FULL-REFERENCE STEREOSCOPIC IMAGE QUALITY ASSESSMENT ACCOUNTING FOR BINOCULAR COMBINATION AND DISPARITY INFORMATION

Yu Fan [1,2], *Mohamed-Chaker Larabi* [1], *Faouzi Alaya Cheikh* [2], *Christine Fernandez-Maloigne* [1]

[1] XLIM, University of Poitiers, France
[2] Faculty of Computer Science and Media Technology, NTNU, Gjovik, Norway

ABSTRACT

One of the most challenging issues in stereoscopic image quality assessment (SIQA) is how to effectively model the binocular behavior of the human visual system (HVS). The latter has a great impact on the perceptual 3D quality. In this paper, we propose a SIQA metric accounting for binocular combination properties and disparity information. Instead of computing the quality of the left and the right views separately, the proposed metric predicts the quality of a cyclopean image so as to have a good consistency with 3D human perception. The cyclopean image is synthesized based on the local entropy and the visual saliency of each view with the aim to simulate the phenomena of binocular fusion/rivalry. A 2D IQA metric is employed to assess the quality of both the cyclopean image and the disparity map. The obtained scores are used to derive the 3D quality score thanks to a pooling stage. Experimental results on three public 3D IQA databases show that the proposed method outperforms many other state-of-the-art SIQA methods, and achieves high prediction accuracy on these databases.

Index Terms— Stereoscopic image quality assessment, cyclopean image, binocular fusion/rivalry, visual saliency.

1. INTRODUCTION

In recent years, three-dimensional (3D) multimedia has become popular thanks to new sensations of immersion. With the rapid development of stereoscopic 3D (S3D) technologies, sources of 3D content and 3D display are more common nowadays. As a result, the perceptual quality assessment of 3D images and videos is quite important in order to guarantee the visual quality of experiences (QoE) at every processing stage ranging from 3D acquisition, compression, transmission and display. While 2D image quality assessment (IQA) has greatly advanced over the last decade, SIQA is still in its early stage and hence challenging [1], especially for asymmetrically distorted S3D images. This is mainly because 3D perceptual quality is affected by both monocular and binocular factors including 2D quality, disparity/depth quality and visual comfort. Although 3D quality can be measured using subjective experiments, these are costly, time-consuming, and thus impractical for real-time applications. Consequently, objective SIQA metrics are needed to automatically predict the perceptual quality of S3D images.

According to the availability of the reference stereo pair, SIQA metrics can be generally classified into three groups: full-reference (FR) [2, 3], reduced-reference (RR) [4, 5] and no-reference (NR) [6, 7] methods. While FR-SIQA metrics use the whole reference S3D images to measure the 3D quality, RR-SIQA metrics make use of a set of features extracted form the reference images. NR-SIQA metrics measure the image quality without using any specific information of the reference images. Our SIQA metric presented in this paper belongs to the FR group.

Meanwhile, FR-SIQA methods can also be categorized into three classes [2] based on the type and amount of information extracted from stereo pairs. The SIQA methods of the first class [8, 9] employ 2D IQA metrics to measure the quality of left and right views separately, and then combine both scores into an overall 3D quality score. This class of methods does not correlate well with human quality judgments, since 2D metrics do not take into account binocular depth cues playing a critical role in 3D perception. Methods of the second class [10, 11] assess the 3D quality using depth/disparity information in addition to both views of a stereo pair. It is worth noting that the performance of the methods in this class depends on the accuracy of the depth/disparity maps estimated by stereo matching algorithms.

In fact, the left and right views of a stereo pair may suffer from the same distortion type and level (namely symmetric distortion) or different distortion levels and/or types (namely asymmetric distortion). Symmetric distortions lead to binocular fusion (BF), whereas asymmetric distortions result in either binocular rivalry (BR) [12] or binocular suppression (BS) [13] according to the difference of distortion strength. These latter have a great impact on the perceptual 3D quality. The FR-SIQA methods of the two above-described classes can perform quite well in the case of symmetric distortion, but are much less effective for asymmetrically distorted S3D images that are very common in real application such as 3D coding. Thus, to improve the SIQA performance for asymmetric distortions, a 3D metric should accurately model the stimulus strength and account for the binocular combination. The third class of SIQA methods takes into account the monocular and/or binocular visual properties in addition to 2D image quality and disparity/depth information.

