Image Quality Assessment with Degradation on Spatial Structure

Jinjian Wu, Weisi Lin*, *Senior Member, IEEE*, and Guangming Shi, *Senior Member, IEEE*

***Abstract*—In this letter, we introduce an improved structural degradation based image quality assessment (IQA) method. Most of the existing structural similarity based IQA metrics mainly consider the spatial contrast degradation but have not fully considered the changes on the spatial distribution of structures. Since the human visual system (HVS) is sensitive to degradations on both spatial contrast and spatial distribution, both factors need to be considered for IQA. In order to measure the structural degradation on spatial distribution, the local binary patterns (LBPs) are first employed to extract structural information. And then, the LBP shift between the reference and distorted images is computed, because noise distorts structural patterns. Finally, the spatial contrast degradation on each pair of LBP shifts is calculated for quality assessment. Experimental results on three large benchmark databases confirm that the proposed IQA method is highly consistent with the subjective perception.**

***Index Terms*—Image Quality Assessment, Structural Degradation, Spatial Distribution, Local Binary Patterns**

I. Introduction

We are living in a multimedia era, and objective image quality assessment (IQA) plays an important role in multimedia information processing, such as in signal transmission, restoration, and display [1]. In the last decade, a lot of IQA algorithms have been introduced for automatic and intelligent quality measurement. The peak signal-to-noise ratio (PSNR) metric is the most popular one, which directly measures the signal error. Though PSNR is well defined with a clear physical meaning, it does not consistent well with the human visual perception [2]. Therefore, researchers try to mimic the subjective perception of the human visual system (HVS) for quality assessment.

Among all of these HVS-oriented IQA algorithms, structural similarity based quality metric (SSIM) [3] is the most accepted one. Since the HVS is highly adapted to extract structural information for visual content understanding, the SSIM metric measure the structural degradation (e.g., similarity on luminance, variance, and covariance) to acquire image quality. Moreover, the SSIM metric is improved by further investigation on image structure [4]. In [5], both phase congruency and contrast are considered for structural degradation computation. In [6], as the HVS is highly sensitive to image edge, the structural information is represented by gradient, and image quality is measured based on gradient degradation. These metrics promote the performance of quality measurement.

Image structure appears as the relationships among pixels [7], which includes both spatial contrast and spatial distribution [8], and degradations on any of them will change the characteristic of image structure. However, the existing structural similarity based IQA metrics [3], [4], [5], [6] mainly consider the degradation on spatial contrast, the change on spatial distribution of structure is not fully considered. So we suggest to take the structural degradation on spatial distribution into account for quality assessment. How to describe the degradation on spatial distribution of structure is still an open problem. Since the local binary patterns (LBPs) [9] can effectively represent the spatial distributions of joint pixels, we adopt LBPs to describe the spatial structural information. Moreover, distortions will cause LBP shifts between the reference image and the distorted image, e.g., blur distortion can change an edge pattern into a flat pattern. Hence, the LBP shift is proposed to measure the degradation on spatial distribution of structure.

In this letter, we introduce a novel IQA method by considering structural degradation on both structural intensity and spatial distribution. First, the structural intensity degradation is computed as the contrast change, and the spatial distribution degradation is calculated as the LBP shift between the reference and distorted images. And then, for each pair of LBP shifts, all of the corresponding structural intensity degradation is cumulated. Finally, the support vector regression (SVR) procedure [10] is employed to pool the degradation on all pairs of LBP shifts and return the quality score. Experimental results on three large benchmark databases demonstrate that the proposed IQA method is highly consistent with the human perception.

The rest of this paper is organized as follows. In Section II, structural degradation is measured based on LBP shift and contrast change. Experimental results of the proposed IQA method are presented in Section III. Finally, conclusions are drawn in Section IV.

Jinjian Wu and Guangming Shi are with Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education of China, School of Electronic Engineering, Xidian University. E-mail: jinjian.wu@mail.xidian.edu.cn; gmshi@xidian.edu.cn.

Weisi Lin (corresponding author) is with the School of Computer Engineering, Nanyang Technological University, Nanyang 639798, Singapore. E-mail: wslin@ntu.edu.sg.

This work is supported by the Major State Basic Research Development Program of China (973 Program) (No.2013CB329402), NSF of China (No. 61227004, 61100155, 61372131, and 61372071).

