Quality Prediction of Asymmetrically Distorted Stereoscopic 3D Images

Jiheng Wang, *Student Member, IEEE*, Abdul Rehman, Kai Zeng, Shiqi Wang, *Member, IEEE*, and Zhou Wang, *Fellow, IEEE*

***Abstract*— Objective quality assessment of distorted stereoscopic images 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 paper, we first build a database that contains both single-view and symmetrically and asymmetrically distorted stereoscopic images. We then carry out a subjective test, where we find that the quality prediction bias of the asymmetrically distorted images could lean toward opposite directions (overestimate or underestimate), depending on the distortion types and levels. Our subjective test also suggests that eye dominance effect does not have strong impact on the visual quality decisions of stereoscopic images. Furthermore, we develop an information content and divisive normalization-based pooling scheme that improves upon structural similarity in estimating the quality of single-view images. Finally, we propose a binocular rivalry-inspired multi-scale model to predict the quality of stereoscopic images from that of the single-view images. Our results show that the proposed model, without explicitly identifying image distortion types, successfully eliminates the prediction bias, leading to significantly improved quality prediction of the stereoscopic images.[1]**

***Index Terms*— Image quality assessment, stereoscopic image, 3D image, asymmetric distortion, SSIM, divisive normalization, contrast sensitivity function.**

I. Introduction

OVER the past years, we have observed an exponential increase in the demand for 3D image and video services. High-quality 3D movies can now be seen in thousands of new generation 3D theaters all around the world. Meanwhile, 3D TV has become technologically mature and won an increasing market share in the consumption market since 2011, where non-cinematic 3D contents could be from Blu-ray 3D or 3D broadcasting [1]. It is expected that mobile phones will be the largest 3D display application on a unit shipment basis in 2018, when 71 million units will have 3D capability [2]. Nevertheless, automatically assessing the quality of 3D visual experience is a challenging problem due to the complex and non-intuitive interactions between multiple 3D visual cues including image quality, depth quality and visual comfort [3], [4]. As a result, recent progress on 3D image quality assessment (IQA) remains limited. This lack of successful objective IQA methods for 3D visual experience has limited the development of 3D imaging applications and services.

In this work, we focus on how to predict the quality of a stereoscopic 3D image from that of the 2D single-view images. First, we carry out a subjective quality assessment experiment on a database that contains both single-view images and stereoscopic images with symmetric and asymmetric distortion types and levels. This database allows us to directly study the quality prediction performance from single-view images to stereoscopic images, for which we observe that simply averaging the quality of both views creates substantial bias on asymmetrically distorted stereoscopic images, and interestingly, the bias could lean towards opposite directions, largely depending on the distortion types. We then develop an information content and divisive normalization based pooling scheme that improves upon SSIM in estimating the quality of single-view images. Furthermore, by incorporating spatial frequency tuned mechanisms of human visual system (HVS), we propose a binocular rivalry inspired model to account for the bias, which not only results in better quality prediction of stereoscopic images with asymmetric distortion levels, but also well generalizes to the case of asymmetric distortions with mixed distortion types.

II. Review of Previous 3D-IQA Studies

A. Previous Subjective 3D-IQA Studies

To the best of our knowledge, there are currently 8 subject-rated image databases that are commonly recognized in the 3D-IQA research community. Table I lists these databases with detailed descriptions. Among them LIVE 3D Image Quality Database Phase I, LIVE 3D Image Quality Database Phase II, IRCCyN/IVC 3D Images Database, and MMSPG 3D Image Quality Assessment Database are publicly available.

Subjective data is essential in understanding the impact of various distortions on the perceptual quality of stereoscopic images. Ideally, we would need a complete set of subjective

Manuscript received May 30, 2014; revised October 26, 2014 and April 13, 2015; accepted June 12, 2015. Date of publication June 17, 2015; date of current version July 7, 2015. This work was supported by the Natural Sciences and Engineering Research Council of Canada within the Strategic, Discovery and Steacie Memorial Award Programs. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Ivana Tosic.

J. Wang, K. Zeng, S. Wang and Z. Wang are with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, N2L 3G1, Canada (e-mail: j237wang@uwaterloo.ca; kzeng@uwaterloo.ca; s269wang@uwaterloo.ca; zhou.wang@uwaterloo.ca).

A. Rehman is with the SSIMWave, Waterloo, ON N2L 3G1, Canada (email: abdul.rehman@ssimwave.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2015.2446942

[1]Some partial preliminary results of this work were presented at International Workshop on Video Processing and Quality Metrics for Consumer Electronics, Chandler, AZ, Jan., 2014. and IEEE International Conference on Multimedia and Expo, Chengdu, China, July, 2014.

1057-7149 © 2015 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

TABLE I
SUMMARY OF EXISTING 3D IMAGE QUALITY DATABASES

Database	# of subjects	Protocol	Display	# of images	Image Sizes	Distortions
LIVE 3D Image Quality Database Phase I [5]	32	SSCQS	Passive	385	640×360	JPEG2000, JPEG, white noise, gaussian blur, fast fading
LIVE 3D Image Quality Database Phase II [6]	33	SSCQS	Active	368	640×360	JPEG2000, JPEG, white noise, gaussian blur, fast fading
IRCCyN/IVC 3D Images Database [7]	17	SAMVIQ	Active	96	512×448	JPEG2000, JPEG, gaussian blur
MICT 3D Image Quality Evaluation Database [8]	24	SSCQS	Auto	500	640×480	JPEG
Ningbo University 3D Image Quality Assessment Database Phase I [9]	20	DSCQS	Passive	410	1252×1110 to 1390×1110	JPEG2000, JPEG, white noise, gaussian blur
Ningbo University 3D Image Quality Assessment Database Phase II [10]	26	DSCQS	Passive	324	480×270 to 1024×768	JPEG2000, JPEG, white noise, gaussian blur, H.264 compression
Tianjin University 3D Image Quality Assessment Database [11]	N/A	DSCQS	Auto	300	320×240 to 1024×768	JPEG2000, JPEG, white noise
MMSPG 3D Image Quality Assessment Database [12]	17	SSCQS	Passive	100	1920×1080	Different camera distances

test on an image database that contains both 2D (single-view) and stereoscopic 3D images, both symmetrically and asymmetrically distorted images at different distortion levels, as well as both single- and mixed-distortion images. The above-mentioned existing 3D image quality databases are highly valuable but limited in one aspect or another. Specifically, IRCCyN/IVC 3D Images Database, Tianjin University Database, Ningbo University Database Phase II, and LIVE 3D Image Quality Database Phase I only include symmetrically distorted stereoscopic images. Ningbo University Database Phase I only includes asymmetrically distorted stereoscopic images. MICT 3D Image Quality Evaluation Database contains both cases but only for JPEG compressed images. The most recent LIVE 3D Image Quality Database Phase II includes both symmetric and asymmetric cases as well as five distortion types. Unfortunately, 2D-IQA of single-view images are missing, making it difficult to directly examine the relationship between the perceptual quality of single-views and stereoscopic images. In addition, asymmetric distortions with mixed distortion types are missing in all existing databases, making it hard to validate the generalization capability of 3D quality prediction models.

B. Previous Objective 3D-IQA Studies

Existing objective 3D-IQA or 3D video quality assessment (3D-VQA) models may be grouped into two categories. The first type of approaches are built directly upon successful 2D-IQA methods. These approaches can be further divided into two subcategories, depending on the use of depth or disparity information. Methods in the first subcategory do not explicitly use depth information. In [13] and [14], 2D-IQA measures, including peak-signal-to-noise ratio (PSNR), structural similarity (SSIM) [15] and video quality metric (VQM) [16], were applied to the left- and right-view images of 3D videos separately and then combined to a 3D quality score. Both experimental results showed that VQM performs better than PSNR and SSIM possibly due to the temporal considerations in VQM. In [17], PSNR and VSSIM [18], which is a version of SSIM adapted for video, were compared to measure the perceptual 3D quality and the VSSIM was found to be closer to the subjective evaluation results. In [19], four 2D-IQA metrics, namely SSIM, universal quality index (UQI) [20], C4 [21] and RRIQA [22] as well as three approaches, called average approach, main eye approach, and visual acuity approach, were tested for measuring the perceptual quality of stereoscopic images. The experimental results show that C4 outperforms the other three metrics on IRCCyN/IVC 3D Images Database. The second subcategory of methods incorporates depth information with 2D-IQA. In [7] and [23], disparity maps between left- and right-views were estimated, followed by 2D quality assessment of disparity quality using SSIM and C4, which was subsequently combined with 2D image quality to produce an overall 3D image quality score. The results claimed that C4 outperforms SSIM on both evaluating stereoscopic image pairs and disparity maps on IRCCyN/IVC 3D Images Database and also suggested that the 3D-IQA performance of SSIM can be improved when adding depth quality. You *et al.* [24] investigated the capabilities of evaluating stereopairs as well as disparity maps with respect to ten well-known 2D-IQA metrics, i.e., PSNR, SSIM, multi-scale SSIM (MS-SSIM) [25], UQI, visual information fidelity (VIF) [26], etc. Their results suggested that an improved performance can be achieved when stereo image quality and depth quality are combined appropriately. Similarly, Yang *et al.* [11], [27] proposed a 3D-IQA algorithm based on the average PSNR of left- and right-views and the absolute difference with respect to disparity map. However, none of these more sophisticated 3D-IQA models perform better than or in most cases, even as good as, direct averaging 2D-IQA measures of both views [6].

The second type of 3D-IQA or 3D-VQA approaches focus on building 3D quality models directly without relying on existing 2D-IQA algorithms. In [28], Gorley and Holliman computed quality scores on matched feature points delivered by SIFT [29] and RANSAC [30]. The experimental results showed that the stereo band limited contrast model performs better than PSNR. In [31], an estimation of stereo image quality was proposed based on a multiple channel HVS model. In [32], Zhu amd Wang proposed a 3D-VQA model by

considering depth perception and their experimental results showed that it performs better than MSE and PSNR. In [33], Jin *et al.* proposed a 3D-VQA model based on 3D-DCT transform. Similar blocks from left- and right-views are found by block-matching, grouped into 3D stack and then analyzed by 3D-DCT. The experimental results showed that the model outperforms PSNR, SSIM, MS-SSIM and UQI on 3D video database [34].

