Image Quality Assessment Based on Measuring the Degradation in Geometrical Structural Information

Chun-Hsien Chou and Yun-Hsiang Hsu
Tatung University, Taipei
E-mail: chou@ttu.edu.tw Tel: +886-022182-2928
Tatung University, Taipei
E-mail: tommy910203048@yahoo.com.tw Tel: +886-022182-2928

***Abstract* — A new image quality metric based on measuring the loss of the structural information inherent in digital images is presented in this paper. The structural information is defined statistically and geometrically in a block basis. Statistical structure information includes luminance means and contrast. Geometrical structure information is extracted from the binary quantization that preserves the first two moments of the image block. To verify the validity of the proposed metric is evaluated against a large amount of test images in LIVE database and compared with that of the famous MSSIM. The cross-distortion test results show that the proposed metric outperforms MSSIM in judging the distorted images corrupted by JPEG2000, Gaussian blurring and fast fading and has the performance close to MSSIM in judging the distorted images corrupted by JPEG and white noises.**

I. Introduction

Image quality metric is an important technique in the development of image-processing systems. For systems and applications where processed images are to be viewed by human beings, the most reliable evaluation in image quality is obtained through subjective judgment. However, subjective evaluation is a cumbersome judgment process which is usually time-consuming and costly. Therefore, the objective image quality metric that can automatically predict the image quality with a high agreement with the perceptual assessment of human beings has always been pursued and taken as the goal of the related research. Traditionally, image quality is assessed by using simple quantitative metrics such as the PSNR and MSE. Nevertheless, these traditional metrics do not faithfully reflect the human visual perception [1-5]. Therefore, the objective image quality metric that can automatically predict the image quality with a meaningful quantitative score showing high consistency with the perceptual assessment of human beings has always been the goal of the related research.

Natural images are considered to possess common statistical properties [8] and in general highly structured in content [14], where pixels within local areas exhibit strong dependencies. This high dependency carries important information manifesting the texture and structure of the objects in the scene. Structural information is therefore one of the important components that affect the visual quality of digital images. In the literature, many approaches using statistical properties of natural images and structural information for image quality assessment were proposed [6-10]. In [8, 9], natural scene statistics is used to assess the image quality by quantifying the mutual information between distorted and reference images. In [6, 7], the structural similarity metric (SSIM) was proposed to evaluate the image fidelity by measuring the loss of image structure. The loss of image structural information can be attributed to the distortions of luminance mean, contrast and cross correlation. In [10], the characteristic matrix obtained by singular value decomposition (SVD) is taken as the structural information for evaluating the image quality.

Among these approaches, structural information is mostly defined by the statistics of the image. These statistical approaches have the drawback that the metric score can be still high as the perceptual quality may become objectionable due to sparse local distortion that contaminates the geometrical structure of the image, or the distortion due to mean shift.

To reflect geometrical distortion in the objective metric score, the image quality assessment is better designed to take not only the statistical structural information but also the geometrical structural information into account. In this paper, the proposed image quality metric integrates the measure in the loss of geometrical structure information with the measure of distortion in local luminance and contrast. The local geometrical structure information is extracted through a binary quantization designed to statistically preserve the first two moments of the local image data. The test images contaminated by various types of distortion (JPEG, JPEG2000, white noise, Gaussian Blur, Fast Fading), as can be found in the LIVE database of 982 images [13], are used in the simulation for verifying the validity of the proposed metric.

II. The Proposed Image Quality Metric

The proposed image quality metric is designed to measure the degradation in structural information extracted from both the image to be assessed and its original. The structural information consists of three separate components: mean intensity of luminance, luminance contrast and the geometrical distribution of classified image data. The geometrical structure information is extracted from the quantization of local image data by binary quantizers designed to preserve the first two moments of the image data [11, 12].

