2014 *International Workshop on Video Processing and Quality Metrics for Consumer Electronics* (VPQM2014)
Scottsdale, AZ, USA, Jan.-Feb. 2014.

PERCEPTUAL QUALITY OF ASYMMETRICALLY DISTORTED STEREOSCOPIC IMAGES: THE ROLE OF IMAGE DISTORTION TYPES

Jiheng Wang and Zhou Wang

Dept. of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, N2L 3G1, Canada
Emails: {jiheng.wang, zhou.wang}@uwaterloo.ca

ABSTRACT

Objective quality assessment of asymmetrically distorted stereoscopic 3D images is a challenging problem, for which direct simple combinations of the quality of the left and right views fail to produce adequate predictions. We carried 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, where the distortion types include noise contamination, blur and JPEG compression. Our results suggest that simply averaging the quality of the left and right view images well predicts the quality of symmetrically distorted 3D images, but generates substantial bias when applied to asymmetrically distorted stereoscopic images. More interestingly, we find that such bias could lean towards opposite directions, largely depending on the distortion types. We propose a computational model that accounts for the bias, leading to significantly improved quality prediction of stereoscopic images.

Index Terms— image quality assessment, stereoscopic image, 3D image, asymmetric distortion, SSIM

1. INTRODUCTION

With the fast development of 3D acquisition, communication, processing and display technologies, automatic quality assessment of 3D images and videos has become ever important. Nevertheless, recent process on 3D image quality assessment (IQA) remains limited. In the literature, a majority of studies have been focused on evaluating the extension of existing 2D-IQA methods to stereoscopic images. These methods can be grouped into two categories based on whether depth or disparity information is explicitly employed. In [1, 2], 2D-IQA measures are applied to the left- and right-view images separately and then combined to a 3D quality score. In [3, 4, 5], disparity maps between left- and right-views are estimated, followed by 2D quality assessment of disparity quality, which is subsequently combined with 2D image quality to produce an overall 3D image quality score.

However, 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 quality predictions of stereoscopic images. In particular, in [6], it was shown that averaging peak-signal-to-noise ratio (PSNR), structural similarity (SSIM) [7] and multi-scale SSIM (MS-SSIM) [8] measurements of left- and right-views performs equally well or better than the advanced 3D-IQA models [2, 3, 4]. Similar results were also observed in [9], where averaging universal quality index (UQI) [10], MS-SSIM [8] and visual information fidelity (VIF) [11] of both views all outperformed 3D-IQA models [2, 3, 4, 5].

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. Previous studies exhibit somewhat conflict observations and opinions regarding the effect of asymmetric distortions. For image blur, evidence in [12] shows that asymmetric blur distorted images is largely dominated by the higher quality view, a result generally agrees with [13]. For image blockiness, it was reported in [14] that 3D image quality has a tendency towards the lower quality view, while in [12], it was claimed that it should be approximated by averaging the quality of both views. In [13], it was suggested that the best strategy of asymmetric quality assessment for compressed images should be content and texture dependent.

Subjective data is essential in understanding the impact of asymmetric distortion on the perceptual quality of stereoscopic images. Ideally, we would need a complete set of subjective 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. Existing 3D image quality databases are highly valuable but limited in one aspect or another. Specifically, IRCCyN/IVC 3D Images Database [4], Tianjin University Database [5], Ningbo University Database Phase II [15], and LIVE 3D Image Quality Database Phase I [9] only include symmetrically distorted stereoscopic images. Ningbo University Database Phase I [13] only includes asymmetrically distorted stereoscopic images. MICT 3D Image Quality Evaluation Database [16] contains both cases but only for JPEG compressed images. The most recent LIVE 3D Image Quality Database Phase II [6] includes both symmetric and asymmetric cases as well as five distortion types. Unfortunately, 2D-IQA of single view images are still missing, making it difficult to directly examine the relationship between the perceptual quality of single-view 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.

