Presented at ICIP 2017, Beijing China

LEARNING NATURAL STATISTICS OF BINOCULAR CONTRAST FOR NO REFERENCE QUALITY ASSESSMENT OF STEREOSCOPIC IMAGES

Yi Zhang, Damon M. Chandler

Laboratory of Computational Perception and Image Quality
Department of Electrical and Electronic Engineering
Shizuoka University, Hamamatsu, Japan

ABSTRACT

Algorithms for no-reference (NR) stereoscopic image quality assessment (SIQA) aim to evaluate the perceptual quality of a stereoscopic/3D image without the assistance of its reference. Current NR SIQA models often require training on 3D distorted images and their associated human opinion scores, which ultimately restrict their further application. In this paper, we present a simple yet effective NR SIQA model that does not require training on existing 3D image databases. Instead, we train our model on a large dataset of natural stereoscopic images based on learning the local statistics of the Cyclopean contrast maps, and then use the existing 2D NR IQA model to help guide the NR SIQA task. Experimental results demonstrate the efficacy of our proposed method.

Index Terms— Stereoscopic image quality assessment, binocular contrast, log-derivative statistics, multivariate Gaussian

1. INTRODUCTION

Research in image quality assessment (IQA) has grown tremendously over the last two decades leading to numerous powerful algorithms for evaluating 2D image quality with or without the reference information (see [1] for reviews). IQA of stereoscopic/3D images, however, is relatively new and less mature, due in large part to the difficulties and complexities in mimicking binocular visual processing in the human visual system (HVS). As with 2D IQA, the SIQA problem can be broadly classified into three main categories based on the availability of a reference image: full-reference (FR), reduced-reference (RR), and no-reference (NR). The vast majority of SIQA algorithms have addressed the FR and RR scenarios (e.g., [2, 3, 4], etc.).

In this paper, we address NR SIQA (or *blind* SIQA), in which only the distorted stereoscopic image is available to the IQA algorithm. Blind assessment of a 3D image quality is extremely challenging not only because of the difficulty in modeling the HVS in blindly evaluating a single image quality, but also because of the complex binocular fusion and rivalry behaviors that will occur when the human eyes are presented with asymmetrically distorted views. The former one has been widely mentioned in many 2D IQA work (e.g., [1, 5, 6], etc.). The latter one has been elaborated in [7] that binocular combination under different distortion types should be considered: the high quality view that contains sufficient information will help suppress the low quality view with information-loss distortion (e.g., blurring) [8], while for the information-additive distortion (e.g., blockiness), the low quality view cannot be compensated [9].

To model the complex behaviors of the HVS in NR SIQA task, various models have been proposed by using different quality-related features (e.g., [7, 10, 11, 12], etc.). However, most of these approaches share a common thread: they require training on 3D images with anticipated distortion information as well as their associated human opinion scores, and consequently, a limited number of existing 3D image databases potentially restrict their wider applicability. To the best of our knowledge, only a few NR SIQA algorithms (e.g., [12, 13]) require no training on distorted stereoscopic images, but their performance is less competitive. To release the NR SIQA algorithms from the dependence on prior knowledge of distorted stereopairs while still maintaining a higher quality predicting performance, here, we present a simple, yet effective NR SIQA algorithm which operates by learning the local natural statistics (NS) of the Cyclopean contrast maps.

The proposed SIQA algorithm, called 3DLN (QA of *3D* images via Learning Natural statistics), is inspired by three previous works: the log-derivative-statistics-based DESIQUE [14] model, the cyclopean-feature-image-based 3D-MAD model [3], and the multivariate Gaussian (MVG)-based NIQE model [6]. To blindly assess the stereoscopic image quality without depending on training the existing 3D image databases, 3DLN explores the *log*-derivative-based statistical features computed from Cyclopean contrast maps of the natural and distorted stereoscopic images, and meanwhile utilizes the 2D-stereopair distortion information to help guide the NR SIQA task. Consequently, the qualities obtained from both the stereopairs and the Cyclopean-contrast-maps are combined to yield an overall quality estimate of the stereoscopic image.