II. Structural Degradation and Quality Measurement

Structural information convey the main visual contents of an image, and structural degradation will directly impact on image perception. Here, we measure image quality via its visual content degradations on both structural distribution and intensity.

First, structural distribution is analyzed with LBPs. The spatial distribution of pixels jointly provide the shape characteristic of image structure [9], [11]. According to the intensity differences between the central pixel x_c and its circularly symmetric neighbor pixels x_i, the classic uniform rotation invariant LBPs procedure is introduced [8], as

$$\mathrm{LBP}(x_i) = \begin{cases} \sum_{i=0}^{P-1} s(g_i - g_c) & \text{if } \mathcal{U}(\mathrm{LBP}_{P,R}) \leq 2 \\ P+1 & \text{else,} \end{cases} \tag{1}$$

$$\begin{aligned} \mathcal{U}(\mathrm{LBP}_{P,R}) = & |s(g_{P-1} - g_c) - s(g_0 - g_c)| \\ & + \sum_{i=1}^{P-1} |s(g_i - g_c) - s(g_{i-1} - g_c)|, \end{aligned} \tag{2}$$

$$s(g_i - g_c) = \begin{cases} 1, & g_i - g_c \geq 0 \\ 0, & g_i - g_c < 0, \end{cases} \tag{3}$$

where g_c and g_i represent the gray values of the central pixel x_c and its neighbor x_i, respectively; P represents the number of neighbors, and R represents the radius of the neighborhood. By considering the computational complexity and accuracy, P and R are usually set as P=8 and R=1, respectively, and there are 10 different types of uniform LBPs (see [8] for more detail).

With the degradation from distortion, the LBP value of a pixel can shift from one type to another ((as to be elaborated next)). Moreover, different kinds of distortion result in different LBP shifts. For example, blur noise mainly degrades the edge of an image, and as a result, an edge pattern may be shifted (distorted) into a flat pattern. In the contrary, JPEG compression mainly causes blockness artifact, and as a result, a flat pattern may be shifted (distorted) into an edge pattern. Therefore, we suggest to measure the degradation on structural distribution based on LBP shift. Since there are 10 different types of LBPs, there can be 10×10 different pairs of LBP shifts caused by distortions.

And then, we measure the degradations on structural intensity. The structural intensity of each pixel is usually computed as the edge height (strength), and the sobel operator is employed to calculate the edge height $\mathcal{E}$. With the structural intensity, its degradation between the reference and distorted images is computed as their similarity,

$$\mathrm{S}_e(x_i, y_i) = \frac{2 \times \mathcal{E}(x_i) \cdot \mathcal{E}(y_i) + C}{(\mathcal{E}(x_i))^2 + (\mathcal{E}(y_i))^2 + C}, \tag{4}$$

where x_i and y_i represent pixels in the reference and distorted images, respectively; and $\mathcal{E}(x_i)$ and $\mathcal{E}(y_i)$ are their corresponding edge heights; C is a small constant to avoid the denominator being zero and is set as $C = (0.05 \times L)^2$ [3], where L is the number of gray levels of the input image.

So far, the visual content degradations on both structural distribution and intensity are computed with LBP shift and edge height similarity, respectively. For each pair of LBP shifts, its corresponding structural intensity degradation can be calculated as the cumulation of all of the edge height similarities associated with current LBP shift. In order to highlight the pixels with serious structural degradations and suppress the pixels with slight structural degradations, the intensity degradation is normalized with their average value. For a distorted image $I^d \in R^{D_1 \times D_2}$ and its corresponding reference image $I^r \in R^{D_1 \times D_2}$, the structural degradation on each pair of LBP shifts (e.g., the m-th structural pattern (LBP^r_m) from the reference image shifts into the n-th structural pattern (LBP^d_n) from the distorted image) is calculated as

$$\mathrm{S}(\mathrm{LBP}^r_m, \mathrm{LBP}^d_n) = \frac{1}{D_1 \times D_2} \sum_{j=1}^{D_1 \times D_2} \mathcal{V}(x_j, y_j, m, n), \tag{5}$$

$$\mathcal{V}(x_j, y_j, m, n) = \begin{cases} \left(\mathrm{S}_e(x_j, y_j) - \bar{\mathrm{S}}_e\right)^2, & \text{if } \mathrm{LBP}(x_j){=}m \\ & \quad\ \mathrm{LBP}(y_j){=}n \\ 0, & \text{else} \end{cases} \tag{6}$$

where $\bar{\mathrm{S}}_e$ is the mean value of $\mathrm{S}_e(\cdot)$, x_j (y_j) represents the pixel in the reference (distorted) image, and k^d (k^r) represents the index of the LBP pattern.