Of particular interests are several models that consider binocular visual characteristics. The quality metric presented in [35] is based on binocular energy contained in the left- and right-views calculated by complex wavelet transform and Bandelet transform and the results showed that it outperforms the no-reference 3D-IQA model proposed in [8]. In [36], a stereo-version of SSIM based on binocular quality perception was proposed. Three components, luminance, contrast, and structural similarities, are combined into a quality index using a binocular quality perception model. The experimental results showed that the binocular perception model performs better than PSNR, SSIM as well as 3D-IQA models [7], [27] on IRCCyN/IVC 3D Images Database. In [37], a quality assessment algorithm of stereoscopic image compression based on binocular combination and binocular frequency integration was proposed. The experimental results showed that the binocular integration model performs better than 3D-IQA models [28], and [38], but not good as [7], [24], and [39] on LIVE 3D Image Quality Database Phase II. In [40], a stereoscopic image is separated into different binocular regions, each evaluated independently by considering their visual properties, followed by an integration step to produce an overall quality score. The experimental results showed that the region-based model outperforms PSNR, MS-SSIM and VIF as well as 3D-IQA models [7], [24] on Ningbo University 3D Image Quality Assessment Database Phase II. In [39], a "cyclopean" 3D-IQA model accounting for binocular rivalry was proposed and the experimental results showed that the framework significantly outperforms conventional 2D-IQA metrics PSNR, SSIM, MS-SSIM and VIF as well as [7], [24], [28], and [38] on LIVE 3D Image Quality Database Phase II.

C. Observations

Recent subjective studies suggested that in the case of symmetric distortion of both views (in terms of both distortion types and levels), simply averaging state-of-the-art 2D-IQA measures of both views is sufficient to provide reasonably accurate image quality predictions of stereoscopic images. In particular, in [5], it was shown that averaging PSNR, SSIM, MS-SSIM, UQI and VIF measurements of left- and right-views performs equally well or better than the advanced 3D-IQA models [7], [11], [24], [27], [28], [31], [32], [38], [41] on LIVE 3D Image Quality Database Phase I. Similar results were also observed in [6], where averaging SSIM and MS-SSIM measurements of both views outperformed advanced 3D-IQA models [7], [24], [28], [37], [38], [41] on LIVE 3D Image Quality Database Phase II. In [40], it was reported that directly averaging MS-SSIM outperformed 3D-IQA models [7], [24] on Ningbo University 3D Image Quality Assessment Database Phase II.

Fig. 1. The 6 pristine images in Waterloo-IVC 3D database Phase I. Only the right-views are shown here. (a) Art. (b) Books. (c) Dolls. (d) Moebius. (e) Laundry. (f) Reindeer.

Compared with the case of symmetric distortions, quality assessment of asymmetrically distorted stereoscopic images is a much more challenging problem. In [6], it was reported that there is a large drop in the performance of both 2D-IQA and 3D-IQA models from quality predictions of symmetrically to asymmetrically distorted stereoscopic images on LIVE 3D Image Quality Database Phase II.

It is worth noting that previous studies exhibit somewhat conflicting observations and opinions regarding the effect of asymmetric distortions. For image blur, evidence in [42] shows that the quality of asymmetrically blurred images is more affected by the higher quality view, which is generally consistent with the results given in [9]. For image blockiness, it was reported in [43] that 3D image quality should be approximated by averaging the quality of high quality and low quality views but there is a tendency towards the low quality view and this tendency becomes stronger when compression levels are high and images contain homogeneous areas. In [42], an under-weighting when direct averaging the quality of both views is found for monocular blockiness from MPEG-2 codec. In [9], it was suggested that the best strategy of asymmetric quality assessment for JPEG compressed images should be content and texture dependent.

III. Subjective Study

A. Waterloo-IVC 3D Image Quality Database Phase I

The new Waterloo-IVC 3D Image Quality database Phase I is created from 6 pristine stereoscopic image pairs (and thus their corresponding single-view images) shown in Fig. 1, all collected from the Middlebury Stereo 2005 Datasets [44]. The original resolution of single-view images is 1390×1100 or 1342×1100. All single-view images and stereopairs were slightly cropped to fit a display of 1920×1080 resolution. Each single-view image was altered by three types of distortions: additive white Gaussian noise contamination, Gaussian blur, and JPEG compression. Each distortion type had four distortion levels, where the distortion control parameters were decided to ensure a good perceptual separation between distortion levels as reported in Table II. More specially,

TABLE II
VALUE RANGES OF CONTROL PARAMETERS FOR DISTORTION SIMULATION

Distortion	Control Parameter	Range
Noise	Variance of Gaussian	[0.10 0.53]
Blur	Variance of Gaussian	[2 20]
JPEG	Quality Parameter	[3 10]

TABLE III
CATEGORIES OF TEST IMAGES ON WATERLOO-IVC DATABASE PHASE I

Group	# of images	Description
2D.0	6×1	Pristine single-view images
2D.1	6×12	Distorted single-view images
3D.0	6×1	Pristine stereopairs
3D.1	6×12	Symmetrically distorted stereopairs with the same distortion type and distortion level
3D.2	6×12	Asymmetrically distorted stereopairs with distortion on one view only
3D.3	6×18	Asymmetrically distorted stereopairs with the same distortion type but different levels
3D.4	6×12	Asymmetrically distorted stereopairs with mixed distortion types and levels

additive white Gaussian noise was applied equally across the R, G and B color channels. Similarly, Gaussian blur was simulated by applying a Gaussian low-pass filter to each of the color channels. For both noise and blur, the control parameter was the variance of the Gaussian. JPEG Compression was simulated using MATLAB®'s JPEG compression utility `imwrite` (Write True Color Image to JPEG). The single-view images were employed to generate distorted stereopairs, either symmetrically or asymmetrically. Altogether, there are totally 78 single-view images and 330 stereoscopic images. Table III categorizes these images into seven groups with detailed descriptions. Group 3D.1, Group 3D.2 and Group 3D.3 cover all combinations while Group 3D.4 includes a random subset from all possible fusions.

To the best of our knowledge, there are two unique features of the current database when compared with existing publicly known 3D-IQA databases. First, this is the only database that allows us to perform subjective test on both 2D and 3D images. The inclusion of 2D images allows us to directly examine the relationship between the perceptual quality of stereoscopic images and that of its single-view images. This is advantageous against previous studies which do not have ground truth of 2D image quality but have to rely on objective 2D-IQA measures to provide estimates. Second, this is the only database that contains mixed distortion types in asymmetrically distorted images.

The motivation of including different asymmetrical distortion levels and various mixed distortion types is threefold. First, purely for scientific curiosity, we are interested in knowing how the HVS behaves in the cases of asymmetrical/mixed distortions. Second, asymmetrical/mixed distortions are realistic in practice. For example, in the case of multi-exposure stereo images [45], because of the different exposure levels being used on different views, the amount of noise coming into the left- and right-view image sensors is different. For another example, asymmetric blur distortions and asymmetric blocking artifacts can be found in the case of mixed-resolution coding and asymmetric transform-domain quantization coding, and such distortions could have mixed types when postprocessing techniques (deblocking or blurring) are employed. Moreover, many 3D images are captured by a texture image view plus a depth map, where the texture image, which could contain noise, is used as one view, and the other view can be synthesized by combining the texture image with the depth map. Such a stereoscopic image contains both noise and mixed types of distortions. Third, the inclusion of these images provides the potential of a much stronger test on 3D-IQA models on their generalizability. Such test has been largely lacking in previous studies where the development of objective 3D-IQA models only took into account asymmetric distortions of specific and very limited distortion types such as compression only.

TABLE IV
VIEWING CONDITIONS OF THE SUBJECTIVE TEST

Parameter	Value	Parameter	Value
Subjects Per Monitor	1	Screen Resolution	1920×1080
Screen Diameter	27.00"	Viewing Distance	45.00"
Screen Width	23.53"	Viewing Angle	29.3°
Screen Height	13.24"	Pixels Per Degree	65.5 pixels

B. Subjective Test

The subjective test was conducted in the Lab for Image and Vision Computing at University of Waterloo. The test environment has no reflecting ceiling walls and floor, and was not insulated by any external audible and visual pollution. An ASUS 27" VG278H 3D LED monitor with NVIDIA 3D Vision™2 active shutter glasses is used for the test. The default viewing distance was 3.5 times the screen height. In the actual experiment, some subjects did not feel comfortable with the default viewing distance and were allowed to adjust the actual viewing distance around it. The details of viewing conditions are given in Table IV.

Twenty-four näive subjects, 14 males and 10 females aged from 22 to 45, participated in the study. A 3D vision test was conducted first to verify their ability to view stereoscopic 3D content. Three of them (1 male, 2 females) failed the vision test and did not continue with the subsequent experiment. As a result, a total of twenty-one naive subjects proceeded to the formal test.[2] Following previous works [3], [4], and [46], the subjects were asked to evaluate four aspects of their 3D viewing experience, including the perception of 3D image quality (3DIQ), depth quality (DQ), visual comfort (VC) and overall 3D quality of experience (3DQoE). The detailed descriptions of each aspects of visual experience including 2D image quality are elaborated in Table V. Since to visualize every 3D stereoscopic image, the subjects need to readjust their eyes so as to adapt to the content of the scene and establish 3D perception, using a double stimulus approach leads to interruptions of the viewing experience. Therefore, to

[2]While a visual acuity test was not performed in this study, a verbal confirmation was obtained prior to the experiment and subjects were asked to use their eyeglasses or contact lenses to correct their visual acuities.

TABLE V
DESCRIPTION OF VISUAL EXPERIENCE CRITERIA

Criterion	Description
2DIQ	The image content quality
3DIQ	The image content quality without considering 3D viewing experience
DQ	The amount, naturalness and clearness of depth perception experience
VC	The comfortness when viewing stereoscopic images
3DQoE	The overall 3D viewing experience

reduce this effect, we choose to use the single stimulus procedure using an 11-grade numerical categorical scale (SSNCS) protocol.

Our pilot tests showed that one-pass experiment (where a subject gives 3DIQ, DQ, VC, and 3DQoE scores to each stereoscopic image in one trial) may cause significant visual fatigue of the human subjects within a short period of time. To avoid this problem, we resorted to a multi-pass approach [4] in the formal test, where within each pass, the subject gives one of the four scores. We also found that the 2D perceptual quality of left- and right-views are so close to each other at the same distortion types and levels, so that the difference in their mean opinion scores (MOS) is negligible. Thus in order to control the scale of this subjective experiment, only one of the views were tested (randomly picked) in Group 2D.0 and Group 2D.1 in the formal test.

The test was scheduled on two consecutive days for each subject. Day 1 (2 hours) was dedicated to 2DIQ, VC and 3DIQ tests, and Day 2 (2 hours) to DQ and 3DQoE tests. First, a general introduction was given after the 3D vision test. All 3D visual experience criteria (3DIQ, DQ, VC and 3DQoE) were introduced and their definitions were given to subjects in both written and oral forms. After this general introduction, a preliminary understanding of four 3D visual experience criteria was expected for the subjects. Second, specific instructions and training sessions were given before each sub-test (2DIQ, 3DIQ, DQ, VC and 3DQoE). In each sub-test, the corresponding rating strategy was first introduced and the subjects were then required to practice by giving scores to training images until they fully understood the criteria and built their own scoring strategies. Note that the training processes were different depending on the characteristic of each criterion:

For both 2DIQ and 3DIQ sub-tests, we use three types of images in the training phase: pristine images, moderately distorted images, and highly-distorted images. The subjects were told to give scores at the high end (close to 10 pts) to the pristine images, at the mid-range to the moderately distorted images, and at the low end (close to 0 pts) to the highly-distorted images. For both DQ and VC sub-tests, self-training processes were employed to help the subjects establishing their own strategies. Previous works reported that the perception of DQ and VC are both highly content and texture dependent [43] and subject dependent [4], [46]. We agree with these observations and believe that it is not desirable to educate the subjects to use the same given rating strategy. For the overall 3DQoE sub-test, there is no training session. The subjects were asked to rate each stereopair based on their overall impression. They were asked to consider the previously introduced visual experience criteria 3DIQ, DQ and VC and were encouraged to use their own strategies to use these criteria.

