Stereoscopic Image Quality Assessment by Analyzing Depth and Local Texture Information

Yang Zhao, Yong Ding, *Member, IEEE,* and Xiaogang Xu

***Abstract*—Stereoscopic image quality assessment (SIQA) plays an important role in modern image and video processing systems, but remains a difficult task due to various complex binocular effects. This Letter presents an SIQA method that considers properties of human visual system including binocular effects and the interests of primary and secondary visual cortex. By taking both image depth and low- and high-level local texture information into account, a comprehensive description of the stereoscopic image quality is obtained. The dissimilarity between the features extracted from a tested image that complete the quality description and that from its reference is then mapped to a quality score by support vector machine. Experimental results on large-scale databases prove that the method is accurate and robust in predicting human's subjective estimation of stereoscopic image quality.**

***Index Terms*—Image quality assessment, stereoscopic image, feature extraction**

I. Introduction

WITH the rapid development of digital technology, especially the progress made in stereo display techniques, there emerges a heavy demand of high-quality 3D images and videos [1]. Serving as evaluator and monitor for the quality of 3D images and videos, stereoscopic image quality assessment (SIQA) is becoming an important research field and a necessary part for almost all modern image and video processing systems, yet the study on it still remains immature due to the limited understanding upon human binocular perception [2]. A major problem brought about by stereoscopic images is that their quality tends to be influenced by complex inherent factors, such as depth information, binocular rivalry, visual fatigue, etc. [1-5]. Therefore, the solution to SIQA is much more mysterious comparing to its 2D counterpart. Same as in 2D IQA, SIQA methods are categorized into subjective and objective ones. Subjective SIQA employs people to judge image quality according to their sensation, and is therefore consuming in time and labor and impossible to be equipped in practical systems. So, the research of SIQA mainly focuses on designing objective methods in the pursuit of predicting stereoscopic image quality automatically and accurately.

This work was supported by National High Technology Research and Development Program of China (Grant No. 2015AA016704c) and Zhejiang Provincial Natural Science Foundation (LY14F020028).

Y. Zhao, Y. Ding and X. Xu are with College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, China (email: yang_z@zju.edu.cn; dingy@vlsi.zju.edu.cn; xiaogangxu@zju.edu.cn).

Thanks to the findings from 2D IQA research, the process how human visual system (HVS) perceives image quality is becoming clearer, and it is logical that these findings should be well making use of when designing SIQA methods [2, 3]. However, straightforwardly introducing 2D methods into SIQA without considering binocular effects is highly insufficient [2, 3, 6]. Especially, these approaches fail in dealing with asymmetrically distorted image pairs. Also, depth information should be understandably not negligible. Estimating a disparity map to acquire depth information, the combinational quality evaluation on both views as well as the disparity map is another straightforward yet more promising solution [7-8]. It is currently acknowledged that binocular fusion and rivalry exist widely and simultaneously when people viewing things [1-4]. More complex models have been established by considering the binocular perception seriously. For example, it is claimed that stereoscopic images can be divided into monocular occlusion and binocular regions and treated differently [1]. Another way is to develop an intermediate image with the two views instead of evaluating them directly, using relative energy of left and right views to model the effects of binocular perception and generate weights for intermediate image construction from both views [3].

Aside from binocular effects, another fact worth addressing is that the features extracted should be attractive to HVS so that they affect subjective evaluation upon image quality. For the sake of effectiveness, we are inspired to utilize the multi-scale property of HVS, and extract both high- and low-level local texture features from constructed intermediate images. Moreover, depth information is also analyzed for supplement. The main contribution of this Letter lies in 1) presenting a comprehensive image quality descriptor that makes use of both high- and low-level quality-aware local texture features as well as image depth information, and 2) developing an SQIA method based on the descriptor, which is proven accurate and robust on public SIQA databases.

II. The Proposed SIQA Method

As is mentioned, the proposed method extracts quality-aware features in two aspects: depth information, which is estimated according to the disparity between both views; and local texture information, which is captured from constructed intermediate images. The dissimilarity of the features extracted from a tested stereoscopic image and that from its reference image is then mapped to its objective quality score by support vector machine (SVM).

