Asymmetrically Compressed Stereoscopic 3D Videos: Quality Assessment and Rate-Distortion Performance Evaluation

Jiheng Wang, *Member, IEEE*, Shiqi Wang, *Member, IEEE*, and Zhou Wang, *Fellow, IEEE*

Abstract—Objective quality assessment of stereoscopic 3D video is challenging but highly desirable, especially in the application of stereoscopic video compression and transmission, where useful quality models are missing, that can guide the critical decision making steps in the selection of mixed-resolution coding, asymmetric quantization, and pre- and post-processing schemes. Here we first carry out subjective quality assessment experiments on two databases that contain various asymmetrically compressed stereoscopic 3D videos obtained from mixed-resolution coding, asymmetric transform-domain quantization coding, their combinations, and the multiple choices of post-processing techniques. We compare these asymmetric stereoscopic video coding schemes with symmetric coding methods and verify their potential coding gains. We observe a strong systematic bias when using direct averaging of 2D video quality of both views to predict 3D video quality. We then apply a binocular rivalry inspired model to account for the prediction bias, leading to a significantly improved full reference quality prediction model of stereoscopic videos. The model allows us to quantitatively predict the coding gain of different variations of asymmetric video compression, and provides new insight on the development of high efficiency 3D video coding schemes.

Index Terms—Video quality assessment, stereoscopic video, 3D video, asymmetric compression, mixed-distortion.

I. INTRODUCTION

W ITH the fast development of 3D acquisition, communication, processing and display technologies, automatic quality assessment of 3D images and videos has become ever important. 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 (in terms of both distortion types and levels), simply averaging state-of-the-art 2D image quality assessment (IQA)

or video quality assessment (VQA) measures of both views is sufficient to provide reasonably accurate quality predictions of stereoscopic images [2]–[5] and stereoscopic videos [6]. In particular, in [2], it was shown that averaging peak-signal-to-noise ratio (PSNR), structural similarity (SSIM) [7], multi-scale SSIM (MS-SSIM) [8], universal quality index (UQI) [9], and visual information fidelity (VIF) [10] measurements of left- and right-views performs equally well or better than the advanced 3D-IQA or 3D-VQA models [11]–[19] on LIVE 3D Image Database Phase I [2]. Similar results were also observed in [3], where averaging SSIM and MS-SSIM measurements of both views outperformed advanced 3D-IQA models [11], [12], [15], [18]–[20] on LIVE 3D Image Database Phase II [21]. Compared with the case of symmetric distortions, quality assessment of asymmetrically distorted stereoscopic images is much more challenging. In [3] and [22], 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 Database Phase II and Waterloo-IVC 3D Image Database Phase I and Phase II. On the other hand, our previous work [5] revealed a strong distortion type dependent prediction bias when predicting quality of asymmetrically distorted stereoscopic images from single-views.

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-resolution 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, with regard to transform-domain quantization coding, Saygili *et al.* found that asymmetric coding can perform better than symmetric coding when the lower quality view is encoded above a threshold value [23]. The subjective studies in [24] showed that stereoscopic asymmetry introduced by way of asymmetric blurriness is preferred over asymmetric blockiness, which is agreed by [25], where low-pass filtering shows no negative effect on the perceived 3D quality, sharpness and depth. In 1992, Perkins [26] introduced the idea of mixed-resolution coding for stereoscopic video and implemented a

TABLE I
SUMMARY OF EXISTING 3D VIDEO QUALITY DATABASES

Database	# of subjects	Protocol	Display	# of videos	Resolutions	Frame/Second	Controlling Parameters
LIVE 3D Video Database [30]	27	DSCQS	Active	54	720 × 480	25	H.264 compression
StSD 3D Video Database [31]	16	DSCQS	Passive	116	1920 × 1080	30	H.264 & HEVC compressions
Tampere 3D Video Database [32]	30	ACR(11)	Auto	60	1920 × 1080	30	Depth levels & H.264 compressions
MMSPG 3D Video Quality Assessment Database [33]	17	SSCQS	Passive	60	1920 × 1080	25	Different camera distances
NAMA3DS1-COSPAD1 [34]	29	ACR(5)	Active	110	1920 × 1080	25	H.264 & JPEG2000 compressions
UBC Digital Multimedia Lab 3D Video Database [35], [36]	16	DSCQS	Passive	64	1920 × 1080	24,30,48,60	HEVC compressions & Frame rates
3DVCL@FER Video Database [37]	35	ACR(5)	Active	184	1920 × 1080	25	H.264, JPEG2000 compressions, Geometric distortion, Packet losses, Frame rates & Frame-freeze

mixed-resolution coding scheme with a subsampling factor of 4, which can reduce the bit rate by 46% and resulted in little subjective degradation in picture quality and only moderate degradation in perceive depth. Brust *et al.* conducted subjective and objective tests on full and mixed-resolution stereo video coding with a downsampling factor of 2 [27]. Experimental results showed that at low bit rates mixed-resolution coded sequences have better perceptual qualities and the optimal bit rate allocation strategy is 30% to 35% of the total bit rate for the lower quality view. In [28], different vertical and horizontal spatial low-pass filtering on the right-view video were applied and subjective results showed that the perceived spatial quality and sharpness have a strong tendency towards the higher quality view and the perceived depth was unaffected. Aflaki *et al.* investigated the extent of downsampling ratios that can be applied to a low quality view without a noticeable degradation on the 3D quality [29]. They also compared the coding efficiency of symmetric quantization coding, asymmetric quantization coding and mixed-resolution coding and found that mixed-resolution coding can pronounce a similar 3D quality to that of symmetric coding with a significantly reduced computational complexity. Several observations in [22] provide useful implications on stereoscopic image/video coding. Specially, for JPEG compression, 3D image quality is more affected by the poorer quality view; while for blur, 3D image quality is more affected by the better quality view. Such distortion type dependency 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 dominates quality judgement regardless of their distortion levels. These observations suggest that simply coding one view at high rate and the other at low rate may not be a wise choice. This also suggests that a significant coding gain may be achieved by mixed-resolution coding, followed by postprocessing techniques. However, in the literature, systematic studies on subjective and objective quality assessment of these variations of asymmetric stereoscopic video coding are still lacking, making it difficult to directly compare different coding strategies, nor to derive 3D-VQA models to guide asymmetrical 3D video coding.

In this work, we first carry out subjective quality assessment experiments on two databases that contain various asymmetrically compressed stereoscopic 3D videos created by mixed-resolution coding, asymmetric transform-domain quantization

coding, their combinations, and multiple choices of post-processing techniques. We compare different variations of asymmetric stereoscopic video coding schemes with symmetric coding methods and verify their potential coding gains. We also observe a strong systematic bias when using direct averaging of 2D video quality of both views to predict 3D video quality. We then apply a binocular rivalry inspired model to account for the prediction bias, leading to a significantly improved full reference quality prediction model for stereoscopic videos. The model allows us to quantitatively predict the coding gain of different variations of asymmetric video compression, and provides new insight on the development of high efficiency 3D video coding schemes.

II. REVIEW OF PREVIOUS 3D-VQA STUDIES

To the best of our knowledge, there are currently 7 subject-rated video databases that are commonly recognized in the 3D-VQA research community. Table I lists these databases with detailed descriptions.

Existing objective 3D-VQA methods may be grouped into two categories. The first type of approaches are built directly upon successful 2D-IQA/VQA methods. In [38] and [39], 2D-IQA measures, including PSNR, SSIM, and video quality metric (VQM) [40], were applied to the left- and right-view images/videos of 3D videos separately and then combined to a 3D quality score. Both experimental results showed that VQM performs better than PSNR and SSIM. In [41], PSNR and VSSIM [42], 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 [30], PSNR and MS-SSIM were applied to estimate 3D image quality and overall 3D quality-of-experience. The subjective testing results showed that MS-SSIM slightly outperforms PSNR with respect to both 3D visual experience criteria.

The second type of 3D-VQA approaches focus on building 3D quality models directly without relying on existing 2D-IQA/VQA algorithms. Zhu and Wang [17] proposed a 3D-VQA model by considering depth perception and their experimental results showed that it performs better than MSE and PSNR. In [43] and [44], 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 stacks and then analyzed by 3D-DCT. The experimental results showed that the model outperforms PSNR, SSIM,
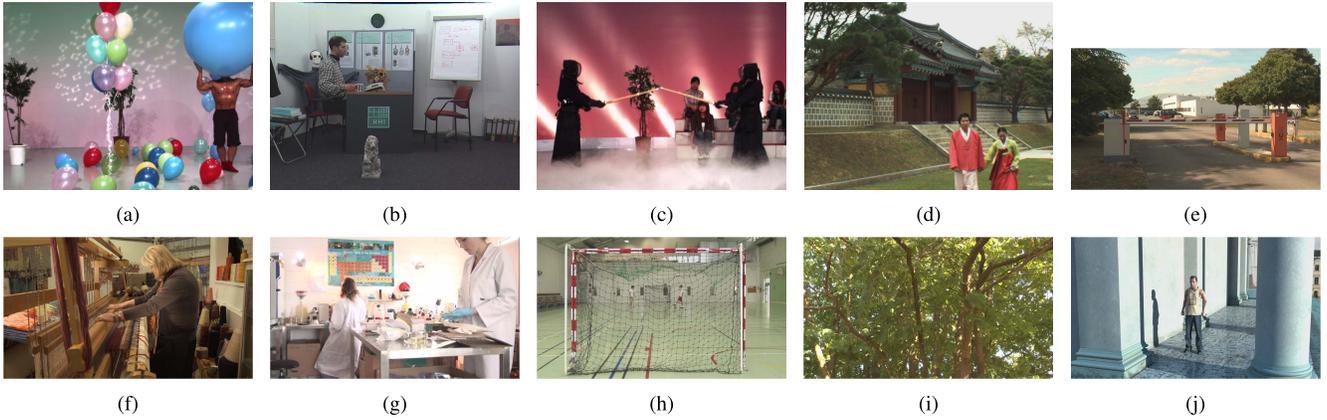
Fig. 1. Sample frames from the pristine videos used in the subjective study. Only the right-views are shown here. Phase I: (a) Balloons, (b) Book, (c) Kendo, (d) Lovebird; Phase II: (e) Barrier, (f) Craft, (g) Laboratory, (h) Soccer, (i) Tree, (j) Dancer.

