Reduced Reference Image Quality Assessment Based on Statistics of Edge

Min Zhang, Wufeng Xue, Xuanqin Mou

Institute of Image Processing and Pattern Recognition, Xi'an Jiaotong University,
Xi'an, Shaanxi 710049, China
xqmou@mail.xjtu.edu.cn

ABSTRACT

Objective Image Quality Assessment (IQA) model investigation is a hot topic in recent times. This paper proposed a novel and efficient universal Reduced Reference (RR) image quality assessment method based upon the statistics of edge discrimination. Firstly, binary edge maps created from the multi-scale wavelet transform modulus maxima were used as the low level feature to discriminate the difference between the reference and distorted image for IQA purpose. Then the gradient operator was applied on the binary map to produce the so called edge pattern map. The histogram of edge pattern map was used to verify the pattern of the edges of reference and distorted image, respectively. The RR features extracted from the histogram was used to discriminate the difference of edge pattern maps, and then form a new RR IQA model. Comparing to the typical RR model (Zhou Wang's method, 2005), only 12 features (96 bits) are needed instead of 18 features (162 bits) in Zhou Wang et al.'s method with better overall performance.

Keywords: wavelet, image quality assessment, modulus maxima, reduced reference

1. INTRODUCTION

With the increasing use of applications based on visual data, objective image quality assessment has become an essential issue. Objective image quality models are mainly categorized into three types: Full reference (FR) models, for which the original image and the distorted image are required. Reduced reference (RR) models, for which a description of the original image into some parameters and the distorted image are both required. No reference (NR) models, which only require the distorted image.

FR models have been investigated for a long time, and their performance has been validated across existed distortion types in realistic world. At present, several models have been used in the applications, such as Sarnoff JNDmetrix visual discrimination model (VDM) [1, 2] and Wang and Bovik's SSIM [3]. Lately, Sheikh and Bovik proposed a competitive FR model named Visual Information Fidelity (VIF) [4,5]. It outperforms all state-of-the-art FR IQA models by a sizeable margin based on their report.

However, in most cases, such as in-service visual quality monitoring, reference images cannot be achieved, and the application of FR IQA models is restricted greatly. It is important to develop more practical IQA models which need not use full reference image information. Concerning to NR models, there is still no such efficient model that could deal with types of distortions. Even to specific distortion type, the performance of these NR models is not very satisfied [6,7,8]. RR models can be considered as the compromise between the unfeasibility of FR models and unavailability of NR models. It means that only partial information of image is demanded. The amount of this partial information is supposed to be much smaller than the reference image, and the distorted image quality should be predicted with it. According to the report of video quality experts groups (VQEG), the investigation and validation of RR methods have been listed as one of the important issues for future directions [9, 10].

In general, for a typical RR system, features relate to the image perception quality is extracted at the sender side with limited bit rate. Then, these RR features are transmitted through an ancillary channel or are hidden in the reference

Digital Photography VII, edited by Francisco H. Imai, Feng Xiao, Jeffrey M. DiCarlo,
Nitin Sampat, Sebastiano Battiato, Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 7876, 787611
 · doi: 10.1117/12.873075

image before they are processed by the distorting system [11]. There is still another RR features extraction from the distorted image at the receiver side. Finally, these RR features are analyzed for IQA purpose.

But limited work has been done in RR IQA research. Zhou Wang and Eero P. Simoncelli proposed a reduced reference image quality assessment using a wavelet-domain natural image statistics model [11]. This RR model could be used as a general-purpose image quality assessment and uses very low data rate (162 bit) for representing RR features of reference image.

The low-level features, such as edges, of image play an important role in visual perception. For FR IQA purpose, binary edge maps created from the multi-scale wavelet transform modulus maxima have been used as the low level feature to discriminate the difference between the reference and distorted image[12]. While in RR architecture, the whole edge maps of images cannot be achieved at the image receiver side, so a statistical method of edge pattern is proposed in this paper. We consider that the degradation of image will result in the change of edge point distribution and the perceptual image distortion is highly correlated with this change of edge point distribution plane hence to form a RR IQA metric.