All stimuli were shown once in each sub-test. However, there were 6 repetitions for single-view images and 12 repetitions for stereopairs, which means that for each subject, her/his first 6 single-view images and first 12 stereopairs were shown twice. The order of stimuli was randomized and the consecutive testing single-view images or stereopairs were from different source images. The 2DIQ sub-test, including 84 testing single-view images with 6 repetitions, was finished under 10 minutes. For 3DIQ , DQ , VC and 3DQoE sub-tests, 342 testing stereopairs with 12 repetitions were partitioned into two sessions and each single session (171 stereopairs) was finished in 15 to 20 minutes. Sufficient relaxation periods (5 minutes or more) were given between sessions.

Moreover, we found that repeatedly switching between viewing 3D images and grading on a piece of paper or a computer screen is a tiring experience. To overcome this problem, we asked the subject to speak out a score between 0 and 10, and a customized graphical user interface on another computer screen was used by the instructor to record the score. All these efforts were intended to reduce visual fatigue and discomfort of the subjects and to reduce the interference between different visual experience criteria.

The rest of the paper focuses on the relationship between the single-view 2D image quality (2DIQ scores) and the 3D image quality (3DIQ scores). More detailed descriptions of our database and analysis of the other aspects of the subjective experiments will be reported in future publications.

C. Waterloo-IVC 3D Image Quality Database Phase II

The new Waterloo-IVC 3D Image Quality database Phase II with more diverse image content is created from 10 pristine stereoscopic image pairs (and thus their corresponding single-view images) shown in Fig. 2. All images were collected from previous subjective 3D quality studies [47], [48] and the resolution of all images is 1920×1080. Each single-view image was altered by the same three types of distortions and each distortion type had the same four distortion levels. The single-view images were employed to generate distorted stereopairs, either symmetrically or asymmetrically. Altogether, there are totally 130 single-view images and 460 stereoscopic images.

The subjective test was conducted with the same test settings and viewing conditions as described in Section III-B. Here we only describe some important differences. Twenty-two näive subjects, 11 males and 11 females aged from 21 to 34, participated in the study and no one failed the vision test. As a result, all twenty-two subjects proceeded to the formal test. There were 20 repetitions for single-view images and 20 repetitions for stereopairs. All single-view images or around 160 stereopairs were evaluated in one session. Currently, only 2DIQ and 3DIQ sub-tests were conducted on Waterloo-IVC Phase II.

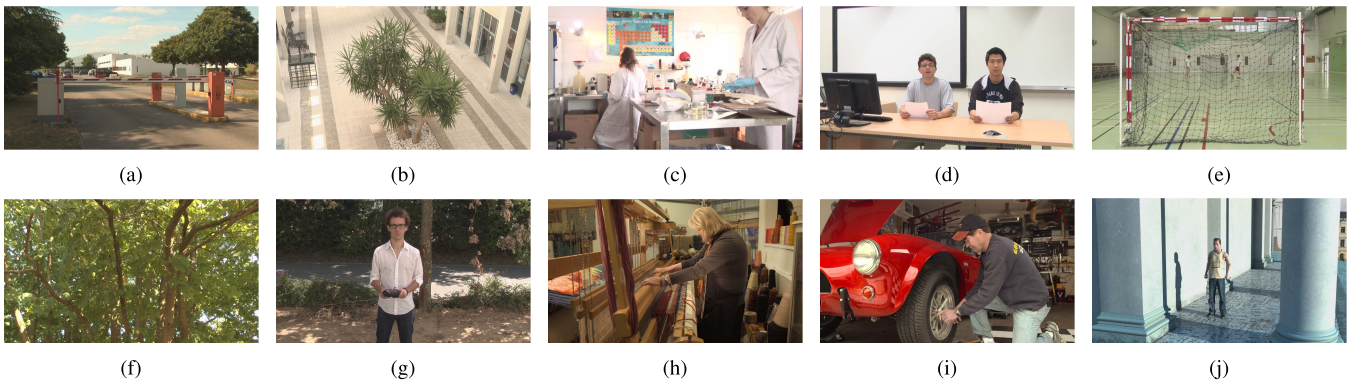


Fig. 2. The 10 pristine images in Waterloo-IVC 3D database Phase II. Only the right-views are shown here. (a) Barrier. (b) Hall. (c) Laboratory. (d) Persons. (e) Soccer. (f) Tree. (g) Umbrella. (h) CraftLoom. (i) OldTownCar. (j) Dancer.

TABLE VI
PERFORMANCE COMPARISON OF 2D-TO-3D QUALITY PREDICTION MODELS ON WATERLOO-IVC 3D DATABASE

		PLCC		SRCC		KRCC	
Dataset	Method	Direct Average	Proposed Weighting	Direct Average	Proposed Weighting	Direct Average	Proposed Weighting
Waterloo-IVC Phase I	All 3D	0.8835	0.9561	0.8765	0.9522	0.7161	0.8162
	Group 3D.1	0.9801	0.9801	0.9657	0.9657	0.8482	0.8482
	Group 3D.2	0.6247	0.9177	0.5433	0.9160	0.4406	0.7556
	Group 3D.3	0.9667	0.9719	0.9164	0.9307	0.7597	0.7789
	Group 3D.4	0.9224	0.9672	0.8271	0.9357	0.6390	0.7822
Waterloo-IVC Phase II	All 3D	0.8769	0.9571	0.8820	0.9477	0.7145	0.8080
	Group 3D.1	0.9802	0.9802	0.9696	0.9696	0.8557	0.8557
	Group 3D.2	0.6123	0.9486	0.5874	0.9497	0.4524	0.8070
	Group 3D.3	0.9485	0.9657	0.8898	0.9318	0.7176	0.7745
	Group 3D.4	0.9261	0.9550	0.8798	0.9320	0.7047	0.7803

D. Analysis and Key Observations

The raw 2DIQ and 3DIQ scores given by each subject were converted to Z-scores, respectively. Then the entire data sets were rescaled to fill the range from 1 to 100 and the MOS scores for each 2D and 3D image was computed after removing outliers. Given the subjective data, the main question we would like to ask in the current paper is how the single-view 2D image quality predicts the 3D image quality (3DIQ scores in the subjective test), especially for the case of asymmetric distortions. The most straightforward 2D-to-3D quality prediction method is to average the MOSs of the left- and right-view images. The first column of Fig. 3 shows the corresponding scatter plots for Waterloo-IVC database Phase I while the third column of Fig. 3 shows the scatter plots for Waterloo-IVC database Phase II. Table VI reports Pearson's linear correlation coefficient (PLCC), Spearman's rank-order correlation coefficient (SRCC) and Kendall's rank-order correlation coefficient (KRCC) between 3DIQ-MOS scores and the average 2DIQ-MOS scores, including the results for all stereoscopic images and for each test image group. PLCC are adopted to evaluate prediction accuracy [49] and SRCC and KRCC are employed to assess prediction monotonicity [49]. Higher PLCC, SRCC and KRCC indicate better consistency with human opinions of quality. PLCC is usually computed after a nonlinear mapping between the subjective and objective scores and the results may be sensitive to the choice of the mapping function. SRCC and KRCC are nonparametric rank order-based correlation metrics, independent of any monotonic nonlinear mapping between subjective and objective scores but do not explicitly estimate the accuracy of quality prediction.

From Table VI and the first and the third columns of Fig. 3, it can be observed that the best prediction occurs in Group 3D.1, which is the category for symmetrically distorted 3D images (consistent with the literature [5], [6]). By contrast, the PLCC, SRCC and KRCC values drop significantly in other test groups (corresponding to asymmetrical distortions) as well as in the all-image group. The drops of correlation coefficient values are also reflected in the scatter plots shown in Fig. 3, where this simple averaging prediction model generates substantial bias of many stereopairs. Most interestingly, this bias leans towards opposite directions, largely depending on the distortion types. In particular, for noise contamination and JPEG compression, average prediction overestimates 3D quality of many images (or 3D image quality is more affected by the poorer quality view), while for blur, average prediction often underestimates 3D image quality (or 3D image quality is more affected by the better quality view). Furthermore, Table VI suggests that the worst performance occurs in Group 3D.2, where only one view image is distorted and thus the quality difference between two views is maximized.

It is interesting to compare our observations regarding distortion type dependency with those published in the literature. For image blur, it was reported in [9] and [42] that 3D image quality is less affected by the view with lower quality, which is consistent with our result. For image blockiness from JPEG compression, in [43], it is claimed that 3D image quality is approximately the average of the higher quality and the lower quality but there is a tendency towards the lower quality view, which is consistent with our observations especially when one of view is highly compressed