MS-SSIM, and UQI on Tampere 3D Video Database [32]. In [45], a SSIM-inspired 3D-VQA model using depth map segmentation was proposed followed by an extensive subjective test. The results indicated that the model can predict perceived 3D video quality effectively. In [46], an objective 3D-VQA algorithm using blocking artifacts, blurring in edge regions, and video quality difference between two views was proposed. The subjective testing results showed that the model outperforms SSIM and VQM. In [31], a binocular suppression inspired StSD metric was proposed based on a comprehensive subjective study. The results indicated that the StSD model significantly outperforms SSIM and the aforementioned 3D-VQA models [43], [46] on StSD 3D Video Database [31].

III. SUBJECTIVE STUDY

A. WATERLOO-IVC 3D Video Quality Databases Phase I and Phase II

The new Waterloo-IVC 3D Video Quality Database Phase I [6] is created from 4 pristine multi-view 3D videos, i.e., Balloons, Book, Kendo, and Lovebird, which are commonly used 3D HEVC testing sequences. The new Waterloo-IVC 3D Video Quality Database Phase II is created from 6 pristine stereoscopic 3D videos, i.e., Barrier, Craft, Laboratory, Soccer, Tree, and Dancer, which were collected from previous subjective 3D video quality studies [34], [47]. All source video sequences in Phase I and Phase II were examined and selected by multiple members of the Lab for Image and Vision Computing at University of Waterloo to make sure that they have good 3D effects and do not cause any obvious visual discomfort [48]–[51]. The details of the all test videos are given in Table II. All videos are in YUV4:2:0 format. Sample frames for each test sequence are shown in Fig. 1.

Waterloo-IVC 3D Video database Phase I include stereoscopic 3D videos obtained from symmetric and asymmetric transform-domain quantization coding followed by different levels of low-pass filtering. Each single-view video was compressed using an HEVC encoder [52] by five levels of transform-domain quantization with QP = {25, 35, 40, 45, 50} in low-delay main profile. The single-view videos were employed to generate compressed stereoscopic videos,

TABLE II
TEST VIDEOS IN WATERLOO-IVC 3D VIDEO DATABASES
PHASE I AND PHASE II

	Resolution	Length	Frames/Second	Views
Book	1024×768	6s	16.67	View 6 & View 8
Balloons	1024×768	10s	30.00	View 1 & View 3
Kendo	1024×768	10s	30.00	View 1 & View 3
Lovebird	1024×768	10s	30.00	View 4 & View 6
Barrier	1920×1080	10s	30.00	N/A
Craft	1920×1080	10s	30.00	N/A
Laboratory	1920×1080	10s	30.00	N/A
Soccer	1920×1080	10s	30.00	N/A
Tree	1920×1080	10s	30.00	N/A
Dancer	1920×1088	10s	30.00	View 1 & View 5

either symmetrically or asymmetrically. There are 11 different kinds of combinations as listed in Table III. The lower and higher QP views are assigned to the left-view or the right-view randomly. Moreover, for each QP combination, four levels of Gaussian low-pass filtering with $\sigma = \{0, 3.5, 7.5, 11.5\}$ are applied to the higher QP (lower quality) views. Altogether, there are totally 176 3D videos in the database.

Waterloo-IVC 3D Video database Phase II include various stereoscopic 3D videos obtained from mixed-resolution coding, asymmetric transform-domain quantization coding, their combinations, and different levels of low-pass filtering. Three choices of pre-processing, i.e., pre-downsampling by 2, pre-downsampling by 4, and pre-processed Gaussian low-pass filtering with $\sigma = 2.5$, are applied to each single-view video. Then the single-view video was compressed using an HEVC encoder [52] by different levels of transform-domain quantization with QP = {25, 30, 35, 40, 45} in low-delay main profile. The single-view videos were employed to generate compressed stereoscopic videos after upsampling if needed, either symmetrically or asymmetrically. Table III categorizes all combinations into nine groups with detailed descriptions. The lower- and higher-quality views are assigned to the left- or right-view randomly. Moreover, for each combination, two levels of Gaussian low-pass filtering with $\sigma = \{3.5, 5.5\}$ are applied to the lower-quality views. Altogether, there are totally 222 2D videos and 528 3D videos in the database.

There are two unique features of the new databases (including both Phases I and II) when compared with existing publicly known 3D-VQA databases. First, these are the only

TABLE III

TEST VIDEOS IN WATERLOO-IVC 3D VIDEO DATABASES PHASE I AND PHASE II

		Waterloo-IVC 3D Video Database Phase I	
Group	# of Videos	Description	Combinations
3D.1.a	4×4	Symmetrically compressed stereoscopic videos	
3D.1.b	4×4	Postprocessing P1: $\sigma = 3.5$	(QP25,QP25), (QP35,QP35), (QP40,QP40) and
3D.1.c	4×4	Postprocessing P3: $\sigma = 7.5$	(QP50,QP50)
3D.1.d	4×4	Postprocessing P4: $\sigma = 11.5$	
3D.2.a	4×7	Asymmetrically compressed stereoscopic videos	
3D.2.b	4×7	Postprocessing P1: $\sigma = 3.5$	(QP25,QP35), (QP25,QP40), (QP25,QP45), (QP25,QP50),
3D.2.c	4×7	Postprocessing P3: $\sigma = 7.5$	(QP35,QP45), (QP35,QP50) and (QP40,QP50)
3D.2.d	4×7	Postprocessing P4: $\sigma = 11.5$	

		Waterloo-IVC 3D Video Database Phase II	
Group	# of Videos	Description	Combinations
2D.1.a	6×4	Compressed single-view videos	
2D.1.b	6×4	Postprocessing P1: $\sigma = 3.5$	QP30, QP35, QP40 and QP45
2D.1.c	6×4	Postprocessing P2: $\sigma = 5.5$	
2D.2	6×4	Compressed single-view videos with pre-processing: S1 $\sigma = 2.5$	S1-QP30, S1-QP35, S1-QP40 and S1-QP45
2D.3.a	6×4	Compressed single-view videos with pre-downsampling by 2: D2	
2D.3.b	6×4	Postprocessing P1: $\sigma = 3.5$	D2-QP25, D2-QP30, D2-QP35 and D2-QP40
2D.3.c	6×4	Postprocessing P2: $\sigma = 5.5$	
2D.4.a	6×3	Compressed single-view videos with pre-downsampling by 4: D4	
2D.4.b	6×3	Postprocessing P1: $\sigma = 3.5$	D4-QP20, D4-QP25 and D4-QP30
2D.4.c	6×3	Postprocessing P2: $\sigma = 5.5$	
3D.3.a	6×4	Symmetrically compressed stereoscopic videos	
3D.3.b	6×4	Postprocessing P1: $\sigma = 3.5$	(QP30,QP30), (QP35,QP35), (QP40,QP40) and
3D.3.c	6×4	Postprocessing P2: $\sigma = 5.5$	(QP45,QP45)
3D.4.a	6×6	Asymmetrically compressed stereoscopic videos	
3D.4.b	6×6	Postprocessing P1: $\sigma = 3.5$	(QP30,QP35), (QP30,QP40), (QP30,QP45), (QP35,QP40),
3D.4.c	6×6	Postprocessing P2: $\sigma = 5.5$	(QP35,QP45) and (QP40,QP45)
3D.5	6×10	Asymmetrically compressed stereoscopic videos with pre-processing: S1 $\sigma = 2.5$	(QP30,S1-QP30), (QP30,S1-QP35), (QP30,S1-QP40), (QP30,S1-QP45), (QP35,S1-QP35), (QP35,S1-QP40), (QP35,S1-QP45), (QP40,S1-QP40), (QP40,S1-QP45) and (QP45,S1-QP45)
3D.6.a	6×8	Asymmetrically compressed stereoscopic videos with pre-downsampling by 2: D2	(QP30,D2-QP25), (QP30,D2-QP30), (QP30,D2-QP35), (QP35,D2-QP30), (QP35,D2-QP35), (QP35,D2-QP40),
3D.6.b	6×8	Postprocessing P1: $\sigma = 3.5$	(QP40,D2-QP35) and, (QP40,D2-QP40)
3D.6.c	6×8	Postprocessing P2: $\sigma = 5.5$	
3D.7.a	6×8	Asymmetrically compressed stereoscopic videos with pre-downsampling by 4: D4	(QP30,D4-QP20), (QP30,D4-QP25), (QP30,D4-QP30), (QP35,D4-QP25), (QP35,D4-QP30), (QP35,D4-QP35),
3D.7.b	6×8	Postprocessing P1: $\sigma = 3.5$	(QP40,D4-QP25) and (QP40,D4-QP30)
3D.7.c	6×8	Postprocessing P2: $\sigma = 5.5$	

databases that allow us to perform subjective test on both 2D and 3D videos. The inclusion of 2D videos allows us to directly examine the relationship between the perceptual quality of stereoscopic video and that of its single-view videos. This is advantageous against previous studies which do not have ground truth of 2D video quality but have to rely on objective 2D-VQA measures to provide estimates. Second, these are the only databases that contain asymmetrically compressed stereoscopic videos from mixed-resolution coding, asymmetric transform-domain quantization coding and their combinations, followed by different levels of low-pass filtering. This provides the potential of a much stronger test on 3D-VQA models on their generalizability to real world applications. Such test has been largely lacking in previous studies where the development of objective 3D-VQA models only took into account asymmetric distortions of specific and very limited distortion types such as compression only. Meanwhile, a broader variety of test scenarios allows us to perform a more comprehensive comparison on different variations of asymmetric stereoscopic video coding schemes with symmetric coding methods and thus to evaluate their potential coding gains.