This paper presents a low-level feature inspired universal RR IQA metric. We consider that the low-level features, such as edges, of image play an important role in visual perceptual distortion. While in RR architecture, these low-level features cannot be achieved at the receiver side. So we proposed a statistical method of this edge information aiming to form a RR IQA metric. This metric is based on the statistics of edge pattern. Firstly, binary maps created from the multi-scale wavelet transform modulus maxima were used as the low-level feature. Secondly, the gradient operator was applied on the binary map to produce the so called edge pattern map. Then the histogram of edge pattern map was used to verify the pattern of the edges of reference and distorted image, respectively. Finally, the RR features extracted from the histogram were used to discriminate the difference of edge pattern maps, and then form a new RR IQA metric.

The rest of the paper is organized as follows. Section 2 presents the proposed RR IQA model. Section 3 describes the experiment setup. Section 4 gives the experimental results and discussions. Section 5 concludes the paper.

2. THE PROPOSED MODEL

A perfect RR quality assessment model must achieve a good balance between the data rate of RR features and the accuracy of image quality prediction. In this section, a general RR IQA model is proposed. Binary edge maps created from the multi-scale wavelet transform modulus maxima are used as the low level feature to discriminate the difference between the reference and distorted image for IQA purpose. While in RR architecture, since the binary edge maps of reference image are not available, the gradient operator are applied on the binary edge map to describe different patterns of the binary edge map, which is so called edge pattern map. The discrete feature histogram which comes from the statistics of edge pattern map is considered as the final RR features for IQA purpose. The assumption is that the degradation for the image will change the distribution of this histogram. So this histogram is used to qualify the visual degradation of image and result in a new RR IQA model.

2.1. RR feature extraction

The proposed RR IQA metric is based on the statistics of edge pattern. The RR features for reference and distorted image are extracted at the sender side or received side separately. Figure 1 shows the RR feature extraction architecture. According to the RR feature extraction architecture shown in Figure 1, firstly, the input image is decomposed into wavelet transform domain. Cki and Dki $(0 \le i \le N, 1 \le k \le M)$ denote the kth sub-bands images of scale i for the reference and distorted image, respectively. Then the binary edge map of scales i is described by the wavelet transform maxima modulus which are represented as C_i^E and D_i^E for the reference and distorted image individually. The detail about the wavelet transform maxima modulus detection is out of this paper, and it can be found in [13, 14]. We consider that the degradation of image will result in the change of edge point distribution and the perceptual image distortion is highly correlated with this change of edge point distribution plane.

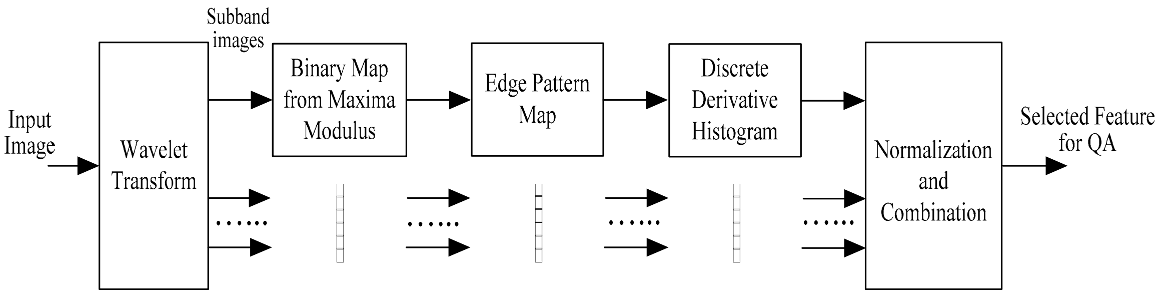


Figure1. Feature extraction architecture for an input image

0	-1	0
-1	4	-1
0	-1	0

Figure 2. Laplacian template

While in RR architecture, it is difficult to transmit the whole edge maps because of the data rate limitation. Here a rough and simple edge pattern verification method is proposed. It is 4-adjacent, 3×3 discrete Laplacian template *H*, as is shown in Figure 2.

