Reduced-Reference Image Quality Assessment with Local Binary Structural Pattern

Jinjian Wu*, Weisi Lin†, Guangming Shi*, and Long Xu†
*School of Electronic Engineering, Xidian University, Xi'an, Shannxi, China
†School of Computer Engineering, Nanyang Technological University, Singapore
Email: jinjian.wu@mail.xidian.edu.cn, WSLin@ntu.edu.sg, gmshi@xidian.edu.cn

***Abstract*—Reduced-reference (RR) image quality assessment (IQA) aims to use less reference data and achieve higher quality prediction accuracy. Recent researches confirm that the human visual system (HVS) is adapted to extract structural information and is sensitive to structure degradation. Therefore, in this paper, we try to represent image contents with several structural patterns, and measure image quality according to the structural degradation on these patterns. The classic local binary patterns (LBPs) are firstly employed to extract image structures and create LBP based structural histogram. And then, the structural degradation is computed as the histogram distance between the reference and distorted images. Experimental results on three large databases demonstrate that the proposed RR IQA method greatly improved the quality prediction accuracy.**

***Index Terms*—Reduced-Reference, Image Quality Assessment, Visual Structural Degradation, Local Binary Pattern**

I. INTRODUCTION

Objective image/video quality assessment (IQA) plays an important role in signal processing, such as in image/video transmission, compression, restoration and display [1]. During the last decade, a large number of IQA methods have been proposed to predict image quality. Most of them are full-reference (FR) methods which need the whole data of the reference image. However, the reference image is always not available, and a no-reference (NR) IQA metric is expected in this condition. Due to the varied image contents and the individual distortion types, NR quality prediction is extremely difficult when no prior knowledge is available [2]

As a compromise between FR and NR, reduced-reference (RR) IQA metrics are designed, which use partial information of the reference image for quality prediction. RR IQA metrics expect to use less data of the reference image and achieve higher prediction accuracy [3]. Therefore, some global features, within which the quality degradation can be effectively represented, are extracted in RR IQA metrics. According to the assumption that most real-world distortions disturb the stable statistical properties of natural images [4], Wang et al. [2] counted the change on marginal distribution of wavelet coefficients of an image for its quality prediction. Analogously, Gao et al. [5] suggested to measure the quality based on the statistical correlations of wavelet coefficients in different subbands. In the recent, Soundararajan and Bovik [6] suggested to measure the quality degradation according to the scaled entropies of wavelet coefficients. Wu et al. [3] measured the energy change of the visual contents caused by distortion for RR IQA. However, these RR IQA metrics either perform poorly with little data of the reference image, or need a large amount of reference data to achieve good performance.

Recent science findings on visual perception indicate that the human visual system (HVS) possesses an internal generative mechanism (IGM) [7], within which the primary visual information is actively predicted and the residual uncertainty is ignored [8]. Moreover, the HVS is highly adapted to extract structural information for image perception and understanding. Therefore, we try to separatively measure the structural degradation on the primary visual information and the residual uncertainty of an input image for quality prediction. According to the IGM theory, an input scene is decomposed into two portions (i.e., the predicted portion which mainly composes of the primary visual information, and the residual portion which includes the disorderly uncertainty) for separatively processing. And then, the famous and well accepted structural descriptor, local binary pattern (LBP) [9], is employed to extract structural information of the two portions. And the LBP histograms are created to represent the structural information of the image. Finally, the structural degradation is computed as the distance of the structural histograms between the distorted and reference images, and the support vector regression (SVR) procedure is adopted to pool the features and return the quality score of the distorted image. Experimental results on three large image databases demonstrate that the proposed RR IQA method outperforms the state-of-the-art RR IQA methods.

The rest of this paper is organized as follows. In Section II, image structures are analyzed and extracted with LBP for quality assessment. Experimental results of the proposed RR IQA method are presented in Section III. Finally, conclusions are drawn in Section IV.

II. QUALITY MEASUREMENT WITH VISUAL STRUCTURAL DEGRADATION

In this section, the visual structural information of an image is analyzed and extracted with the classical LBP procedure for image quality assessment. Firstly, according to the IGM theory, an image is decomposed into predicted portion and

This work was supported the Research Fund for the Doctoral Program of Higher Education (No. 20130203130001), NSFC(Nos. 61372131, 61100155, 11204014, 61301288, 61372071).

residual portion for separatively processing. And then, the LBP procedure is briefly introduced for visual structural extraction. Finally, the SVR model is employed for feature pooling and quality prediction.