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

TABLE IV

VIEWING CONDITIONS IN 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

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 the viewing conditions are given in Table IV.

In Phase I, twenty-two naïve subjects, 12 males and 10 females aged between 22 and 35, participated in the study. In Phase II, thirty-two naïve subjects, 20 males and 12 females aged between 24 and 37, participated in the study. A 3D vision test (Random dot stereo test) was conducted first to verify their ability to view stereoscopic 3D content and no one failed the vision test. As a result, a total of twenty-two and thirty-two subjects proceeded to the formal test in Phase I and Phase II, respectively. 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.
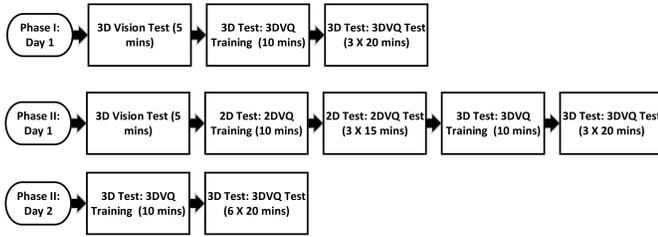
Fig. 2.   The procedure of the subjective test in Waterloo-IVC 3D Video Database Phase I and Phase II.

TABLE V
p-VALUES FROM THE ONE-SAMPLE t-TEST FOR DIFFERENT TEST VIDEO GROUPS

Group	p-values
Group 3D.3.a	0.5901
Group 3D.4.a	0.4255
Group 3D.3.b-c and Group 3D.4.b-c	0.1401
Group 3D.5	0.6637
Group 3D.6.a	0.6833
Group 3D.6.b-c	0.3172
Group 3D.7.a	0.4379
Group 3D.7.b-c	0.3495

The subjects were asked to evaluate their overall 3D viewing experience − 3D Video Quality (3DVQ) in this study. Since to visualize every stereoscopic 3D video, the subjects need to make readjustment 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. To reduce this effect, we opt to the single stimulus procedure using an 11-grade numerical categorical scale (SSNCS) protocol. A general introduction was given at the beginning of the whole test, and more specific instructions and training session were given afterwards. The rating strategy was introduced and the subjects were required to practice by giving scores to training 2D/3D videos until they fully understood the criteria and built up their own scoring strategies. For both Phase I and Phase II tests, we use three types of videos in the training phase: pristine 2D/3D videos, moderately distorted 2D/3D videos, and highly-distorted 2D/3D videos. The subjects were told to give scores at the high end (close to 10 pts) to the pristine 2D/3D videos, at the mid-range to the moderately distorted 2D/3D videos, and at the low end (close to 0 pts) to the highly-distorted videos. With regard to Phase II, we also found that the 2D perceptual quality of left- and right-view videos are very close to each other at the same compression or postprocessing levels, and 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.1 to Group 2D.4 in the formal test.

Most stimuli were shown once in each test. However, there were 12 repetitions for single-view or stereoscopic videos, which means that for each subject, her/his first 12 single-view or stereoscopic videos were shown twice. The order of stimuli was randomized and the consecutive testing stereoscopic videos were from different source contents. There are three sessions for the 3D test in Phase I, while in Phase II, there are three sessions for the 2D test and nine sessions for the 3D test. Each single session, where around 80 single-view or 60 stereoscopic videos were evaluated, was finished in 15 to 20 minutes. Sufficient relaxation periods (5 minutes or more) were given between sessions. Thus in Phase I all sessions were finished in 2 hours in one day. In Phase II, the test was scheduled on two consecutive days for each subject. Day 1 (2 to 2.5 hours) was dedicated to all 2D sessions and the first three 3D sessions and Day 2 (2 to 2.5 hours) to the remaining six 3D sessions. Fig. 2 shows the detailed procedure of our formal subjective test.

Moreover, we found that repeatedly switching between viewing 3D videos 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 scores. All these efforts were intended to reduce visual fatigue and discomfort of the subjects.

C. 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 [53]. When studying visual quality of asymmetrically compressed stereoscopic videos, 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 analysis on the impact of eye dominance in the perception of asymmetrically compressed stereoscopic videos in Phase II. The side of the dominant eye under static conditions was checked first by Rosenbach's test [54]. This test examines which eye determines the position of a finger when the subject is asked to point to an object. Among thirty-two subjects who finished the formal test in Phase II, thirteen subjects (8 males, 5 females) are left-eye dominant, and the others (12 males, 7 females) are right-eye dominant.

The 3DVQ MOS scores for each video in Phase II were computed for left-eye dominant subjects and right-eye dominant subjects, denoted as $3DVQ_L$ and $3DVQ_R$, respectively. We employed the one-sample t-test to obtain a test decision for the null hypothesis that the difference between $3DVQ_L$ and $3DVQ_R$, i.e., $3DVQ_D = 3DVQ_L - 3DVQ_R$, comes from a normal distribution 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 different test video groups are reported in Table V. From Table V, it can be seen that the null hypothesis cannot be rejected at the 5% significance level, which indicates that the impact of eye dominance in the perception of asymmetrically compressed stereoscopic videos is insignificant.

It is worth noting that similar conclusions were reached in our earlier studies on the impact of eye dominance on the quality of asymmetrically distorted stereoscopic images [22] and on the depth perception induced by stereo cues of asymmetrically distorted stereograms [55]. These observations are

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

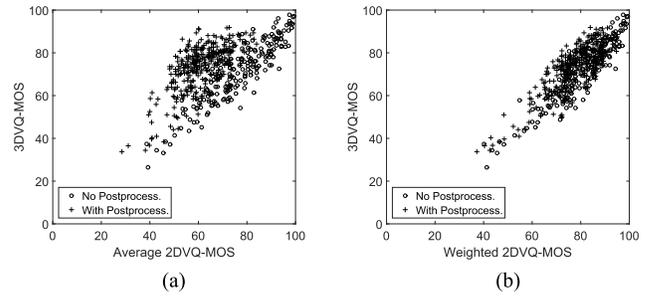| Waterloo-IVC 3D Video Database Phase I			
Method	PLCC	SRCC	RMSE
PSNR (average)	0.7085	0.5336	15.4507
PSNR (proposed weighting)	0.8980	0.8366	9.6344
SSIM (average)	0.3964	0.2872	20.1010
SSIM (proposed weighting)	0.8905	0.8393	9.9615
MS-SSIM (average)	0.4072	0.2969	19.9978
MS-SSIM (proposed weighting)	0.8838	0.8287	10.2448
IW-SSIM (average)	0.4833	0.2787	19.1683
IW-SSIM (proposed weighting)	0.8942	0.8364	9.8035
VQM (average)	0.7912	0.6321	13.3905
VQM (proposed weighting)	0.9191	0.8655	8.6273
Waterloo-IVC 3D Video Database Phase II			
Method	PLCC	SRCC	RMSE
2DVQ-MOS (average)	0.6912	0.6277	8.9039
2DVQ-MOS (proposed weighting)	0.8829	0.8727	5.7849
PSNR (average)	0.3699	0.3414	11.4465
PSNR (proposed weighting)	0.5590	0.5109	10.2154
SSIM (average)	0.3303	0.2589	11.6291
SSIM (proposed weighting)	0.7571	0.7309	8.0487
MS-SSIM (average)	0.3034	0.2503	11.7395
MS-SSIM (proposed weighting)	0.6813	0.6377	9.0188
IW-SSIM (average)	0.3243	0.2459	11.6545
IW-SSIM (proposed weighting)	0.7677	0.7423	7.8943
VQM (average)	0.7019	0.6287	8.7759
VQM (proposed weighting)	0.8496	0.8042	6.4976

Fig. 3. 3DVQ-MOS versus predictions from 2DVQ-MOS values of 2D left- and right-views on Waterloo-IVC Phase II. (a) Average of 2DVQ-MOS. (b) Weighted average of 2DVQ-MOS by the proposed method.

consistent with the "stimulus" view of rivalry that is widely accepted in the field of visual neuroscience [56]. A comprehensive review and discussion on "stimulus" rivalry versus "eye" rivalry can be found in [56] and [57].