A. Disparity Map Dissimilarity Measurement

It is the ability to perceive the depth of objects that allows HVS to generate stereo vision. Depth information in stereo image pairs reflects in the different pixel positions of same objects in left and right views. The position that an object appears in the right view is always to the left of that in the left view. This position shift, denoted as disparity, gives out the depth information, and is derived by a structural similarity based scheme [3]. The value of each element of the disparity map *D* of a stereo image pair represents the position shift of the pixel in corresponding position.

The disparity map serves for two purposes in the proposed method. Firstly, assessing the quality of a stereoscopic image is equivalent to quantifying how severe it sustains distortion, which is partly reflected in the destruction in disparity maps. Therefore, the dissimilarity between the disparity maps of the tested stereo image pair and its reference gives a quantitative representation of its quality, in a certain aspect. The dissimilarity is measured by correlation coefficient (CC),

$$CC(D_t, D_r) = \frac{\sum_i^{N_D}[D_t(i) - \overline{D_t}][D_r(i) - \overline{D_r}]}{\sqrt{\sum_i^{N_D}[D_t(i) - \overline{D_t}]^2 \sum_i^{N_D}[D_r(i) - \overline{D_r}]^2}} \quad (1)$$

where D_t and D_r are the disparity maps of the tested and reference stereoscopic images, respectively, $\overline{D_t}$ and $\overline{D_r}$ are the average element values of them, and N_D is the number of element in D_t or D_r.

On the other hand, the disparity map helps in further SIQA stages, where dissimilarity between more complex features between the tested and reference images is computed. We fuse the left and right view of a stereo image pair into one intermediate image to capture such features. In the primary visual cortex (V1) of HVS, a cyclopean image is formed from binocular stimulus, so that we see only one image in our minds, instead of two separate views, but how the process is completed in our minds remains a puzzle [9]. We employ a simple linear model to construct intermediate images to serve as cyclopean images,

$$\begin{aligned} I_C(x,y) &= W_L(x,y)I_L(x,y) \\ &\quad + W_R(x, y-D(x,y))I_R(x, y-D(x,y)) \end{aligned} \quad (2)$$

where I_L, I_R, and I_C denotes the left and right view and the constructed intermediate image, W_L and W_R defines the weights put on the pixels for the left and right view, and *D* is the disparity map, which in our case is calculated according to the left view. Coordinates *x* and *y* here are respectively vertical and horizontal, starting from the top left pixel.

Local energy is widely adopted to derive the weights, where solutions including mature wavelet and pyramid representation, e.g. steerable pyramid, Difference of Gaussian, Gabor wavelet, can be found in the literature [1, 3, 10]. Log Gabor wavelet is employed in the proposed method, as it is an alternative to Gabor representation which better models the receptive field profile of simple cells in V1 [11]. Since multi-scale is another important property of HVS, we compute the weights using log Gabor responses on 5 scales and therefore develops 5 intermediate images for each stereo image pair [2, 12]. On each scale, the log Gabor responses are normalized by

$$\begin{aligned} W_L(x,y) &= LG_L(x,y) / [LG_L(x,y) + LG_R(x, y-D(x,y))] \\ W_R(x, y-D(x,y)) &= 1 - W_L(x,y) \end{aligned} \quad (3)$$

where LG_L and LG_R are the log Gabor responses of left and right views, respectively.

B. Local Texture Dissimilarity Measurement

After binocular vision formed in V1, the visual signal is then sent to the secondary visual cortex (V2) [9]. On the front end of visual cortex and occupying two largest areas in visual cortex, V1 and V2 are crucial to the generation of early vision, and is highly responsible to human's sensation of visual quality [13]. It is discovered that higher hierarchy in visual cortex tends to focus on more complex and higher-level features [9, 11]. Concretely, V1 is only sensitive to the lowest-level features, e.g. edges, bars, and spatial frequency information, V2 focuses on these features as well to some extent, while is also interested in higher-level features such as more complex texture and shape information [11]. Therefore, binocular vision developed in V1 is far from adequate for SIQA.