A. IGM based Image Decomposition

Recent research on the brain science indicated that the HVS possesses an IGM for visual perception [8]. Within the IGM, the primary visual content of the input retinal stimuli is actively predicted according to the inherent priori knowledge, and the remaining uncertainty is ignored for further processing [7]. In summary, an image I should be decomposed into two portions, the predicted portion I_p and the residual portion I_r, for separatively processing. Moreover, a Bayesian brain theory is introduced to mimic the active prediction [8]. The core of the Bayesian brain theory is Bayesian probabilistic prediction, which optimizes an input image by minimizing its prediction error. For example, the value of a pixel x can be predicted with its surrounding $\mathcal{X}$. According to the correlation between the central pixel x and its neighbors $x_i \in \mathcal{X}$, the conditional probability $p(x/\mathcal{X})$ is maximizd for error minimization. By taking the mutual information $I(x;x_i)$ as the autoregressive coefficient, an autoregressive model is created to mimic the active prediction in IGM of the HVS [10],

$$g' = \sum_{g_i \in \mathcal{X}} \mathcal{C}_i g_i + \varepsilon, \tag{1}$$

where g' is the predicted value of pixel x, g_i is the value of the neighbor x_i, $\mathcal{C}_i = I(x;x_i)/\sum_k I(x;x_k)$ is the normalized coefficient, and ε is white noise. With (1), the primary visual content of an input scene is actively predicted and the predicted portion I_p is acquired. And the residual portion I_r (i.e., prediction error) of the original image (I) can be computed as $I_r = I - I_p$.

Since the two portions (I_p and I_r) contain different visual information and play different roles for image perception, distortions on the two portions will result in different quality degradations. We will discriminatively analyze the information fidelities on the two portions for quality prediction in the next two subsections.

B. LBP based Structure Extraction

Since image structures convey the main visual information of a scene for perception and understanding [10], [11], we extract image structures for quality prediction. In the past, the first-order statistics of local property values, such as the variance and covariance, are usually adopted to simply analyze the character of image structure. These statistics values can not characterize the spatial distribution of structure [9]. To this end, the classic LBP procedure is employed for structure extraction in this work. the LBP value of a pixel x_c is computed as the difference with its circularly symmetric neighborhood [9],

$$\mathrm{LBP}_{P,R} = \sum_{i=0}^{P-1} s(g_i - g_c) 2^i, \tag{2}$$

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

where g_c (g_i) is the gray value of the central pixel x_c (the neighbor x_i), P is the number of neighbors, and R is the radius of the neighborhood. By considering the computational complexity and accuracy, we set P=8 and R=1 in our experiments.

Further researches on LBP verified that the uniform LBP patterns provide the vast majority structural information (sometimes over 90%) [9]. According to the uniform patterns, a locally rotation invariant pattern can be defined as:

$$\mathrm{LBP}^{riu2}_{P,R} = \begin{cases} \sum_{i=0}^{P-1} s(g_i - g_c) & \text{if } \mathcal{U}(\mathrm{LBP}_{P,R}) \leq 2 \\ P+1 & \text{else}, \end{cases} \tag{4}$$

where

$$\begin{aligned} \mathcal{U}(\mathrm{LBP}_{P,R}) = & \ |s(g_{P-1} - g_c) - s(g_0 - g_c)| \\ & + \sum_{i=1}^{P-1} |s(g_i - g_c) - s(g_{i-1} - g_c)|. \end{aligned} \tag{5}$$

With (4), the $\mathrm{LBP}^{riu2}_{P,R}$ value of each pixel from image I can be calculated. And then, the LBP based structural historgam is mapped into $P+2$ bins ($P+1$ bins for all of the uniform LBP patterns, and 1 bin for the other patterns) to represent the structure of image I.

C. Quality Prediction

Distortions always damage the structures of images, and we try to measure image quality based on visual structural degradation. For a test image I^d (and its reference image I^o), it is firstly decomposed with (1) to return the predicted portion I^d_p (I^o_p) and the residual portion I^d_r (I^o_r). And then, with the LBP procedure (4), their structural histograms are acquired, which are expressed as H^d_p and H^r_d (H^o_p and H^r_o for the reference image), respectively.

In order to measure the structural degradation, the similarity formulation [11] is adopted to compute the changes on each bins of the histogram,

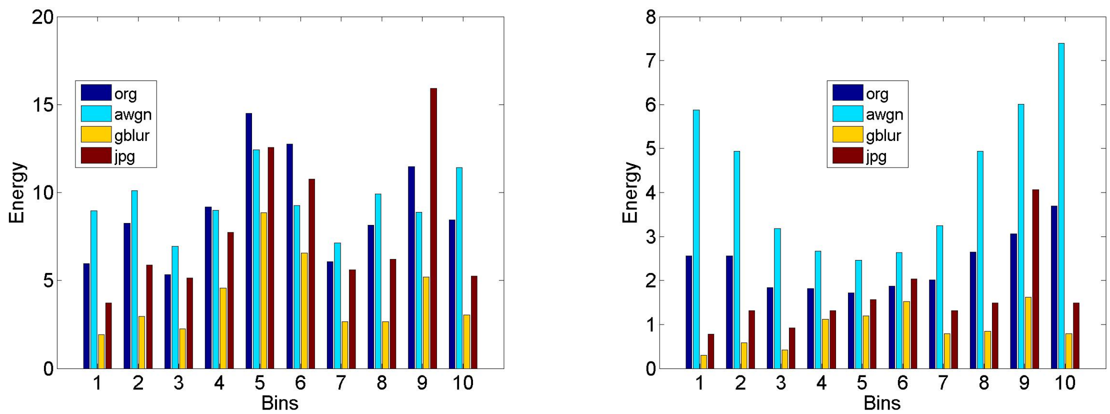
$$\mathrm{HC}(I^d_i, I^o_i) = \frac{2 \times H^d_i \cdot H^o_i}{(H^d_i)^2 + (H^o_i)^2}, \tag{6}$$

