Reduced Reference Image Quality Assessment Using Entropy of Primitives

Shiqi Wang, Xiang Zhang, Siwei Ma, Wen Gao
School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China

***Abstract*—In this paper, we propose a new reduced reference image quality assessment algorithm based on the recent advances in sparse coding and representation, particularly, the entropy of primitives (EoP). The EoP is defined in terms of the distribution of the primitives, which form an overcomplete dictionary to represent the natural scene by linear combination. Constructively, we develop a reduced reference EoP based distortion metric (EoPM). EoPM has the property that it is nearly invariant to the geometry distortions, which hardly affect the visual quality but are often wrongly predicted by the existing image quality assessment metrics with severe distortion. Experimental results show that the accuracy of EoPM is highly competitive to the popular reduced reference image quality assessment algorithm on the public dataset.**

***Index Terms*—Reduced reference, image quality assessment, sparse coding, entropy of primitives.**

I. Introduction

Nowadays, developing accurate and efficient objective image quality assessment (IQA) algorithms are crucial as they can be used to evaluate, control, and optimize the visual quality of the multimedia systems. Most of the existing IQA algorithms are full reference (FR), indicating that they are based on the full access of the original image, such as the structural similarity (SSIM) [1] and visual signal-to-noise ratio [2]. However, in many applications the reference images are not available, and thus creating IQA algorithms that depend on much less information from the reference image is another focus of research. These algorithms can be categorized into reduced reference (RR) and no reference (NR) algorithms. The NR algorithms are usually designed with the prior knowledge of the distortion process [3][4]. However, they are less efficient in providing a high correlation with the subjective quality evaluations because of the absence of the reference image information. Fortunately, the RR IQA can achieve a good tradeoff between the FR and NR algorithms, as they can predict the image quality in terms of a few extracted features, which require significantly less data than transmitting the original image.

The General framework of an RR IQA system is shown in Fig. 1 [5]. Taking the instance of a typical visual communication system, the selected features are extracted and transmitted to the receiver side. The data rate of the side information in transmitting the features is typically much lower than that of the visual information channel. At the receiver, the difference between the features that are extracted from the original and distorted images is employed in evaluating the image quality. Generally, the selected features should summary the whole image and exhibit strong correlations with image distortions, such that the perceptual quality of natural images can be reflected in the feature comparison process.

Recently, many RR IQA algorithms have been proposed. When the prior knowledge of image distortion type is known, the RR features can be defined to quantity specific artifacts, such as blurring, blocking and ringing [6][7]. However, because of the limitations of the prior image distortion type, this approach cannot be extended to general purposes. Another attempt is evaluating the image quality based on the modeling of natural image statistics [8]. The intrinsic principle behind this approach is that image distortions may make the image unnatural. Therefore, the performance of this approach is not very much affected by the distortion process, which is practical for real-world applications.

Recently, the Sparseland model has been proposed to describe signals based on the sparsity and redundancy of signal representations [9]. The basic idea is to perform redundant transform of the input signal with the constraint of sparest representation. Sparse representation has been shown to be efficient in many image processing applications, such as denosing [10] and super-resolution [11]. In this paper, inspired by the concept of sparse representation, we propose a RR IQA algorithm based on the entropy of primitives (EoP). The proposed EoP based metric (EoPM) has been shown to effectively quantify various types of distortion.

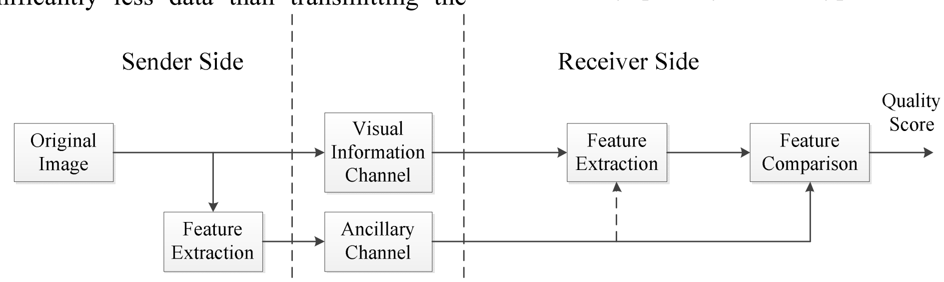

Fig. 1. General framework of an RR IQA system [5].

The rest of this paper is organized as follows. In section II we introduce the image primitive coding and the concept of EoP. In section III, we propose the RR IQA algorithm. Section IV shows the performance of the proposed RR IQA algorithm. Finally, we conclude this paper in section V.