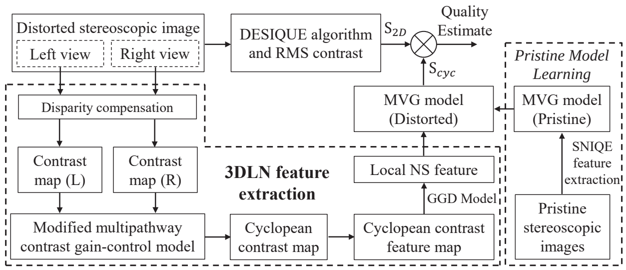

Fig. 1. A block diagram of the 3DLN algorithm.

This paper is organized as follows: Section 2 provides details of the proposed 3DLN algorithm. In Section 3, we analyze performance of 3DLN on 3D image quality databases. General conclusions are presented in Section 4.

2. ALGORITHM

The proposed 3DLN algorithm consists of two main stages: (1) the 2D-DESIQUE-based quality estimate on stereopairs, and (2) the local NS-based quality estimate on Cyclopean contrast maps, and its block diagram is shown in Figure 1. In the following subsections, we provide details for each stage.

2.1. 2D-DESIQUE on Stereopairs

In the 2D-DESIQUE-based QA stage, the conventional DESIQUE algorithm [14] is applied to the stereopairs (i.e., the left and right view images) to estimate the perceived distortion corresponding to each monocular view. Then, the overall 2D-DESIQUE quality is computed as the weighted sum of both stereopair distortion measures, where the weights are computed based on the block-based RMS contrast measure.

Specifically, given an image, we first divide it into blocks of 16×16 pixels (with 75% overlap), and then compute the RMS contrast of each block via

$$C(b) = \tilde{\sigma}(b)/\mu(b) \tag{1}$$

where $\mu(b)$ represents the average luminance value of block b, and $\tilde{\sigma}(b)$ represents the minimum standard deviation among the four 8×8 subblocks within b (see Appendix A in [15]). We compute block-based contrast maps for the left and right views of the distorted stereoscopic image (denoted by C_L and C_R, respectively), and then the 2D-DESIQUE score is given by

$$\mathrm{S}_{2D} = (\bar{C}_L \cdot Q_L + \bar{C}_R \cdot Q_R)/(\bar{C}_L + \bar{C}_R) \tag{2}$$

where $\bar{C}$ is the mean value of C; Q_L and Q_R denote the DESIQUE quality estimates of the left and right views.

2.2. Local NS-based Cyclopean IQA

The local NS-based Cyclopean IQA explores the idea in [6] that image quality can be measured locally, but extends this idea from 2D to 3D by using an efficient binocular fusion/rivalry model, and different quality-related features. In addition, important distortion information obtained in the first stage of 3DLN is used to guide the Cyclopean IQA process.

2.2.1. Cyclopean Contrast Maps

Motivated by [3], we propose two types of contrast as the algorithm's raw inputs for the Cyclopean IQA stage:

$$f_1\left(x,y\right) = \frac{2\sqrt{L\left(x,y\right) \cdot \bar{L}_B\left(x,y\right)}}{L\left(x,y\right) + \bar{L}_B\left(x,y\right) + K}, \tag{3}$$

$$f_2\left(x,y\right) = L\left(x,y\right)/[\bar{L}_B\left(x,y\right) + K], \tag{4}$$

where L denotes the luminance value; $\bar{L}_B\left(x,y\right)$ denotes the average luminance value of a 9×9 block centered around pixel (x,y); $K = 0.001$ is a small constant that prevents division by zero.

Given the two contrast maps for each view of the stereoscopic image, we next build disparity-compensated Cyclopean contrast maps based on a modified multipathway contrast-gain control model (MCM)[3]. In our previous FR SIQA work [3], we argue that the quality of a monocular scene with higher contrast will play a more dominant role in determining the HVS's judgment of a 3D image quality. In this paper, we further argue that if the two monocular scenes have similar contrast, then the scene with more quality degradations weighs more on determining the overall 3D image quality. This statement is also in accord with findings in [9] that low quality view caused by information-additive distortion (e.g., blockiness) cannot be compensated by the high quality view. Thus, to better model the binocular interactions between the two monocular views, we modify MCM in [3] as follows: if the two views have significantly different quality ratings but similar averaged RMS contrast values, then the view with lower quality will be compensated by a larger weight in the Cyclopean-view-build process.