Several SIQA methods that simulate the binocular visual phenomena have been proposed. These methods assess the quality of a single-view separately, and then combine both quality scores into a 3D quality with the help of weights modeling stimulus strength. For instance, Wang *et al.* [3] proposed an information content and distortion weighted SSIM metric for left and right views, and employed a BR inspired multi-scale model to predict the perceived 3D quality from the 2D images based on image local variance. Recently, Cao *et al.* [14] developed a FR-SIQA method based on several visual characteristics of the human visual system (HVS). The patch-based image gradient entropy was used for modeling the stimulus strength. When a stereo pair is observed by a human subject, the HVS merges both views of the stereo pair to yield a single mental view (namely, cyclopean perceptual image) according to the binocular combination behavior [15]. The cyclopean perceptual image can be used to model BF and BR properties. Therefore, based on different binocular combination strategies, many other SIQA approaches [2, 16–19] in the literature combine left and right views into one cyclopean image,

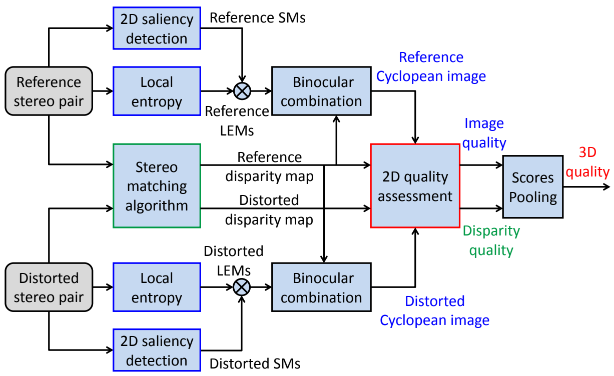

Fig. 1: Flowchart of the proposed SIQA method.

and the final 3D quality is evaluated by analyzing this merged image. For example, Chen *et al.* [17] developed a metric by assessing the quality of the cyclopean images constructed by a linear model. The weights of this model are derived from Gabor filter magnitude responses, which simulate the BR. Recently, Zhang and Chandler [2] presented a FR-SIQA metric based on monocular image quality estimated from left and right views, and cyclopean image quality measured using lightness distance and pixel-based contrast. Although these methods achieve much progress, various characteristics of the HVS have not been deeply explored, which limit the prediction accuracy. Therefore, to design more reliable and accurate SIQA metrics, it is important to account for the different perceptual processes.

In this paper, we propose a new SIQA method based on binocular combination properties and disparity information, combining quality scores of the cyclopean image and the disparity map. Specifically, the major contribution lies in the development of a 3D perceptual quality prediction framework by modeling the BF/BR phenomena, and accounting for disparity distortion as well as monocular visual saliency in the binocular combination. Besides, we provide a comprehensive experimental evaluation for our proposed method and a comparison with other SIQA methods on three databases. The rest of this paper is organized as follows. In Sec. 2, we detail the proposed SIQA method. Sec. 3 gives experimental results and comparative analysis. Finally, we conclude this paper in Sec. 4.

2. PROPOSED SIQA METHOD

As mentioned previously, the HVS does not account for left and right stimuli separately. Instead, it perceives distortions of the cyclopean image as 2D impairments, and depth/disparity distortion as 3D ones. Inspired by this, our proposed SIQA method predicts the overall 3D quality by combining the cyclopean image quality with disparity quality. Fig.1 illustrates the flowchart of the proposed method.

As shown in Fig.1 , the first step is to determine the disparity images for reference and distorted stereo-pairs. To achieve this, we use the stereo matching algorithm proposed in [20] for low-resolution S3D images and SSIM-based stereo algorithm [17] for high-resolution S3D images. These two algorithms can efficiently achieve good performance in disparity estimation and deal with the issue of occlusion and depth discontinuities. Next, inspired by the linear combination model proposed in [17, 21], by modeling the BF/BR phenomena when a stereo stimulus is presented, we generate the synthesized cyclopean image I_c as follows:

$$I_c(i,j) = \frac{LE_l(i,j+d_r)}{LE_T(i,j)} \times I_l(i,j+d_r) + \frac{LE_r(i,j)}{LE_T(i,j)} \times I_r(i,j), \quad (1)$$

$$LE_T(i,j) = LE_l(i,j+d_r) + LE_r(i,j), \quad (2)$$

where I_l and I_r represent the left and right views respectively. LE_l and LE_r denote the local energy maps for their corresponding images, and used to describe the stimuli to left and right eyes. In addition, (i,j) is the pixel coordinate. The left image I_l and its local energy map LE_l are warped to their corresponding locations in the right view using the disparity of right image d_r that corresponds to the horizontal shift of the pixel from the right to the left view. As shown in Eq.(1), the BR phenomenon is correlated to the relative stimulus strength of each view instead of the absolute stimulus strength [17].