IV. ANALYSIS AND FINDINGS

A. Relationship Between 2D and 3D Video Quality

Following the previous work [7], the raw 2DVQ and 3DVQ scores given by each subject were converted to Z-scores [58], 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 video, i.e., MOS-2DVQ and MOS-3DVQ, were computed after removing outliers [59]. Given the subjective 2D and 3D data, we are interested in how single-view 2D video quality predicts stereoscopic 3D video quality, especially for the case of asymmetrically compressed and post-processed stereoscopic videos. The most straightforward 2D-to-3D quality prediction method is to average the qualities of the left- and right-view videos. Table VI shows the Pearson's linear correlation coefficient (PLCC), Spearman's rank-order correlation coefficient (SRCC), and Root Mean Squared Error (RMSE) between 3DVQ-MOS scores and average 2DVQ-MOS scores (Phase II Only), where the 2D-IQA/VQA measurements include PSNR, SSIM, MS-SSIM, information content weighted SSIM (IW-SSIM) [60], and VQM for all stereoscopic videos in Phase I and Phase II. Table VII and VIII reports PLCC, SRCC, and RMSE values for different test video groups in Phase I and Phase II, respectively. PLCC and RMSE are adopted to evaluate prediction accuracy [61] and SRCC is employed to assess prediction monotonicity [61]. Higher PLCC and SRCC and lower RMSE indicate better consistency with human opinions of quality. PLCC and RMSE are 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 is 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.

In Table VII and VIII, we first compare the performance of symmetrically compressed 3D videos without postprocessing against asymmetrically compressed 3D videos without postprocessing. Unsurprisingly, accurate predictions are obtained in the category of symmetrically compressed 3D videos. By contrast, the performance drops for asymmetrically compressed 3D videos. In [62], we reported that for JPEG compression, average prediction overestimates 3D quality (or 3D quality is more affected by the poorer quality view). More importantly we found that for blockiness, the bias of the averaging prediction model increases with the level of distortions, and thus whether the bias is pronounced depends on the quality range being investigated. With respect to blockiness created from HEVC compression, this overestimated prediction bias is still pronounced, but not as strong as JPEG compression, which is likely due to the reduction of blocking artifacts in HEVC.

We then compare the performance of compressed 3D videos without postprocessing against compressed 3D videos with postprocessing. From Table VII and VIII, it can be observed that the direct averaging model performs well for 3D videos without postprocessing (by Gaussian blurring). By contrast, the correlation values drop significantly for videos with postprocessing. The Fig. 3 (a) and the first and the third columns of Fig. 4 show the corresponding scatter plots between 3DVQ-MOS scores and 2DVQ-MOS scores or 2D-IQA/VQA measurements, where the simple averaging prediction model generates substantial bias on many stereoscopic videos. In [62], we reported that for blurriness, average prediction often underestimates 3D quality (or 3D quality is more affected by the better quality view). Here the same kind of prediction bias is clearly observed, as direct averaging of state-of-the-art 2D-IQA/VQA metrics always underestimates 3D video quality for these post-processed videos.

B. Quality of Asymmetric Stereoscopic Video With Postprocessing

Given the subjective data, the second question we would like to ask is how Gaussian low-pass post-filtering affects the perceptual 3D quality of asymmetrically compressed stereoscopic videos. Table IX and X report 3DVQ-MOS changes

TABLE VII

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

	PLCC		SRCC		RMSE	
Method	Sym. Compress.	Asym. Compress.	Sym. Compress.	Asym. Compress.	Sym. Compress.	Asym. Compress.
PSNR (average)	0.9839	0.8217	0.9581	0.8226	5.3453	10.8551
PSNR (proposed weighting)	0.9839	0.8798	0.9581	0.8311	5.3437	9.0531
SSIM (average)	0.9876	0.8560	0.9478	0.7960	4.6806	9.8468
SSIM (proposed weighting)	0.9877	0.8843	0.9478	0.8335	4.6782	8.8952
MS-SSIM (average)	0.9851	0.8445	0.9478	0.7990	5.1417	10.2018
MS-SSIM (proposed weighting)	0.9851	0.8746	0.9478	0.8363	5.1391	9.2357
IW-SSIM (average)	0.9937	0.9304	0.9581	0.9017	3.3373	6.9836
IW-SSIM (proposed weighting)	0.9937	0.9322	0.9581	0.9012	3.3379	6.8942
VQM (average)	0.9876	0.8831	0.9669	0.8283	4.6878	8.9408
VQM (proposed weighting)	0.9876	0.8770	0.9669	0.8527	4.6893	9.1511
Symmetric compression: Group 3D.1.a; Asymmetric compression: Group 3D.2.a.						
	PLCC		SRCC		RMSE	
Method	No Postprocess.	With Postprocess.	No Postprocess.	With Postprocess.	No Postprocess.	With Postprocess.
PSNR (average)	0.9131	0.7382	0.8922	0.6580	9.6120	18.7419
PSNR (proposed weighting)	0.9146	0.8750	0.8831	0.8165	9.5338	10.1255
SSIM (average)	0.9193	0.4824	0.9028	0.2621	9.2752	18.3211
SSIM (proposed weighting)	0.9358	0.8876	0.9197	0.8056	8.3072	9.6338
MS-SSIM (average)	0.9132	0.5315	0.9009	0.2701	9.6036	17.7168
MS-SSIM (proposed weighting)	0.9299	0.8798	0.9202	0.7896	8.6678	9.9415
IW-SSIM (average)	0.9601	0.5935	0.9446	0.3262	6.5882	16.8340
IW-SSIM (proposed weighting)	0.9641	0.8977	0.9491	0.8135	6.2567	9.2176
VQM (average)	0.9265	0.8733	0.9076	0.7140	8.8679	10.1899
VQM (proposed weighting)	0.9282	0.9341	0.9041	0.8414	8.7715	7.4652
No Postprocessing: Group 3D.1.a and Group 3D.2.a; With Postprocessing: Group 3D.1.b-d and Group 3D.2.b-d.						

TABLE VIII

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

	PLCC		SRCC		RMSE	
Method	Sym. Compress.	Asym. Compress.	Sym. Compress.	Asym. Compress.	Sym. Compress.	Asym. Compress.
2DVQ-MOS (average)	0.9676	0.8645	0.9470	0.8387	5.2708	6.3878
2DVQ-MOS (proposed weighting)	0.9676	0.9002	0.9470	0.8984	5.2708	5.5350
PSNR (average)	0.7717	0.3718	0.7191	0.3323	13.2946	11.7976
PSNR (proposed weighting)	0.7716	0.5027	0.7191	0.4468	13.2957	10.9864
SSIM (average)	0.9236	0.5454	0.8974	0.5332	8.0089	10.6524
SSIM (proposed weighting)	0.9236	0.7625	0.8974	0.7610	8.0091	8.2224
MS-SSIM (average)	0.8733	0.4658	0.8609	0.4431	10.1749	11.2459
MS-SSIM (proposed weighting)	0.8733	0.6697	0.8609	0.6534	10.1751	9.4382
IW-SSIM (average)	0.9335	0.5460	0.9209	0.5253	7.4910	10.6476
IW-SSIM (proposed weighting)	0.9335	0.7836	0.9209	0.7828	7.4923	7.8957
VQM (average)	0.9586	0.8107	0.9226	0.7731	5.9446	7.4397
VQM (proposed weighting)	0.9586	0.8633	0.9226	0.8457	5.9448	6.4137
Symmetric compression: Group 3D.3.a; Asymmetric compression: Group 3D.4.a, Group 3D.5, Group 3D.6.a and Group 3D.7.a.						
	PLCC		SRCC		RMSE	
Method	No Postprocess.	With Postprocess.	No Postprocess.	With Postprocess.	No Postprocess.	With Postprocess.
2DVQ-MOS (average)	0.8769	0.7800	0.8540	0.7394	6.6768	6.9184
2DVQ-MOS (proposed weighting)	0.9144	0.8603	0.9097	0.8414	5.6234	5.6356
PSNR (average)	0.4205	0.4204	0.3811	0.3772	12.6039	10.0312
PSNR (proposed weighting)	0.5411	0.5773	0.4905	0.5385	11.6823	9.0271
SSIM (average)	0.5841	0.2596	0.5694	0.2459	11.2760	10.6766
SSIM (proposed weighting)	0.7890	0.7493	0.7835	0.7084	8.5352	7.3217
MS-SSIM (average)	0.5111	0.2911	0.4868	0.2286	11.9400	10.5769
MS-SSIM (proposed weighting)	0.7040	0.6708	0.6864	0.6014	9.8656	8.1992
IW-SSIM (average)	0.5710	0.3875	0.5520	0.3427	11.4040	10.1920
IW-SSIM (proposed weighting)	0.8068	0.7588	0.8002	0.7159	8.2076	7.2011
VQM (average)	0.8231	0.8051	0.7911	0.7155	7.8901	6.5569
VQM (proposed weighting)	0.8840	0.8326	0.8674	0.7590	6.4932	6.1242
No Postprocessing: Group 3D.3.a, Group 3D.4.a, Group 3D.5, Group 3D.6.a and Group 3D.7.a; With Postprocessing: Group 3D.3.b-c, Group 3D.4.b-c, Group 3D.6.b-c and Group 3D.7.b-c.						

after applying different levels of Gaussian low-pass filtering with respect to different QP combinations and blurring levels for Phase I and Phase II, respectively. It can be observed that for symmetrically compressed 3D videos, blurring reduces perceptual 3D video quality in most cases. By contrast, for asymmetrically compressed 3D videos, blurring on the lower quality views *improves* the perceptual 3D video quality when the quality difference of left- and right-view is high. Generally, the improvement increases with the level of blurring and with the quality difference between the higher view and the lower view. Table X also includes the cases of asymmetrically compressed stereoscopic videos with pre-downsampling by

factors of 2 and 4, where it can be seen that this 3DVQ-MOS improvement is less pronounced especially for the case of pre-downsampling by 4. This analysis verifies that the adoption of certain postprocessing techniques such as blurring could improve the efficiency of stereoscopic video coding but may not always work well for the cases of pre-downsampling.