Specifically, we use DESIQUE scores as the approximate quality estimates of the two monocular views, and their quality difference is computed by $r = 2Q_LQ_R/(Q_L^2 + Q_R^2)$. We assume that $r < 0.75$ indicates a significant quality difference between the two views, and consequently their corresponding Cyclopean contrast maps (denoted by $C_{f_i} (i = 1, 2)$) are computed via Eq.(5). Note that all symbols in Eq.(5) have the same meanings and values as that defined in [3] except the two compensation factors E_L and E_R, which are given by

$$E_L = e^{1-r} S\left(\omega\right) u\left(Q_L - Q_R\right), \tag{6}$$

$$E_R = e^{1-r} S\left(\omega\right) u\left(Q_R - Q_L\right), \tag{7}$$

where $u(\cdot)$ is a step function; $S\left(\omega\right)$ is a sigmoid transducer function:

$$S\left(\omega\right) = \frac{A}{1 + e^{t_1(\omega - t_2)}} + B. \tag{8}$$

Here, $\omega = 2\bar{C}_L\bar{C}_R/(\bar{C}_L^2 + \bar{C}_R^2)$ represents the RMS contrast similarity between the two monocular views. Four free parameters are selected as follows: $A = 9$, $B = 1$, $t_1 = -8$,

$$C_{f_i}(x,y)=\frac{\left[\left(\eta_L f_{i,L}(x-d_{x,y},y)\frac{E_L}{1+\frac{\epsilon_R(x-d_{x,y},y)}{1+\beta\epsilon_L(x-d_{x,y},y)}}\right)^{\gamma_2}+\left(\eta_R f_{i,R}(x,y)\frac{E_R}{1+\frac{\alpha\epsilon_L(x,y)}{1+\epsilon_R(x,y)}}\right)^{\gamma_2}\right]^{\frac{1}{\gamma_2}}}{\left[\left(\frac{\eta_L E_L}{1+\frac{\epsilon_R(x-d_{x,y},y)}{1+\beta\epsilon_L(x-d_{x,y},y)}}\right)^{\gamma_2}+\left(\frac{\eta_R E_R}{1+\frac{\alpha\epsilon_L(x,y)}{1+\epsilon_R(x,y)}}\right)^{\gamma_2}\right]^{\frac{1}{\gamma_2}}},(i=1,2) \quad (5)$$

and $t_2 = 0.6$. We found that changing these parameter values does not affect the algorithm performance significantly. For those $r > 0.75$ stereopairs, the same MCM-based binocular model [3] is used to compute the Cyclopean contrast maps.

2.2.2. *Local NS Feature Extraction*

Based on the two Cyclopean contrast maps, the next step is to extract efficient local statistical features that are representative of the quality degradation. Towards this end, for each Cyclopean contrast map, we first apply the eight derivative-related image operators to build the corresponding eight Cyclopean contrast feature (CCF) maps. Then, the log-derivative statistics are applied to the mean-subtracted contrast-normalized (MSCN) coefficients of each CCF map in a block-based manner to extract the local features.

Let C_f denote one of the two Cyclopean contrast maps, and its elements are denoted by $C_f(i,j)$. Then, for each Cyclopean contrast map, its eight CCF maps are computed by

$$\nabla^1 C_f(i,j) = C_f(i',j') - C_f(i,j) \quad (9)$$

$$\nabla^2 C_f(i,j) = \frac{2C_f(i',j')C_f(i,j)}{C_f^2(i',j') + C_f^2(i,j)} \quad (10)$$

where ∇^1 and ∇^2 denote the derivative-related image operators; $C_f(i',j')$ denotes the four neighboring pixels around $C_f(i,j)$ [i.e., $C_f(i',j') = C_f(i,j+1), C_f(i+1,j), C_f(i+1,j+1)$, or $C_f(i+1,j-1)$].

We found that the MSCN coefficients of these CCF maps fit quite well to the generalized Gaussian distribution (GGD) model when the four-orientation, spatial-domain log-derivative statistics are applied (i.e., D_1 through D_4 in Section 3.2.1 in [14]). Thus, similar to [6], we extract local log-derivative statistical features from the selected CCF map patches, which correspond to the sharper regions of the synthesized Cyclopean view.