In this study, we carried 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, where the distortion types include noise contamination, blur and JPEG compression. This database allows us to directly observe the quality prediction performance from single-view to stereoscopic images, for which we observe that simply averaging the quality of both views creates substantial bias on asymmetrically distorted stereoscopic images and the bias could lean toward opposite directions, largely depending on the distortion types. Furthermore, we propose a 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.

2. SUBJECTIVE STUDY

Before the subjective experiment, we built an image database created from 6 pristine stereoscopic image pairs (and thus their corresponding single-view images), all collected from the Middlebury Stereo 2005 Datasets [17]. 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 1. The single-view images are employed to generate distorted stereopairs, either symmetrically or asymmetrically. Altogether, there are totally 78 single-view images and 330 stereoscopic images in the database. Table 2 categorizes these images into seven groups with detailed descriptions.

Table 1. Value ranges of control parameters for distortion simulation

Distortion	Control Parameter	Range
Noise	Variance of Gaussian	[0.10 0.40]
Blur	Variance of Gaussian	[2 20]
JPEG	Quality parameter	[3 10]

Table 2. Categories of test images

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

To the best of our knowledge, there are two unique features of the current database when compared with existing publicly available 3D-IQA databases. First, this is the only database that performs subjective test on both 2D and 3D images. The including 2D images allows us to directly examine the relationship between the perceptual quality of stereoscopic images and that of the 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 an objective 2D-IQA measure as an estimate. Second, this is the only database that contains mixed distortion types in asymmetrically distorted images. This provides the potential of a much stronger test on 3D-IQA models on their generalizability, especially those developed to account for asymmetric distortions of specific distortion types.

(a) Art (b) Books (c) Dolls (d) Moebius (e) Laundry (f) Reindeer

Fig. 1. The 6 Pristine Images used in the subjective study. Only the right-views are shown here.

The subjective test was conducted in the Lab for Image and Vision Computing at University of Waterloo. The single stimulus continuous quality scale (SSCQS) protocol was adopted to obtain subjective quality ratings for all of the single-view images and stereopairs. The test environment had 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$^{\text{TM}}$2 active shutter glasses is used for the test. The default viewing distance was 3.5 times the screen height. The details of viewing conditions are given in Table 3.

Table 3. Viewing conditions of the subjective test

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

Twenty-four naive subjects (14 males, 10 females) – all university graduate students – took part in the study. They are aged from

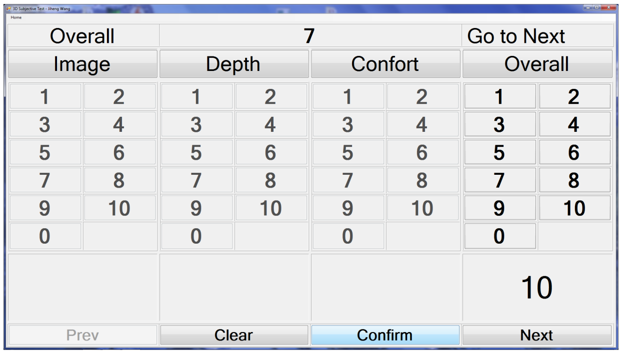


Fig. 2. Customized GUI application for score recording

22 to 45 years old. A 3D vision test was conducted first to verify their ability to view stereoscopic 3D content and three of them (1 male, 2 females) were rejected and did not continue with the test. As a result, a total of twenty-one naive subjects (13 males, 8 females) proceeded to the formal test. Following previous works [18, 19, 20], the subjects were asked to evaluate four aspects of their 3D viewing experience, including the perception of 3D image quality (IQ), depth quality (DQ), visual comfort (VC) and overall 3D quality of experience (3DQoE). The detailed descriptions of each aspects of visual experience are elaborated in Table 4.

Table 4. Description of visual experience criteria

Criterion	Description
IQ	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

Our pilot test showed that one-pass experiment (where a subject gives 3DIQ, DQ, VC, and 3DQoE scores to each stereoscopic image in one trial) causes severe visual fatigue of the human subjects within a short period of time. To avoid this problem, we resorted to a multi-pass approach [20] in the formal test, where within each pass, the subject only needs to give one of the four scores. The test was scheduled on two consecutive days for each subject. Day 1 was dedicated to 2DIQ, VC and 3DIQ tests, and Day 2 to DQ and 3DQoE tests. A 3D vision test and a general introduction were given at the beginning of the whole test, and more specific instructions and training session were given before each sub-test. Each session is controlled to be within 20 minutes and sufficient relaxation periods 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 ask the subject to speak out a score between 0 and 10, and a customized graphical user interface shown in Fig. 2 on another computer screen is 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 interactive effects between different visual experiences criteria.