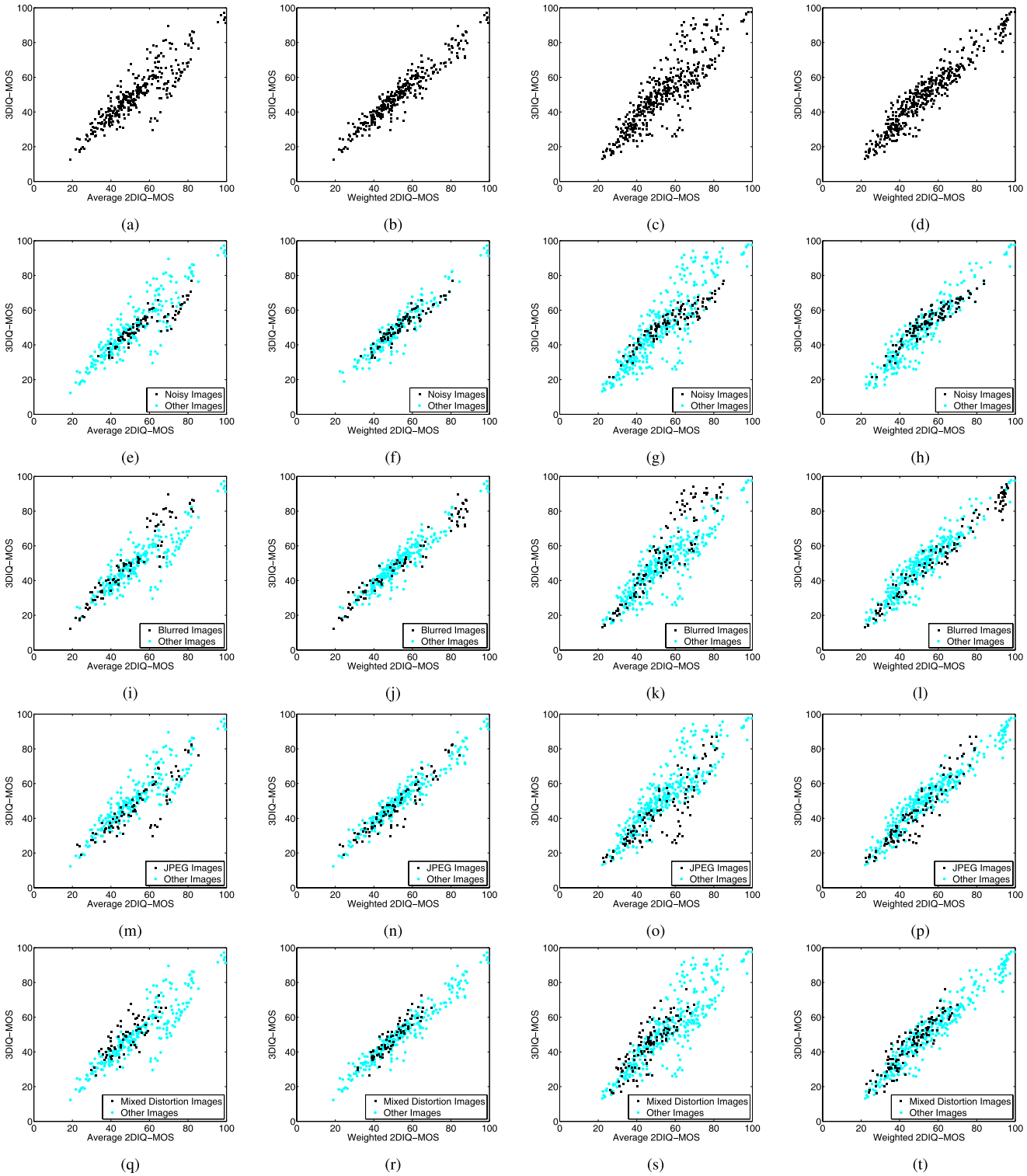


Fig. 3. 3DIQ-MOS versus predictions from 2DIQ-MOS of 2D left- and right-views. First column, predictions by averaging the 2DIQ-MOS scores of both views on Waterloo-IVC Phase I; Second column, predictions by weighting the 2DIQ-MOS scores of both views on Waterloo-IVC Phase I; Third column, predictions by averaging 2DIQ-MOS scores of both views on Waterloo-IVC Phase II; Fourth column, predictions by weighting 2DIQ-MOS scores of both views on Waterloo-IVC Phase II. (a) All Images, Average 2DIQ-MOS. (b) All Images, Weighted 2DIQ-MOS. (c) All Images, Average 2DIQ-MOS. (d) All Images, Weighted 2DIQ-MOS. (e) Noisy Images, Average 2DIQ-MOS. (f) Noisy Images, Weighted 2DIQ-MOS. (g) Noisy Images, Average 2DIQ-MOS. (h) Noisy Images, Weighted 2DIQ-MOS. (i) Blurred Images, Average 2DIQ-MOS. (j) Blurred Images, Weighted 2DIQ-MOS. (k) Blurred Images, Average 2DIQ-MOS. (l) Blurred Images, Weighted 2DIQ-MOS. (m) JPEG Images, Average 2DIQ-MOS. (n) JPEG Images, Weighted 2DIQ-MOS. (o) JPEG Images, Average 2DIQ-MOS. (p) JPEG Images, Weighted 2DIQ-MOS. (q) Mixed Distortion, Average 2DIQ-MOS. (r) Mixed Distortion, Weighted 2DIQ-MOS. (s) Mixed Distortion, Average 2DIQ-MOS. (t) Mixed Distortion, Weighted 2DIQ-MOS.

and the other keeps uncompressed. Meanwhile, in [9], no bias was discovered when lower levels of asymmetric JPEG compression were evaluated. These seemingly controversial results are well explained by the scatter plots shown in Fig. 3 and the 2D-line plot shown in Fig. 4, which shows 3DIQ-MOS versus average 2DIQ-MOS minus 3DIQ-MOS for the case of strong

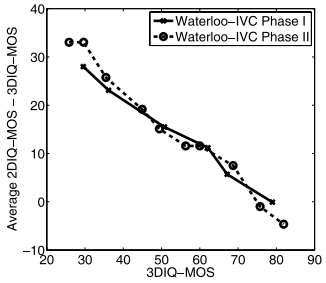


Fig. 4. 3DIQ-MOS versus average 2DIQ-MOS - 3DIQ-MOS for Group 3D.2 on Waterloo-IVC Phase I and Phase II.

asymmetric compressions (Group 3D.2) on Waterloo-IVC Phase I and Phase II. From these figures, it can be observed that the bias of the averaging prediction model increases with the level of distortions, and thus the strength of the bias is pronounced depending on the quality range being investigated.

E. Impact of Eye Dominance

Eye dominance is a common visual phenomenon, referring to the tendency to prefer the input from one eye to the other, depending on the human subject [50]. When studying visual quality of asymmetrically distorted images, it is important to understand if eye dominance plays a significant role in the subjective test results. For this purpose, we carried out a separate study on the impact of eye dominance in the perception of asymmetrically distorted stereoscopic images.

Twenty subjects (12 males and 8 females) participated in the experiment. The side of the dominant eye under static conditions was checked first by Rosenbach's test [51]. This test examines which eye determines the position of a finger when the subject is asked to point to an object. Ten subjects (7 males, 3 females) had a dominant left eye, and the others (5 males, 5 females) are right-eye dominant. A subjective test was conducted with the same test settings and viewing conditions as described in Section III-B. All test images are selected from Waterloo-IVC 3D database Phase I, which are the case of strong asymmetric distortions (Group 3D.2) and mixed distortions (Group 3D.4). Each asymmetric image creates two test cases, with the left- and right-views exchanged. Altogether, there are totally 78 symmetric stereoscopic images and 144 pairs (288 singles) of asymmetrically distorted stereoscopic images.

The MOS scores for each image were computed for left-eye dominant subjects and right-eye dominant subjects, denoted as $\mathrm{MOS_L}$ and $\mathrm{MOS_R}$, respectively. We employed the one-sample t-test to obtain a test decision for the null hypothesis that the difference between $\mathrm{MOS_L}$ and $\mathrm{MOS_R}$, i.e., $\mathrm{MOS_D} = \mathrm{MOS_L} - \mathrm{MOS_R}$, comes from a normal distribution with of zero-mean and unknown variance. The alternative hypothesis is that the population distribution does not have a mean equaling zero. The result h is 1 if the test rejects the null hypothesis at the 5% significance level, and 0 otherwise. The returned p-values for symmetric and asymmetric images are 0.3801 and 0.1322, respectively, thus the null hypothesis cannot be rejected at the 5% significance level, which indicates that the impact of eye dominance in the perception of asymmetrically distorted stereoscopic images is not considered significant. This is consistent with the "stimulus" view of rivalry that is widely accepted in the field of visual neuroscience [52]. A comprehensive review and discussion on the question of "stimulus" rivalry versus "eye" rivalry can also be found in [52] and [53].

IV. Objective Study: 2D-to-3D Quality Prediction

We opt to use a two-stage approach in the design of an objective 3DIQ predictor. The first stage aims to evaluate the perceptual quality of single-view images, while in the second stage, a binocular rivalry inspired multi-scale model is developed to combine 2D image quality of both views into a quality estimation of 3D image quality.

A. Objective 2D Quality Assessment

In the literature, the SSIM index [15] as well as its derivatives MS-SSIM [25] and information content weighted SSIM (IW-SSIM) [54] have demonstrated competitive performance in 2D objective IQA tests [54]. An advantage of the SSIM approach is that it provides a quality map that indicates the variations of image quality over space [15]. It was shown that spatial pooling built upon the quality map based on information content weighting or distortion weighting further improves the performance [54]. Here we build our 2D-IQA model upon SSIM, but improve it further by incorporating an information content and divisive normalization based pooling scheme.

A general form of spatially weighted pooling is given by

$$Q^{2D} = \frac{\sum_{i=1}^{N} w_i q_i}{\sum_{i=1}^{N} w_i}, \tag{1}$$

where q_i and w_i are the local quality value (e.g., local SSIM value) and the weight assigned to the i-th spatial location (i-th pixel), respectively. The assumption behind information content weighted pooling is that the spatial locations that contain more information are more likely to attract visual attention, and thus should be given larger weights. Let $\mathbf{x}_i$ and $\mathbf{y}_i$ be the local image patches extracted around the i-th spatial location from the reference and the distorted images, respectively. Following the information content evaluation method in [55], we compute the weighting factor by

$$w_i^{ic} = \log\left[\left(1 + \frac{\sigma_{x_i}^2}{C}\right)\left(1 + \frac{\sigma_{y_i}^2}{C}\right)\right], \tag{2}$$

where σ_{x_i} and σ_{y_i} are the standard deviations of $\mathbf{x}_i$ and $\mathbf{y}_i$, respectively, and C is the noisy visual channel power.

Another useful pooling strategy is distortion weighted pooling, which is based on the intuitive idea that the spatial locations that contain more distortions are more likely to attract visual attention, and thus should be given more weights. Since the local quality has been gauged by q_i (e.g., the SSIM value at location i), it is straightforward to convert it to a local distortion measure, for example, let $d_i = 1 - \mathrm{SSIM}_i$. Divisive normalization has been recognized as a perceptually and

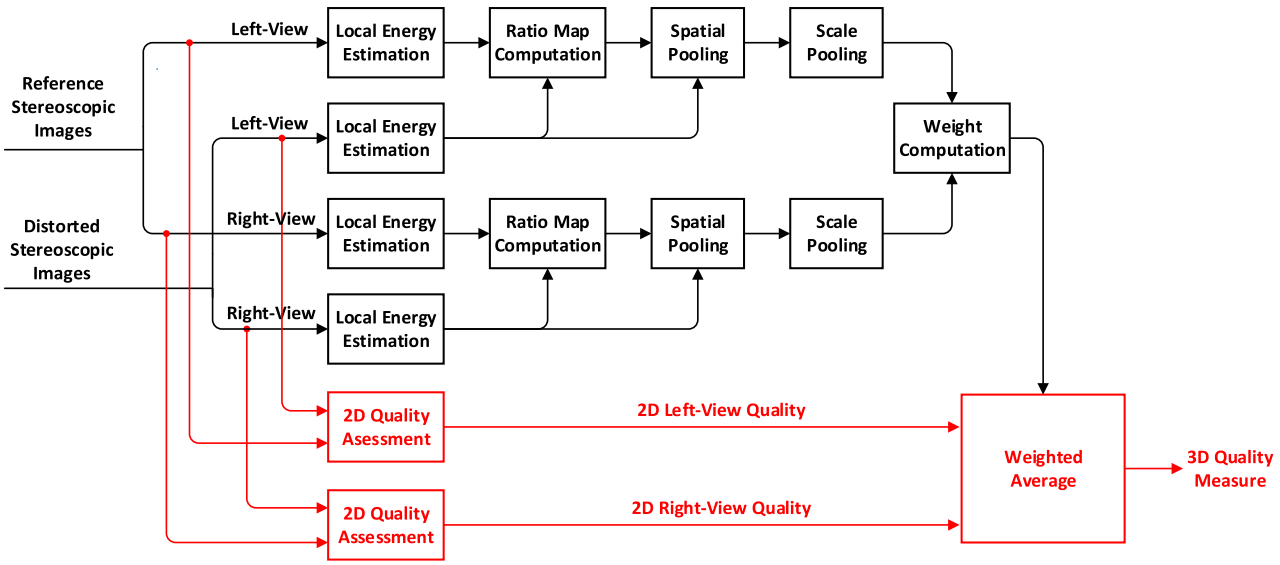


Fig. 5. Diagram of the proposed 2D-to-3D quality prediction model.