The next step is to compute the local energy maps of two views to model the stimulus strength. The study in [22] found that the 3D human perception is dominated by the view of high contrast or rich contours. In other words, the perceptual 3D quality follows the quality of the view containing a higher amount of information. Therefore, the local entropy is used to determine the stimulus strength of each view. Moreover, we assume that the local energy of one view depends on the visual importance of the stimulus corresponding to the 2D visual saliency. The local energy $LE(i,j)$ of one view is defined by:

$$LE(i,j) = (EN(i,j) \times VS(i,j))^2, \quad (3)$$

where $EN(i,j)$ denotes the local entropy of a pixel (i,j) in one view of a stereo pair, and VS is the visual saliency map of this view. On the one hand, we use the method proposed in [23] to estimate the saliency map, because it performs well in terms of saliency prediction accuracy and computational efficiency. On the other hand, the image entropy is related to the amount of information that can be coded in the compression process. For example, a low entropy image contains very little contrast. The local entropy of a pixel computed based on 11-by-11 neighborhood with specific shape around this pixel is described as follows:

$$EN(i,j) = -\sum_{s=g_{min}}^{g_{max}} p(x_s) \times \log_2(p(x_s)), \quad (4)$$

where g_{min} and g_{max} are the minimum and maximum values respectively in the corresponding neighborhood pixels. $p(x_s)$ denotes the probability that the difference between two adjacent pixels is equal to s. Based on Eq.(1)-(4), the proposed SIQA metric tries to simulate the BF/BR phenomena. Specifically, different local energies in both views lead to BR, and the 3D quality of a region is more affected by the view containing higher contrast energies.

Given the cyclopean images (I_{rc}, I_{dc}) and the disparity maps (Dp_r, Dp_d) of the reference and distorted stereo pairs, we separately measure the cyclopean quality and disparity quality by using 2D IQA metric. In [11], You *et al.* found that universal image quality index (UQI) [24] performs best for 3D quality prediction among all tested 2D IQA metrics. Furthermore, the study in [26] revealed that the visual information fidelity (VIF) [25] metric can achieve an accurate quality prediction for 2D IQA database consisting of 2D high-resolution images such as CSIQ database [28]. On the other hand, UQI metric provided the best performance for IQA on the disparity map. In fact, UQI used in disparity quality estimation is based on comparing the structural information, and the disparity can express such information of the original images. Thereby the qualities

Table 1: Performance of SIQA methods on LIVE 3D IQA database (phase I). Italicized entries denote 2D-based IQA, and the results of the best-performing SIQA method are highlighted in boldface.