The raw 3DIQ scores given by each subject were converted to Z-scores and the entire data set was rescaled to fill the range from 1 to 100. The mean opinion score (MOS) for each image was then computed. The rest of the paper focuses on the relationship between single-view image quality and the 3DIQ scores. More detailed descriptions of our database and analysis of the other aspects of the subjective experiments will be reported in future publications.

3. 2D-TO-3D IMAGE QUALITY PREDICTION

3.1. Distortion Type Dependency

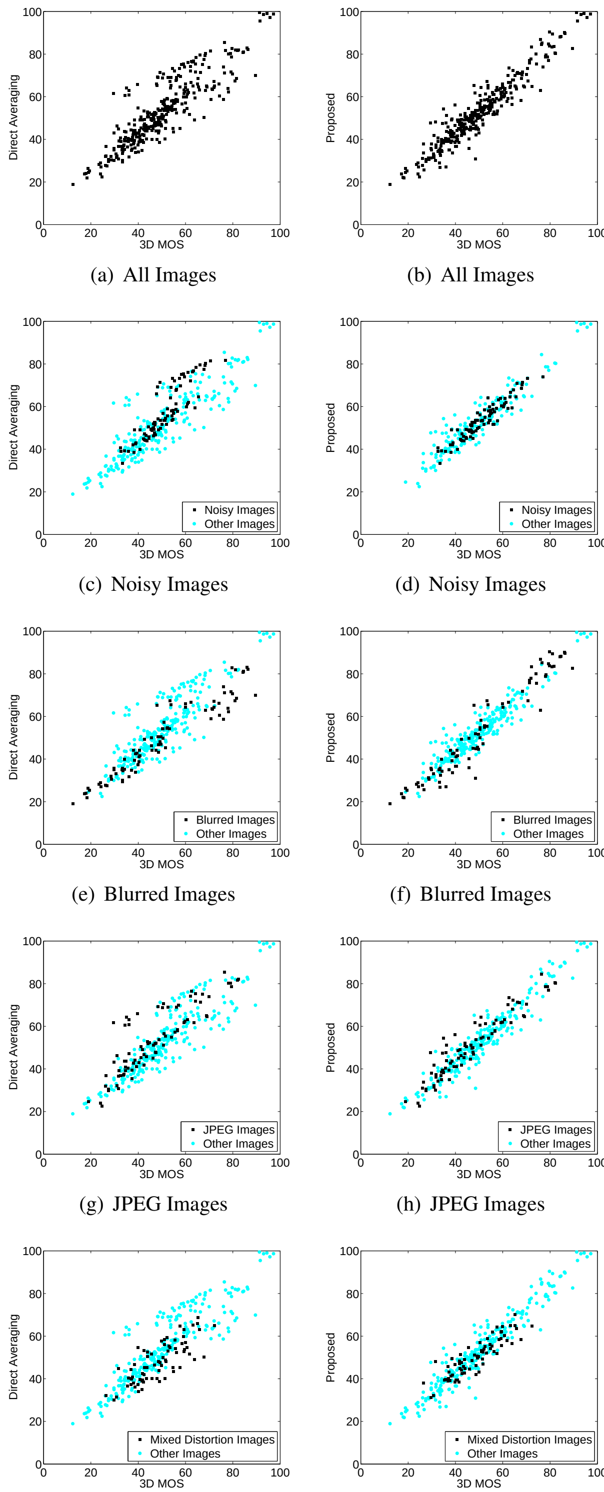
The main question we would like to ask in the current paper is how the single-view 2D image quality predicts the 3D image quality (3DIQ scores in the subjective test), especially for the case of asymmetric distortions. The most straightforward 2D-to-3D quality prediction method is to average the MOSs of the left- and right-view images. Table 5 reports Pearson's linear correlation coefficient (PLCC), Spearman's rank-order correlation coefficient (SRCC) and Kendall's rank-order correlation coefficient (KRCC) between 3D-MOS scores and the average 2D-MOS scores, including the results for all stereoscopic images, for each test image group, and for each distortion type. The left column of Fig. 3 shows the corresponding scatter plots.