II. Entropy of Primitive

A. Image Primitive Coding

In the Sparse-Land model [9], it is assumed that the natural signals can be decomposed over a redundant dictionary with sparse representation. In this model, for each signal $x \in \mathbb{R}^d$, x can be represented by a linear composition of the primitives from the dictionary D ($D \in \mathbb{R}^{d\times k}$), which is an over-complete dictionary containing k primitives. Such that $\forall\, x$, $\exists \alpha \in \mathbb{R}^k$ satisfying $x \approx D\alpha$ and $\|\alpha\|_0 \ll d$, where the notation $\|\bullet\|_0$ represents the l_0 norm. Typically, the dictionary is redundant to the signal x and $k > d$.

The overcomplete dictionary can either be a pre-defined set or learned adaptively from the input signal. In this work, we employ the second approach by training the dictionary according to the input image. Assuming the input image to be X, which can be partitioned into many overlapped patches x_1, x_2,…, x_i, i=1, 2,…,M. These patches are collected and used as training samples. The K-SVD algorithm [12] is applied to the set of patches to obtain the content adaptive dictionary D, and the objective function is formulated as follows,

$$\min_{D,\{a_i\}} \sum \|x_i - D a_i\|_2^2 \quad \text{s.t. } \forall i, \|a_i\|_0 < L \tag{1}$$

The dictionary trained from 8x8 patches partitioned from *Lena* image is shown in Fig. 2.

Sparse coding is referred to calculating appropriate coefficients α with respected to the trained dictionary D. Thus it can be formulated as

$$a_i = \underset{a_i}{\operatorname{argmin}} \|x_i - D a_i\|_2^2 \quad \text{s.t. } \|a_i\|_0 < L \tag{2}$$

This is a NP-hard problem in general, which can be approximated by a wide range of techniques [9]. In this paper, we employ the orthogonal matching pursuit (OMP) [13] algorithm to solve this problem because of its simplicity and efficiency.

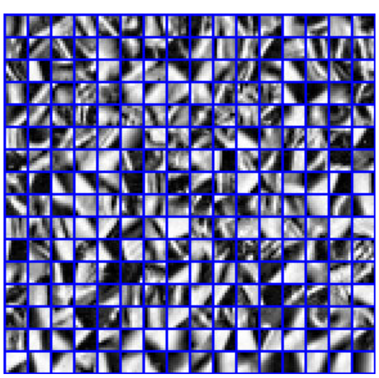

Fig. 2. Dictionary trained with overlapped image patches from *Lena* image.

B. Entropy of Primitive

Based on the image primitive coding, the entropy of primitive (EoP) is proposed to measure the amount of visual information [14]. The algorithm of EoP calculation is shown in Algorithm 1. Particularly, the K-SVD algorithm is applied to adaptively generate the dictionary, and OMP algorithm is employed to perform sparse coding for each non-overlapped patch. After obtaining the probability of each primitive, the EoP can be finally calculated.

Algorithm 1. Summary of the EoP calculation algorithm.

Step 1: Generate the image primitives with K-SVD scheme.

(8x8 overlapped image patch set, image primitive number: 256)

Step 2: Divide image into s non-overlapped patches.

Step 3: For each patch, perform sparse coding.

l: the number of image primitives to represent each patch

T: the number of image primitives used in representing the image

$$T = s \cdot l \tag{3}$$

N_j: the number of primitive j used in sparse coding

Step 4: Compute the EoP.

p_j: the probability of each primitive

$$p_j = N_j / T \tag{4}$$

EoP is calculated by

$$\text{EoP} = -\Sigma_j\, p_j \log(p_j) \tag{5}$$

III. Reduced Reference Image Quality Assessment

Given each l, there will be a corresponding EoP value. The EoP value for each image is shown in Fig. 3. It is observed that the EoP value increases with the number of primitives l, and tends to converge to stable values beyond a threshold. From the image statistics point of view, the sparse coding can be treated as the best explanation of the input image data with the sparest constraint, and EoP could well quantify the visual information based on the sparse representation [15]. When the EoP converges, little perceptual loss is observed compared to the original input image.

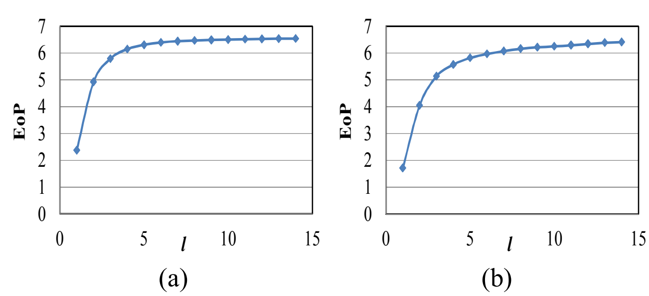

Fig. 3. EoP curves of the original images for (a) *Lena* and (b) *Airplane*.