Distortion type	Criteria	*UQI [24]*	*VIF [25]*	*GMSD [26]*	Benoit [10]	You [11]	Fezza [18]	Chen [17]	Shao [27]	Lin [19]	Proposed
WN	LCC	*0.927*	*0.930*	*0.950*	0.925	0.941	0.947	**0.955**	0.945	0.927	0.932
	SROCC	*0.926*	*0.931*	*0.943*	0.929	0.940	0.944	**0.948**	0.941	0.929	0.927
	RMSE	*6.240*	*6.103*	*5.197*	6.308	5.622	5.351	**4.963**	-	6.257	6.038
JPEG	LCC	*0.769*	*0.603*	*0.664*	0.641	0.487	0.706	0.527	0.520	0.755	**0.781**
	SROCC	*0.737*	*0.580*	*0.620*	0.603	0.439	0.657	0.521	0.495	0.716	**0.748**
	RMSE	*4.178*	*5.216*	*4.888*	5.022	5.710	4.632	5.557	-	4.291	**4.086**
JP2K	LCC	*0.944*	*0.888*	*0.933*	0.940	0.878	0.937	0.920	0.921	0.952	**0.954**
	SROCC	*0.910*	*0.902*	*0.906*	0.910	0.860	0.896	0.887	0.895	0.913	**0.915**
	RMSE	*4.270*	*5.959*	*4.676*	4.427	6.207	4.532	5.070	-	3.963	**3.868**
GB	LCC	*0.952*	***0.962***	*0.960*	0.949	0.920	0.934	0.943	0.959	0.958	0.958
	SROCC	*0.925*	*0.934*	*0.939*	0.931	0.882	0.909	0.924	**0.940**	0.933	0.926
	RMSE	*4.451*	*3.955*	***4.051***	4.571	5.680	5.173	4.813	-	4.137	4.182
FF	LCC	*0.879*	*0.862*	*0.839*	0.747	0.730	0.783	0.776	0.859	0.862	**0.891**
	SROCC	*0.833*	*0.804*	*0.791*	**0.889**	0.583	0.693	0.700	0.796	0.829	0.844
	RMSE	*5.925*	*6.306*	*6.755*	8.258	8.492	7.730	7.832	-	6.299	**5.644**
ALL	LCC	*0.943*	*0.925*	*0.944*	0.903	0.881	0.821	0.922	0.935	0.937	**0.944**
	SROCC	*0.937*	*0.920*	*0.936*	0.889	0.879	0.922	0.914	0.925	0.931	**0.940**
	RMSE	*5.478*	*6.230*	*5.404*	7.062	7.746	9.358	6.351	5.816	5.744	**5.404**

of the cyclopean image and the disparity map are calculated as follows:

$$Q_c = UQI/VIF(I_{rc}, I_{dc}), Q_d = UQI(Dp_r, Dp_d), \quad (5)$$

where Q_c is the quality score of the test cyclopean image, and Q_d denotes the quality score of the disparity map. To estimate the cyclopean image quality, we use the UQI metric for LIVE 3D IQA databases (phase I [29] and phase II [17]), and the VIF metric for Waterloo-IVC 3D database (phase I) [3]. Finally, the S3D quality score Q_{3D} is calculated by a linear model:

$$Q_{3D} = \alpha \times Q_c + (1 - \alpha) \times Q_d \quad (6)$$

where α is the weight for adjusting the relative importance of Q_c and Q_d. In the implementation, we set $\alpha = 0.65$ for LIVE 3D phase I and II databases, and $\alpha = 0.6$ for Waterloo-IVC 3D phase I database. This is because the overall disparity of the stereo pairs in Waterloo-IVC 3D database is generally larger than that in LIVE 3D databases, and the disparity quality plays more important role in Waterloo-IVC 3D phase I database.

3. EXPERIMENTAL RESULTS AND ANALYSES

In this section, we evaluate the performance of the proposed and other SIQA methods on three publicly available 3D IQA databases providing subjective scores (DMOS values): LIVE 3D IQA databases (phase I [29] and phase II [17]), and the recently created Waterloo-IVC 3D database (phase I) [3]. LIVE 3D phase I database contains 20 reference stereo pairs and 365 symmetrically distorted stereo pairs, including five distortion types: additive white gaussian noise (WN), JPEG, JPEG 2000 compression (JP2K), gaussian blur (GB), and fast fading (FF). The LIVE 3D phase II database is composed of 8 reference stereo pairs and 360 symmetrically and asymmetrically distorted stereo pairs corresponding to the same distortion types. Waterloo-IVC 3D phase I database consists of 6 reference stereo pairs and 330 distorted stereo pairs with symmetric and asymmetric distortion levels and types including WN, JPEG and GB. The image resolution per view is 640×360 in LIVE 3D databases, and 1920×1080 in Waterloo-IVC 3D phase I database.

We compare the proposed method with other representative FR-3D-IQA methods [10, 11, 17–19, 27]. Besides, we further explored the performance of SIQA methods using only 2D IQA metrics including UQI [24], VIF [25], GMSD [26] and FSIM [30]. We choose these metrics in this paper since they yield promising results on 3D databases compared to other 2D IQA metrics. The performance of the SIQA metrics has been evaluated using three well-known measures: the Linear Correlation Coefficient (LCC), the Spearman Rank Order Correlation Coefficient (SROCC) and Root-Mean-Square Error (RMSE). The larger LCC and SROCC values, and the smaller RMSE value indicate better performance in terms of correlation with human opinion. The three performance measures were computed between DMOS and the predicted scores after a non-linear regression with a five-parameter logistic function described in [31].

