereopairs, IEEE Trans. Image Process., 22, 9, 3379-3391, (2013).

[4] Z. Wang and A. C. Bovik, Reduced- and no-reference image quality assessment, IEEE Signal Processing Magazine, 28, 6, 29-40 (2011).

[5] J. Wang and Z. Wang, Perceptual quality of asymmetrically distorted stereoscopic images: the role of image distortion types, in Proc. Int. Workshop Video Process. Quality Metrics Consum. Electron, Chandler, AZ, USA, Jan. 2014, 1-6.

[6] L. Kaufman, Sight and Mind: An Introduction to Visual Perception. London, U.K.: Oxford Univ. Press, (1974).

[7] B. Julesz, Foundations of Cyclopean Perception. Chicago, IL, USA: Univ. Chicago Press, (1971).

[8] W. J. M. Levelt, The alternation process in binocular rivalry, Brit. J. Psychol., 57. 3-4, 225-238, (1966).

[9] R. Blake, Threshold conditions for binocular rivalry, J. Experim. Psychol., Human Perception Perform., 3, 2, 251-257, (1977).

[10] M. Fahle, Binocular rivalry: Suppression depends on orientation and spatial frequency, Vis. Res., 22, 7, 787-800, (1982).

[11] J. Ding and G. Sperling, A gain-control theory of binocular combination, Proc. Nat. Acad. Sci. USA, 103, 4, 1141-1146, (2006).

[12] J. Wang, A. Rehman, K. Zeng, S. Wang, and Z. Wang, Quality prediction of asymmetrically distorted stereoscopic 3D images, IEEE Trans. Image Process., 24, 11, 3400–3414, (2015).

[13] A. K. Moorthy and A. C. Bovik, A two-step framework for constructing blind image quality indices, IEEE Signal Processing Letters, 17, 5, 513-516 (2010).

[14] M. A. Saad, A. C. Bovik, and C. Charrier, Blind image quality assessment: A natural scene statistics approach in the DCT domain, IEEE Trans. Image Process., 21, 8, 3339-3352 (2012).

[15] A. Mittal, A. K. Moorthy, and A. C. Bovik, No-reference image quality assessment in the spatial domain, IEEE Trans. Image Process., 21, 12, 4695-4708 (2012).

[16] P. Ye, J. Kumar, L. Kang, and D. Doermann, Unsupervised feature learning framework for no-reference image quality assessment, in Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition, Providence, RI, USA, Jun. 2012, 1098-1105.

[17] A. K. Moorthy and A. C. Bovik, Blind image quality assessment: From natural scene statistics to perceptual quality, IEEE Trans. Image Process., 20(12), 33503364, (2011).

[18] Q. Wu, Z. Wang, and H. Li, A highly efficient method for blind image quality assessment, in Proc. IEEE Int. Conf. Image Process., Quebec City, Canada, Sep. 2015, 339–343.

[19] W. Xue, X. Mou, L. Zhang, A. C. Bovik, and X. Feng, Blind image quality assessment using joint statistics of gradient magnitude and laplacian features, IEEE Trans. Image Process., 23, 11, 4850-4862, (2014).

[20] A. Mittal, R. Soundararajan, and A. C. Bovik, Making a completely blind image quality analyzer, IEEE Signal Processing Letters, 20, 3, 209-212, (2013).

[21] W. Xue, L. Zhang, and X. Mou, Learning without human scores for blind image quality assessment, in Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition, Jun. 2013, 9951002.

[22] Q. Wu, H. Li, F. Meng, K. N. Ngan, B. Luo, C. Huang, and B. Zeng, Blind image quality assessment based on multi-channel features fusion and label transfer, IEEE Trans. Circuits and Systems for Video Tech., 20, 3, 209-219, (2015).

[23] D. J. Heeger, Normalization of cell responses in cat striate cortex, Visual Neuroscience, 9, 2, 181–197, (1992).

[24] E. P. Simoncelli and D. J. Heeger, A model of neuronal responses in visual area MT, Vision Research, 38, 5, 743–761, (1998).

[25] J. M. Foley, Human luminance pattern-vision mechanisms: masking experiments require a new model, Journal of the Optical Society of America A, 11, 6, 1710–1719, (1994).

[26] H. R. Sheikh, M. F. Sabir, and A. C. Bovik, A statistical evaluation of recent full reference image quality assessment algorithms, IEEE Trans. Image Process., 15, 11, 3440-3451, (2006).

[27] Q. Li and Z. Wang, Reduced-reference image quality assessment using divisive normalization-based image representation, IEEE J. of Selected Topics in Signal Process., 3, 2, 202–211, (2009).