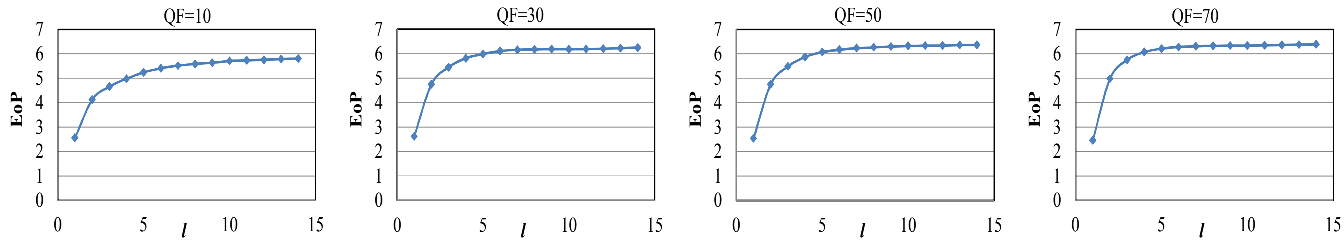

Fig. 4. EoP curves for the distorted images with different quality factors.

Another observation on EoP is that it is related to the visual quality of natural images. In Fig. 4, the EoP of the distorted images compressed with different quality factors (QF) are demonstrated. It is observed that though the curve shapes are similar, the EoP value is highly correlated with the image quality. Based on the observation, the EoP based IQA metric is proposed by comparing the EoP between the original and distorted images. Specifically, assume the EoP value at the point l is denoted to be EoP_l, and then the point t at which the EoP converges to a constant value is obtained as follows,

$$t = \arg\min\{l\},\ s.t. \frac{EoP_l - EoP_{l-1}}{\max(EoP) - \min(EoP)} \le \varepsilon \tag{6}$$

where ε is a threshold. The EoP_t of different QFs are shown in Fig. 5, which illustrates that the EoP_t increase monotonously with the image quality. Then the proposed quality measure is defined by comparing this quantity of the original image EoP_t and the distorted image EoP_t' as follows,

$$EoPM = | EoP_t - EoP_t{'} | \tag{7}$$

In this scheme, only the data EoP_t of the original image is required to transmit in evaluating the image quality, which is negligible compared to the original image, and even much less than the side information of the selected features in the state-of-the-art RR algorithms.

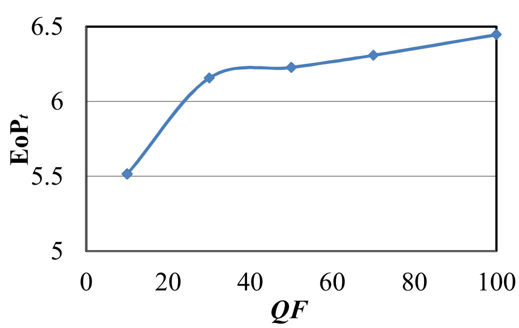

Fig. 5. EoP_t for the distorted images of different QFs.

IV. Experimental Results

In this paper, the popular LIVE image database [16] which was developed at The University of Texas at Austin is used to evaluate the performance of the proposed metric. Specifically, 29 original high-resolution 24 bits/pixel RGB color images are used in the database. Different distortion types are generated, such as JPEG2000, JPEG, white noise, Gaussian blur, and fast fading channel distortion of the JPEG2000 bitstream. Five evaluation metrics that are used to access the performance of the quality metrics are as follows.

a) PLCC: Pearson linear correlation coefficient
b) MAE: Mean absolute error
c) RMS: Root mean-squared
d) SRCC: Spearman rank order correlation coefficient
e) KRCC: Kendall's rank correlation coefficient

These metrics can efficiently evaluate the prediction accuracy, monotonicity, and consistency of the IQA algorithms. In this experiment, the popular full reference IQA metric SSIM [1] and reduced reference IQA metric Wavelet marginal [8] are used for comparison.

The performance in terms of the five indicators described above is shown in Table 1. It is observed that the proposed method outperforms the state-of-the-art RR metric Wavelet marginal in terms of all the five indicators. The performance is not as good as the full reference metric SSIM, as only the EoP of the reference image is used in the quality evaluation.

However, the state-of-the-art quality metrics such as SSIM and PSNR are not accurate in quantifying the geometric distortions, as shown in Figs. 6&7. The translation and rotation have little effects on the image quality, while both SSIM and PSNR drop drastically. In contrast, EoPM can accurately predict the geometric distortions, as the EoPM for the images that undergo the geometric distortion are smaller, indicating that the distorted images appear similar as the original image.

Table 1. Performance comparison with the SSIM and Wavelet marginal methods on the LIVE database.