Table 1 shows the performance of SIQA methods on LIVE 3D IQA phase I database. Overall, the proposed method outperforms most other 2D/3D IQA methods. Lin's [19] and Chen's [17] methods achieve better performance than Benoit's [10] and You's [11] methods thanks to consideration of the binocular vision properties. However, Lin's and Chen's methods are slower than the proposed method due to using 2D Gabor filter in their methods. Interestingly, all 2D-based IQA metrics perform quite well on the symmetrically distorted databases, and GMSD and UQI metrics perform even better than certain 3D IQA methods. Specifically, we also examine the performance of the SIQA metrics on each individual distortion type. As shown in Table 1, the proposed SIQA method provides better predictions on most distortion types in comparison with other methods except for white noise and gaussian blur. However, the obtained performance for the latter distortions remain competitive and in a very acceptable level. For the WN distortion, Chen's method [17] performs best since the MS-SSIM metric used in this method can yield a high prediction for WN distorted images. This observation indicates that the performances of some SIQA methods highly depend on the performance of the used 2D metric. Generally, all 2D-based or 3D IQA methods achieve reasonably accurate prediction results on LIVE 3D phase I database.

The quality prediction on LIVE 3D phase II database, which partially contains asymmetrically distorted stereo pairs, is more challenging than on LIVE 3D phase I database. For each SIQA method,

Table 2: Performance of SIQA methods on LIVE 3D IQA database (phase II). The symbols As and S are respectively the asymmetric and symmetric distortions.

Method	LCC			SROCC			RMSE		
	S	As	All	S	As	All	S	As	All
UQI [24]	*0.941*	*0.795*	*0.864*	*0.939*	*0.755*	*0.842*	*4.213*	*6.154*	*5.677*
VIF [25]	*0.928*	*0.777*	*0.837*	*0.916*	*0.732*	*0.819*	*4.652*	*6.382*	*6.182*
GMSD [26]	*0.920*	*0.738*	*0.803*	*0.910*	*0.716*	*0.783*	*4.897*	*6.842*	*6.723*
Benoit [10]	0.921	0.746	0.764	0.910	0.732	0.748	5.712	6.976	7.281
You [11]	0.911	0.659	0.721	0.898	0.604	0.721	7.128	8.009	7.141
Fezza [18]	0.930	0.820	0.871	0.921	0.796	0.862	4.576	5.801	5.553
Chen [17]	0.935	0.870	0.902	0.923	**0.851**	0.895	4.438	4.991	4.866
Shao [27]	-	-	0.863	-	-	0.849	-	-	5.706
Proposed	**0.945**	**0.883**	**0.912**	**0.941**	0.848	**0.898**	**4.076**	**4.754**	**4.643**

Table 2 shows the overall performance and the performance on separate subsets of symmetrically and asymmetrically S3D images in LIVE 3D phase II database. These results demonstrate that the proposed SIQA method delivers the best performance compared to the others methods. Moreover, the proposed method is particularly effective for asymmetric distortions. As expected, 2D-based IQA methods achieve high performance for the symmetric distortions, but they generally perform worse than most 3D IQA methods for asymmetric distortions. This is mainly due to the fact that 2D-based SIQA methods assess the perceptual 3D quality considering neither the depth/disparity information nor the binocular vision characteristics. It is worth noting that UQI-based SIQA method performs best within all 2D-based SIQA methods. Despite the consideration of the disparity distortion, the performance of Benoit's [10] and You's [11] metrics are much lower than the proposed metric, and particularly for asymmetric distortions. This is because these methods have not accounted for binocular vision properties such as BF. The methods based on cyclopean image (*i.e.*, Chen's [17], Fezza 's [18] and our proposed methods) achieve better performance than the other 3D IQA methods.

Table 3: Performance of SIQA methods on Waterloo-IVC 3D database (phase I).