C. Rate-Distortion Performance of Mixed Distortion Asymmetric Stereoscopic Video

The third question we would like to ask is what is the rate-distortion (R-D) performance of different mixed-distortion asymmetric stereoscopic video coding schemes.
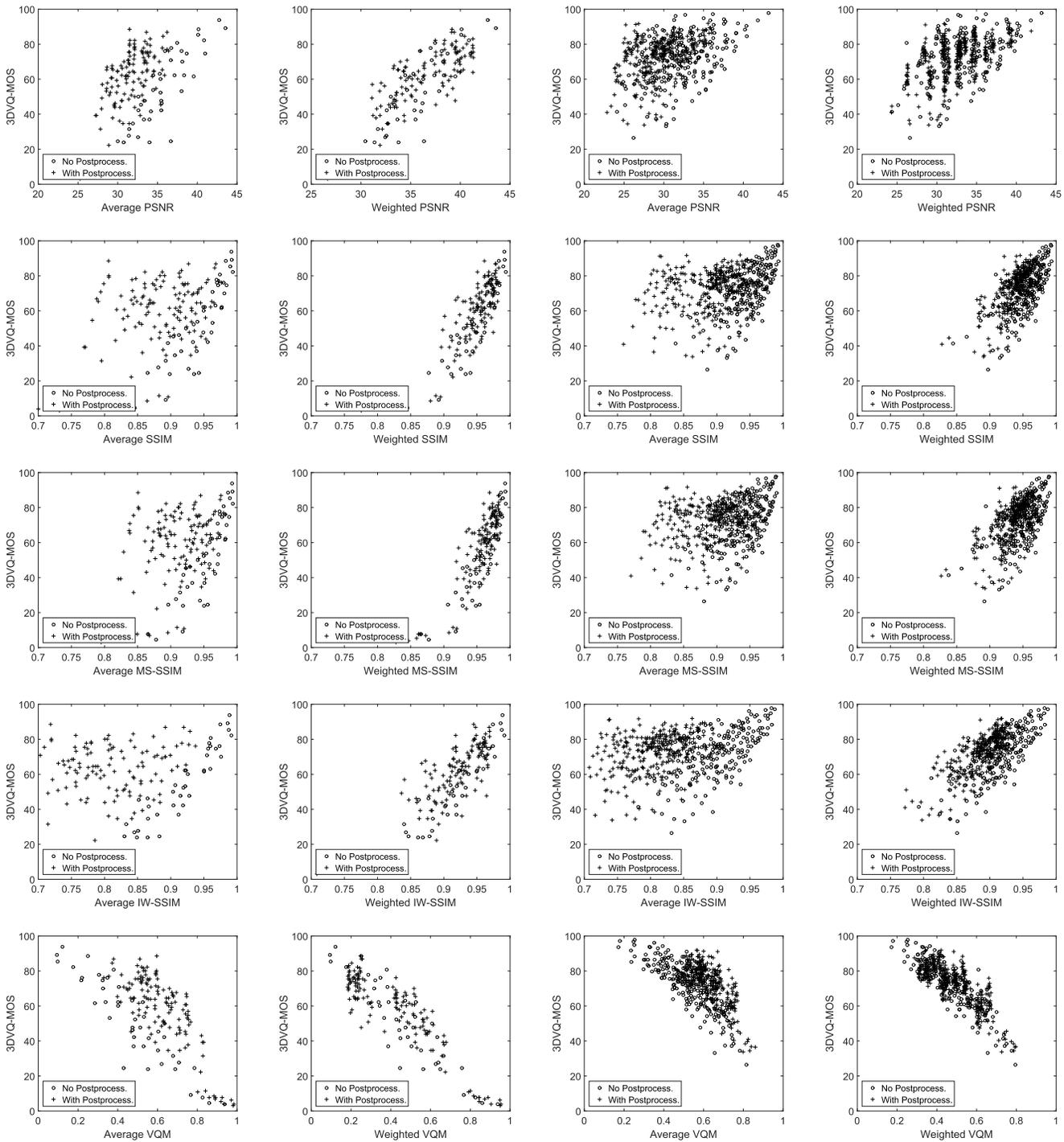
Fig. 4. 3DVQ-MOS versus predictions from 2D-IQA/VQA estimations of 2D left- and right-views. First column: predictions by averaging 2D-IQA/VQA estimations of both views on Waterloo-IVC Phase I; Second column: predictions by weighted averaging 2D-IQA/VQA estimations of both views on Waterloo-IVC Phase I using the proposed method; Third column: predictions by averaging 2D-IQA/VQA estimations of both views on Waterloo-IVC Phase II; Fourth column: predictions by weighted averaging 2D-IQA/VQA estimations of both views on Waterloo-IVC Phase II using the proposed method. First row: PSNR as the base 2D-IQA/VQA model; Second row: SSIM as the base 2D-IQA/VQA model; Third row: MS-SSIM as the base 2D-IQA/VQA model; Fourth row: IW-SSIM as the base 2D-IQA/VQA model; Fifth row: VQM as the base 2D-IQA/VQA model.

Table XI and XII report the so-called Bjontegaard delta rate (BD-Rate) in terms of 3DVQ-MOS by comparing the mixed-distortion asymmetric stereoscopic video coding schemes with the symmetric coding method for each test sequence in Phase I and Phase II, respectively [63]. It can be seen that most mixed-distortion asymmetric stereoscopic

video coding schemes achieve better R-D performance over the symmetric coding method (Group 3D.1.a in Phase I and Group 3D.3.a in Phase II).

For Phase I, the asymmetric compression-only scheme (Group 3D.2.a) degrades the R-D performance significantly due to the fact that 3D quality had a tendency towards the

TABLE IX

3DVQ-MOS CHANGES AFTER APPLYING DIFFERENT LEVELS OF
GAUSSIAN BLURRING AS POSTPROCESSING ON WATERLOO-IVC
3D VIDEO DATABASE PHASE I

Combinations		3DVQ-MOS Changes		
QP_l	QP_h	P1 $\sigma = 3.5$	P3 $\sigma = 7.5$	P4 $\sigma = 11.5$
25	25	-10.96	-16.15	-11.35
35	35	-0.19	-7.31	-4.81
40	40	-3.46	+0.19	-1.35
50	50	+1.54	+1.15	-0.96
25	35	-2.50	-3.27	-1.35
25	40	+7.69	+6.92	+7.12
25	45	+2.12	+16.92	+14.23
25	50	+14.81	+31.35	+27.12
35	45	+11.92	+13.08	+17.50
35	50	+11.35	+24.23	+29.62
40	50	+7.31	+16.35	+19.42

TABLE X

3DVQ-MOS CHANGES AFTER APPLYING DIFFERENT LEVELS OF
GAUSSIAN BLURRING AS POST-PROCESSING ON WATERLOO-IVC
3D VIDEO DATABASE PHASE II

Original Resolution					
Combinations		2DVQ-MOS		3DVQ-MOS Changes	
QP_l	QP_h	$2DVQ_h$	$2DVQ_l$	P1 $\sigma = 3.5$	P2 $\sigma = 5.5$
30	30	97.85	97.85	-12.58	-11.61
35	35	90.04	90.04	-7.40	-7.40
40	40	69.10	69.10	-1.67	-5.32
45	45	48.75	48.75	+9.68	+6.40
30	35	97.85	90.04	-1.93	-3.76
30	40	97.85	69.10	+5.72	+3.62
30	45	97.85	48.75	+7.80	+11.40
35	40	90.04	69.10	+1.66	+2.52
35	45	90.04	48.75	+5.54	+5.48
40	45	69.10	48.75	+2.35	+0.90
Downsample by 2					
Combinations		2DVQ-MOS		3DVQ-MOS Changes	
QP_l	QP_h	$2DVQ_h$	$2DVQ_l$	P1 $\sigma = 3.5$	P2 $\sigma = 5.5$
30	D2-25	97.85	85.75	-1.34	-3.66
30	D2-30	97.85	76.83	-3.76	-5.22
30	D2-35	97.85	61.18	+0.43	-0.91
35	D2-30	90.04	76.83	+1.18	-1.99
35	D2-35	90.04	61.18	+4.14	+3.39
35	D2-40	90.04	39.73	+6.45	+6.88
40	D2-35	69.10	61.18	+1.40	-0.81
40	D2-40	69.10	39.73	+2.10	+1.94
Downsample by 4					
Combinations		2DVQ-MOS		3DVQ-MOS Changes	
QP_l	QP_h	$2DVQ_h$	$2DVQ_l$	P1 $\sigma = 3.5$	P2 $\sigma = 5.5$
30	D4-25	97.85	57.58	+2.47	+0.81
30	D4-30	97.85	50.75	-0.97	-2.58
30	D4-35	97.85	39.73	+1.72	-1.02
35	D4-25	90.04	57.58	+0.59	-1.05
35	D4-30	90.04	50.75	-1.02	-2.69
35	D4-35	90.04	39.73	+1.94	-0.05
40	D4-25	69.10	57.58	+3.12	+3.49
40	D4-30	69.10	50.75	+5.97	+1.94

TABLE XI

R-D PERFORMANCE COMPARISONS OF ASYMMETRIC STEREOSCOPIC
VIDEO CODING IN TERMS OF 3DVQ-MOS ON WATERLOO-IVC
3D VIDEO DATABASE PHASE I (ANCHOR: SYMMETRIC
STEREOSCOPIC VIDEO CODING)

3DVQ-MOS					
Group	Balloons	Book	Kendo	Lovebird	Average
3D.2.a	+88.3%	+37.5%	+55.8%	+164.2%	+86.4%
3D.2.b	+60.2%	+16.0%	+39.4%	+50.6%	+41.6%
3D.2.c	-16.8%	-0.6%	+23.2%	-26.1%	-5.1%
3D.2.d	-25.7%	-30.0%	-14.4%	-9.5%	-19.9%

lower quality view with respect to blockiness [22]. However, this degradation is less pronounced and *is even reverted* when we increase the level of Gaussian low-pass post-filtering (Group 3D.2.b-d), but the optimal level of this postprocessing is content dependent.