With (5), the structural intensity degradations for the 10×10 pairs of LBP shifts are acquired. Since each pair of LBP shifts represents individual degradation of structural distribution, we cannot simply pool them with an equal weight for quality assessment. In this letter, the support vector regression (SVR) is employed to learn a feature pooling method. SVR is proved to be an effective procedure for high dimensional data pooling, and is widely used in machine learning [12]. We adopt the LibSVM package [10] to implement feature pooling,

$$\mathcal{Q}_s(I^r, I^d) = \mathrm{SVR}\left(\mathcal{S}, \mathcal{MOD}\right), \tag{7}$$

where $\mathcal{Q}_s$ is the quality score, $\mathcal{S}$ is the structural feature set defined as $\mathcal{S} = \{\mathrm{S}(\mathrm{LBP}^r_m, \mathrm{LBP}^d_n) | m = 1, \cdots 10; n = 1, \cdots 10\}$, and $\mathcal{MOD}$ is a trained model for regression.

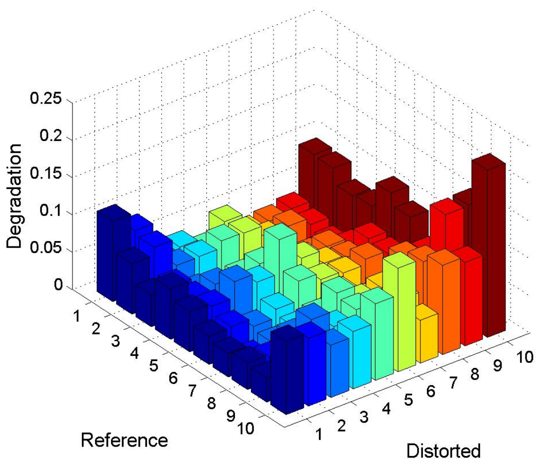
III. Experimental Results

In this section, the structural degradations on both structural distribution and intensity are first demonstrated. And then, we compare the proposed IQA method with the state-of-the-art IQA metrics in three large benchmark databases to verify the effectiveness of the proposed method.

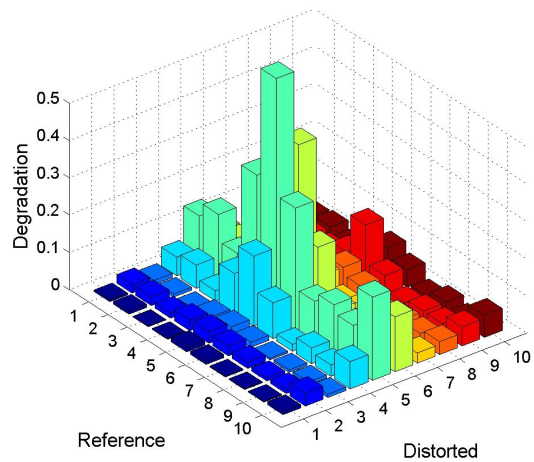
Noise will degrade image structures on both structural intensity and distribution. As shown in Fig. 1 (a) and (b), the hats image is distorted by white noise and JPEG2000 compression, respectively. Though the energy of noise in the two distorted images is almost same (Fig. 1 (a) is with MSE=127 and Fig. 1 (b) is with MSE=126), their subjective quality is quite different (The MOS (mean opinion score) of subjective viewing for Fig. 1 (a) is 5.27 and that for Fig. 1 (b) is 2.87; the higher of MOS, the better perceived quality). By analyzing the degradation on structural information, we have found that the changes on the structural intensity (represented by edge height) of the two distorted images are much similar, where the average edge height difference is $\Delta\mathcal{E}$=5.63 for Fig. 1 (a) and $\Delta\mathcal{E}$=5.78 for Fig. 1 (b). Therefore, degradation on structural intensity cannot accurately represent the quality degradations of the two contaminated images. With further analysis we have found their degradations on structural distribution are

(a)

(b)



(c)



(d)