	LIVE database (full 982 images)[16]					
IQA measure	**Type**	**PLCC**	**MAE**	**RMS**	**SRCC**	**KRCC**
SSIM	FR	0.9648	6.1507	8.2263	0.9729	0.8635
Wavelet marginal	RR	0.8342	12.8153	17.2557	0.8517	0.6683
EoPM	RR	0.8801	11.5777	14.8589	0.9007	0.7326

V. CONCLUSION

In this paper, we propose a EoP based reduced reference image quality metric. The image quality is evaluated in terms of the difference of the EoP, which has been shown to efficiently quantify the visual information. As a reduced reference metric, only the EoP value of the reference image is required to be transmitted, and this is negligible compared to the size of the reference image. The proposed metric is invariant to the geometric distortion, and performs competitively against the well know quality assessment metric on the public dataset. In the future work, we will investigate more about the EoP based image processing algorithms, from both the theoretical and applicable aspects.

VI. ACKNOWLEDGEMENT

This work was supported in part by National High-tech R&D Program of China (863 Program, 2012AA010805) and National Science Foundation (61210005 and 61103088).

REFERENCES

[1] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. “Image quality assessment: From error visibility to structural similarity,” IEEE Trans.on Image Processing, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[2] D. M. Chandler and S. S. Hemami, “VSNR: A Wavelet-Based Visual Signal-to-Noise Ratio for Natural Images ”, IEEE Transactions on Image Processing, Vol. 16 (9), pp. 2284-2298, 2007.

[3] L. Liang, S. Wang, J. Chen, S. Ma, D. Zhao and Wen Gao, No-reference perceptual image quality metric using gradient profiles for JPEG2000. Sig. Proc.: Image Comm. 25(7): 502-516 (2010).

[4] H. Liu, I. Heynderickx, Issues in the design of a no-reference metric for perceived blur. In SPIE vol. 7867 (2011).

[5] Z. Wang and A. C. Bovik, “Reduced- and no-reference image quality assessment: The natural scene statistic model approach,” IEEE Signal Processing Magazine, vol. 28, Nov. 2011.

[6] T. M. Kusuma and H.-J. Zepernick, “A reduced-reference perceptual quality metric for in-service image quality assessment,” in Proc. Joint 1st Workshop Mobile Future and Symp. Trends Commun., Oct. 2003, pp. 71–74.

[7] I. P. Gunawan and M. Ghanbari, “Reduced reference picture quality estimation by using local harmonic amplitude information,” in Proc. London Commun. Symp., Sep. 2003, pp. 137–140.

[8] Z. Wang, G. Wu, H. R. Sheikh, E. P. Simoncelli, E.-H. Yang, and A. C. Bovik, “Quality-aware images,” *IEEE Trans. Image Process.*, vol. 15, no. 6, pp. 1680–1689, Jun. 2006.

[9] M. Elad, "Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing", Springer, 2010.

[10] M. Elad, and M. Aharon, “Image denoising via sparse and redundant representations over learned dictionaries,” IEEE Trans. on Image Processing, vol. 15, no. 12, pp. 3736 –3745, 2006.

[11] J. Yang, J. Wright, T. Huang, and Y. Ma, “Image super-resolution via sparse representation”, IEEE Transactions on Image Processing, vol. 19, pp. 2861-2873, Nov. 2010.

[12] M. Aharon, M. Elad, and A. M. Bruckstein, "The K-SVD: An Algorithm for Designing of Overcomplete Dictionaries for Sparse Representation," the IEEE Trans. on Signal Processing, vol. 54, no. 11, pp. 4311-4322, Nov. 2006.

[13] J. A. Tropp and A. A. Gilber, “Signal Recovery from Random Measurements via Orthogonal Matching Pursuit,” IEEE Trans. on Information Theory, vol. 53, no. 12, pp. 4655–4666, 2007.

[14] J. Zhang, S. Ma, R. Xiong, D. Zhao and W. Gao, “Image Primitive Coding and Visual Quality Assessment,” PCM 2012, vol. 7674, pp. 674-685, 2012.

[15] X. Zhang, S. Wang, S. Ma, S. Liu, W. Gao, “Entropy of Primitive: A Top-Down Methodology for Evaluating the Perceptual Visual Information”, submitted to VCIP 2013.

[16] H. R. Sheikh, 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/

(a) Original image

(b) Translation (3,0)
PSNR=20.33, SSIM=0.553
EoPM=0.2739

(d) Rotation
PSNR=18.12, SSIM=0.5707
EoPM=0.4807

(c) JPEG Compression
PSNR=30.41, SSIM=0.8179
EoPM=0.7489

Fig. 6. *Lena* image undergone various degradations and corresponding values of different quality metrics.

(a) Original image

(b) Translation (3,0)
PSNR=20.45, SSIM=0.722
EoPM=0.1432

(c) Rotation
PSNR=16.35, SSIM=0.6558
EoPM=0.4968

(d) JPEG Compression
PSNR=29.68, SSIM=0.8435
EoPM=1.1704

Fig. 7. *Airplane* image undergone various degradations and corresponding values of different quality metrics.