For Phase II, similarly, the asymmetric compression plus postprocessing scheme (Group 3D.4.b-c) improves the R-D performance while the asymmetric compression-only scheme (Group 3D.4.a) reduces the R-D performance. It can also be observed that the proposed pre-downsampling plus postprocessing asymmetric coding schemes perform even better than the asymmetric compression plus postprocessing scheme, which is consistent with the previous subjective studies [29] that mixed-resolution stereoscopic video coding can achieve the best coding efficiency. By comparing different post-processing levels, it is found that the scheme that applies pre-downsampling by a factor of 2 followed by postprocessing P1 ($\sigma = 3.5$), (i.e., Group 3D.6.b), provides the best overall performance. It is also interesting to note that for the case of no pre-downsampling or pre-downsampling by 2, the proposed postprocessing step increases the R-D performance significantly; while for the case of pre-downsampling by 4, the proposed postprocessing does not help in improving the R-D performance.

In general, this trend may be explained as follows. First, in asymmetric video coding, the blockiness, which is usually more severe in the lower quality view, is often the dominant effect that determines the overall quality. Second, in the case of no pre-downsampling or pre-downsampling by a factor of 2, the blockiness in the lower quality view is very obvious. Third, low-pass post-filtering largely reduces blockiness, and thus such a postprocessing step significantly improves visual quality in the case of no pre-downsampling or pre-downsampling by a factor of 2. Fourth, in the case of pre-downsampling by a factor of 4, since the low resolution video frame is significantly smaller (1/16 of original size), the downsampled frame is easier to encode for the target bit rate and thus the blockiness is not as strong. Furthermore, the interpolation step that expands the frame four times in each dimension creates strong blurring effect, which further reduces blockiness. As a result, additional low pass post-filtering following interpolation makes little impact on further removing blockiness or improving visual quality.

V. A MODEL FOR 2D-TO-3D QUALITY PREDICTION

A. 2D Video Quality Prediction

We first examine the capabilities of state-of-the-art 2D-IQA/VQA methods to predict perceptual quality of single-view 2D videos with different pre- and post- processing procedures. The tested full reference 2D-IQA/VQA methods include PSNR, SSIM, MS-SSIM, IW-SSIM, and VQM. Table XIII reports PLCC, SRCC, and RMSE results between 2DVQ-MOS scores and 2D-IQA/VQA measurements, where it can be observed that IW-SSIM and VQM provide the most accurate quality predictions.

B. 2D-to-3D Quality Prediction

The diagram of the proposed method is shown in Fig. 5. Let ($I_{i,r,l}$, $I_{i,r,r}$) and ($I_{i,d,l}$, $I_{i,d,r}$) be the i-th left and right frames of the reference and compressed stereoscopic videos, respectively. We first create their local energy maps by computing the local variances at each spatial location,

TABLE XII

R-D PERFORMANCE COMPARISONS OF ASYMMETRIC STEREOSCOPIC VIDEO CODING ON WATERLOO-IVC 3D VIDEO DATABASE PHASE II (ANCHOR: SYMMETRIC STEREOSCOPIC VIDEO CODING)

			3DVQ-MOS				
Group	Barrier	Craft	Laboratory	Soccer	Tree	Undo	Average
3D.4.a	+8.4%	+10.1%	+11.3%	+3.9%	+7.9%	+16.8%	+9.7%
3D.4.b	-7.2%	-8.3%	-17.1%	-7.7%	-5.4%	-11.1%	-9.5%
3D.4.c	-10.5%	-12.9%	-14.4%	+0.6%	-9.1%	-12.6%	-9.8%
3D.5	-22.0%	-6.8%	-11.5%	-8.6%	-14.6%	-8.4%	-12.0%
3D.6.a	-13.8%	-8.2%	-9.0%	-17.9%	-9.9%	-9.7%	-11.4%
3D.6.b	-32.3%	-17.3%	-16.1%	-16.3%	-31.8%	-19.4%	-22.2%
3D.6.c	-17.8%	-22.1%	-11.3%	-13.7%	-25.4%	-13.9%	-17.4%
3D.7.a	-28.5%	-13.4%	-11.6%	-6.5%	-28.8%	-15.9%	-17.5%
3D.7.b	-27.0%	-14.5%	-17.2%	-14.3%	-25.6%	-14.3%	-18.8%
3D.7.c	-27.2%	-14.4%	-16.6%	-10.2%	-19.0%	-7.2%	-15.7%
			Average 2DVQ-MOS				
Group	Barrier	Craft	Laboratory	Soccer	Tree	Undo	Average
3D.4.a	+12.3%	+13.4%	+11.2%	+23.8%	+19.0%	+22.3%	+17.0%
3D.4.b	+35.6%	+33.7%	+24.3%	+103.3%	+109.9%	+43.6%	+58.4%
3D.4.c	+148.9%	+79.0%	+56.9%	+208.5%	+405.1%	+97.1%	+165.9%
3D.5	-7.0%	-10.9%	-10.3%	+4.5%	-4.4%	-3.7%	-5.3%
3D.6.a	+5.9%	+7.0%	+2.7%	+34.5%	+17.3%	+19.9%	+14.5%
3D.6.b	+42.4%	+39.3%	+16.7%	+110.0%	+110.2%	+55.0%	+62.3%
3D.6.c	+132.0%	+98.7%	+68.7%	+194.3%	+366.2%	+97.7%	+159.6%
3D.7.a	+23.6%	+20.9%	+5.4%	+92.4%	+45.8%	+26.9%	+35.8%
3D.7.b	+64.4%	+62.3%	+38.9%	+124.9%	+149.1%	+60.3%	+83.3%
3D.7.c	+146.0%	+106.6%	+91.3%	+199.8%	+438.4%	+111.4%	+182.3%
			Weighted 2DVQ-MOS				
Group	Barrier	Craft	Laboratory	Soccer	Tree	Undo	Average
3D.4.a	+10.1%	+10.0%	+9.6%	+19.4%	+13.0%	+17.3%	+13.2%
3D.4.b	-15.4%	-9.4%	-10.6%	-7.4%	-13.8%	-8.3%	-10.8%
3D.4.c	-9.6%	-5.2%	-12.6%	-2.8%	-12.6%	-1.9%	-7.4%
3D.5	-13.7%	-12.6%	-9.6%	-10.6%	-13.1%	-9.5%	-11.5%
3D.6.a	-16.1%	-10.1%	-10.0%	-8.2%	-16.3%	-5.4%	-11.0%
3D.6.b	-23.9%	-14.6%	-20.0%	-14.3%	-24.0%	-14.0%	-18.5%
3D.6.c	-21.0%	-10.0%	-18.0%	-12.5%	-21.9%	-11.0%	-15.7%
3D.7.a	-23.3%	-13.3%	-16.1%	-12.2%	-24.2%	-13.7%	-17.2%
3D.7.b	-24.0%	-10.8%	-15.5%	-16.2%	-24.0%	-12.8%	-17.2%
3D.7.c	-21.2%	-9.1%	-15.3%	-14.6%	-21.6%	-8.8%	-15.1%

TABLE XIII

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

Method	PLCC	SRCC	RMSE
PSNR	0.7122	0.6488	16.8383
SSIM	0.8371	0.7660	13.1235
MS-SSIM	0.7787	0.7116	15.0502
IW-SSIM	**0.8821**	**0.8474**	**11.2982**
VQM	**0.8997**	**0.8665**	**10.4689**

i.e., the variances of local image patches extracted around each spatial location, 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 by $E_{i,r,l}$, $E_{i,r,r}$, $E_{i,d,l}$, and $E_{i,d,r}$, respectively. We then compute the local energy ratio maps in both views:

$$R_{i,l} = \frac{E_{i,d,l}}{E_{i,r,l}} \quad \text{and} \quad R_{i,r} = \frac{E_{i,d,r}}{E_{i,r,r}}. \tag{1}$$

The energy ratio maps provide useful local binocular rivalry information, which may be combined with the qualities of single-view frames to predict 3D quality. A pooling stage is necessary for this purpose. High-energy image regions are likely to contain more information content. Based on the principle exploited in [60], if the ultimate goal of visual perception is to efficiently extract useful information from the visual scene, then the more informative regions are more likely to attract visual attention, and thus should be given more importance. The modeling in [60] suggests more informative regions typically have higher energy. To emphasize on the importance of high-energy image regions in binocular rivalry, we adopt an energy weighted pooling method [64] given by

$$g_{i,l} = \frac{\sum E_{i,d,l} R_{i,l}}{\sum E_{i,d,l}} \quad \text{and} \quad g_{i,r} = \frac{\sum E_{i,d,r} R_{i,r}}{\sum E_{i,d,r}}, \tag{2}$$

where the summations are over the full energy and ratio maps. Here $g_{i,l}$ and $g_{i,r}$ are estimations of the level of dominance of the i-th left and right frames, respectively. Let N denotes the frame number of the entire 3D video sequence, we compute

$$g_l = \frac{1}{N} \sum_{i=1}^{N} g_{i,l} \quad \text{and} \quad g_r = \frac{1}{N} \sum_{i=1}^{N} g_{i,r}, \tag{3}$$

where g_l and g_r denote the level of dominance of the left- and right-view video, respectively.