[28] M. Carandini and D. J. Heeger, Normalization as a canonical neural computation, Nature Reviews Neuroscience, 13, 1, 51-62, (2012).

[29] C. P. Said and D. J. Heeger, A model of binocular rivalry and crossorientation suppression, PLoS Comput. Biol, 9, 3, (2013).

[30] E. P. Simoncelli, W. T. Freeman, E. H. Adelson, and D. J. Heeger, Shiftable multiscale transforms, IEEE Trans. Information Theory, 38, 2, 587-607, (1992).

[31] J. Wang, K. Zeng, and Z. Wang, Quality prediction of asymmetrically distorted stereoscopic images from single views, in Proc. IEEE Int. Conf. Multimedia and Expo, Chengdu, China, July 2014, 1-6.

[32] R. Akhter, Z. P. Sazzad, Y. Horita, and J. Baltes, No-reference stereoscopic image quality assessment, in Proc. SPIE 7524, Stereoscopic Displays and Appli. XXI, San Jose, CA, USA, Jan. 2010.

[33] 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, (2012).

[34] 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, IEEE Trans. Broadcasting, 61, 2, 154-165, (2015).

Author Biography

Jiheng Wang received the M.Math. degree in Statistics-Computing from the University of Waterloo, Waterloo, ON, Canada, in 2011, and the Ph.D. degree in electrical and computer engineering from the University of Waterloo, in 2016. He is currently a Post-Doctoral Fellow with the Department of Electrical and Computer Engineering from the University of Waterloo. In 2013, he was with the Video Compression Research Group, Blackberry, Waterloo. His current research interests include 3D image and video quality assessment, perceptual 2D and 3D video coding, biomedical signal processing, statistical learning and dimensionality reduction.

Qingbo Wu received the B.E. degree in Education of Applied Electronic Technology from Hebei Normal University in 2009, and the Ph.D. degree in signal and information processing in University of Electronic Science and Technology of China in 2015. From February 2014 to May 2014, he was a Research Assistant with the Image and Video Processing (IVP) Laboratory at Chinese University of Hong Kong. Then, from October 2014 to October 2015, he served as a visiting scholar with the Image & Vision Computing (IVC) Laboratory at University of Waterloo. He is currently a lecturer in the School of Electronic Engineering, University of Electronic Science and Technology of China. His research interests include image/video coding, quality evaluation, and perceptual modeling and processing.

Abdul Rehman holds a PhD degree in Electrical & Computer Engineering from the University of Waterloo, Canada and a Masters degree in Communications Engineering from Technical University Munich, Germany. Currently, he is the President & CEO of SSIMWave, a company he co-founded in 2013, dedicated to delivering excellence in visual quality-of-experience. He leads the development of SSIMWaves state-of-the-art video QoE measurement and optimization products geared towards the media, communication, and entertainment industry. Prior to SSIMWave, he worked at Blackberry, Advanced Micro Devices, Novel Mobile Radio Research, University of Waterloo, and Technical University Munich. His research interests include image and video processing, coding and quality assessment, and multimedia communications.

Shiqi Wang received the B.S. degree in computer science from the Harbin Institute of Technology in 2008, and the Ph.D. degree in computer application technology from the Peking University, in 2014. He was a Postdoc Fellow with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, Canada. He is currently with the Rapid-Rich Object Search Laboratory, Nanyang Technological University, Singapore, as a Research Fellow. His research interests lie in image and image/video coding, processing, quality assessment and analysis.

Zhou Wang is currently a Professor in the Department of Electrical and Computer Engineering, University of Waterloo, Canada. His research interests include image processing, coding, and quality assessment; computational vision and pattern analysis; multimedia communications; and biomedical signal processing. He has more than 100 publications in these fields with over 30,000 citations (Google Scholar). Dr. Wang serves as a Senior Area Editor of IEEE Transactions on Image Processing (2015-present), and an Associate Editor of IEEE Transactions on Circuits and Systems for Video Technology (2016-present). Previously, he served as a member of IEEE Multimedia Signal Processing Technical Committee (2013-2015), and an Associate Editor of IEEE Transactions on Image Processing (2009-2014), Pattern Recognition (2006-present) and IEEE Signal Processing Letters (2006-2010). He is a Fellow of Canadian Academy of Engineering, and a recipient of 2015 Primetime Engineering Emmy Award, 2014 NSERC E.W.R. Steacie Memorial Fellowship Award, 2013 IEEE Signal Processing Magazine Best Paper Award, and 2009 IEEE Signal Processing Society Best Paper Award.