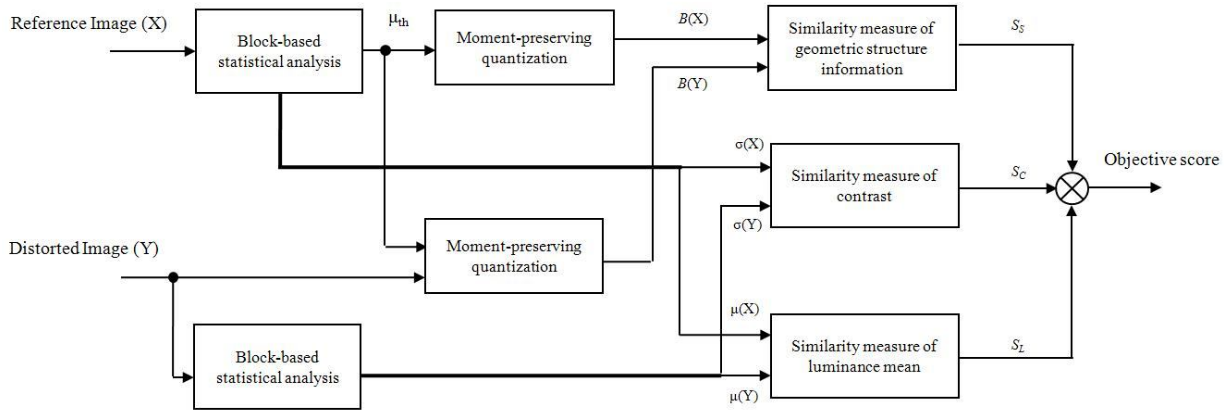


Fig. 1 Block diagram for evaluating the objective score of the proposed metric

The functional block diagram for evaluating the objective score of the proposed metric is shown in Fig.1 and the mean metric score can be expressed as

$$S_{MPM}(\mathbf{X},\mathbf{Y}) = \frac{1}{M}\sum_{l=0}^{M-1} S_{MP}(\mathbf{x}_l,\mathbf{y}_l) \tag{1}$$

where **X** and **Y** denote image to be assessed and the reference image, respectively. M is the number of image blocks involved in the assessment of image quality. $S_{MP}(\mathbf{x}_l,\mathbf{y}_1)$ denotes the metric score obtained from measuring the similarity between two corresponding blocks within the image to be evaluated, $\mathbf{x}_l$, and within the reference image, $\mathbf{y}_l$, in terms of three measurements. That is

$$S_{MP}(\mathbf{x}_l,\mathbf{y}_l) = S_L(\mathbf{x}_l,\mathbf{y}_l)\cdot S_C(\mathbf{x}_l,\mathbf{y}_l)\cdot S_S(\mathbf{x}_l,\mathbf{y}_l) \tag{2}$$

where $S_L(\mathbf{x}_l,\mathbf{y}_l)$ measures the similarity in the structure information of luminance, $S_C(\mathbf{x}_l,\mathbf{y}_l)$ the similarity in the structure information of contrast, and $S_S(\mathbf{x}_l,\mathbf{y}_l)$ the similarity in spatial distribution of classified image data. Since the human visual system (HVS) is sensitive to the change in averaged luminance, the similarity measure in luminance structure is included and is defined, as similar to the counterpart defined in SSIM [7], as

$$S_L(\mathbf{x}_l,\mathbf{y}_l) = \left(\frac{2\mu(\mathbf{x}_l)\mu(\mathbf{y}_l)+K_0}{\mu^2(\mathbf{x}_l)+\mu^2(\mathbf{y}_l)+K_0}\right) \tag{3}$$

where

$$\mu(\mathbf{x}_l) = \frac{1}{N}\sum_{i=1}^{N} x_l(i) \quad \text{and} \quad \mu(\mathbf{y}_l) = \frac{1}{N}\sum_{i=1}^{N} y_l(i) \tag{4}$$

denote the mean luminance of the image block $\mathbf{x}_l$ and $\mathbf{y}_l$. N determines the size of the image block involved in the similarity measurement. The constant K_0 is used to avoid the instability due to zero mean luminance. The contrast of an image block reveals the dynamic range information of the local image data. The dissimilarity in the dynamic range of signals between the two image blocks in comparison is measured as

$$S_C(\mathbf{x}_l,\mathbf{y}_l) = \left(\frac{2\sigma(\mathbf{x}_l)\sigma(\mathbf{y}_l)+C_0}{\sigma^2(\mathbf{x}_l)+\sigma^2(\mathbf{y}_l)+C_0}\right) \tag{5}$$

where

$$\sigma(\mathbf{x}_l) = \left[\frac{1}{N}\sum_{i=1}^{N} x_l^2(i) - \mu(\mathbf{x}_l)^2\right]^{\frac{1}{2}}$$
$$\sigma(\mathbf{y}_l) = \left[\frac{1}{N}\sum_{i=1}^{N} y_l^2(i) - \mu(\mathbf{y}_l)^2\right]^{\frac{1}{2}} \tag{6}$$

represent the deviations of the image blocks $\mathbf{x}_l$ and $\mathbf{y}_l$, and constant C_0 is also used to avoid the instability when both image blocks in comparison are uniform.