In [6], sharp image patches are selected based on their averaged local variances, and this process is applied only to the natural images. In our work, we use FISH_{bb} algorithm [16] to more accurately measure the local sharpness of a Cyclopean image, and the sharp CCF patches are selected from both the natural and distorted images for feature extraction. We found that even in distorted images with loss of sharpness, those relatively sharper regions still weigh more on determining the overall image quality. Also, a less number of computing blocks allows less computational complexity.

Finally, the four-orientation, spatial-domain log-derivative statistics are applied to the selected CCF map patch, from which the two GGD parameters (α and σ) are estimated based on [17]. We use $\log(1+\alpha)$ as the features extracted from C_{f_1} map, and σ from C_{f_2} map. Consequently, a total of 64 features (4 log-derivative orientations $\times$ 8 CCF maps $\times$ 2 contrast maps) are extracted from each patch.

2.2.3. *Cyclopean Quality Estimate*

The quality estimation of the two Cyclopean contrast maps [defined in Eqs.(3) and (4)] follows the same procedures as in [6]. First, the MVG distribution was employed to model the patch-based features extracted from both the pristine and distorted CCF maps, which is given by

$$f_X(x_1,...,x_k) = \frac{exp\left[-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right]}{(2\pi)^{k/2}|\Sigma|^{1/2}}, \quad (11)$$

where $x_1,...,x_k$ are the computed feature vector for each patch; μ and Σ denote the mean and covariance matrix of the MVG model. Then the quality of each Cyclopean contrast map is given by

$$d_i = \sqrt{\left(\mu_{f_i}^d - \mu_{f_i}^p\right)^T \left(\frac{\Sigma_{f_i}^d + \Sigma_{f_i}^p}{2}\right)^{-1} \left(\mu_{f_i}^d - \mu_{f_i}^p\right)}. \quad (12)$$

Here, $\mu_{f_i}^{p/d}$ and $\Sigma_{f_i}^{p/d}(i=1,2)$ denote the estimated mean and covariance matrix (corresponding to C_{f_i}) of the pristine and distorted MVG models, respectively. Finally, the Cyclopean quality is computed by

$$\mathrm{S}_{cyc} = \ln(1 + d_1 \times d_2). \quad (13)$$

Note that for the noise-corrupted images, quality prediction using Eq.(13) may very likely produce an unexpected deviation from the normal quality range maintained by other distortions. To address this problem, we employ an Gaussian-filter which smooths partial noise in stereopairs before the two contrast maps are computed. In our implementation, we use DESIQUE algorithm [14] to detect noise, and then the Gaussian filter variance is determined by the 2D DESIQUE score of each monocular view through another sigmoid transducer function, which has the same form as Eq.(8). Here, the variable ω represents the smaller value of the two DESIQUE scores (i.e., $\omega = \min\{Q_L, Q_R\}$), and the Gaussian-filter size is 5×5 pixels. Also, we set the following parameter values: $A = 0.55$, $B = 0$, $t_1 = -0.2$, and $t_2 = 60$. These parameter values are empirically selected to help achieve the best performance on the LIVE3D image database [18].

Table 1. Performance of 3DLN and other FR/NR SIQA algorithms on the LIVE3D phase I and phase II databases.

	JP2K	JPEG	WN	GBLUR	FF	ALL
			LIVE3D phase I			
Chen [2]	*0.896*	*0.558*	*0.948*	*0.926*	*0.688*	*0.916*
Lin [19]	*0.839*	*0.207*	*0.928*	*0.935*	*0.658*	*0.856*
Shao [20]	*0.883*	*0.599*	*0.930*	*0.910*	*0.793*	*0.927*
QAC [21]	0.886	0.682	0.938	0.871	0.558	0.839
NIQE [6]	0.632	0.599	0.907	0.861	0.519	0.797
IL_NIQE [5]	0.861	0.544	0.920	0.873	0.536	0.860
Akhter [13]	0.866	0.675	0.914	0.555	0.640	0.383
Zhou [12]	0.837	0.638	0.931	0.833	0.649	0.892
3DLN-2D	0.881	0.592	0.933	0.815	0.713	0.908
3DLN-cyc	0.847	0.579	0.858	0.867	0.759	0.877
3DLN	0.883	0.616	0.933	0.851	0.759	0.915
			LIVE3D phase II			
Chen [2]	*0.833*	*0.840*	*0.955*	*0.910*	*0.889*	*0.901*
Lin [19]	*0.719*	*0.613*	*0.907*	*0.711*	*0.701*	*0.638*
Shao [20]	*0.788*	*0.745*	*0.807*	*0.939*	*0.935*	*0.819*
QAC [21]	0.781	0.821	0.621	0.843	0.891	0.785
NIQE [6]	0.599	0.637	0.600	0.851	0.775	0.707
IL_NIQE [5]	0.593	0.491	0.617	0.880	0.758	0.673
Akhter [13]	0.724	0.649	0.714	0.682	0.559	0.543
Zhou [12]	0.553	0.593	0.893	0.869	0.828	0.825
3DLN-2D	0.787	0.826	0.884	0.812	0.910	0.858
3DLN-cyc	0.696	0.657	0.865	0.864	0.817	0.831
3DLN	0.796	0.832	0.954	0.858	0.890	0.884