If we consider a video signal as 3D volume data, then it can also be viewed from the side or the top. This has been explored by the poly-view fusion method, which has been shown as a simple and effective strategy to account for the temporal correlation and motion information contained in video signals [65], [66]. Instead of only estimating the level of dominance form the front-view, here we apply a poly-view fusion strategy to estimate the overall level of dominance from the front-view, the top-view, and the side-view together. We first compute the levels of dominance using Eq. (1) to Eq. (3) for the front-view, the top-view, and the side-view, separately, and denote them as g_l^F, g_l^T, and g_l^S for
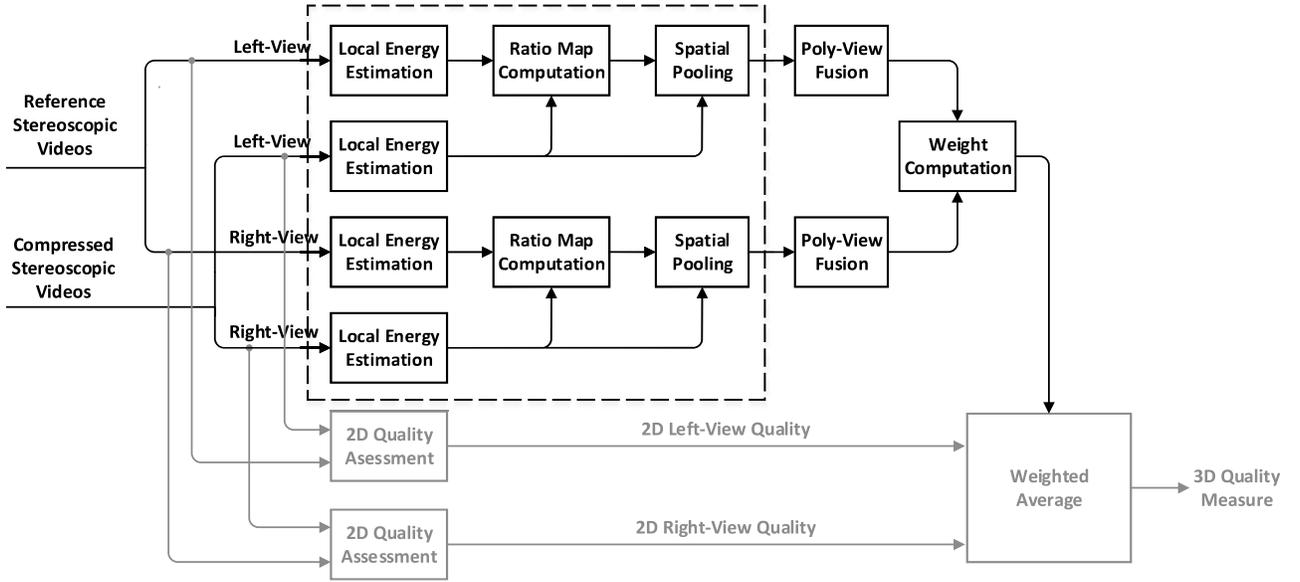
Fig. 5.   Diagram of the proposed 2D-to-3D quality prediction model.

TABLE XIV

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

Waterloo-IVC 3D Video Database Phase II		
Method		SRCC
2DVQ-MOS (average)	− − − − −	0.6277
2DVQ-MOS (proposed weighting)	Front-view Only	0.8643
2DVQ-MOS (proposed weighting)	Front + Top + Side	**0.8727**
VQM (average)	− − − − −	0.6287
VQM (proposed weighting)	Front-view Only	0.7976
VQM (proposed weighting)	Front + Top + Side	**0.8042**
Waterloo-IVC 3D Video Database Phase I		
Method		SRCC
VQM (average)	− − − − −	0.6321
VQM (proposed weighting)	Front-view Only	0.8552
VQM (proposed weighting)	Front + Top + Side	**0.8655**

TABLE XV

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

Waterloo-IVC 3D Video Database Phase I			
Method	PLCC	SRCC	RMSE
PSNR (proposed weighting)	0.8980	0.8366	9.6344
SSIM (proposed weighting)	0.8905	0.8393	9.9615
MS-SSIM (proposed weighting)	0.8838	0.8287	10.2448
IW-SSIM (proposed weighting)	0.8942	0.8364	9.8035
VQM (proposed weighting)	0.9191	0.8655	8.6273
Silva [31]	0.7416	0.6856	14.6893
Lin [20]	0.2986	0.1103	20.8960
Benoit [11]	0.3457	0.2588	20.5455
Yang [13]	0.5886	0.3769	17.6998
You [12]	0.3986	0.2919	20.0805
Chen [3]	0.3828	0.2988	20.2271
Waterloo-IVC 3D Video Database Phase II			
Method	PLCC	SRCC	RMSE
PSNR (proposed weighting)	0.5590	0.5109	10.2154
SSIM (proposed weighting)	0.7571	0.7309	8.0487
MS-SSIM (proposed weighting)	0.6813	0.6377	9.0188
IW-SSIM (proposed weighting)	0.7677	0.7423	7.8943
VQM (proposed weighting)	0.8496	0.8042	6.4976
Silva [31]	0.5566	0.5184	10.1051
Lin [20]	0.3675	0.2115	11.3123
Benoit [11]	0.4074	0.3163	11.1085
Yang [13]	0.3828	0.3252	11.2372
You [12]	0.4074	0.3164	11.1084
Chen [3]	0.3682	0.2180	11.3092

the left-view video, and g_r^F, g_r^T, and g_r^S for the right-view video, respectively. Then the overall level of dominance of the left- and right-view video after poly-view fusion is computed as

$$g_l^O = g_l^F + g_l^T + g_l^S \quad \text{and} \quad g_r^O = g_r^F + g_r^T + g_r^S, \quad (4)$$

respectively.

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

$$w_l = \frac{g_l^{O2}}{g_l^{O2} + g_r^{O2}} \quad \text{and} \quad w_r = \frac{g_r^{O2}}{g_l^{O2} + g_r^{O2}}, \quad (5)$$

respectively.

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

$$Q^{3D} = w_l Q_l^{2D} + w_r Q_r^{2D}, \quad (6)$$

where Q_l^{2D} and Q_r^{2D} denote the 2D video quality of the left- and right-view videos, respectively.

C. Validation

The proposed 2D-to-3D video quality prediction model with and without including the top- and side-views is firstly tested on all stereoscopic 3D videos in Waterloo-IVC 3D video databases. SRCC values between 3DVQ-MOS and the predicted Q^{3D} value are reported in Table XIV. Note that in our Phase II database, both MOS scores of stereoscopic 3D videos and single-view 2D videos are available. This allows us to apply the 2D-to-3D prediction models directly on 2DVQ-MOS scores (without worrying about the accuracy of the base 2D-IQA/VQA scores when they are produced by objective models). Our Phase I database, however, does not include a 2D subjective experiment on single-view videos, and thus

we choose a high performance 2D objective model, VQM, in our current test. Meanwhile, we have used VQM to replace 2DVQ-MOS in our test with Phase II database. The results in Table XIV suggest that the improvement by including top- and side-views is consistent in all test cases. In practice, depending on the affordable computational cost, we may seek the best compromise between accuracy and speed, and choose whether to include the front- and side-views.

Then the proposed 2D-to-3D prediction model (Front + Top + Side) is tested on all 3D videos and each test video group in Waterloo-IVC 3D video databases by applying it to the ground truth 2DVQ-MOS scores (Phase II only) and different base 2D-IQA/VQA approaches (Phase I and Phase II). The PLCC, SRCC, and RMSE values between 3DVQ-MOS and the predicted Q^{3D} value are given in Table VI (all videos), Table VII (different groups in Phase I) and Table VIII (different groups in Phase II). The corresponding scatter plots are shown in Fig. 3 (b) and the second and the fourth columns of Fig. 4. It can be observed that the proposed 2D-to-3D model outperforms the direct averaging method significantly with respect to 2DVQ-MOS scores and all tested 2D-IQA/VQA approaches. For different levels of compressions, pre- and post- processing, the proposed method, which does not attempt to recognize the distortion types or give any specific treatment, removes or significantly reduces the 2D-to-3D quality prediction biases.

We have also compared the proposed method with state-of-the-art 3D-IQA/VQA methods using both databases. PLCC, SRCC, and RMSE values between 3DVQ-MOS and the predicted Q^{3D} value are reported in Table XV. From Tables VI and XV, it can be observed that most of these methods produce similar performance to the cases of directly averaging PSNR, SSIM, MS-SSIM and IW-SSIM of the left- and right-views. The only exception is the StSD algorithm [31], which performs better than directly averaging 2D-IQA methods, but worse than applying the proposed weighting scheme on 2D-IQA/VQA methods. There may be two explanations. First, the StSD method was developed based on the StSD 3D Video Database, which contains weak asymmetrically compressed stereoscopic 3D videos; Second, the StSD method is a 3D-VQA algorithm, while all the other methods being tested are purely 3D-IQA algorithms.