Local tetra pattern is utilized for high-level texture analysis for its defining both "direction" and "magnitude" when describing the local texture information [14]. For an image pixel g_c, the first-order derivatives along 0º and 90º is determined by its relative value with the pixels next to it to the top, g_v, and to the right, g_h,

$$\begin{aligned} D_{0^\circ}(g_c) &= I(g_h) - I(g_c) \\ D_{90^\circ}(g_c) &= I(g_v) - I(g_c) \end{aligned} \quad (4)$$

and the direction of g_c is defined as

$$Dir(g_c) = \begin{cases} 1, D_{0^\circ}(g_c) \geq 0 \,\&\, D_{90^\circ}(g_c) \geq 0 \\ 2, D_{0^\circ}(g_c) < 0 \,\&\, D_{90^\circ}(g_c) \geq 0 \\ 3, D_{0^\circ}(g_c) < 0 \,\&\, D_{90^\circ}(g_c) < 0 \\ 4, D_{0^\circ}(g_c) \geq 0 \,\&\, D_{90^\circ}(g_c) < 0 \end{cases} \quad (5)$$

There are four direction candidates for each pixel, and four texture direction maps, TD_i (i = 1, 2, 3, 4), are used for describing each of it, in the form of binary images. Each element value in a direction map is denoted as

$$TD_i(g_c) = \begin{cases} 1, Dir(g_c) = i \\ 0, Dir(g_c) \neq i \end{cases} \quad (6)$$

Evidently, direction information alone is inadequate for texture analysis, thus a texture magnitude map, *TM*, is added as the fifth texture descriptor,

$$TM(g_c) = \sum_{i=1}^{8} 2^{(i-1)} \times \mathrm{sgn}(M(g_i) - M(g_c)) \quad (7)$$

where g_i are the eight neighboring pixels to g_c, $M(g)$ is the root sum square of $D_{0^\circ}(g)$ and $D_{90^\circ}(g)$, and sgn(x) is a sign function returning 1 when x is positive and 0 otherwise.

The dissimilarity between the extracted texture information of the tested intermediate image and that of its reference is measured using histogram intersection distance (HID) between their distribution histograms. On each scale, 4 indices describing the texture direction and 1 describing the magnitude is compute. HID is defined as

$$HID(T_t, T_r) = \sum_{i}^{B} \min[T_t(i), T_r(i)] \Big/ \sum_{i}^{B} T_r(i) \qquad (8)$$

where T_t and T_r denote the distributions of texture direction (or magnitude) maps of the intermediate image constructed from the tested stereo pair and its reference, respectively, and B is the number of possible pattern codes (2 for direction maps and 256 for magnitude maps).

Unfortunately, texture extraction by local pattern analysis may abandon some quality-aware information with substantial importance. Contrast information is largely lost during the binary coding process, for one instance. To make up for it, edge information, the lower-level texture, is introduced in our method. On one hand, edge information well reflects contrast in local regions of an image. On the other hand, the edge information is essential to both V1 and V2 as a low-level feature. Thus, edge information makes a good compensation for the local pattern description. In this Letter, Sobel operator is adopted for edge detection due to its efficiency and stability. Because the elements of edge detection results are vector carrying both magnitude and orientation information, we represent them using separated magnitude and orientation maps. Moreover, the process is operated on the multi-scale intermediate images as well. The dissimilarity of the edge magnitude (or orientation) maps between the tested and the reference intermediate images on each scale is measured by correlation coefficient, same as in (1),

$$CC(E_t, E_r) = \frac{\sum_{i}^{N_E}[E_t(i) - \overline{E_t}][E_r(i) - \overline{E_r}]}{\sqrt{\sum_{i}^{N_E}[E_t(i) - \overline{E_t}]^2 \sum_{i}^{N_E}[E_r(i) - \overline{E_r}]^2}} \qquad (9)$$

where E_t and E_r are the edge magnitude (or orientation) maps of the tested and reference stereoscopic images, respectively, $\overline{E_t}$ and $\overline{E_r}$ are the average element values of them, and N_E is the total element number in E_t or E_r.

