Quality Assessment of Gaussian Blurred Images Using Symmetric Geometric Moments

Chong-Yaw Wee, Raveendran Paramesran
Dept. of Electrical Engineering,
Faculty of Engineering,
University of Malaya,
50603 Kuala Lumpur, Malaysia.
ravee@um.edu.my

R. Mukundan
Dept. of Computer Science
and Software Engineering,
University of Canterbury,
Private Bag 4800, Christchurch, New Zealand.
mukundan@canterbury.ac.nz

Abstract

A novel objective full-reference image quality assessment metric based on symmetric geometric moments (SGM) is proposed. SGM is used to represent the structural information in the reference and test images. The reference and test images are divided into (8×8) blocks and the SGM up to fourth order for each block is computed. SGM of the corresponding blocks of the reference and test images are used to form the correlation index or quality metric of each block. The correlation index of the test image is then obtained by taking the average of all blocks. The performance of the proposed metric is validated through subjective evaluation by comparing with objective methods (PSNR and MSSIM) on a database of 174 Gaussian blurred images. The proposed metric performs better than PSNR and MSSIM by providing larger correlation coefficients and smaller errors after nonlinear regression fitting.

1 Introduction

One of the most challenging tasks in visual form analysis is the quality assessment (QA) of image or video acquired by a real, imperfect video or imaging system against a set of known, but diverse, reference image database. Image quality measurement metrics developed for this task normally incorporating human visual system (HVS) characteristics to improve their performance [6]. Objective image quality metrics can be divided into three categories based on their formulation approach as full-reference [11, 15], reduced-reference [16] and no-reference [7, 14].

In this paper, we are focusing on the full-reference (FR) image quality assessment where it deliver a single consistent metric value that generalizes the quality of an image or video sequence in various viewing conditions against a 'reference' image or video of 'perfect quality' (no loss in fidelity with the original scene) [6, 11, 12, 15]. It can automatically predict the quality of perceived image and video in fast and cost effective manner.

Most prevalently used FR objective metrics include the Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Signal-to-Noise Ratio (SNR) and Peak Signal-to-Noise Ratio (PSNR). These metrics were derived based on the statistical error sensitivity approach. The major advantage of this approach is its simplicity. Nevertheless, these FR objective metrics do not correlate well with perceived quality measurement because human perception of image/video distortions and artifacts is not taken into consideration. It is shown that by integrating with a very simple HVS model, the performance of the approach can be slightly improved [3].

We propose a novel objective FR image quality metric using symmetric geometric moments (SGM). SGM is applied in forming quality metric due to its capability in representing structural information of image content and evaluating correlation between moment kernels across scales and orientations of the natural scene of an image. The quality of an image is evaluated by dividing the image into (8×8) blocks, where the moment values of each block are computed before they are combined together to provide the quality perception of that image. The purpose of using a (8×8) image block is to effectively represent the structural information without interference from other regions of the image.

The organization of this paper is given as follow. The commonly used image quality assessment metrics such as PSNR and MSSIM as well as subjective evaluation is presented in section 2. The definition and characteristics of symmetric geometric moments (SGM) and the formulation of the SGM based image quality metric is presented in section 3. The performance of the proposed image quality met-

ric is validated through subjective evaluation experiment for Gaussian blurred images is shown in section 4. Section 5 concludes the study.

2 Background

The most important assumption in error sensitivity approach is that an image whose quality is being evaluated can be considered as a linear combination of an undistorted reference image and an error image. This approach attempts to objectively weigh or quantify the different aspects of strength of the error image. The simplest implementation of this approach is the mean square error (MSE) and peak signal-to-noise ratio (PSNR) which uses a Minkowski norm as their error pooling operator.