TABLE VII
PERFORMANCE COMPARISON OF 2D-IQA MODELS ON WATERLOO-IVC 3D DATABASE (SINGLE-VIEW IMAGES)

	Waterloo-IVC Phase I			Waterloo-IVC Phase II		
2D-IQA	PLCC	SRCC	KRCC	PLCC	SRCC	KRCC
PSNR	0.7878	0.4800	0.3552	0.7238	0.3885	0.2832
SSIM	0.8775	0.7726	0.5715	0.7824	0.6232	0.4441
MS-SSIM	0.8440	0.6402	0.4821	0.7455	0.5049	0.3590
IW-SSIM	0.8790	0.7696	0.5909	0.7960	0.6310	0.4640
IDW-SSIM	0.9634	0.9430	0.7865	0.9540	0.9341	0.7780

statistically motivated non-linear transformation [56]. We apply divisive normalization to the local distortion map and define a normalized distortion based weighting factor by

$$w_i^d = \frac{d_i}{\sqrt{\sum_{j\in\mathcal{N}_i} d_j^2 + D_0}}, \tag{3}$$

where $\mathcal{N}_i$ denotes the set of neighboring pixels surrounding the i-th spatial location, and D_0 is a stability constant.

The final weighting factor is obtained by combining information content and divisive normalization-based distortion weighting factors

$$w_i = \max\left\{\left(w_i^{ic}\right)^2, \left(w_i^d\right)^2\right\}. \tag{4}$$

where the max operation is based on the strategy to choose either w^{ic} or w^d, depending on which one is more significant.

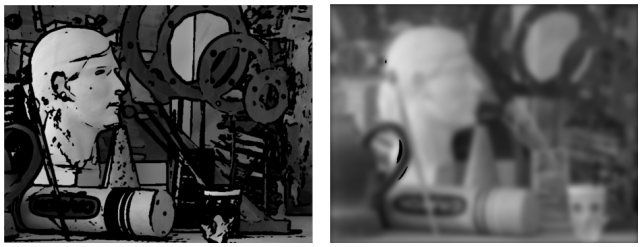
Applying this weighted pooling approach to the SSIM map, we obtain an information content and distortion weighted SSIM (IDW-SSIM) measure. This has led to significant performance improvement when tested using the single-view images in our new Waterloo-IVC database Phase I and Phase II. Quantitative measures of PLCC, SRCC and KRCC can be found in Table VII.

B. 2D-to-3D Quality Prediction Model

The competition between binocular fusion and binocular rivalry [57] provides a potential theory to develop 2D-to-3D quality prediction models. When the left- and right-view images are consistent, they are fused in the visual system to a single percept of the scene, known as binocular fusion. On the other hand, when the images of the two views are inconsistent, instead of the two images being seen superimposed, one of them may dominate or two images may be seen alternately, known as binocular rivalry [57]. Although there is a rich literature on binocular fusion and rivalry in biological vision science [57], [58] (where simple and ideal visual stimuli are often used), how to apply the principle to 3D-IQA remains an active research topic. Since in 3D-IQA we need to work on complicated scenes and distortions, simplifications are essential to create practical solutions.

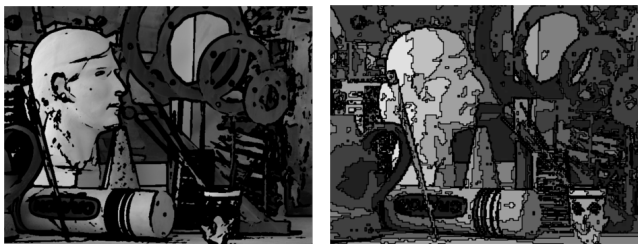
Our work is motivated by existing vision studies on binocular rivalry [59]–[62], where it was found that for simple ideal stimuli, an increasing contrast increases the predominance of one view against the other. Also note that in complicated scenes the contrast of a signal increases with its signal strength measured using energy. This inspires us to hypothesize that the strength of view dominance in binocular rivalry of stereoscopic images is related to the relative energy of the two views.

The diagram of the proposed method is shown in Fig. 5. Let $(I_{r,l}, I_{r,r})$ and $(I_{d,l}, I_{d,r})$ be the left- and right-view image pairs of the reference and distorted stereoscopic images, respectively. We first create their local energy maps by computing the local variances at each spatial location, i.e., the variances of local image patches extracted around each spatial location from the reference or the distorted images 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 energy maps are denoted as $E_{r,l}$, $E_{r,r}$, $E_{d,l}$ and $E_{d,r}$, respectively. Examples are given in Fig. 6 and Fig. 7, where the reference or distorted images are used as the background, and the pixels with local energy larger than 50 are highlighted as black.



(a) (b)

Fig. 6. Binocular Rivalry Example 1: The left-view is original and the right-view is blurred. The structural consistency between two-views is affected. The left-view may dominate the right-view at any time instance. (a) Left-view: Reference. (b) Right-view: Blur.



(a) (b)

Fig. 7. Binocular Rivalry Example 2: The left-view is original and the right-view is JPEG compressed. The structural consistency between two-views is affected. The right-view may dominate the left-view at any time instance. (a) Left-view: Reference. (b) Right-view: JPEG.

Assume that the reference stereopair has perfect quality with strong 3D effect, where binocular fusion prevails. When at least one of the single-view images is distorted at some spatial locations, the distortion may affect the consistency between the image structures from the two views, and thus binocular rivalry prevails. As a result, one view may dominate the other at any time instance. Based on our hypothesis, we compute the local energy ratio maps in both views:

$$R_l = \frac{E_{d,l}}{E_{r,l}} \quad \text{and} \quad R_r = \frac{E_{d,r}}{E_{r,r}}. \tag{5}$$

The energy ratio maps provide useful local binocular rivalry information, which may be combined with the qualities of single-view images to predict 3D image quality. A pooling stage is necessary for this purpose. High-energy image regions are likely to contain more information. If the ultimate goal of visual perception is to efficiently extract useful information from the visual scene, then the high-energy regions are more likely to attract visual attention, and thus should be given more importance. To emphasize on the importance of high-energy image regions in binocular rivalry, we adopt an energy weighted pooling method given by

$$g_l = \frac{\sum E_{d,l} R_l}{\sum E_{d,l}} \quad \text{and} \quad g_r = \frac{\sum E_{d,r} R_r}{\sum E_{d,r}}, \tag{6}$$

where the summations are over the full energy and ratio maps. Here g_l and g_r are estimations of the level of dominance of the left- and right-views, respectively.

Meanwhile, the study presented in [61] suggests that the dominance of one view over the other in binocular rivalry depends on the spatial frequency content of the stimuli used. Psychophysical experiments have shown that the spatial frequency sensitivity of human stereopsis behaves similar to the visual contrast sensitivity function (CSF) [63], which accounts for the visual contrast sensitivity as a function of spatial frequency [64]. We take advantage of this similarity and treat different spatial frequency subbands based on CSF. Specifically, we divide an image into multiple scales, by employing an iterative low-pass filtering and downsampling procedure, and subsequently calculate the level of view dominance for every subband of the left- and right-view images. We then combine the scale level dominance values of the left- and right-view images and the overall levels of dominance g_l and g_r are given by:

$$g_l = \sum_{i=1}^{N_s} \alpha_i g_{i,l} \quad \text{and} \quad g_r = \sum_{i=1}^{N_s} \alpha_i g_{i,r}, \tag{7}$$

where $g_{i,l}$ and $g_{i,r}$ denote the levels of dominance of the i^{th} scale of the left- and right-view images, respectively. N_s is the number of scales and α_i denotes the perceptual importance of the i^{th} scale determined using the CSF formula given by [64]:

$$S(u) = \frac{5200 e^{(-0.0016u^2(1+100/L)^{0.08})}}{\sqrt{\left(1+\frac{144}{X_o^2}+0.64u^2\right)\left(\frac{63}{L^{0.83}}+\frac{1}{1-e^{(-0.02u^2)}}\right)}}, \tag{8}$$

where u, L, and X_o^2 denote spatial frequency in cycles/degree, luminance in cd/m^2, and angular object area in square degrees, respectively. α_i values are calculated using

$$\alpha_i = S(f_i^{\text{c}}), \tag{9}$$

where f_i^{c} denotes the center spatial frequency of the i^{th} scale.

Given the values of g_l and g_r, the weights assigned to the left- and right-view images are given by

$$w_l = \frac{g_l^2}{g_l^2+g_r^2} \quad \text{and} \quad w_r = \frac{g_r^2}{g_l^2+g_r^2}, \tag{10}$$

respectively.

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

$$Q^{3D} = w_l Q_l^{2D} + w_r Q_r^{2D}, \tag{11}$$

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

In our previous work, preliminary results on simplified single-scale model of the proposed approach that ignores the variation of visual sensitivity across scales were reported in [65] and [66].

V. Validation

We use two 3D image quality databases to test the proposed algorithm, which are the new Waterloo-IVC 3D Image Quality database (Phase I and Phase II) and the LIVE 3D database Phase II [6]. The latter is a recent database that contains both symmetrically and asymmetrically distorted images. Note that the parameters of the proposed 2D-to-3D quality

TABLE VIII
PERFORMANCE COMPARISON OF 2D-TO-3D PREDICTION MODELS

		PLCC		SRCC		KRCC	
Database	2D-IQA Method	Direct Average	Proposed Weighting	Direct Average	Proposed Weighting	Direct Average	Proposed Weighting
Waterloo IVC Phase I	2D-MOS	0.8835	0.9561	0.8765	0.9522	0.7161	0.8162
	PSNR	0.6926	0.7419	0.5209	0.5192	0.3775	0.3830
	SSIM	0.7105	0.8134	0.5963	0.6712	0.4354	0.5047
	IDW-SSIM	0.7943	0.9300	0.7695	0.9177	0.5902	0.7537
Waterloo IVC Phase II	2D-MOS	0.8769	0.9571	0.8820	0.9477	0.7145	0.8080
	PSNR	0.6391	0.7187	0.4964	0.4948	0.3521	0.3605
	SSIM	0.5498	0.7278	0.4684	0.5640	0.3316	0.4090
	IDW-SSIM	0.7818	0.8918	0.7740	0.8687	0.5900	0.6901
LIVE Phase II	PSNR	0.7546	0.7764	0.7303	0.7528	0.5372	0.5531
	SSIM	0.8024	0.8433	0.7925	0.8408	0.6016	0.6444
	IDW-SSIM	0.8165	0.9159	0.7973	0.9188	0.6184	0.7441

prediction method are selected empirically when working with Waterloo-IVC database Phase I, but are completely independent of the Waterloo-IVC database Phase II and the LIVE database.