C. Objective Score Mapping

With the quantified dissimilarity indices of different aspects calculated by (1), (8) and (9). The final step is to map these indices into an objective quality score, i.e. to build a regression function that outputs the objective quality score with the indices as inputs. Recently, machine learning technique is widely adopted as the regression strategy, including linear regression, K-nearest neighbor, SVM, neural networks, etc., in which SVM is employed in the proposed method due to its effectiveness and moderate computational complexity [15].

Before training, all the dissimilarity indices are normalized to the same scale and the regression targets are defined as the corresponding subjective quality scores. During the training process, the regression function is continuously optimized so that the outputs gradually approach the training targets. In experiments, the five-fold cross-validation scheme is adopted to show how well the proposed method performs. When a dataset is used for training, it is randomly divided into five subsets with equivalent sizes, four of them are employed for training and the rest one is used for testing. It should be noted that any result of our method presented in this Letter is the average value of 1,000 times of cross-validation because of the random division procedure.

TABLE I
OVERALL PERFORMANCE IN DATABASES LIVE 3D DATABASES PHASE I AND II

	LIVE phase I			LIVE phase II		
	PLCC	SRCC	RMSE	PLCC	SRCC	RMSE
Benoit	0.8899	0.8901	7.4786	0.7450	0.7280	7.4900
You	0.9303	0.9247	6.0161	0.7970	0.7840	6.7720
Bensalma	0.8874	0.8747	7.5585	0.7699	0.7513	7.2035
Shao (1)	0.9245	0.9217	6.2522	0.7585	0.7451	7.3554
Chen	0.9167	0.9157	6.5503	0.9010	0.8930	4.9870
Shao (2)	0.9350	0.9251	5.8155	0.8628	0.8494	5.7058
Lin	0.9366	0.9314	5.7440	0.9113	0.8935	4.6477
Geng	0.9430	0.9320	5.5140	0.9210	0.9190	5.4000
Proposed	**0.9599**	**0.9489**	**4.5631**	**0.9376**	**0.9312**	**3.8980**

TABLE II
RESULTS OF SIGNIFICANCE TEST AGAINST METHODS LISTED IN TABLE I

	LIVE phase I			LIVE phase II		
	PLCC	SRCC	RMSE	PLCC	SRCC	RMSE
Benoit	1	1	1	1	1	1
You	1	1	1	1	1	1
Bensalma	1	1	1	1	1	1
Shao (1)	1	1	1	1	1	1
Chen	1	1	1	1	1	1
Shao (2)	1	1	1	1	1	1
Lin	1	0	1	0	1	1
Geng	1	0	1	0	0	1

III. EXPERIMENTAL RESULTS AND DISCUSSIONS

Experiments are conducted on LIVE SIQA database, Phase I [16] and Phase II [3], both containing stereo images distorted by JPEG 2000 compression (JP2K), JPEG compression, white noise (WN), Gaussian blur (GB) and Rayleigh channel fast-fading (FF), and their corresponding human subjective ratings. There are only symmetrically distorted images in Phase I. Considering that a stereoscopic image pair may be distorted symmetrically or asymmetrically in the real world, Phase II consists of both symmetrically and asymmetrically distorted pairs. Three metrics, Pearson's linear correlation coefficient (PLCC), Spearman's rank-order correlation coefficient (SRCC), and root mean squared error (RMSE), are employed to measure the consistency of the objective quality scores predicted by an SIQA method and the subjective scores provided by the databases. The subjective scores also serve as the training targets of the regression function, as explained in Section II.

Existing successful SIQA methods, proposed by Benoit *et al.* [7], You *et al.* [8], Bensalma *et al.* [17], Shao *et al.* (1) [18], Chen *et al.* [3], Shao *et al.* (2) [4], Lin *et al.* [19], and Geng *et al.* [20], respectively, are referred for comparison. The overall performance comparison is presented in Table I. Clearly, the proposed method outperforms other methods on every metric. Since the performance varies during experiments, to further investigate how significant the superiority is, a T-test with significance level at 0.05 is conducted and the results are given in Table II, where 1 denotes a significant divergence and 0 otherwise. From Table II, the lead of the proposed method against the others is quite significant.