For binary edge map C_i^E and D_i^E, the edge pattern map C_i^P and D_i^P are defined as the convolution between the binary edge map and the template *H*.

$$\mathrm{C}_i^P = \mathrm{C}_i^E \otimes H \quad \text{and} \quad \mathrm{D}_i^P = \mathrm{D}_i^E \otimes H \tag{1}$$

Since we aim to verify the patterns of edges, it means that this template only convolute with edge point in the binary edge maps. Hence numeric values of the each point in edge pattern maps range 1 to 4. They represent four different patterns. While in this paper, we found the amount of edge pattern point valued 1 is small. So numeric values of the each point in edge pattern maps range 2 to 4 are involved to form the RR feature histogram, which is denote as $p_{C_i}(x_k)$ and $p_{D_i}(x_k)$ $(x_k = [2,3,4])$ for the discrete edge pattern maps of scale *i*. Each bin of the histogram is considered as the one RR feature. In this paper, 4 scale wavelet transform are used to form 12 scalar RR features for one image.

Finally, normalized histogram comes from the statistics of the edge pattern maps C_i^P and D_i^P, which is denote as $p_{C_i}(x_k)$ and $p_{D_i}(x_k)$ $(x_k = [1,2,3])$ for the discrete edge pattern maps of scale *i*.

2.2. Distortion measure

The RR features for the reference image are extracted at the sender side. It is firstly quantized and transmitted to the receiver side. Then the RR features for the distorted image are extracted at the receiver side in the same way. D_M is the Minkowski distance (L-p norms) between the features of reference and distorted images. It is defined as

$$D_M = \sum_{i=1}^{4}\sum_{k=1}^{3} \log_{10}\left[p_{C_i}(x_k) - p_{D_i}(x_k) \right]^2 \tag{2}$$

D_M represents the predicted objective score for one image.

3. IMPLEMENTATION ISSUES

3.1. Simulation details

The proposed RR model and the model purposed by Zhou Wang and Eero P. Simoncelli are compared. They all operated upon the luminance component of images. Four scale wavelet transform is employed in the RR features extraction architecture for the proposed method. Zhou Wang's RR model is used as the default usage provided [11].

All the performance validation metrics were computed after a nonlinear mapping between the objective and the subjective scores [10]. A five-parameter nonlinearity (a logistic function with additive linear term constrained to be monotonic) is selected. It is shown as follows

$$\text{Quality}(x) = \beta_1 \text{logistic}(\beta_2, (x - \beta_3)) + \beta_4 x + \beta_5 \tag{3}$$

$$\text{logistic}(\tau, x) = \frac{1}{2} - \frac{1}{1 + \exp(\tau x)}$$

3.2. The test image database

The image database [15] used in the paper is developed by the Laboratory of Image and Video Engineering (LIVE), the University of Texas at Austin to validate the performance of the proposed quality model. 29 high-resolution 24-bits/pixel RGB color images were distorted using five distortion types: JPEG2000, JPEG, white noise, Gaussian blur, fast fading. A total of 982 images, out of which 203 were the reference images, were evaluated by human subjects in seven experiments. Finally, a Difference Mean Opinion Score (DMOS) value for each distorted image was computed.

4. RESULTS AND DISCUSSION

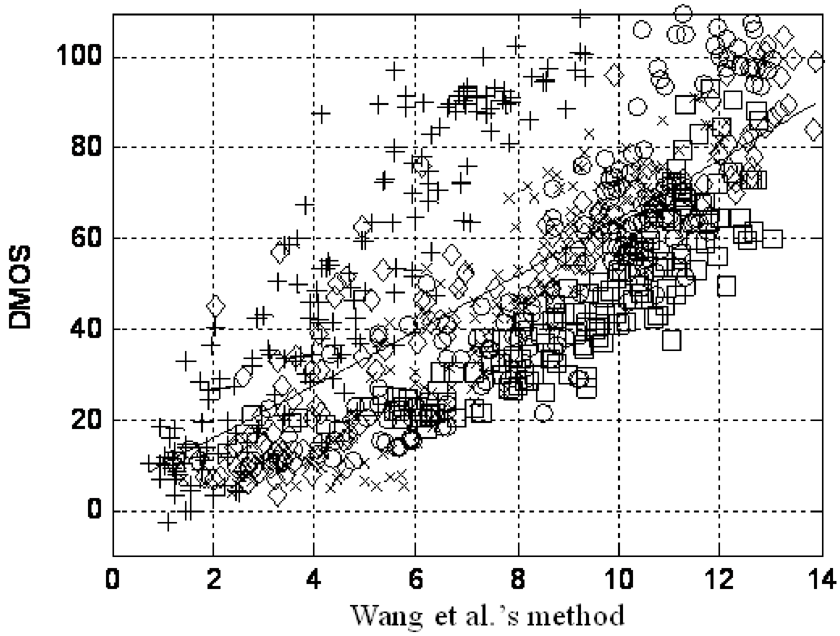
In this section, the results of the proposed model on the database mentioned in section 3 are presented. Zhou Wang's RR model is chosen for performance comparison. Because it is a typical RR model and the code is available.