We test the proposed 2D-to-3D quality prediction model on all 3D images in Waterloo-IVC database Phase I and Phase II by applying it to the ground truth 2DIQ-MOS scores. The PLCC, SRCC and KRCC values between 3DIQ-MOS and the predicted Q^{3D} value for all stereoscopic images and for each test image group are given in Table VI. The corresponding scatter plots are shown in the second and the fourth columns of Fig 3. From Table VI and Fig 3, it can be observed that 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 Fig. 3, we observe how the proposed 2D-to-3D prediction model affects each image distortion type. For different distortion types, although the direct averaging method produces different levels of quality prediction biases towards different directions, 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, as mentioned earlier, the mixed distortion case provides the strongest test on the generalization ability of the model, for which the proposed method maintains consistent performance.

We also test the proposed 2D-to-3D quality prediction model by applying it to different base 2D-IQA approaches on both databases. Note that exactly the same 2D-to-3D quality prediction model obtained from 2DIQ-MOS and 3DIQ-MOS scores with Waterloo-IVC database Phase I is used and thus the model is completely independent of any tested objective 2D-IQA approaches including PSNR, SSIM and IDW-SSIM. The comparison results with the direct averaging method are shown in Table VIII, where it can be seen that the proposed method significantly improves most base 2D-IQA methods. The only exception is PSNR, which might be due to its poor performance in 2D image quality assessment, and thus merely changing 2D to 3D prediction method would not lead to any meaningful result. The scatter plots of 3DIQ-MOS scores versus predictions by averaging IDW-SSIM and weighting IDW-SSIM are shown in Fig 8, the first and second columns are the corresponding scatter plots for Waterloo-IVC database Phase I while the third and fourth columns are the scatter plots for Waterloo-IVC database Phase II. From Table VIII and Fig 8, it can be observed that the proposed 2D-to-3D model produces the most significant performance improvement from symmetric to asymmetric distortions in the case of using IDW-SSIM as the base 2D-IQA approach.

We have also compared the proposed method with state-of-the-art 3D-IQA approaches [6], [7], [11], [24], [39] using both databases, and the results are shown in Tables IX and X, respectively. The proposed method achieves the best performance in both databases among all objective IQA methods. The highly competitive performance in the Waterloo-IVC database Phase II and LIVE database Phase II is a more convincing result because no parameter has been determined using the Waterloo-IVC Phase II and the LIVE database. Another important observation is that there is a large performance drop in all other objective methods from symmetric to asymmetric distortions, whereas the drop is much smaller in the proposed method.

VI. DISCUSSION ON MIXED-DISTORTION STEREOSCOPIC VIDEO CODING

Studying the impact of asymmetric distortions on the quality of stereoscopic images not only has scientific values in understanding the HVS, but is also desirable in the practice of 3D video compression and transmission. The distortions involved in 3D video coding/communication are not only compression artifacts. The practical encoder/decoder also needs to decide on whether deblocking filters need to be turned on, and whether mixed-resolutions of the left/right-views should be used. Mixed-resolution coding, asymmetric transform-domain quantization coding, and postprocessing techniques (deblocking or blurring) can be employed individually or collectively. Previously in [67]–[69], the extent of the downsampling ratio that can be applied to a low quality view without a noticeable degradation on the 3D quality has been investigated. In [69], symmetric stereoscopic video coding, asymmetric quantization coding and mixed-resolution coding have been compared and the results suggested that mixed-resolution coding achieves the best coding efficiency.

In this work, our key observations provide some useful implication on stereoscopic image/video coding. For JPEG compression, 3D image quality is more affected by the poorer quality view, thus the poorer quality view deserves a higher weight; while for blur, 3D image quality is more affected by

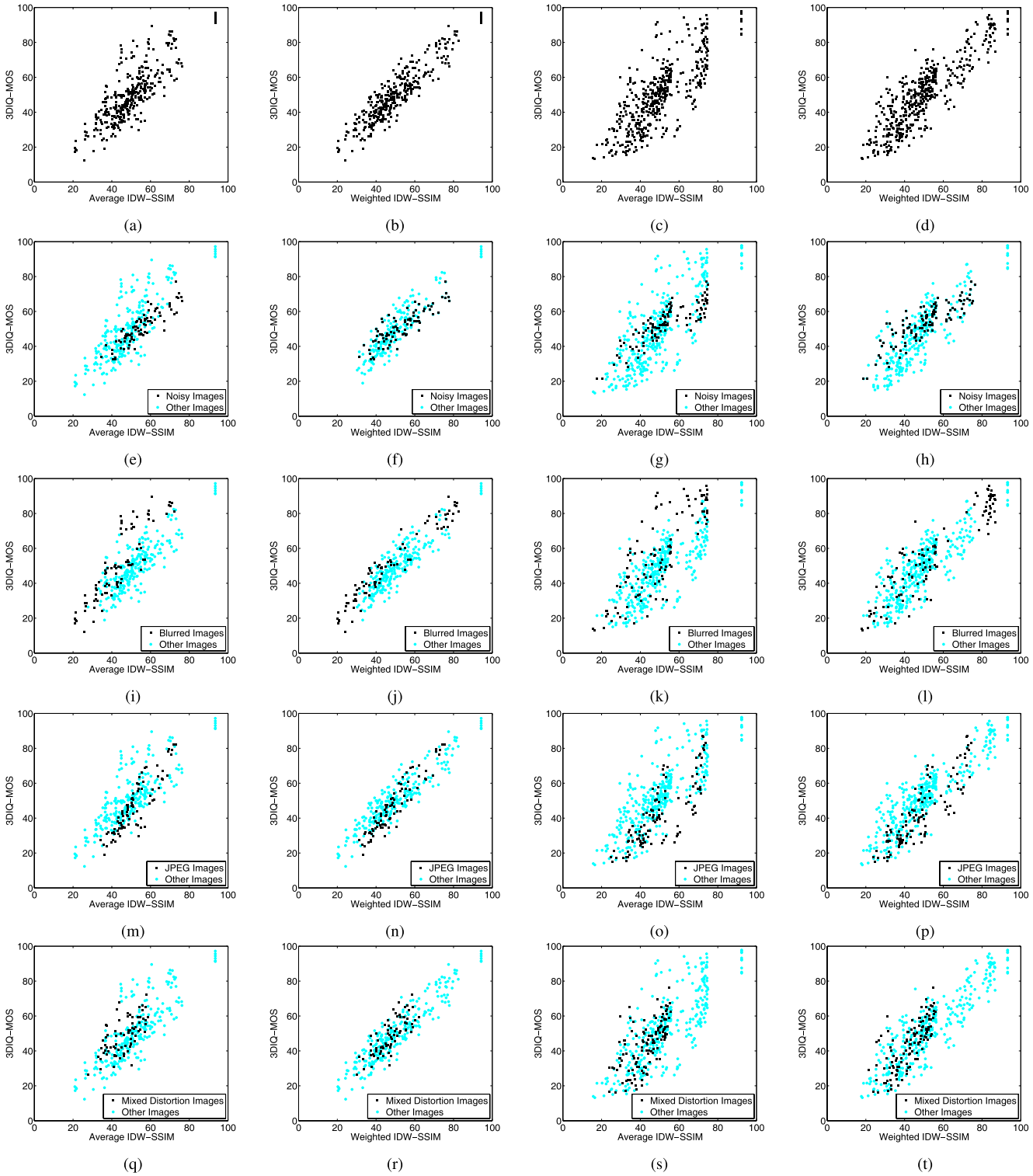


Fig. 8. 3DIQ-MOS versus predictions from IDW-SSIM of 2D left- and right-views. First column, predictions by averaging the IDW-SSIM estimations of both views on Waterloo-IVC Phase I; Second column, predictions by weighting the IDW-SSIM estimations of both views on Waterloo-IVC Phase I; Third column, predictions by averaging IDW-SSIM estimations of both views on Waterloo-IVC Phase II; Fourth column, predictions by weighting IDW-SSIM estimations of both views on Waterloo-IVC Phase II. (a) All Images, Average IDW-SSIM. (b) All Images, Weighted IDW-SSIM. (c) All Images, Average IDW-SSIM. (d) All Images, Weighted IDW-SSIM. (e) Noisy Images, Average IDW-SSIM. (f) Noisy Images, Weighted IDW-SSIM. (g) Noisy Images, Average IDW-SSIM. (h) Noisy Images, Weighted IDW-SSIM. (i) Blurred Images, Average IDW-SSIM. (j) Blurred Images, Weighted IDW-SSIM. (k) Blurred Images, Average IDW-SSIM. (l) Blurred Images, Weighted IDW-SSIM. (m) JPEG Images, Average IDW-SSIM. (n) JPEG Images, Weighted IDW-SSIM. (o) JPEG Images, Average IDW-SSIM. (p) JPEG Images, Weighted IDW-SSIM. (q) Mixed Distortion, Average IDW-SSIM. (r) Mixed Distortion, Weighted IDW-SSIM. (s) Mixed Distortion, Average IDW-SSIM. (t) Mixed Distortion, Weighted IDW-SSIM.

the better quality view, thus the better quality view gains a higher weight. Such unbalanced weighting is more pronounced for strong asymmetric distortions. Moreover, for mixed distortion types, when one view is JPEG compressed and the other is blurred, the JPEG compressed view always obtains a higher weight regardless of their distortion levels.