From Table 5 and 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 [9, 6]). By contrast, the PLCC, SRCC and KRCC values drop significantly in other test groups (corresponding to asymmetrical distortions) as well as for different distortion types and 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 5 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 [12, 13] that the 3D quality is less affected by the view with lower quality, which is consistent with our result. For image blockiness, [14] and [12] reported somewhat conflicting results. The former concluded that the 3D image quality is mainly dependent on the view with lower quality, and the latter suggested that quality averaging of both views is a better choice. These seemingly controversial results are well explained by the scatter plot shown in Fig. 3(a), where 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.

3.2. 2D-to-3D Quality Prediction Model

The competition between binocular fusion and binocular rivalry [21, 22] provides a potential theory for 2D-to-3D quality prediction. When the left- and right-view images are sufficiently similar, 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 sufficiently different, instead of the two images being seen superimposed, one of them may dominate or two images may be seen alternatively, known as binocular rivalry [21, 22]. Although



(a) All Images (b) All Images (c) Noisy Images (d) Noisy Images (e) Blurred Images (f) Blurred Images (g) JPEG Images (h) JPEG Images (i) Mixed Distortion Images (j) Mixed Distortion Images

Fig. 3. Scatter plots of 3D image quality MOS score versus predictions from MOS scores of 2D left- and right-views. Left column, prediction by averaging the MOS scores of both views; right column, prediction by the proposed method.

there is a rich literature on binocular fusion and rivalry in biological vision science [21, 22] (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 [23, 24], 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 level of view dominance in binocular rivalry of stereoscopic images is monotonically increasing with the relative energy of the two views.

The diagram of the proposed method is shown in Fig. 4. 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 a sliding Gaussian window with standard deviation of 1.5 is employed. The resulting energy maps are denoted as $E_{r,l}$, $E_{r,r}$, $E_{d,l}$ and $E_{d,r}$, respectively. 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:

$$R_l = \frac{E_{d,l}}{E_{r,l}} \qquad \text{and} \qquad R_r = \frac{E_{d,r}}{E_{r,r}} \,. \tag{1}$$

The energy ratio maps provide useful local 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. To emphasize on the importance of high-energy image regions, we adopt an energy weighted pooling method given by

$$g_l = \frac{\sum E_{d,l} R_l}{\sum E_{d,l}} \qquad \text{and} \qquad g_r = \frac{\sum E_{d,r} R_r}{\sum E_{d,r}} \,, \tag{2}$$

where the summations are over the full energy and ratio maps. The weights assigned to the left- and right-view images are then given by

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

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

$$Q_{3D} = w_l Q_l + w_r Q_r \,, \tag{4}$$

where Q_l and Q_r denote the 2D image quality of the left- and right-views, respectively. In the current paper, Q_l and Q_r are the MOS scores obtained from our subjective experiments on the 2D images in Group 2D.0 and Group 2D.1 image sets.

3.3. Validation

The proposed 2D-to-3D quality prediction model is tested on all 3D images in our database. The PLCC, SRCC and KRCC values between 3D-MOS and the predicted Q_{3D} value for each category and

Table 5. Performance of 2D-to-3D quality prediction models: 2D-MOS with direct average and 2D-MOS with proposed weighting