Method	LCC			SROCC		
	S	As	All	S	As	All
UQI [24]	*0.814*	*0.724*	*0.753*	*0.635*	*0.631*	*0.640*
VIF [25]	*0.918*	***0.788***	*0.839*	***0.914***	***0.755***	***0.801***
FSIM [30]	*0.839*	*0.668*	*0.767*	*0.918*	*0.625*	*0.704*
Benoit [10]	0.850	0.697	0.680	0.728	0.577	0.585
You [11]	0.868	0.709	0.713	0.752	0.571	0.600
Fezza [18]	0.881	0.611	0.692	0.782	0.484	0.553
Chen [17]	0.837	0.536	0.657	0.649	0.496	0.382
Wang [32]	0.833	0.609	0.677	0.683	0.500	0.552
Proposed	**0.949**	0.774	**0.841**	**0.914**	0.729	0.790

In addition to performance evaluation on LIVE 3D databases, the performance comparison between the proposed and other SIQA methods on Waterloo-IVC 3D database (phase I) is shown in Table 3. The proposed method performs better than the others and ranks second for asymmetrically distorted S3D image. The decrease of performance in our method is mostly due to the use of asymmetrically mixed distortion types in this database making it more challenging to be assessed by most SIQA metrics. The scatter plots of DMOS vs. our quality scores are shown on Fig. 2.

4. CONCLUSION

In this paper, we proposed a full-reference quality assessment method for stereoscopic images accounting for binocular combination and disparity distortion. The proposed method models the human stereo vision by fusing the left and right views to generate a cyclopean image based on local entropy and monocular visual saliency. Then, a 2D quality metric is employed to separately evaluate the quality of both the cyclopean image and disparity map derived from a stereo matching algorithm. Finally, two quality scores are combined to yield an overall 3D quality score. An extensive performance comparison of the proposed method with some 2D-based IQA and 3D QA methods is conducted on three databases. The experimental results demonstrate that the proposed method achieves better performance than other SIQA methods for most of the databases and distortions. The is less true for Waterloo 3D database (phase I) because of the use of mixed asymmetric distortions types. This latter case, will be explored in the future in order to improve the quality prediction of our metric.

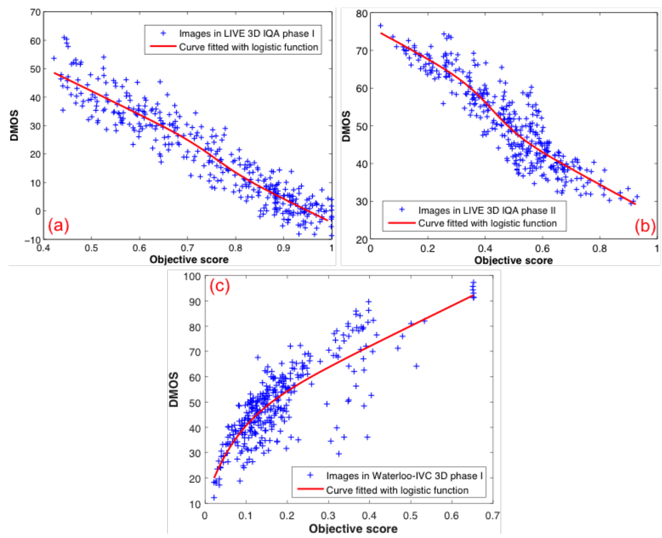

Fig. 2: Scatter distribution of predicted scores obtained by proposed SIQA metric vers DMOS for three databases, (a) LIVE 3D phase I, (b) LIVE 3D phase II, (c) Waterloo-IVC 3D phase I.

5. ACKNOWLEDGMENT

This research has been funded by both the Research Council of Norway through project no. 221073 "HyPerCept Colour and quality in higher dimensions" and the Region "Nouvelle Aquitaine".

6. 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*. Springer, 2015, pp. 185–212.

[2] Y. Zhang and D. M. Chandler, "3d-mad: A full reference stereoscopic image quality estimator based on binocular lightness and contrast perception," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 3810–3825, 2015.

[3] 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, 2015.

[4] F. Qi, D. Zhao, and W. Gao, "Reduced reference stereoscopic image quality assessment based on binocular perceptual information," *IEEE Transactions on Multimedia*, vol. 17, no. 12, pp. 2338–2344, 2015.

[5] L. Ma, X. Wang, Q. Liu, and K. N. Ngan, "Reorganized dct-based image representation for reduced reference stereoscopic image quality assessment," *Neurocomputing*, vol. 215, pp. 21–31, 2016.

[6] W. Zhang, C. Qu, L. Ma, J. Guan, and R. Huang, "Learning structure of stereoscopic image for no-reference quality assessment with convolutional neural network," *Pattern Recognition*, vol. 59, pp. 176–187, 2016.