The PSNR between reference image $\mathbf{X}$ and test image $\mathbf{Y}$ of size $(M \times N)$ is determined by

$$\text{PSNR} = 20\log_{10}\left(\frac{255}{\sqrt{\text{MSE}}}\right) \tag{1}$$

where MSE is defined as

$$\text{MSE} = \frac{\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}[\mathbf{X}(x,y) - \mathbf{Y}(x,y)]^2}{(M \times N)} \tag{2}$$

Nevertheless, this Minkowski norm-based image quality metrics could not predict the quality of a distorted image with precision due to their inherent constraints for spatially pooling errors which explicitly assumed to be statistically independent at different locations. The visual perception quality of an image is highly dependent on the image content. Therefore, the assumption does not hold as pixels with the same value will give different image perception to the viewers in different environment [4], [13].

In the structural similarity approach, natural images are assumed to be highly structured, that is, their pixels exhibit strong dependencies, especially when they are spatially proximate. These dependencies carry important information about the structure of the objects in the visual scene [11], [13]. In addition, HVS is assumed to be highly adaptable to represent structural information from the viewing field. This implies that a measure of structural information change can provide a good approximation for perceived image distortion.

The structural similarity based image quality assessment metric (SSIM) which was proposed recently by Wang *et al.* between image blocks $\mathbf{x}$ and $\mathbf{y}$ is given as [13]

$$\text{SSIM}(\mathbf{x}, \mathbf{y}) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{3}$$

where $\{\mu_x, \sigma_x\}$ and $\{\mu_y, \sigma_y\}$ denote the mean intensity and standard deviation set of image blocks $\mathbf{x}$ and $\mathbf{y}$, respectively

while σ_{xy} denotes their correlation coefficient. C_1 and C_2 are constants of small values which are included to avoid instability when $(\mu_x^2 + \mu_y^2)$ and $(\sigma_x^2 + \sigma_y^2)$ are very close to zero respectively. This metric is implemented locally within a (8×8) square window, which moves pixel by pixel over the entire image. Through a mean operator, local quality indexes are then combined to provide a single index value for the image which is denoted as MSSIM.

Validation is an important step towards the successful development of practical image quality assessment models. It is essentially performed through a subjective evaluation associated with the prediction on perceived image quality by human vision system. Nevertheless, the subjective evaluation experiments are complicated by many aspects of human psychology and, also, of viewing conditions such as observer vision ability, translation of quality perception into ranking score, preference for content, adaptation, display devices and ambient light levels [15].

Two most prevalently used subjective evaluation methods are the Single Stimulus Continuous Quality Evaluation (SSCQE) and the Double Stimulus Continuous Quality Scale (DSCQS). These methods show repeatable and stable results as well as providing consistent viewing configuration and subjective tasks. Consequently, they have been adopted as parts of an international standard by the International Telecommunications Union (ITU) [1]. The raw scores obtained from multiple subjects in the subjective evaluation are averaged to yield the Mean Opinion Score (MOS). As an alternative to the MOS, the difference scores between test and reference images are averaged to yield the Different Mean Opinion Score (DMOS) [2].

3 SGM Based Image Quality Assessment Metric

Geometric moment (GM) of order $(p+q)$ of image function $f(x,y)$ is defined as [5], [9], [10]

$$M_{pq} = \int_0^\infty \int_0^\infty x^p y^q f(x,y)\mathrm{d}x\mathrm{d}y \tag{4}$$

where $x, y, p, q \in [0, \infty]$. The above definition shows the projection of the function $f(x,y)$ onto the domain of monomials $x^p y^q$. However, the basis set $\{x^p y^q\}$ is not orthogonal and the M_{pq} computed from it contain a certain degree of redundancy. SGM is obtained from (4) by restricting the limit of x, y within the interval of $[-1, 1]$.

The FR image quality metric proposed in this paper is based on the aforementioned SGM which measure the structural correlation between reference and test images. It has been inspired by the recent works done in [2], [11], [13]. The quality metrics proposed by Wang *et al.*, the Universal Quality Index (UQI) and Mean Structural Similarity

based Image quality assessment Metric (MSSIM), uses low order statistics such as mean and variance (standard deviation) together with correlation coefficients to represent the structural information between the test and reference images [11], [13], [15]. The main advantage of using low order statistics is their computational simplicity. However, they may not be sensitive enough to provide accurate yet consistent image quality scores due to their limited representation of structural information of images.