TABLE IX

PERFORMANCE COMPARISON OF 2D-TO-3D QUALITY PREDICTION MODELS ON WATERLOO-IVC 3D DATABASE

Waterloo-IVC 3D Phase I									
	PLCC			SRCC			KRCC		
3D-IQA Method	All images	Symmetric	Asymmetric	All images	Symmetric	Asymmetric	All images	Symmetric	Asymmetric
2D-MOS with Direct Average	0.8835	0.9801	0.8641	0.8765	0.9657	0.8471	0.7161	0.8482	0.6780
2D-MOS with Proposed Weighting	0.9561	0.9801	0.9523	0.9522	0.9657	0.9452	0.8162	0.8482	0.8026
IDW-SSIM with Direct Average	0.7943	0.9638	0.7607	0.7695	0.9480	0.7214	0.5902	0.7979	0.5426
IDW-SSIM with Proposed Weighting	0.9300	0.9638	0.9286	0.9177	0.9481	0.9097	0.7537	0.7989	0.7416
You [24]	0.7125	0.8681	0.7089	0.5968	0.7517	0.5706	0.4351	0.5615	0.4175
Benoit [7]	0.6797	0.8503	0.6970	0.5852	0.7275	0.5766	0.4212	0.5488	0.4135
Yang [11]	0.7061	0.8359	0.7150	0.6106	0.6668	0.6108	0.4428	0.4821	0.4436
Chen [39]	0.7337	0.9553	0.7324	0.6815	0.9241	0.6428	0.5212	0.7551	0.4854
Waterloo-IVC 3D Phase II									
	PLCC			SRCC			KRCC		
3D-IQA Method	All images	Symmetric	Asymmetric	All images	Symmetric	Asymmetric	All images	Symmetric	Asymmetric
2D-MOS with Direct Average	0.8769	0.9802	0.8422	0.8820	0.9696	0.8501	0.7145	0.8557	0.6693
2D-MOS with Proposed Weighting	0.9571	0.9802	0.9517	0.9477	0.9696	0.9424	0.8080	0.8557	0.7977
IDW-SSIM with Direct Average	0.7818	0.9377	0.7509	0.7740	0.9056	0.7454	0.5900	0.7309	0.5548
IDW-SSIM with Proposed Weighting	0.8918	0.9377	0.8799	0.8687	0.9052	0.8482	0.6901	0.7319	0.6655
You [24]	0.6817	0.7634	0.6857	0.5873	0.5602	0.5997	0.4181	0.4006	0.4225
Benoit [7]	0.5507	0.7549	0.5548	0.4595	0.5713	0.4539	0.3209	0.4006	0.3170
Yang [11]	0.6388	0.7920	0.6413	0.5875	0.6627	0.5946	0.4136	0.4848	0.4154
Chen [39]	0.6130	0.8371	0.6330	0.5781	0.7581	0.5627	0.4165	0.5635	0.4055

TABLE X

PERFORMANCE COMPARISON OF 2D-TO-3D QUALITY PREDICTION MODELS ON LIVE 3D DATABASE PHASE II

	PLCC			SRCC			KRCC		
3D-IQA Method	All images	Symmetric	Asymmetric	All images	Symmetric	Asymmetric	All images	Symmetric	Asymmetric
IDW-SSIM with Direct Average	0.8165	0.9368	0.7365	0.7973	0.9229	0.6874	0.6184	0.7548	0.5200
IDW-SSIM with Proposed Weighting	0.9159	0.9372	0.8981	0.9188	0.9234	0.9019	0.7441	0.7562	0.7237
You [24]	0.8015	0.8245	0.7832	0.7924	0.8030	0.7721	0.6015	0.6210	0.5784
Benoit [7]	0.7624	0.7339	0.7701	0.7436	0.6959	0.7474	0.5684	0.5434	0.5639
Yang [11]	0.7346	0.7752	0.7088	0.7210	0.7608	0.6960	0.5327	0.5746	0.5080
Chen [39]	0.9073	0.9384	0.8753	0.9013	0.9252	0.8538	0.7307	0.7599	0.6783
Chen [6]	0.8950	N/A	N/A	0.8800	0.9180	0.8340	N/A	N/A	N/A

These observations reject the hypothesis that only one of the two views need to be coded at high rate, and thus significant bandwidth can be saved by coding the other view with low rate. This also suggests that a significant coding gain may be achieved by mixed-resolution coding, followed by postprocessing techniques such as deblocking filtering.

We are currently building our WATERLOO-IVC 3D Video Quality Database [70] including various stereoscopic 3D videos obtained from mixed-resolution coding, asymmetric transform-domain quantization coding, their combinations, and multiple choices of postprocessing techniques, aiming to move further along the direction, which would allow us to investigate how to quantitatively predict the potential coding gain of asymmetric video compression, and to provide new insights on the development of high efficiency 3D video coding schemes to maintain a good tradeoff between perceptual 3D image quality, depth quality and/or visual discomfort.

VII. CONCLUSION

The major contributions of the current paper are as follows: First, we create a new subjective 3D-IQA database that has two unique features − the inclusion of both 2D and 3D images, and the inclusion of mixed distortion types. Second, we observe strong distortion type dependent bias when using the direct average of 2D image quality of both views to predict 3D image quality. Third, we observe that eye dominance does not have strong impact on visual quality evaluations of asymmetrically distorted stereoscopic images. Fourth, we develop an information content and divisive normalization based pooling scheme that improves upon SSIM in estimating the quality of single-view images. Fifth, we propose a binocular rivalry inspired multi-scale model to predict the quality of stereoscopic images from that of its single-view 2D images. Our results show that the proposed model, without explicitly identifying image distortion types, successfully eliminates the prediction bias, leading to significantly improved quality prediction of stereoscopic 3D images. The performance gain is most pronounced in the case of asymmetric distortions. In the future, we will extend our study to understand human opinions on depth quality, visual comfort and the overall 3D QoE, aiming to develop a complete objective quality assessment models for 3D QoE.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their valuable comments.

REFERENCES

[1] *List of 3D Channels*. [Online]. Available: http://en.wikipedia.org/wiki/3D_television, accessed Jun. 15, 2015.

[2] C. Hsieh. *3D Display Technology and Market Forecast Report*. [Online]. Available: http://www.displaysearch.com/cps/rde/xchg/displaysearch/hs.xsl/3d_display_technology_market_forecast_report.asp, accessed May 30, 2014.

[3] P. J. H. Seuntiëns, "Visual experience of 3D TV," M.S. thesis, Faculty Technol. Manage., Eindhoven Univ. Technol., Eindhoven, The Netherlands, 2006.

[4] W. Chen, F. Jérôme, M. Barkowsky, and P. Le Callet, "Exploration of quality of experience of stereoscopic images: Binocular depth," in *Proc. Int. Workshop Video Process. Quality Metrics Consum. Electron.*, Scottsdale, AZ, USA, Jan. 2012, pp. 116–121.

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

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

[7] A. Benoit, P. Le Callet, P. Campisi, and R. Cousseau, "Quality assessment of stereoscopic images," *EURASIP J. Image Video Process.*, vol. 2008, p. 659024, Oct. 2008.

[8] Z. M. P. Sazzad, S. Yamanaka, Y. Kawayokeita, and Y. Horita, "Stereoscopic image quality prediction," in *Proc. Workshop Quality Multimedia Exper.*, San Diego, CA, USA, Jul. 2009, pp. 180–185.

[9] X. Wang, M. Yu, Y. Yang, and G. Jiang, "Research on subjective stereoscopic image quality assessment," *Proc. SPIE*, vol. 7255, p. 725509, Jan. 2009.

[10] J. Zhou *et al.*, "Subjective quality analyses of stereoscopic images in 3DTV system," in *Proc. IEEE Int. Conf. Vis. Commun. Image Process.*, Tainan, Taiwan, Nov. 2011, pp. 1–4.

[11] J. Yang, C. Hou, Y. Zhou, Z. Zhang, and J. Guo, "Objective quality assessment method of stereo images," in *Proc. 3DTV Conf., True Vis.-Capture, Transmiss. Display 3D Video*, Potsdam, Germany, May 2009, pp. 1–4.

[12] L. Goldmann, F. De Simone, and T. Ebrahimi, "Impact of acquisition distortions on the quality of stereoscopic images," in *Proc. Int. Workshop Video Process. Quality Metrics Consum. Electron.*, Scottsdale, AZ, USA, Jan. 2010, pp. 45–50.

[13] C. T. E. R. Hewage, S. T. Worrall, S. Dogan, and A. M. Kondoz, "Prediction of stereoscopic video quality using objective quality models of 2D video," *Electron. Lett.*, vol. 44, no. 16, pp. 963–965, Jul. 2008.

[14] S. L. P. Yasakethu, C. T. E. R. Hewage, W. A. C. Fernando, and A. M. Kondoz, "Quality analysis for 3D video using 2D video quality models," *IEEE Trans. Consum. Electron.*, vol. 54, no. 4, pp. 1969–1976, Nov. 2008.

[15] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[16] M. H. Pinson and S. Wolf, "A new standardized method for objectively measuring video quality," *IEEE Trans. Broadcast.*, vol. 50, no. 3, pp. 312–322, Sep. 2004.

[17] A. Tikanmaki, A. Gotchev, A. Smolic, and K. Miller, "Quality assessment of 3D video in rate allocation experiments," in *Proc. IEEE Int. Symp. Consum. Electron.*, Vilamoura, Portugal, Apr. 2008, pp. 1–4.

[18] Z. Wang, L. Lu, and A. C. Bovik, "Video quality assessment based on structural distortion measurement," *Signal Process., Image Commun.*, vol. 19, no. 2, pp. 121–132, Feb. 2004.

[19] P. Campisi, P. Le Callet, and E. Marini, "Stereoscopic images quality assessment," in *Proc. Eur. Signal Process. Conf.*, Poznan, Poland, Sep. 2007, pp. 2110–2114.

[20] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, Mar. 2002.

[21] M. Carnec, P. Le Callet, and D. Barba, "An image quality assessment method based on perception of structural information," in *Proc. IEEE Int. Conf. Image Process.*, vol. 3. Barcelona, Spain, Sep. 2003, pp. III-185–III-188.

[22] Z. Wang and E. P. Simoncelli, "Reduced-reference image quality assessment using a wavelet-domain natural image statistic model," *Proc. SPIE*, vol. 5666, pp. 149–159, Mar. 2005.

[23] A. Benoit, P. Le Callet, P. Campisi, and R. Cousseau, "Using disparity for quality assessment of stereoscopic images," in *Proc. IEEE Int. Conf. Image Process.*, San Diego, CA, Oct. 2008, pp. 389–392.

[24] 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. Int. Workshop Video Process. Quality Metrics Consum. Electron.*, Scottsdale, AZ, USA, Jan. 2010, pp. 61–66.

[25] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *Proc. IEEE 37th Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, USA, Nov. 2003, pp. 1398–1402.

[26] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Trans. Image Process.*, vol. 15, no. 2, pp. 430–444, Feb. 2006.

[27] J. Yang, C. Hou, R. Xu, and J. Lei, "New metric for stereo image quality assessment based on HVS," *Int. J. Imag. Syst. Technol.*, vol. 20, no. 4, pp. 301–307, Dec. 2010.

[28] P. Gorley and N. Holliman, "Stereoscopic image quality metrics and compression," *Proc. SPIE*, vol. 6803, p. 680305, Feb. 2008.

[29] D. G. Lowe, "Object recognition from local scale-invariant features," in *Proc. IEEE Int. Conf. Comput. Vis.*, vol. 2. Kerkyra, Greece, Sep. 1999, pp. 1150–1157.

[30] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, Jun. 1981.

[31] L. Shen, J. Yang, and Z. Zhang, "Stereo picture quality estimation based on a multiple channel HVS model," in *Proc. IEEE Int. Congr. Image Signal Process.*, Tianjin, China, Oct. 2009, pp. 1–4.

[32] Z. Zhu and Y. Wang, "Perceptual distortion metric for stereo video quality evaluation," *WSEAS Trans. Signal Process.*, vol. 5, no. 7, pp. 241–250, Jul. 2009.