2.3. 3DLN Quality Index

The final stage of 3DLN is to combine the quality estimates obtained from the two previous stages into an overall quality index, which is given by

$$\mathrm{3DLN} = \mathrm{S}_{2D} \times \mathrm{S}_{cyc}. \tag{14}$$

Smaller values denote predictions of better stereoscopic image quality.

3. RESULTS AND ANALYSIS

We tested and compared 3DLN with other FR/NR SIQA algorithms on the LIVE3D phase I and phase II databases [18] in terms of the Spearman rank-order correlation coefficient (SROCC) values (the Pearson linear correlation coefficient (CC) values follow the similar trend). The three FR SIQA algorithms include the Cyclopean MS-SSIM proposed by Chen *et al.* [2], the frequency-integrated PSNR (FI-PSNR) proposed by Lin *et al.* [19], and the BJND-based method proposed by Shao *et al.* [20]. The five NR SIQA algorithms include three 2D IQA and two SIQA algorithms, all of which do not require training on distorted stereoscopic images. Note that for the three 2D IQA metrics (QAC [21], NIQE [6], and IL_NIQE [5]), the predicted quality of a stereoscopic image was taken to be the average quality predicted from the left and right views. The two NR SIQA algorithms are the local-feature-based method proposed by Akhter *et al.* [13] and the MVG-based method proposed by Zhou *et al.* [12].

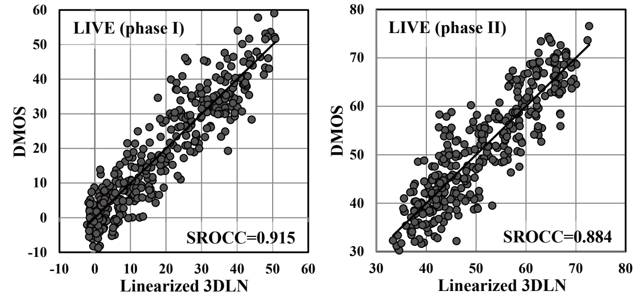

Fig. 2. Scatter-plots of objective scores predicted by 3DLN algorithm after logistic transform versus subjective scores on the LIVE3D phase I and phase II databases.

The experiment results are shown in Table 1, in which italicized entries denote FR SIQA algorithms. Also included are the SROCC values of the first and second stage of 3DLN, denoted by 3DLN-2D and 3DLN-cyc, respectively. Observed from Table 1 that 3DLN outperforms all the other five NR SIQA algorithms. Regarding to the three FR SIQA algorithms, 3DLN achieves better performance than Lin's method [19] on both databases, and challenges the other two, both of which have taken into account the binocular fusion/rivalry properties for analysis. Compared with 3DLN-2D and 3DLN-cyc, we see that the combined 2D and Cyclopean contrast analysis improves upon each individual stage.

Figure 2 shows the scatter-plots of logistic-transformed 3DLN quality predictions vs. subjective ratings (DMOS) on the two testing databases. In both graphs, the y-axis denotes the subjective ratings of the perceived distortions and the x-axis denotes the predicted quality value after the logistic transform as in [3]. Despite the presence of some outliers, the plots are generally heteroscedastic.