Fig. 1: Structural degradations on both intensity and distribution. (a) White noise distorted image with MSE=127, MOS $= 5.27$, and the average edge height difference $\Delta\mathcal{E}$=5.63. (b) JPEG2000 compression distorted image with MSE=126, MOS $= 2.87$, and $\Delta\mathcal{E}$=5.78. (c) and (d) are their sketch maps of structural degradations, where each bar means the structural intensity degradation on the corresponding pair of LBP shift.

quite different. White noise mainly adds random disturbances into the image, which results in random changes on structural distribution (i.e., random LBP shifts). As shown in Fig. 1 (c), there exists degradation on each pair of LBP shifts, while no one is dominant. However, the degradations caused by JPEG2000 compression is quite different from white noise, which mainly brings ringing artifact in the edge regions, as shown in Fig. 1 (b). As a result, the structural degradations are mainly concentrated on several pairs of LBP shifts, as shown in Fig. 1 (d). Under different LBP shifts, the two contaminated images present different quality degradations. Therefore, we should consider structural degradations on both intensity and distribution for quality assessment, and the proposed LBP shift can effectively represent the distortions on structural distribution.

In order to make a comprehensive analysis on the performance, we compare the proposed method with 8 state-of-the-art IQA metrics, namely, IGM [4], FSIM [5], ADM [13], GSIM [6], MAD [14], VIF [15], MSSIM [3], and PSNR. Three performance criteria are adopted, which are Pearson linear correlation (CC), the Spearman rank-order correlation coefficient (SRCC), and the root mean squared error (RMSE) [13]. Meanwhile, the performance of these IQA metrics are verified with three large benchmark databases: TID [16], which is composed with 25 reference images and 1700 distorted images across 17 types of noise; CSIQ [17], which is composed with 30 reference images and 866 distorted images across 6 types of noise; and LIVE [18], which is composed with 29 reference images and 799 distorted images across 5 types of noise.

Since the proposed IQA method employs SVR procedure for quality assessment, we randomly select 80% reference images and their corresponding distorted images for training SVR model and the rest for testing. In order to eliminate the performance bias, the 80%-20% training-testing SVR procedure is repeated for 100 times, and the average performance is calculated for the final result. The performance of the proposed

TABLE I: PERFORMANCE COMPARISON OF IQA METRICS ON 3 LARGE BENCHMARK DATABASES

DB	Algo. / Crit.	Proposed	IGM	FSIM	ADM	GSIM	MAD	VIF	MSSIM	PSNR
TID (1700)	CC	**0.914**	**0.886**	0.874	0.869	0.846	0.831	0.809	0.843	0.531
	SRCC	**0.908**	**0.890**	0.881	0.862	0.855	0.834	0.750	0.853	0.525
	RMSE	**0.543**	**0.623**	0.653	0.662	0.715	0.747	0.789	0.730	1.137
CSIQ (866)	CC	**0.960**	0.928	0.912	0.928	0.898	**0.950**	0.928	0.900	0.800
	SRCC	**0.965**	0.940	0.924	0.933	0.917	**0.947**	0.919	0.914	0.806
	RMSE	**0.073**	0.098	0.108	0.098	0.116	**0.082**	0.098	0.115	0.158
LIVE (799)	CC	**0.961**	0.958	0.960	0.936	0.944	**0.967**	0.960	0.943	0.872
	SRCC	**0.964**	0.958	0.963	0.954	0.955	**0.967**	0.963	0.945	0.876
	RMSE	**7.457**	7.925	7.678	9.627	9.038	**6.924**	7.673	9.096	13.37
Average	CC	**0.937**	**0.914**	0.905	0.900	0.883	0.894	0.876	0.881	0.681
	SRCC	**0.936**	**0.919**	0.912	0.902	0.895	0.895	0.844	0.891	0.681

IQA method and the 8 state-of-the-art IQA metrics on the three databases (TID, CSIQ, and LIVE) are listed in Table I. As can be seen, the proposed method has larger CC and SRCC values and smaller RMSE values than the other state-of-the-art IQA metrics on both TID and CSIQ databases, which means the proposed method outperforms the existing metrics on the two databases. Meanwhile, the proposed method performs almost the same to the best metric (i.e., MAD) on the LIVE database. Furthermore, the weighted average values on CC and SRCC values (no RMSE average value since its range in the three databases are quite different) of the proposed metric are also larger than the existing metrics, which further confirm that the proposed IQA method outperforms the-state-of-the-art metrics.