where $i \in \{p, r\}$.

For $\mathrm{LBP}^{riu2}_{8,1}$, it contains 10 uniform LBP patterns. And therefore, with (6), there are 20 features from the predicted and residual portions. How to pool features for quality assessment is still an open problem. For simplicity, all features are considered to have the same importance [11], and are equally accumulated to acquire the final score. However, it is obvious that the 20 LBP features in this paper are not equally important: 1) the visual contents in the predicted and residual portions are quite different, and they play different roles for image perception; 2) each LBP pattern represents different spatial structure, and its change leads to different quality degradation. Therefore, we employ a support vector machine regressor (SVR) to learn a feature pooling method

(a) MOS=4.26 and MSE=256 (b) MOS=3.40 and MSE=294 (c) MOS=3.25 and MSE=230

Fig. 1: Example of image quality degradations under different types of distortions. (a) AWGN. (b) GBLUR. (c) JPG.



(a) (b)

Fig. 2: LBP based structural change under different types of distortions (i.e., AWGN, GBLUR, and JPG), and org represent the histogram of the reference image. (a) and (b) The LBP histograms of I_p and I_r.

for quality measurement. SVR can effectively handling high dimensional data [12], which has been widely used in feature pooling. In this paper, we adopt the LIBSVM package [13] to implement the SVR for quality assessment with a radial basis function (RBF) kernel.

III. Experimental Results

In this section, the LBP based histogram is firstly analyzed to demonstrate its effectiveness of the proposed method. And then, we compared the proposed method with three state-of-the-art RR IQA metrics in three publicly databases (i.e., LIVE [14], TID [15], CSIQ [16]).

The proposed RR IQA algorithm is based on visual structural degradation, which measures the structural changes on the LBP histograms. For example, the lighthouse distorted with three types of distortions (additive white Gaussian noise (AWGN), Gaussian blur (GBLUR), JPEG (JPG) compression noise) is shown in Fig. 1, and their corresponding LBP histograms are shown in Fig. 2. As can be seen in Fig. 2, different types of distortions result in quite different structure changes. As shown in Fig. 1 (a), the AWGN noise brings in a lot of random disturbs, which mainly effect on the residual portion. As a result, the energy on each bin of the LBP histogram is greatly increased, as the light blue bar shown in Fig. 2 (b). On the contrary, the GBLUR noise erases many visual contents, as shown in Fig. 1 (b). This type of distortion will decrease the structural information, as the yellow bars shown in Fig. 2, the energies on all of these bins are obviously decreased with a similar ratio. The JPG noise degrades the structural information of the image, such as the blockness artifact as shown in Fig. 1 (c), which decrease the energies on most structural bins. Meanwhile, the blockness brings in new edges between patches, which results in remarkable increasing at the 9th bin of the histograms. In summary, different types of distortions cause different damages on visual structures, and the LBP based structural histogram can effectively represent the visual information degradations.

And then, we compare the proposed method with three existing RR IQA metrics (i.e., RRVIF [3], WNISM [2] and RRED [6]) to demonstrate the performance on three large databases. Since the proposed method adpots SVR for feature pooling, which requires a training procedure, we need to divide each database into train and test subsets. In our experiment, 80% of the reference images and their corresponding distorted

TABLE I: PERFORMANCE OF IQA INDICES ON LIVE DATABASE.