TABLE III
PERFORMANCE COMPARISON OF INDIVIDUAL DISTORTION TYPES ON LIVE 3D DATABASES PHASE I

		JP2K	JPEG	WN	GB	FF
PLCC	Benoit	0.8859	0.5766	0.9354	0.9217	0.7477
	You	0.9050	0.6333	0.9351	0.9545	0.8589
	Bensalma	0.8389	0.3803	0.9147	0.9369	0.7339
	Shao (1)	0.9238	0.6563	0.9410	0.9513	0.8403
	Chen	0.9164	0.6344	0.9436	0.9417	0.7580
	Shao (2)	0.9213	0.5200	0.9448	0.9592	0.8594
	Lin	0.9520	0.7546	0.9266	0.9583	0.8620
	Geng	0.9420	0.7190	**0.9630**	**0.9620**	0.8670
	Proposed	**0.9630**	**0.8017**	0.9592	0.9591	**0.8944**
SRCC	Benoit	0.8730	0.4983	0.9369	0.8802	0.6242
	You	0.9051	0.6008	0.9403	0.9300	0.8030
	Bensalma	0.8171	0.3283	0.9055	0.9157	0.6500
	Shao (1)	0.8752	0.6148	0.9431	0.9375	0.7814
	Chen	0.8956	0.5582	0.9481	0.9261	0.6879
	Shao (2)	0.8945	0.4951	0.9405	**0.9403**	0.7963
	Lin	0.9127	0.7164	0.9289	0.9332	0.8286
	Geng	0.9050	0.6530	**0.9560**	0.9310	0.8160
	Proposed	**0.9176**	**0.7647**	0.9353	0.9000	**0.8412**

TABLE IV
PERFORMANCE COMPARISON OF INDIVIDUAL DISTORTION TYPES ON LIVE 3D DATABASES PHASE II

		JP2K	JPEG	WN	GB	FF
PLCC	Benoit	0.6467	0.5328	0.8610	0.8814	0.8472
	You	0.7320	0.6741	0.5464	0.9763	0.8561
	Bensalma	0.6667	**0.8577**	0.9436	0.9077	0.9097
	Shao (1)	0.8377	0.7504	0.8496	0.8270	0.8808
	Chen	0.8426	0.8422	0.9602	0.9650	0.9097
	Shao (2)	0.7823	0.7472	0.9464	0.9580	0.9046
	Geng	0.8510	0.8350	0.9490	0.9790	**0.9480**
	Proposed	**0.9218**	0.8235	**0.9670**	**0.9904**	0.9373
SRCC	Benoit	0.6325	0.5078	0.8569	0.8545	0.8319
	You	0.7309	0.5229	0.4820	0.9227	0.8392
	Bensalma	0.8038	**0.8461**	0.9386	0.8838	0.8743
	Shao (1)	0.8477	0.7195	0.8455	0.8005	0.8509
	Chen	0.8334	0.8396	0.9554	0.9096	0.8890
	Shao (2)	0.7845	0.7330	**0.9651**	0.9204	0.8905
	Geng	0.8360	0.8410	0.9390	0.9200	0.9160
	Proposed	**0.8964**	0.7685	0.9358	**0.9500**	**0.9253**

Moreover, to evaluate the effectiveness and robustness of the methods in detail, the experiments are also operated on the individual distortion types in the two databases, as shown in Table III and IV. It can be observed that the proposed method shows relatively high and stable predicting accuracy, validating its robustness. On one hand, the proposed method performs well on most distortion types in both databases, sometimes with a large leading margin. On the other hand, although the proposed method fails to show superiority on some distortion types, the gap to the leading method is narrow, and it maintains a comparatively stable performance among all distortion types. In comparison, the listed previous methods are only capable to lead on one or two distortion types at most, and often shows obvious weakness on some specific types. Therefore, the proposed method shows high and stable consistency with human's subjective sensation upon stereoscopic image quality, among various distortion types and both databases.