[33] L. Jin, A. Boev, A. Gotchev, and K. Egiazarian, "3D-DCT based perceptual quality assessment of stereo video," in *Proc. 18th IEEE Int. Conf. Image Process.*, Brussels, Belgium, Sep. 2011, pp. 2521–2524.

[34] S. Jumisko-Pyykkö, T. Haustola, A. Boev, and A. Gotchev, "Subjective evaluation of mobile 3D video content: Depth range versus compression artifacts," *Proc. SPIE*, vol. 7881, p. 78810C, Jan. 2011.

[35] R. Bensalma and M.-C. Larabi, "Towards a perceptual quality metric for color stereo images," in *Proc. IEEE Int. Conf. Image Process.*, Hong Kong, Sep. 2010, pp. 4037–4040.

[36] S. Ryu, D. H. Kim, and K. Sohn, "Stereoscopic image quality metric based on binocular perception model," in *Proc. IEEE Int. Conf. Image Process.*, Orlando, FL, USA, Sep./Oct. 2012, pp. 609–612.

[37] Y.-H. Lin and J.-L. Wu, "Quality assessment of stereoscopic 3D image compression by binocular integration behaviors," *IEEE Trans. Image Process.*, vol. 23, no. 4, pp. 1527–1542, Apr. 2014.

[38] C. T. E. R. Hewage and M. G. Martini, "Reduced-reference quality metric for 3D depth map transmission," in *Proc. 3DTV-Conf., True Vis.-Capture, Transmiss. Display 3D Video*, Tampere, Finland, Jun. 2010, pp. 1–4.

[39] M.-J. Chen, C.-C. Su, D.-K. Kwon, L. K. Cormack, and A. C. Bovik, "Full-reference quality assessment of stereopairs accounting for rivalry," *Signal Process., Image Commun.*, vol. 28, no. 9, pp. 1143–1155, Oct. 2013.

[40] F. Shao, W. Lin, S. Gu, G. Jiang, and T. Srikanthan, "Perceptual full-reference quality assessment of stereoscopic images by considering binocular visual characteristics," *IEEE Trans. Image Process.*, vol. 22, no. 5, pp. 1940–1953, May 2013.

[41] R. Akhter, Z. M. P. Sazzad, Y. Horita, and J. Baltes, "No-reference stereoscopic image quality assessment," *Proc. SPIE*, vol. 7524, p. 75240, Jan. 2010.

[42] D. V. Meegan, L. B. Stelmach, and W. J. Tam, "Unequal weighting of monocular inputs in binocular combination: Implications for the compression of stereoscopic imagery," *J. Experim. Psychol., Appl.*, vol. 7, no. 2, pp. 143–153, Jan. 2001.

[43] P. Seuntiens, L. Meesters, and W. Ijsselsteijn, "Perceived quality of compressed stereoscopic images: Effects of symmetric and asymmetric JPEG coding and camera separation," *ACM Trans. Appl. Perception*, vol. 3, no. 2, pp. 95–109, Apr. 2006.

[44] D. Scharstein and C. Pal, "Learning conditional random fields for stereo," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, Minneapolis, MN, USA, Jun. 2007, pp. 1–8.

[45] N. Sun, H. Mansour, and R. Ward, "HDR image construction from multi-exposed stereo LDR images," in *Proc. 17th IEEE Int. Conf. Image Process.*, Hong Kong, Sep. 2010, pp. 2973–2976.

[46] M.-J. Chen, D.-K. Kwon, and A. C. Bovik, "Study of subject agreement on stereoscopic video quality," in *Proc. IEEE Southwest Symp. Image Anal. Interpretation*, Santa Fe, NM, USA, Apr. 2012, pp. 173–176.

[47] M. Urvoy *et al.*, "NAMA3DS1-COSPAD1: Subjective video quality assessment database on coding conditions introducing freely available high quality 3D stereoscopic sequences," in *Proc. Workshop Quality Multimedia Exper.*, Yarra Valley, VIC, Australia, Jul. 2012, pp. 109–114.

[48] M. Pinson, "The consumer digital video library [best of the Web]," *IEEE Signal Process. Mag.*, vol. 30, no. 4, pp. 172–174, Jul. 2013.

[49] *Final Report From the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment*, document VQEG, Mar. 2000.

[50] A. Z. Khan and J. D. Crawford, "Ocular dominance reverses as a function of horizontal gaze angle," *Vis. Res.*, vol. 41, no. 14, pp. 1743–1748, Jun. 2001.

[51] O. Rosenbach, "On monocular prevalence in binocular vision," *Med Wochenschrift*, vol. 30, pp. 1290–1292, 1903.

[52] R. Blake, "A primer on binocular rivalry, including current controversies," *Brain Mind*, vol. 2, no. 1, pp. 5–38, Apr. 2001.
[53] A. P. Mapp, H. Ono, and R. Barbeito, "What does the dominant eye dominate? A brief and somewhat contentious review," *Perception Psychophys.*, vol. 65, no. 2, pp. 310–317, Feb. 2003.
[54] Z. Wang and Q. Li, "Information content weighting for perceptual image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 5, pp. 1185–1198, May 2011.
[55] Z. Wang and X. Shang, "Spatial pooling strategies for perceptual image quality assessment," in *Proc. IEEE Int. Conf. Image Process.*, Atlanta, GA, USA, Oct. 2006, pp. 2945–2948.
[56] M. J. Wainwright and E. P. Simoncelli, "Scale mixtures of Gaussians and the statistics of natural images," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 12. May 2000, pp. 855–861.
[57] L. Kaufman, *Sight and Mind: An Introduction to Visual Perception*. London, U.K.: Oxford Univ. Press, 1974.
[58] B. Julesz, *Foundations of Cyclopean Perception*. Chicago, IL, USA: Univ. Chicago Press, 1971.
[59] W. J. M. Levelt, "The alternation process in binocular rivalry," *Brit. J. Psychol.*, vol. 57, nos. 3–4, pp. 225–238, Nov. 1966.
[60] R. Blake, "Threshold conditions for binocular rivalry," *J. Experim. Psychol., Human Perception Perform.*, vol. 3, no. 2, pp. 251–257, May 1977.
[61] M. Fahle, "Binocular rivalry: Suppression depends on orientation and spatial frequency," *Vis. Res.*, vol. 22, no. 7, pp. 787–800, 1982.
[62] J. Ding and G. Sperling, "A gain-control theory of binocular combination," *Proc. Nat. Acad. Sci. USA*, vol. 103, no. 4, pp. 1141–1146, 2006.
[63] Y. Yang and R. Blake, "Spatial frequency tuning of human stereopsis," *Vis. Res.*, vol. 31, nos. 7–8, pp. 1176–1189, 1991.
[64] P. G. J. Barten, "Formula for the contrast sensitivity of the human eye," *Proc. SPIE*, vol. 5294, pp. 231–238, Jan. 2004.
[65] J. Wang, K. Zeng, and Z. Wang, "Quality prediction of asymmetrically distorted stereoscopic images from single views," in *Proc. IEEE Int. Conf. Multimedia Expo*, Chengdu, China, Jul. 2014, pp. 1–6.
[66] 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, pp. 1–6.
[67] L. Stelmach, W. J. Tam, D. Meegan, and A. Vincent, "Stereo image quality: Effects of mixed spatio-temporal resolution," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, no. 2, pp. 188–193, Mar. 2000.
[68] H. Brust, A. Smolic, K. Mueller, G. Tech, and T. Wiegand, "Mixed resolution coding of stereoscopic video for mobile devices," in *Proc. 3DTV Conf., True Vis.-Capture, Transmiss. Display 3D Video*, Potsdam, Germany, May 2009, pp. 1–4.
[69] P. Aflaki, M. M. Hannuksela, and M. Gabbouj, "Subjective quality assessment of asymmetric stereoscopic 3D video," *Signal, Image Video Process.*, vol. 9, no. 2, pp. 331–345, Feb. 2015.
[70] J. Wang, S. Wang, and Z. Wang, "Quality prediction of asymmetrically compressed stereoscopic videos," in *Proc. IEEE Int. Conf. Image Process.*, Quebec City, QC, Canada, Sep. 2015, pp. 1–5.

Jiheng Wang (S'11) received the M.Math. degree in statistics-computing from the University of Waterloo, Waterloo, ON, Canada, in 2011, where he is currently pursuing the Ph.D. degree in electrical and computer engineering. He has been a Research Assistant with the Department of Electrical and Computer Engineering, University of Waterloo, since 2011. In 2013, he was with the Video Compression Research Group, Blackberry, Waterloo. From 2009 to 2010, he was a Research Assistant with the Department of Statistics and Actuarial Science, University of Waterloo. His current research interests include 3D image and video quality assessment, perceptual 2D and 3D video coding, statistical learning, and dimensionality reduction.

Abdul Rehman received the Ph.D. degree in electrical and computer engineering from the University of Waterloo, Canada, in 2013. He is currently the President of SSIMWave, a company he co-founded in 2013. His research interests include image and video processing, coding and quality assessment, and multimedia communications.

Kai Zeng received the B.E. and M.A.Sc. degrees in electrical engineering from Xidian University, Xi'an, China, in 2006 and 2009, respectively, and the Ph.D. degree in electrical and computer engineering from the University of Waterloo, Waterloo, ON, Canada, where he is currently a Post-Doctoral Fellow with the Department of Electrical and Computer Engineering. His research interests include computational video and image pattern analysis, multimedia communications, and image and video processing (coding, denoising, analysis, and representation), with an emphasis on image and video quality assessment and corresponding applications. He was a recipient of IEEE Signal Processing Society Student Travel Grant at the 2010 and 2012 IEEE International Conference on Image Processing, and the prestigious 2013 Chinese Government Award for Outstanding Students Abroad.

Shiqi Wang (M'15) 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 Peking University, in 2014. He is currently a Post-Doctoral Fellow with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, Canada. In 2011, he was with Microsoft Research Asia, Beijing, as an Intern. He has proposed over 20 technical proposals to ISO/MPEG, ITU-T, and AVS video coding standards. His current research interests include video compression and image/video quality assessment.

Zhou Wang (S'99–M'02–SM'12–F'14) received the Ph.D. degree in electrical and computer engineering from The University of Texas at Austin, in 2001. He is currently a Professor with 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 25 000 citations (Google Scholar). He is a member of IEEE Multimedia Signal Processing Technical Committee (2013-2015). He served as an Associate Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING (2009-2014), *Pattern Recognition* (2006-present), and the *IEEE Signal Processing Letters* (2006-2010), and a Guest Editor of the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING (2013-2014 and 2007-2009), the *EURASIP Journal of Image and Video Processing* (2009-2010), and *Signal, Image and Video Processing* (2011-2013). He was a recipient of the 2014 NSERC E.W.R. Steacie Memorial Fellowship Award, the 2013 *IEEE Signal Processing Magazine* Best Paper Award, the 2009 IEEE Signal Processing Society Best Paper Award, the 2009 Ontario Early Researcher Award, and the ICIP 2008 IBM Best Student Paper Award (as a Senior Author).