4. CONCLUSION

This paper presented an algorithm, called 3DLN, to blindly evaluate the quality of stereoscopic images via learning the local natural statistics of binocular contrast maps, and also by utilizing the 2D distortion information to help guide the SIQA task based on properties of the human binocular vision. The proposed algorithm consists of two main stages. The first stage employs DESIQUE algorithm to assess the quality of each monocular view, and to obtain the basic distortion identification information. The second stage explores the log-derivative statistics of the Cyclopean contrast map, and estimate the Cyclopean quality based on computing the Bhattacharyya distance between the estimated distorted MVG models and the pre-trained natural MVG models. The overall stereoscopic image quality is a combination of the two quality estimates from the two stages. We demonstrated the efficiency of 3DLN on several image databases.

5. REFERENCES

[1] D. M. Chandler, "Seven challenges in image quality assessment: past, present, and future research," *ISRN Signal Processing*, vol. 2013, Article ID 905685, 53 pages, 2013.

[2] 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.

[3] Yi Zhang and Damon 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.

[4] Wujie Zhou, Gangyi Jiang, Mei Yu, Feng Shao, and Zongju Peng, "Reduced-reference stereoscopic image quality assessment based on view and disparity zero-watermarks," *Signal Processing: Image Communication*, vol. 29, no. 1, pp. 167–176, 2014.

[5] Lin Zhang, Lei Zhang, and Alan C Bovik, "A feature-enriched completely blind image quality evaluator," *IEEE Transactions on Image Processing*, vol. 24, no. 8, pp. 2579–2591, 2015.

[6] Anish Mittal, R. Soundararajan, and Alan C. Bovik, "Making a complete blind image quality analyzer," *IEEE Signal Processing Letters*, vol. 20, no. 3, pp. 209–212, 2013.

[7] Feng Shao, Weijun Tian, Weisi Lin, Gangyi Jiang, and Qionghai 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] Daniel V Meegan, Lew B Stelmach, and W James 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.

[9] Pieter Seuntiens, Lydia Meesters, and Wijnand Ijsselsteijn, "Perceived quality of compressed stereoscopic images: Effects of symmetric and asymmetric jpeg coding and camera separation," *ACM Transactions on Applied Perception (TAP)*, vol. 3, no. 2, pp. 95–109, 2006.

[10] Ming-Jun Chen, Lawrence K Cormack, and Alan C Bovik, "No-reference quality assessment of natural stereopairs," *IEEE Transactions on Image Processing*, vol. 22, no. 9, pp. 3379–3391, 2013.

[11] Che-Chun Su, Lawrence K Cormack, and Alan C Bovik, "Oriented correlation models of distorted natural images with application to natural stereopair quality evaluation," *IEEE Transactions on Image Processing*, vol. 24, no. 5, pp. 1685–1699, 2015.

[12] Wujie Zhou, Lu Yu, Weiwei Qiu, Ting Luo, Zhongpeng Wang, and Ming-Wei Wu, "Utilizing binocular vision to facilitate completely blind 3d image quality measurement," *Signal Processing*, 2016.

[13] Roushain Akhter, ZM Parvez Sazzad, Yuukou Horita, and Jacky Baltes, "No-reference stereoscopic image quality assessment," in *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 2010, pp. 75240T–75240T.

[14] Yi Zhang and Damon M. Chandler, "No-reference image quality assessment based on log-derivative statistics of natural scenes," *Journal of Electronic Imaging*, vol. 22, no. 4, pp. 043025, 2013.

[15] 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. 011006, 2010.

[16] P. V. Vu and D. M. Chandler, "A fast wavelet-based algorithm for global and local image sharpness estimation," *Signal Processing Letters, IEEE*, vol. 19, no. 7, pp. 423–426, July 2012.

[17] K. Sharifi and A. Leon-Garcia, "Estimation of shape parameter for generalized gaussian distribution in subband decompositions of video," *IEEE Transactions on Circuits and System for Video Technology*, vol. 5, no. 1, pp. 52–56, 1995.

[18] A. K. Moorthy, C.-C. Su, A. Mittal, and A. C. Bovik, "Subjective evaluation of stereoscopic image quality," *Signal Processing: Image Communication*, vol. 28, pp. 870–873, 2012.

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

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

[21] Wufeng Xue, Lei Zhang, and Xuanqin Mou, "Learning without human scores for blind image quality assessment," in *IEEE International Conference on Computer Vision and Pattern Recognition*, 2013, pp. 995–1002.