As the alternative, image statistics, SGM, where their order is determined by the moment order are used to develop a novel FR image quality metric. By using higher order of SGM, more corresponding image structural information is represented. The image quality metric derived based on the higher order moments (statistics) will provide more sensitive, precise and consistent quality scores in contrast to the metric which uses only the low order statistics such as the UQI and MSSIM.

The proposed SGM based image quality metric is formed by computing the SGM values from images which divided into small square image blocks. The correlation information of the computed moment values for the test and reference images is measured block by block before it is combined together to provide a single quality score. The procedures of the proposed metric are given as follow. The proposed image quality metric for an (8×8) image block, $\mathbf{y}$ from the test image $\mathbf{Y}$ compared to its corresponding block, $\mathbf{x}$, from the reference image $\mathbf{X}$ is given as

$$Q(\mathbf{x}, \mathbf{y}) = \frac{(2M_{\mathbf{x}}M_{\mathbf{y}})}{(M_{\mathbf{x}}^2 + M_{\mathbf{y}}^2)} \qquad (5)$$

where $M_{\mathbf{x}}$ and $M_{\mathbf{y}}$ denote the SGM computed up to $(p+q = 4)$ order for $\mathbf{x}$ and $\mathbf{y}$ image blocks respectively.

The computed score Q for $\mathbf{y}$ is the quality of that particular image block when compared to $\mathbf{x}$. In order to provide a single quality score for the test image, the mean value of Q is obtained as

$$Q(\mathbf{X}, \mathbf{Y}) = \frac{1}{H} \sum_{h=1}^{H} Q(\mathbf{x}_h, \mathbf{y}_h) \qquad (6)$$

where H is the number of (8×8) image blocks in either image $\mathbf{X}$ or image $\mathbf{Y}$. The block by block operation is important in maintaining the local quality since it minimizes the interference of particular region of the image by other regions. This is a crucial step in maintaining the overall quality information since different regions of an image can possess distinct visual effects on HVS. Hence, the block by block operation contributes to the robustness and the feasibility of the proposed image quality metric.

4 Experimental Study

Since the ultimate users of most image processing systems are the human visual system, a quantitative measure on the performance of the objective quality assessment models is required to evaluate their correlation and comparability with the human perception. Hence, an experiment which follows the performance evaluation procedures employed in the Video Quality Experts Group (VQEG) Phase I FR-TV test is performed [2]. The test images dataset used in this subjective evaluation consists of 174 color Gaussian blurred images which excerpted from the LIVE image dataset [8]. The objective quality score of all images is predicted using the PSNR, MSSIM and the proposed model, Q. The details regarding the procedures on how to obtain the subjective scores can be found from [8].

Logistic functions are used in a fitting procedure to provide a nonlinear mapping between the obtained subjective and objective scores. The nonlinear logistic fitted curves for the PSNR, MSSIM and proposed model are shown in Figure 1. The fitted curve obtained for the proposed model is better when compared to the one obtained for the PSNR and MSSIM especially for high quality images. For the proposed model, the logistic functions map the data points fairly and a nice and appropriate fitted curve is obtained. Almost all the data points are located at short and equal distance from and along the fitted curves except for the low quality end. This implies that they provide better quality perception which matches the human visual system.

After the nonlinear curve fitting process, three metrics are used to measure the correlation between the subjective and objective scores of all the tested models. The first metric is the correlation coefficient (CC) after nonlinear regression analysis to evaluate the prediction accuracy of the objective models to the subjective scores. The second metric is the Spearman rank-order correlation coefficient (SROCC) to evaluate the prediction monotonicity and unimodality of the objective models. The last metric is the outlier ratio (OR) which is used to evaluate the prediction consistency of the objective models. The last metric is computed by finding the percentage of the number of predictions which are outside the range of ± 2 times of the standard deviations. The mean absolute error (MAE) and the root mean squared prediction error (RMS) of the objective models are also computed after the nonlinear regression analysis. The evaluation results for all the models being compared are given in Table 1.