Figure 3 shows the scatter plots (in which each data point represents one test image) of difference mean opinion score (DMOS) versus the predicted objective score by the proposed method and Zhou Wang's RR model.

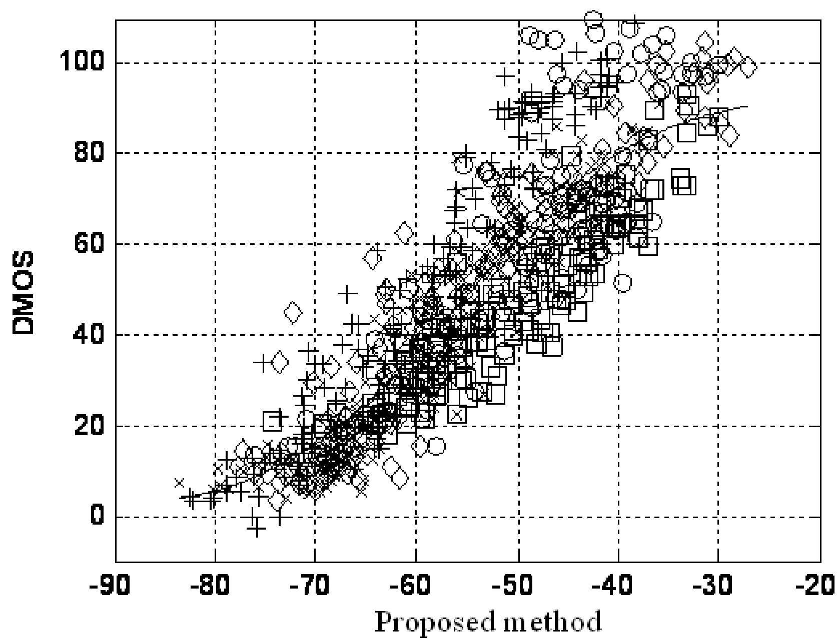
Concerning to the overall performance, the more these curves close to each other, the more stable predict quality across distortion types will be achieved. Obviously, Comparing to Zhou Wang's RR model, there is a noticeable improvement in term of cross distortion. While Zhou Wang's RR model employs 18 scalar features (or 162 bit after quantization) to predict the perceptual distortion of the distorted images, as for the proposed model, only 12 scalar features (or 96 bit after quantization) are used to predict the perceptual distortion The data rate decreases to 14.8% comparing to Zhou Wang's RR model. But there is still a lot of work to improve the overall performance concerning to RR model.

Table 1 and 2 give the performance comparisons of these two quality assessment models on five individual distortion types (including 175 JPEG2000 images, 169 JPEG images, 145 white noise images, 145 Gaussian blur images and 145 transmission errors in JPEG2000 stream over fast-fading Rayleigh channel images). Table 1 shows the correlation coefficient between the objective scores and the subjective after nonlinear regression. Table 2 shows the RMSE performance between the objective scores after compensation and the subjective scores.

It is sometimes interesting from an application perspective to restrict the quality measures to a specific distortion type, concerning to the performance on individual distortion type. As it can be seen from Table 1 and Table 2, the performance of the proposed model is almost the same as Zhou Wang's RR model on JPEG2000 and JPEG image data sets. While the performance of the proposed model on the distortion of Gaussian blur is much more attractive. But the performance for the white noise is slightly worse.