[7] F. Shao, W. Tian, W. Lin, G. Jiang, and Q. Dai, "Toward a blind deep quality evaluator for stereoscopic images based on monocular and binocular interactions," *IEEE Transactions on Image Processing*, vol. 25, no. 5, pp. 2059–2074, 2016.

[8] P. Campisi, P. Le Callet, and E. Marini, "Stereoscopic images quality assessment," in *Signal Processing Conference, 2007 15th European*. IEEE, 2007, pp. 2110–2114.

[9] 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*. IEEE, 2009, pp. 1–4.

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

[11] J. You, L. Xing, A. Perkis, and X. Wang, "Perceptual quality assessment for stereoscopic images based on 2d image quality metrics and disparity analysis," in *Proc. of Int. Workshop on Video Process. and Quality Metrics for Consumer Electronics, Scottsdale, AZ, USA*, 2010.

[12] R. Blake and N. K. Logothetis, "Visual competition," *Nature Reviews Neuroscience*, vol. 3, no. 1, pp. 13–21, 2002.

[13] J. Brascamp, H. Sohn, S.-H. Lee, and R. Blake, "A monocular contribution to stimulus rivalry," *PNAS*, vol. 110, no. 21, pp. 8337–8344, 2013.

[14] Y. Cao, W. Hong, and L. Yu, "Full-reference perceptual quality assessment for stereoscopic images based on primary visual processing mechanism," in *Multimedia and Expo (ICME), 2016 IEEE International Conference on*. IEEE, 2016, pp. 1–6.

[15] B. Julesz, "Foundations of cyclopean perception." 1971.

[16] A. Maalouf and M.-C. Larabi, "Cyclop: A stereo color image quality assessment metric," in *ICASSP*. IEEE, 2011, pp. 1161–1164.

[17] 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 Processing: Image Communication*, vol. 28, no. 9, pp. 1143–1155, 2013.

[18] S. A. Fezza and M.-C. Larabi, "Stereoscopic 3d image quality assessment based on cyclopean view and depth map," in *ICCE*. IEEE, 2014, pp. 335–339.

[19] Y. Lin, J. Yang, L. Wen, Q. Meng, Z. Lv, and H. Song, "Quality index for stereoscopic images by jointly evaluating cyclopean amplitude and cyclopean phase," *IEEE Journal of Selected Topics in Signal Processing*, 2016.

[20] S. Lee, J. H. Lee, J. Lim, and I. H. Suh, "Robust stereo matching using adaptive random walk with restart algorithm," *Image and Vision Computing*, vol. 37, pp. 1–11, 2015.

[21] J. Ding and G. Sperling, "A gain-control theory of binocular combination," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 4, pp. 1141–1146, 2006.

[22] 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, p. 143, 2001.

[23] H. R. Tavakoli, E. Rahtu, and J. Heikkilä, "Fast and efficient saliency detection using sparse sampling and kernel density estimation," in *Scandinavian Conference on Image Analysis*. Springer, 2011, pp. 666–675.

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

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

[26] W. Xue, L. Zhang, X. Mou, and A. C. Bovik, "Gradient magnitude similarity deviation: A highly efficient perceptual image quality index," *IEEE Transactions on Image Processing*, vol. 23, no. 2, pp. 684–695, 2014.

[27] F. Shao, K. Li, W. Lin, G. Jiang, M. Yu, and Q. Dai, "Full-reference quality assessment of stereoscopic images by learning binocular receptive field properties," *IEEE Transactions on Image Processing*, vol. 24, no. 10, pp. 2971–2983, 2015.

[28] E. C. Larson and D. M. Chandler, "Most apparent distortion: full-reference image quality assessment and the role of strategy," *Journal of Electronic Imaging*, vol. 19, no. 1, pp. 011 006–011 006, 2010.

[29] A. K. Moorthy, C.-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.

[30] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "Fsim: A feature similarity index for image quality assessment," *IEEE transactions on Image Processing*, vol. 20, no. 8, pp. 2378–2386, 2011.

[31] H. R. Sheikh, M. F. Sabir, and A. C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," *IEEE Trans. Image Process.*, vol. 15, no. 11, pp. 3440–3451, 2006.

[32] X. Wang, S. Kwong, and Y. Zhang, "Considering binocular spatial sensitivity in stereoscopic image quality assessment," in *VCIP*. IEEE, 2011, pp. 1–4.