IV. CONCLUSION

The fact that the proposed method exhibits high accuracy and robustness in experiments validates that combinational analysis on image depth and high- and low-level local texture is a promising solution to SIQA.

REFERENCES

[1] W. Zhou, G. Jiang, M. Yu, F. Shao, and Z. Peng, "PMFS: a perceptual modulated feature similarity metric for stereoscopic image quality assessment," *IEEE Signal Process. Lett.*, vol. 21, no. 8, pp. 1003-1006, Aug. 2014.

[2] 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.

[3] 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.

[4] F. Shao, K. Li, W. Lin, G. Jiang, and M. Yu, "Using binocular feature combination for blind quality assessment of stereoscopic images," *IEEE Signal Process. Lett.*, vol. 22, no. 10, pp. 1548-1551, Oct. 2015.

[5] C. Fernandez-Maloigne, J. Moreno, A. Rizzi, and C. Bonanomi, "QUALITAS: image quality assessment for stereoscopic images," J. Imag. Sci. Technol., vol. 60, no. 5, Sep. 2016.

[6] Y. Liu, J. Yang, Q. Meng, Z. Lv, Z. Song, and Z. Gao, "Stereoscopic image quality assessment method based on binocular combination saliency model," *Signal Process.*, vol. 125, pp. 237-248, Aug. 2016.

[7] A. Benoit, P. Le Callet, P. Campisi, and R. Cousseau, "Quality assessment of stereoscopic images," *EURASIP J. Image Video Process.*, vol. 2008, pp. 1-13, Jan. 2009.

[8] 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.

[9] N. Krüger, P. Janssen, S. Kalkan, M. Lappe, A. Leonardis, J. Piater, A. J. Rodriguez-Sanchez, and L. Wiskott, "Deep hierarchies in the primate visual cortex: what can we learn for computer vision?" *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1847-1871, Aug. 2013.

[10] S. K. Md, B. Appina, and S. S. Channappayya, "Full-reference stereo image quality assessment using natural stereo scene statistics," *IEEE Signal Process. Lett.*, vol. 22, no. 11, pp. 1985-1989, Nov. 2015.

[11] S. M. Lajevardi and H. Wu, "Facial expression recognition in perceptual color space," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3721-3733, Aug. 2012.

[12] S. K. Md and S. S. Channappayya, "Multiscale-SSIM index based stereoscopic image quality assessment," presented at the 22nd Nat. Conf. Comm., Guwahati, India, Mar. 4-6, 2016.

[13] A. Hyvärinen, J. Hurri, and P. O. Hoyer, Natural Image Statistics: A Probabilistic Approach to Early Computational Vision, Springer, New York, USA, 2009.

[14] S. Murala, R. P. Maheshwari, and R. Balasubramanian, "Local tetra pattern: a new feature descriptor for content-based image retrieval," *IEEE Trans. Image Process.*, vol. 21, no. 5, pp. 2874-2886, May 2012.

[15] C. Chang and C. Lin, "LIBSVM: a library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 1-27, Apr. 2011.

[16] 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.

[17] R. Bensalma and M. C. Larabi, "A perceptual metric for stereoscopic image quality assessment based on the binocular energy," *Multidimensional Syst. Signal Process.*, vol. 24, no. 2, pp. 281-316, Jun. 2013.

[18] 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.

[19] Y. Lin, J. Yang, W. Lu, Q. Meng, Z. Lv, and H. Song, "Quality index for stereoscopic images by jointly evaluating cyclopean amplitude and cyclopean phase," *IEEE J. Select. Topics Signal Process.*, vol. 11, no. 1, pp. 89-101, Feb. 2017.

[20] X. Geng, L. Shen, K. Li, and P. An, "A stereoscopic image quality assessment model based on independent component analysis and binocular fusion property," *Signal Process. Image Comm.*, vol. 52, pp. 54-63, Mar. 2017.