	PLCC		SRCC		KRCC	
Method	direct average	proposed weighting	direct average	proposed weighting	direct average	proposed weighting
All 3D	0.8835	0.9590	0.8765	0.9484	0.7161	0.8162
Group 3D.1	0.9801	0.9801	0.9657	0.9657	0.8482	0.8482
Group 3D.2	0.6247	0.9463	0.5433	0.9374	0.4406	0.7915
Group 3D.3	0.9661	0.9775	0.9164	0.9471	0.7597	0.8034
Group 3D.4	0.9222	0.9660	0.8271	0.9413	0.6390	0.7962
Noise	0.9305	0.9761	0.9370	0.9474	0.8052	0.8137
Blur	0.9564	0.9736	0.9707	0.9714	0.8562	0.8632
JPEG	0.9188	0.9675	0.8865	0.9494	0.7508	0.8247
Mixed	0.9222	0.9660	0.8271	0.9413	0.6390	0.7962

Table 6. Performance of 2D-to-3D quality prediction models

	PLCC			SRCC			KRCC		
	All	Symmetric	Asymmetric	All	Symmetric	Asymmetric	All	Symmetric	Asymmetric
2D-MOS with direct average	0.8835	0.9801	0.8572	0.8765	0.9657	0.8471	0.7161	0.8482	0.6780
2D-MOS with proposed weighting	0.9590	0.9801	0.9544	0.9484	0.9657	0.9405	0.8162	0.8482	0.8038
You [3]	0.5746	0.6416	0.5549	0.5968	0.7517	0.5705	0.4351	0.5615	0.4174
Benoit [4]	0.6276	0.7776	0.6767	0.5435	0.6588	0.5751	0.3946	0.4921	0.4223
Yang [5]	0.6984	0.8233	0.7065	0.6106	0.6668	0.6108	0.4428	0.4821	0.4436

each distortion type are given in Table 5. The corresponding scatter plots are shown in the right column of Fig. 3.

From Table 5 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). Most importantly, 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 improvement.

To the best of our knowledge, no previous method uses 2D single-view MOS values to predict 3D-MOS value, and thus straightforward comparison with existing methods is not possible. The results of some state-of-the-art 3D-IQA approaches [3, 4, 5] are shown in Table 6. Here the main observation is that there is a large performance drop in all these existing objective methods from cases of symmetric to asymmetric distortions. Clearly, the drop is much smaller in the proposed method, which leads to the most significant performance gain in the asymmetric distortion case.

4. CONCLUSION

The major contributions of the current paper are as follows: First, we created 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 direct averaging 2D image quality of both views to predict 3D image quality. Third, we propose a computational model for 2D-to-3D quality prediction, which does not perform any distortion type detection or provides any distortion type specific treatment, significantly reduces the quality prediction bias. The new 3D-IQA database provides many opportunities of our future work, which includes localized 2D-to-3D quality prediction, and objective quality assessment of a more complete set of 3D visual experience.

5. REFERENCES

[1] 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," *Electronics letters*, vol. 44, no. 16, pp. 963–965, 2008.

[2] P. Gorley and N. Holliman, "Stereoscopic image quality metrics and compression," in *Proc. SPIE 6803, Stereoscopic Displays and Applications XIX, San Jose, CA*, Jan. 2008.

[3] 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 *Workshop on Video Processing and Quality Metrics, Scottsdale, AZ*, 2010.

[4] A. Benoit, P. Le Callet, P. Campisi, and R. Cousseau, "Quality assessment of stereoscopic images," *EURASIP journal on image and video processing*, 2008.

[5] J. Yang, C. Hou, Y. Zhou, Z. Zhang, and J. Guo, "Objective quality assessment method of stereo images," in *3DTV Conference: The True Vision-Capture, Transmission and Display of 3D Video*, 2009, pp. 1–4.