IV. Conclusion

In this letter, we have introduced an improved structural degradation based IQA method. The HVS is highly sensitive to distortion on image structure, and structural similarity is widely used for IQA. However, the existing structural similarity based IQA methods mainly consider structural degradation on spatial contrast, without proper addressing the degradation on spatial distribution. We have suggested to take both factors into account for IQA. To this end, we first analyzed the spatial distribution of image structure with LBPs. Since noise distorts structural patterns, the degradation on structural distribution is measured with LBP shift between the reference and distorted images. Next, the structural intensity degradation on each pair of LBP shifts was calculated. Finally, the degradations on all pairs of LBP shifts were pooled with SVR procedure toward the quality score. Experimental results have demonstrated that the proposed IQA method is highly consistent with the HVS perception.

References

[1] W. Lin and C.-C. J. Kuo, “Perceptual visual quality metrics: A survey,” *J. Visual Communication and Image Representation*, vol. 22, no. 4, pp. 297–312, 2011.

[2] Z. Wang and A. C. Bovik, “Mean squared error: Love it or leave it?” *IEEE Signal Processing Magazine*, vol. 26, no. 1, pp. 98–117, Jan. 2009.

[3] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” *Image Processing, IEEE Transactions on*, vol. 13, no. 4, pp. 600–612, 2004.

[4] J. Wu, W. Lin, G. Shi, and A. Liu, “Perceptual quality metric with internal generative mechanism,” *IEEE Transactions on Image Processing*, vol. 22, no. 1, pp. 43 –54, Jan. 2013.

[5] L. Zhang, L. Zhang, X. Mou, and D. Zhang, “FSIM: a feature similarity index for image quality assessment,” *Image Processing, IEEE Transactions on*, vol. 20, no. 8, pp. 2378–2386, Aug. 2011.

[6] A. Liu, W. Lin, and M. Narwaria, “Image quality assessment based on gradient similarity,” *Image Processing, IEEE Transactions on*, vol. 21, no. 4, pp. 1500–1512, Apr. 2012.

[7] J. Wu, F. Qi, and G. Shi, “Self-similarity based structural regularity for just noticeable difference estimation,” *Journal of Visual Communication and Image Representation*, vol. 23, no. 6, pp. 845–852, 2012.

[8] T. Ojala, M. Pietikinen, and T. Menp, “Multiresolution gray-scale and rotation invariant texture classification with local binary patterns,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, Jul. 2002.

[9] T. Ojala, M. Pietikinen, and D. Harwood, “Performance evaluation of texture measures with classification based on kullback discrimination of distributions,” in *Proceedings of the 12th IAPR International Conference on Pattern Recognition*, vol. 1, pp. 582–585.

[10] C. C. Chang and C. J. Lin. (2001) Libsvm: a library for support vector machines. [Online]. Available: http://www.csie.ntu.edu.tw/ cjlin/libsvm/.

[11] T. Ojala, M. Pietikinen, and D. Harwood, “A comparative study of texture measures with classification based on featured distributions,” *Pattern Recognition*, vol. 29, no. 1, pp. 51–59, 1996.

[12] C. J. C. Burges, “A tutorial on support vector machines for pattern recognition,” *Data mining and knowledge discovery*, vol. 2, no. 2, pp. 121–167, 1998.

[13] S. Li, F. Zhang, L. Ma, and K. N. Ngan, “Image quality assessment by separately evaluating detail losses and additive impairments,” *Multimedia, IEEE Transactions on*, vol. 13, no. 5, pp. 935–949, Oct. 2011.

[14] 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, p. 011006, 2010.

[15] H. R. Sheikh and A. C. Bovik, “Image information and visual quality,” *Image Processing, IEEE Transactions on*, vol. 15, no. 2, pp. 430–444, Feb. 2006.

[16] N. Ponomarenko and K. Egiazarian. (2009) Tampere image database 2008 tid2008. [Online]. Available: http://www.ponomarenko.info/tid2008.htm.

[17] E. C. Larson and D. M. Chandler. (2004) Categorical image quality (csiq) database. [Online]. Available: http://vision.okstate.edu/csiq.

[18] H. R. Sheikh, K. Seshadrinathan, A. K. Moorthy, Z. Wang, A. C. Bovik, and L. K. Cormack. (2006) Image and video quality assessment research at live. [Online]. Available: http://live.ece.utexas.edu/research/quality/.