(a)



(b)

Figure 3. Scatter plots for the two reduced referenced objective quality method: (a) Zhou Wang's RR model and (b) the proposed model. The distortion types are: JPEG2000 (x), JPEG (+), white noise in RGB space (○), Gaussian blur (□), and transmission errors in JPEG2000 stream over fast-fading Rayleigh channel (◊). The solid line represents the fitting curve for the data point of each distortion types. The color of the solid line is in accordance with the date point**.**

Table 1 Spearman correlation coefficient of the IQA models on individual distortion types after nonlinear regression

Distortion type\Model	Zhou Wang et al.'s method	Proposed method
JPEG2000(1) (87)	0.9352	0.9134
JPEG 2000(2)(82)	0.9437	0.9495
JPEG(1) (87)	0.8077	0.9105
JPEG(2)(145)	0.8924	0.9294
Noise (145)	0.8688	0.8417
Blur (145)	0.9206	0.9265
FF (145)	0.9221	0.9365
ALL	0.7361	**0.8832**

Table 2 Correlation coefficient of the IQA models on individual distortion types after nonlinear regression

Distortion type\Model	Zhou Wang et al.'s method	Proposed method
JPEG2000(1) (87)	0.9315	0.9087
JPEG 2000(2)(82)	0.9481	0.9511
JPEG(1) (87)	0.8430	0.9094
JPEG(2)(145)	0.9693	0.9777
Noise (145)	0.8903	0.8623
Blur (145)	0.8933	0.9234
FF (145)	0.9223	0.9392
ALL	0.7467	**0.8744**

5. CONCLUSION

Reduced reference image quality assessment is an important issue in image processing applications. This paper proposed a simple edge pattern verification method for RR image quality assessment. Experimental results validated the performance of the proposed metric.

Although, the proposed metric is a simple one, the data rate is dramatically lower than other general-purpose RR model. Only 12 scalar features (or 96 bits after quantization) are needed while the other well-known RR model (Zhou Wang's model) uses 18 scalar features (or 162 bits after quantization).

6. Acknowledgement

This work is partially supported by National Natural Science Foundation of China (NSFC) through No. 90920003 and 60472004.

7. REFERENCES

[1] Sarnoff Corporation. JNDmetrix Technology. http://www.sarnoff.com/productsservices/videovision/jndmetrix/.

[2] J. Lubin, "A visual discrimination model for imaging system design and evaluation," in Visual Models for Target Detection and Recognition, E. Peli, Ed. Singapore: World Scientific, pp. 245–283(1995).

[3] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. "Image quality assessment: From error measurement to structural similarity," IEEE Trans. Image Process., vol. 13, no. 4, pp. 600–612, Apr. 2004.

[4] H.R. Sheikh and A.C. Bovik. "Image information and visual quality," IEEE Trans. Image Process., vol.15, no.2pp. 430- 444, Feb. 2006.

[5] H.R. Sheikh, M.F. Sabir and A.C. Bovik. "A statistical evaluation of recent full reference image quality assessment algorithms," Image Processing, IEEE Transactions on, vol. 15, no. 11, pp. 3440-3451, Nov. 2006.

[6] H. R. Sheikh, A. C. Bovik, and L. Cormack. "No-reference quality assessment using natural scene statistics: JPEG2000," IEEE Trans. Image Process., vol. 14, no. 11, pp. 1918–1927, Nov. 2005.

[7] Z. Wang, H. R. Sheikh, and A. C. Bovik. "No-reference perceptual quality assessment of JPEG compressed images," Proc. IEEE Int. Conf. Image Proc. Sept. 2002.

[8] P. Marziliano et al. "A no-reference perceptual blur metric," Proc. IEEE ICIP, 2002, vol. 3, pp. 57–60 (2002).

[9] VQEG: The Video Quality Experts Group, http://www.vqeg.org/.

[10] VQEG. Final report from the video quality experts group on the validation of objective models of video quality assessment," Apr. 2000. http://www.vqeg.org/.

[11] Z. Wang and E. P. Simoncelli, "Reduced-reference image quality assessment using a wavelet-domain natural image statistic model," Human Vision and Electronic Imaging X Conference, San Jose, CA, January 17-20, 2005,Proceedings of SPIE, Vol. 5666.

[12] M. Zhang and X. Mou, "A psychovisual image quality metric based on multi-scale structure similarity," in ICIP, pp. 381–384, 2008

[13] S. Mallat and W. L. Hwang, "Singularity detection and processing with wavelets," IEEE Trans.Inform. Theory, 38 (1992), pp. 617–643.

[14] S. Mallat and S. Zhong. "Characterization of signals from multiscale edges," IEEE Trans. Pattern Anal. Machine Intell., vol. 14, pp. 710–732, July 1992.

[15] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik. (2005) Live Image Quality Assessment Database Release2. [Online]. vailable:http://live.ece.utexas.edu/research/quality.