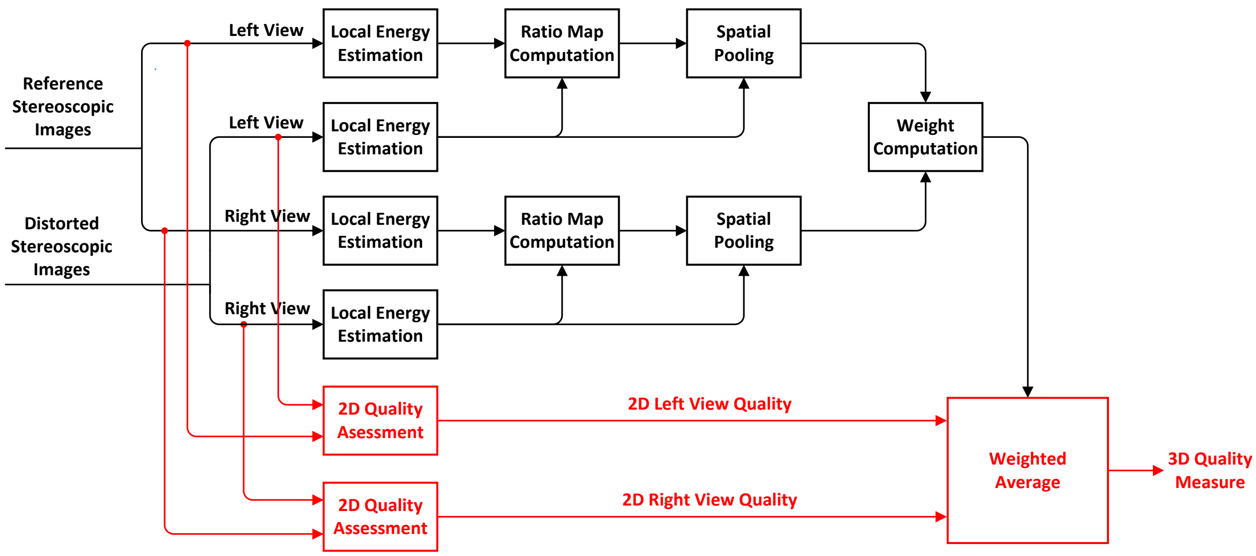


Fig. 4. Diagram of the proposed 2D-to-3D quality prediction model.

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

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

[8] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multi-scale structural similarity for image quality assessment," in *Proc. IEEE Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2003, pp. 1398–1402.

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

[10] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Processing Letters*.

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

[12] 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.," *Journal of Experimental Psychology: Applied*, vol. 7, no. 2, pp. 143, 2001.

[13] X. Wang, M. Yu, Y. Yang, and G. Jiang, "Research on subjective stereoscopic image quality assessment," in *Proc. SPIE 7255, Multimedia Content Access: Algorithms and Systems III, San Jose, CA*, Jan. 2009.

[14] 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. Applied Perception*, vol. 3, no. 2, pp. 95–109, 2006.

[15] J. Zhou, G. Jiang, X. Mao, M. Yu, F. Shao, Z. Peng, and Y. Zhang, "Subjective quality analyses of stereoscopic images in 3DTV system," in *Proc. IEEE Int. Conf. on Visual Communications and Image Process.*, Nov. 2011, pp. 1–4.

[16] Z. M. P. Sazzad, S. Yamanaka, Y. Kawayokeita, and Y. Horita, "Stereoscopic image quality prediction," in *Proc. Workshop on Quality of Multimedia Experience*, July 2009, pp. 180–185.

[17] D. Scharstein and C. Pal, "Learning conditional random fields for stereo," in *Proc. IEEE Int. Conf. on Computer Vision and Pattern Recognition*, June 2007, pp. 1–8.

[18] P. Seuntiens, "Visual experience of 3D TV," *Doctor doctoral Thesis, Eindhoven University of Technology*, 2006.

[19] M. Chen, D. Kwon, and A.C. Bovik, "Study of subject agreement on stereoscopic video quality," in *IEEE Southwest Symp. Image Analysis & Interpretation*, 2012, pp. 173–176.

[20] W. Chen, J. Fournier, 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, Jan. 2012.

[21] L. Kaufman, *Sight and Mind: An Introduction to Visual Perception*, Oxford U. Press, 1974.

[22] B. Julesz, *Foundations of cyclopean perception*, U. Chicago Press, 1971.

[23] W.J.M. Levelt, "The alternation process in binocular rivalry," *British Journal of Psychology*, vol. 57, no. 3-4, pp. 225–238, 1966.

[24] R. Blake, "Threshold conditions for binocular rivalry.," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 3, no. 2, pp. 251, 1977.