The results in Table 1 show the numerical and quantitative measures of the scatter plots observed in Table 1. Based on the values computed from the aforementioned 5 metrics (CC, MAE, RMS, OR and SROCC), the proposed SGM based model, Q performs better than PSNR and MSSIM in all categories except for OR. It provides larger correlation

Figure 1. Scatter plots of DMOS versus model prediction for Gaussian blurred images excerpted from LIVE image dataset.
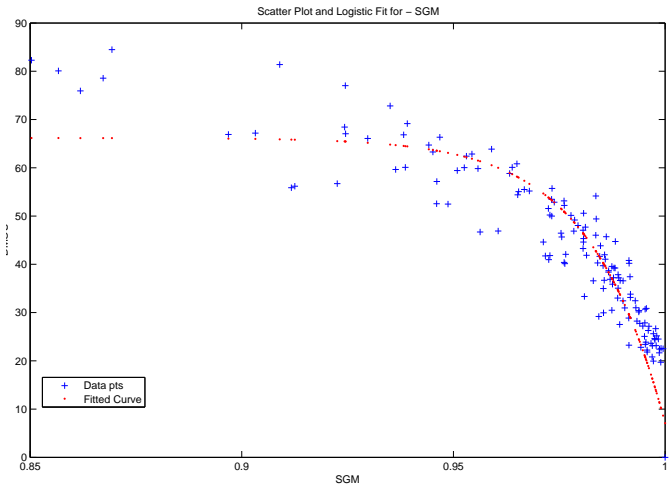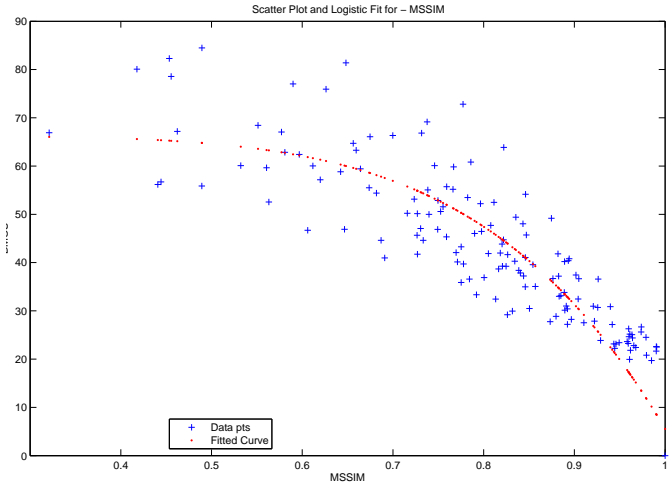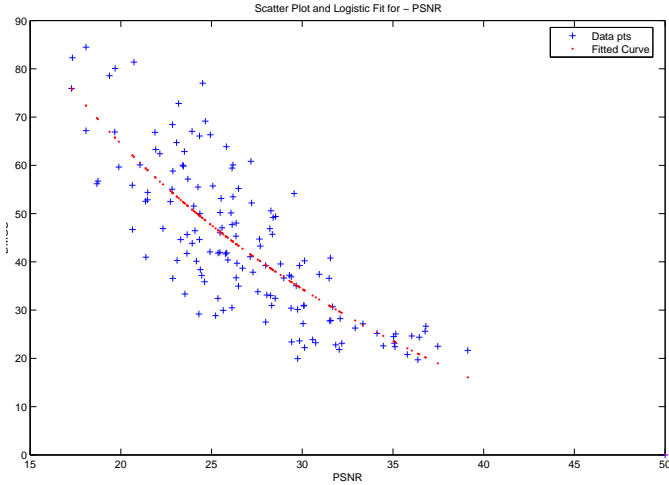
Table 1. Performance comparison of image quality assessment models for Gaussian blurred images.