DB	Algo. / Crit.	Proposed	RRVIF	WNISM	RRED
No. of scalars		20	2	18	20
AWGN	CC	**0.990**	0.957	0.890	0.938
	SRCC	**0.982**	0.946	0.870	0.950
	RMSE	**2.25**	4.66	7.29	9.71
GBLUR	CC	**0.966**	0.955	0.888	0.956
	SRCC	**0.960**	0.961	0.915	0.951
	RMSE	**3.98**	4.66	7.22	5.43
JPEG	CC	0.894	0.895	0.876	**0.962**
	SRCC	0.858	0.885	0.851	**0.956**
	RMSE	6.94	7.15	7.71	**4.7**
J2K	CC	0.927	0.932	0.924	**0.956**
	SRCC	0.913	0.950	0.920	**0.951**
	RMSE	5.96	5.88	6.18	**3.28**
FF	CC	**0.950**	0.944	0.925	0.892
	SRCC	**0.926**	0.941	0.923	0.920
	RMSE	**4.86**	5.42	6.25	12.86
Overall	CC	**0.935**	0.725	0.710	0.831
	SRCC	**0.932**	0.732	0.703	0.834
	RMSE	**9.59**	17.6	18.4	15.2

TABLE II: OVERALL PERFORMANCE OF IQA INDICES ON CSIQ and TID DATABASES.

DB	Algo. / Crit.	Proposed	RRVIF	WNISM	RRED
CSIQ	CC	**0.873**	0.698	0.696	0.780
	SRCC	**0.872**	0.733	0.705	0.780
	RMSE	**0.126**	0.182	0.189	0.164
TID	CC	**0.814**	0.535	0.572	0.725
	SRCC	**0.819**	0.500	0.495	0.709
	RMSE	**0.773**	1.134	1.101	0.924

images are randomly chosen for training, and the rest for testing. Moreover, to eliminate performance bias, we repeat this random train-test procedure 100 times and calculate the average performance for the final result.

The performce on LIVE database, which consists of five types of distortions (additive white Gaussian noise (AWGN), Gaussian blur (GBLUR), JPEG (JPG) compression, JPEG2000 (J2K) compression, and a Rayleigh fast-fading channel simulation (FF) noise), is listed in Table I. As can be seen, the proposed metric uses similar quantity of the reference data as WNISM and RRED, while much more than RRVIF. Meanwhile, the proposed method use LBP based strcuctural histogram, which can effectively represent quality degradation caused by different types of distortions. And therefore, the performance of the proposed method is greatly improved: it performs better than the other three metrics on three types of distortions (i.e., AWGN, GBLUR, and FF), has similar performance with the best metric on the other two types of distortions, and the overall performance of the proposed metric is much better than the other metrics.

In order to further demonstrate the effectiveness of the proposed metric, performance on the other two large databases (i.e., CSIQ and TID) is given in Table II. As can be seen, on both databases, the proposed metric returns much larger CC and SRCC values than that of the other three metrics (i.e., RRVIF, WNISM, and RRED), and also has smaller RMSE values than the other RR IQA metrics. This further confirms that the proposed RR IQA metric outperforms the state-of-the-art RR IQA metrics.

IV. Conclusion

In this paper, we have introduced a novel RR IQA metric according to the visual structural degradation. The RR IQA metrics aim to use less reference data and achieve higher prediction accuracy. Since the HVS is highly adapted to extract structural information and is sensitive to structural degradation, we suggested to represent the image structure with several uniform LBP patterns, and measure image quality based on the structural changes on these patterns. Experimental results on different types of distortions demonstrate that the proposed LBP based structural histogram can effectively represent the structure degradation. Moreover, performance on three publicly available databases demonstrate that the proposed RR IQA metric performs much better than the existing RR IQA metrics, and is highly consistent with the subjective 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 E. P. Simoncelli, “Reduced-reference image quality assessment using a wavelet-domain natural image statistic model,” in *SPIE*, vol. 5666, 2005, pp. 149–159.

[3] J. Wu, W. Lin, G. Shi, and A. Liu, “Reduced-reference image quality assessment with visual information fidelity,” *IEEE Transactions on Multimedia*, vol. 15, no. 7, pp. 1700–1705, Nov. 2013.

[4] H. Sheikh, A. Bovik, and L. Cormack, “No-reference quality assessment using natural scene statistics: JPEG2000,” *IEEE Transactions on Image Processing*, vol. 14, no. 11, pp. 1918 –1927, Nov. 2005.

[5] X. Gao, W. Lu, D. Tao, and X. Li, “Image quality assessment based on multiscale geometric analysis,” *IEEE Transactions on Image Processing*, vol. 18, no. 7, pp. 1409–1423, Jul. 2009.

[6] R. Soundararajan and A. Bovik, “RRED indices: Reduced reference entropic differencing for image quality assessment,” *IEEE Transactions on Image Processing*, vol. 21, no. 2, pp. 517 –526, Feb. 2012.

[7] K. Friston, “The free-energy principle: a unified brain theory?” *Nat Rev Neurosci*, vol. 11, no. 2, pp. 127–138, Feb. 2010.

[8] D. C. Knill and R. Pouget, “The bayesian brain: the role of uncertainty in neural coding and computation,” *Trends In Neuroscience*, vol. 27, pp. 712–719, 2004.

[9] T. Ojala, M. Pietikainen, and T. Maenpaa, “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.

[10] 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.

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

[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] 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/.

[14] 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/.

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

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