Furthermore, the R-D performance of different variations of asymmetric stereoscopic video coding in terms of the average and weighted 2DVQ-MOS for each test sequence in Phase II are reported in Table XII. Again, we use the BD-Rate as the test criterion, which provides a useful quantitative measure to evaluate the R-D performance. From Table XII, it can be seen that, compared with the bit rate savings measured by 3DVQ-MOS, the direct averaging 2DVQ-MOS generates substantial bias for all sequences and all mixed-distortion combinations. On the other hand, the proposed weighting 2DVQ-MOS significantly reduces the biases and indicates highly consistent bit rate savings with 3DVQ-MOS. This demonstrates great potentials of the proposed method to be employed in perceptually inspired R-D optimization of stereoscopic video coding systems.

VI. Conclusion

The major contributions of the current paper are as follows: first, we carried out subjective quality assessment experiments on two databases (Waterloo-IVC 3D video database Phase I and Phase II) that contain various asymmetrically compressed stereoscopic 3D videos obtained from mixed-resolution coding, asymmetric transform-domain quantization coding, their combinations, and multiple choices of post-processing techniques. Second, we compared different mixed-distortion asymmetric stereoscopic video coding schemes with symmetric coding methods and verified their potential coding gains. Third, we observed a strong systematic bias when using direct averaging of 2D video quality of both views to predict 3D video quality. Fourth, we proposed a model to account for the prediction bias, leading to significantly improved full reference quality predictions of stereoscopic videos. Fifth, we showed that the proposed model can help us predict the coding gain of mixed-distortion asymmetric video compression. In the future, we aim to develop novel high efficiency asymmetric 3D video coding schemes, incorporating the key observations and the proposed 3D-VQA model in this work, together with advanced perceptual models of visual discomfort [48]–[51] and depth perception [67].

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*. Cham, Switzerland: Springer, 2015, pp. 185–212.

[2] 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, Dec. 2013.

[3] 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.

[4] 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.

[5] 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.

[6] 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.

[7] 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.

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

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

[10] 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.

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

[12] 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.

[13] J. Yang, J. 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, Gernamy, May 2009, pp. 1–4.

[14] 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.

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

[16] 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.

[17] 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.

[18] 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.

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

[20] 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.

[21] 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.

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

[23] G. Saygılı, C. G. Gürler, and A. M. Tekalp, "Quality assessment of asymmetric stereo video coding," in *Proc. IEEE Int. Conf. Image Process.*, Hong Kong, Sep. 2010, pp. 4009–4012.

[24] V. D. Silva *et al.*, "Psycho-physical limits of interocular blur suppression and its application to asymmetric stereoscopic video delivery," in *Proc. Int. Packet Video Workshop*, Munich, Germany, May 2012, pp. 184–189.

[25] M. Azimi, S. Valizadeh, X. Li, L. E. Coria, and P. Nasiopoulos, "Subjective study on asymmetric stereoscopic video with low-pass filtered slices," in *Proc. Int. Comput., Netw. Commun.*, Maui, HI, USA, Jan./Feb. 2012, pp. 719–723.

[26] M. G. Perkins, "Data compression of stereopairs," *IEEE Trans. Commun.*, vol. 40, no. 4, pp. 684–696, Apr. 1992.

[27] 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.

[28] 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.

[29] 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, Mar. 2015.

[30] 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.

[31] V. De Silva, H. K. Arachchi, E. Ekmekcioglu, and A. Kondoz, "Toward an impairment metric for stereoscopic video: A full-reference video quality metric to assess compressed stereoscopic video," *IEEE Trans. Image Process.*, vol. 22, no. 9, pp. 3392–3404, Sep. 2013.

[32] 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.

[33] L. Goldmann, F. De Simone, and T. Ebrahimi, "A comprehensive database and subjective evaluation methodology for quality of experience in stereoscopic video," *Proc. SPIE*, vol. 7526, p. 75260S, Jan. 2010.

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

[35] A. Banitalebi-Dehkordi, M. T. Pourazad, and P. Nasiopoulos, "Effect of high frame rates on 3D video quality of experience," in *Proc. IEEE Int. Conf. Consumer Electron.*, Las Vegas, NV, USA, Jan. 2014, pp. 416–417.

[36] A. Banitalebi-Dehkordi, M. T. Pourazad, and P. Nasiopoulos, "The effect of frame rate on 3D video quality and bitrate," *3D Res.*, vol. 6, no. 1, pp. 1–13, Mar. 2015.

[37] E. Dumić, S. Grgić, K. Šakic, P. M. R. Rocha, and L. A. da Silva Cruz, "3D video subjective quality: A new database and grade comparison study," *Multimedia Tools Appl.*, vol. 76, no. 2, pp. 2087–2109, Jan. 2017.

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

[39] 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.

[40] 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.

[41] 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.

[42] 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.

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

[44] L. Jin, A. Gotchev, A. Boev, and K. Egiazarian, "Validation of a new full reference metric for quality assessment of mobile 3DTV content," in *Proc. Eur. Signal Process. Conf.*, Barcelona, Spain, Aug./Sep. 2011, pp. 1894–1898.

[45] S. S. Sarikan, R. F. Olgun, and G. B. Akar, "Quality evaluation of stereoscopic videos using depth map segmentation," in *Proc. Int. Workshop Quality Multimedia Exper.*, Mechelen, Belgium, Sep. 2011, pp. 67–71.

[46] J. Seo, X. Liu, D. Kim, and K. Sohn, "An objective video quality metric for compressed stereoscopic video," *Circuits, Syst., Signal Process.*, vol. 31, no. 3, pp. 1089–1107, Jun. 2012.

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

[48] J. Park, S. Lee, and A. C. Bovik, "3D visual discomfort prediction: Vergence, foveation, and the physiological optics of accommodation," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 3, pp. 415–427, Jun. 2014.

[49] J. Park, H. Oh, S. Lee, and A. C. Bovik, "3D visual discomfort predictor: Analysis of disparity and neural activity statistics," *IEEE Trans. Image Process.*, vol. 24, no. 3, pp. 1101–1114, Mar. 2015.

[50] H. Oh, S. Lee, and A. C. Bovik, "Stereoscopic 3D visual discomfort prediction: A dynamic accommodation and vergence interaction model," *IEEE Trans. Image Process.*, vol. 25, no. 2, pp. 615–629, Feb. 2016.

[51] Q. Jiang, F. Shao, G. Jiang, M. Yu, and Z. Peng, "Three-dimensional visual comfort assessment via preference learning," *J. Electron. Imag.*, vol. 24, no. 4, p. 043002, Jun. 2015.

[52] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1649–1668, Dec. 2012.

[53] 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.

[54] O. Rosenbach, "On monocular prevalence in binocular vision," *Med. Wochenschrift*, vol. 50, pp. 1290–1292, 1903.

[55] J. Wang, S. Wang, and Z. Wang, "Depth perception of distorted stereoscopic images," in *Proc. IEEE Int. Workshop Multi. Signal Process.*, Xiamen, China, Oct. 2015, pp. 1–6.

[56] R. Blake, "A primer on binocular rivalry, including current controversies," *Brain Mind*, vol. 2, no. 1, pp. 5–38, Apr. 2001.

[57] 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.

[58] A. M. van Dijk, J.-B. Martens, and A. B. Watson, "Quality asessment of coded images using numerical category scaling," *Proc. SPIE*, vol. 2451, pp. 90–101, Mar. 1995.

[59] M. H. Pinson and S. Wolf, "Comparing subjective video quality testing methodologies," *Proc. SPIE*, vol. 5150, pp. 573–582, Jul. 2003.

[60] 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.

[61] VQEG. (Apr. 2000). *Final Report From the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment.* [Online]. Available: http://www.vqeg.org

[62] 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.

[63] G. Bjontegaard, *Improvements of the BD-PSNR Model*, document ITU-T SG16 Q.6, 2008, p. 35.

[64] 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.

[65] K. Zeng and Z. Wang, "Polyview fusion: A strategy to enhance video-denoising algorithms," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 2324–2328, Apr. 2012.

[66] K. Zeng and Z. Wang, "3D-SSIM for video quality assessment," in *Proc. IEEE Int. Conf. Image Process.*, Orlando, FL, USA, Sep./Oct. 2012, pp. 621–624.

[67] J. Wang, "Perceptual quality-of-experience of stereoscopic 3D images and videos," Ph.D. dissertation, Faculty Eng., Univ. Waterloo, Waterloo, ON, Canada, 2016.

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 was a Post-Doctoral 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.

Jiheng Wang (S'11–M'17) received the M.Math. degree in statistics-computing and the Ph.D. degree in electrical and computer engineering from the University of Waterloo, Waterloo, ON, Canada, in 2011 and 2016, respectively. In 2013, he was with the Video Compression Research Group, Blackberry, Waterloo. He is currently a Post-Doctoral Fellow with the Department of Electrical and Computer Engineering, University of Waterloo. His current research interests include 3-D image and video quality assessment, perceptual 2-D and 3D video coding, biomedical signal processing, statistical learning, and dimensionality reduction.

Zhou Wang (S'99–M'02–SM'12–F'14) received the Ph.D. degree 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 over 100 publications in these fields with over 30,000 citations.

Dr. Wang served as a member of the IEEE Multimedia Signal Processing Technical Committee from 2013 to 2015. He is a recipient of the 2015 Primetime Engineering Emmy Award, 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. He has been a Senior Area Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING since 2015, and an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY since 2016. He served as an Associate Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING from 2009 to 2014, the *Pattern Recognition* from 2006 and the IEEE SIGNAL PROCESSING LETTERS from 2006 to 2010. He served as a Guest Editor of the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING from 2013 to 2014 and from 2007 to 2009. He is a fellow of the Canadian Academy of Engineering.