Model	Non-linear Regression				Rank-order
	CC	MAE	RMS	OR	SROCC
PSNR	0.8398	6.3708	8.7068	0.0920	0.8797
MSSIM	0.8596	6.7206	8.1507	0.1782	0.9377
Q	0.8945	5.7792	7.0639	0.2644	0.9697

coefficients (CC and SROCC) when compared to PSNR and MSSIM to show its performance is highly correlated to human perception. It also provides smaller MAE and RMS values than PSNR and MSSIM after the nonlinear regression fitting. Nevertheless, it provides slightly larger OR value than PSNR and MSSIM. These validation results show the robustness and feasibility of the proposed model since the test image dataset includes a large amount of test images (174 images) of different scenes and blur conditions.

5 Conclusion

A novel full-reference image quality metric based on SGM is proposed to provide consistent yet precise predictions on the perception quality of Gaussian blurred images. This metric can be categorized as structural similarity approach metric since the SGM is used to represent the structural information of images. The correlation index or quality metric is obtained by taking the mean value of the computed index of each block. This operation is crucial in the derived metric for maintaining the local quality of particular region of an image. The performance of the proposed metric is validated through subjective evaluation where the results show that it performs better than PSNR and MSSIM by giving higher correlation coefficients (CC and SROCC) with the human visual perception. It also shows smaller MAE and RMS values but with a slightly larger OR after nonlinear regression fitting.

Acknowledgement

References

[1] Methodology for the subjective assessment of quality for television pictures. Technical Report ITU-R Rec. BT. 500-10, International Telecommunications Union (ITU).

[2] Final report from the video quality experts group on the validation of objective models of video quality assessment. On-line report, Video Quality Experts Group (VQEG), 2000. Available: http://vqeg.org/.

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

[4] B. Girod. What's wrong with mean-squared error. In A. B. Watson, editor, *Digital Images and Human Vision*, pages 207 – 220. MIT Press, 1993.

[5] R. Mukundan and K. R. Ramakrishnan. *Moment functions in image analysis*. World Scientific Publishing, Singapore, 1998.

[6] T. N. Pappas and R. J. Safranek. Perceptual criteria for image quality evaluation. In A. C. Bovik, editor, *Handbook of Image and Video Processing*, pages 669 – 684. Academic Press, San Diego, 2000.

[7] H. R. Sheikh, A. C. Bovik, and L. Cormack. No-reference quality assessment using natural scene statistics: Jpeg2000. *IEEE Trans. on Image Processing*, 14(11):1918 – 1927, Nov. 2005.

[8] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik. Live image quality assessment database release 2, 2005. Available: http//:live.ece.utexas.edu/research/quality.

[9] C. H. Teh and R. T. Chin. On image analysis by the methods of moments. *IEEE Trans. on Patt. Anal. and Mach. Intell.*, 10(4):496 – 512, Jul. 1988.

[10] C. H. Teh and R. T. Chin. On image analysis by the methods of moments. In *Proc. of Conf. on Computer Vision and Pattern Recognition (CVPR'88)*, pages 556 – 561, 1988.

[11] Z. Wang and A. C. Bovik. A universal image quality. *IEEE Signal Processing Letters*, 9(3):81 – 84, Mac. 2002.

[12] Z. Wang, A. C. Bovik, and B. L. Evans. Blind measurement of blocking artifacts in images. In *Proc. of IEEE Int. Conf. of Image Processing*, volume 3, pages 981 – 984, Sept. 2000.

[13] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Processing*, 13(4):600 – 612, April 2004.

[14] Z. Wang, H. R. Sheikh, and A. C. Bovik. No-reference perceptual quality assessment of jpeg compressed images. In *Proc. of IEEE Int.l Conf. on Image Processing*, pages 477 – 480, Sept. 2002.

[15] Z. Wang, H. R. Sheikh, and A. C. Bovik. Objective video quality assessment. In B. Furht and O. Marqure, editors, *The Handbook of Video Databases: Design and Applications*, chapter 41, pages 1041–1078. CRC Press, Boca Raton, FL, Sept. 2003.

[16] Z. Wang, E. P. Simoncelli, B. E. Rogowitz, T. N. Pappas, and S. J. Daly. Reduced-reference image quality assessment using a wavelet-domain natural image statistic model. In *PSIE Proc. of Human Vision and Electronic Imaging X*, volume 5666, pages 149 – 159, San Jose, CA, Jan. 2005.