The spatial distribution of the image data that have common attributes is also considered as important information in judging the similarity between two images in comparison. As the simplest case, the locations of the image data that possess the luminance higher than the average luminance can be regarded as the information indicating the texture of the image content. This structural information can be simply represented by a set of binary signals after an appropriate classification process or binary quantization. For instance, the binary map obtained from quantizing the luminance signals of an image block shown in Fig. 2 by a binary quantizer can be taken as the local information of geometrical structure extracted from the image. The binary quantizer can be designed by preserving the first two moments of the reference image block [11, 12]. The decision level of the binary quantizer can be simply set to the mean luminance μ_{th} ($\mu(x_l)$) of the reference image. The degradation of the geometrical structure inherent in the test image against the reference image can be measured as the percentage of the image data that are not in the same class after the quantization. The similarity in geometrical structure between two image block can be expressed as

$$S_S(\mathbf{x}_l, \mathbf{y}_l) = \left(\frac{N - \sum_{i=1}^{N} Q(x_l(i)) \oplus Q(y_l(i))}{N} \right). \tag{7}$$

where $\{Q(x_l(i))\}$ and $\{Q(y_l(i))\}$ denote the resulting bitmaps of the image blocks $\mathbf{x}_l$ and $\mathbf{y}_l$ after the quantization with the quantizer

$$Q(u_l) = \begin{cases} 1, & if \quad u_l \geq \mu_{th} \\ 0, & if \quad u_l < \mu_{th} \end{cases}. \tag{8}$$

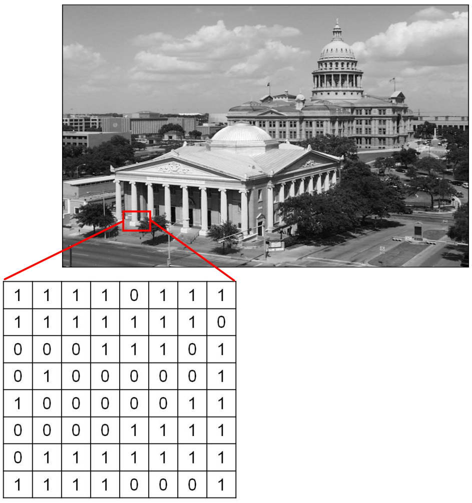


Fig. 2. The geometrical structure information embedded in the image block after the moment-preserving quantization

As can be calculated, the metric score of $S_S(\mathbf{x}_l, \mathbf{y}_l)$ is 0.0 if the geometrical structure of the image block $\mathbf{x}_l$ is totally different from that of the image block $\mathbf{y}_l$.

III. Experimental Results

The prediction accuracy of the proposed image quality metric is inspected by the simulation that employs the test images in the LIVE database [13]. The test images in the database are contaminated by five types of distortion (JPEG, JPEG2000, white noise, Gaussian Blur, Fast Fading) with the associated values of different mean opinion score (DMOS). To avoid the intensive computational complexity in the simulation, non-overlapping 8×8 blocks are used for evaluating the objective score. The objective scores of the proposed metric after the cross-distortion test are compared against the subjective DMOS. The correlation between objective and subjective scores is inspected by Pearson correlation coefficient (CC), Root mean squared error (RMS), and Mean absolute error (MAE), in which nonlinear logistic regression method is adopted to fit the experimental data. The solid curve shown in Fig.3(a) is the result of the logistic fitting. For the purpose of comparison, the prediction performance of the famous SSIM is also inspected by the same set of test images. The associated curve for fitting the correlation between objective and subjective scores is shown in Fig.3(b). The fitting of the correlation curve associated with the proposed metric is better than that associated with the metric of SSIM in terms of the errors of fitting. According to the performance of correlation evaluated by different tools (CC, RMSE and MAE), as demonstrated in Table1 and Table2, The proposed image quality metric is in average superior to SSIM. It is especially the case when assessing the images contaminated by the distortion due to white noises, Gaussian blurring and fast fading.

TABLE I

Performance of the proposed metric on different types of distortion

	jpg2k	jpeg	wn	gblur	fast-fading	overall
Pearson CC	0.9259	0.9375	0.9317	0.9303	0.9506	0.920
RMSE	9.2771	8.4955	8.056	8.0492	6.9181	9.072
MAE	8.6064	7.2174	6.4912	6.4789	4.7859	8.2318

TABLE II

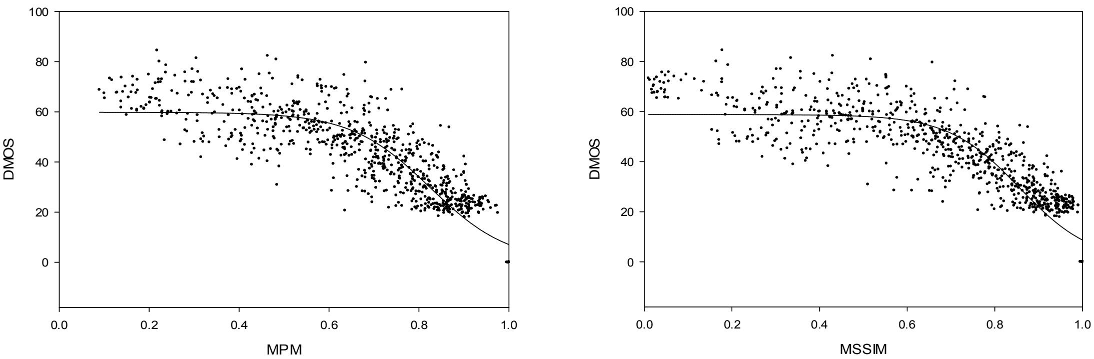
Performance of SSIM measure on different types of distortion

	jpg2k	jpeg	wn	gblur	fast-fading	overall
Pearson CC	0.9364	0.9444	0.9313	0.9221	0.9383	0.9166
RMSE	8.6197	8.0272	8.0756	8.4882	7.7071	9.258
MAE	7.4299	6.4436	6.5215	7.2049	5.9399	8.5711

Further simulation results indicate that the proposed metric is able to reflect the image distortion due to shift in luminance mean in the objective score while the SSIM is less sensitive to the same type of distortion. As an instance shown in Fig. 4, the SSIM scores remain high as the quality of the image degrades due to the shift in luminance. The proposed metric seems in general being able to reflect more fidelity of the image than SSIM.

IV. Conclusion

In this paper, a new metric that incorporates the geometrical structure information in the assessment of image quality is proposed. The structural information is extracted from the binary quantization that preserves the first two moments of the image. Simulation results show that the



(a) (b)

Fig. 3. (a) Scatter plot of the objective scores predicted by the proposed metric versus DMOS for the assessment of the images contaminated by fast fading (fitting error = 8.232) (b) Scatter plot of the objective scores predicted by SSIM versus DMOS for the assessment of the images contaminated by fast fading (fitting error = 8.571)

(a) (b) (c)

Fig. 4. (a)The original image of “coinsinfountain; (b) the same image with higher luminance mean (SSIM=0.8544, proposed metric score=0.4991) (c) the image with luminance mean shift of local space (SSIM=0.8791, proposed metric score =0.6771)

proposed metric is in general superior to SSIM in prediction accuracy, and especially outperforms SSIM for assessing the images contaminated by certain types of distortion. The future research work will focus on the extraction of more dedicated geometrical structure information from color images for visual quality assessment.

REFERENCES

[1] B. Girod, “What’s wrong with mean-squared error,” in *Digital Images and Human Vision* (A. B. Watson, ed.), pp. 207–220, the MIT press, 1993.

[2] P. C. Teo and D. J. Heeger, “Perceptual image distortion,” in *Proc. SPIE*, vol. 2179, pp. 127–141, 1994.

[3] A. M. Eskicioglu, and P. S. Fisher, “Image quality measures and their performance,” *IEEE Trans. Communications*, vol. 43, pp. 2959–2965, Dec. 1995.

[4] M. P. Eckert and A. P. Bradley, “Perceptual quality metrics applied to still image compression,” *Signal Processing*, vol. 70, pp. 177–200, Nov. 1998.

[5] S. Winkler, “A perceptual distortion metric for digital color video,” *Proc. SPIE*, vol. 3644, pp. 175–184, 1999.

[6] Z. Wang and A. C. Bovik, “A universal image quality index,” IEEESignal Processing Letters, vol. 9, no. 3, pp. 81–84, Mar. 2002.

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

[8] H. R. Sheikh, A. C. Bovik, and G. de Veciana, “An information fidelity criterion for image quality assessment using natural scene statistics,” *IEEE Trans. Image Processing*, vol. 14, no. 12, pp. 2117–2128, Dec. 2005.

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

[10] A. Shnayderman, A. Gusev, and A. M. Eskicioglu, “An SVD-based grayscale image quality measure for local and global assessment,” *IEEE Trans. Image Processing*, vol. 15, no. 2, pp. 422–429, Feb. 2006.

[11] Wen-Hsiang Tsai, "Moment-Preserving Thresholding: A New Approach", Computer Vision, Graphics, and Image Processing 29. 377-393., 1985

[12] O. R. Mitchell , E. J. Delp, and S. G. Carlton, “Block truncation coding: a new approach to image compression,” *Proceedings of the IEEE International Conference on Communications,* vol. 1, June 4-7 1978, pp. 12B.1.1-12B.1.4.

[13] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik, “LIVE image quality assessment database release 2,” http://live.ece.utexas.edu/research/quality

[14] Z. Wang, A. C. Bovik, and L. Lu, “Why is image quality assessment so difficult,” in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing*, vol. 4,(Orlando), pp. 3313−3316, May 2002.