EDGE-BASED STRUCTURAL SIMILARITY FOR IMAGE QUALITY ASSESSMENT

Guan-Hao Chen[1], *Chun-Ling Yang*[1], *Lai-Man Po*[2], *Sheng-Li Xie*[1]

[1]School of Electronic and Information Engineering, South China University of Technology, Guangzhou, Guangdong, 510640, China
[2]Department of Electronic Engineering, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon Tong, Hong Kong, China
eecghster@gmail.com, eeclyang@scut.edu.cn, eelmpo@cityu.edu.hk, adshlxie@scut.edu.cn.

ABSTRACT

Objective quality assessment has been widely used in image processing for decades and many researchers have been studying the objective quality assessment method based on Human Visual System (HVS). Recently the Structural Similarity (*SSIM*) is proposed, under the assumption that the HVS is highly adapted for extracting structural information from a scene, and simulation results have proved that it is better than *PSNR* (or *MSE*). By deeply studying the *SSIM*, we find it fails in measuring the badly blurred images. Based on this, we develop an improved method which is called Edge-based Structural Similarity (*ESSIM*). Experiment results show that *ESSIM* is more consistent with HVS than *SSIM* and *PSNR* especially for the blurred images.

1. INTRODUCTION

Image quality assessment plays an important role in image processing systems. Existing image quality evaluation methods can be divided into two categories: Subjective evaluation and objective evaluation. The HVS is our terminal of image processing systems, thus the most correct method of quantifying image quality is through subjective evaluation. In practice, however, subjective evaluation needs to organize the observers to mark the distorted images, which is too inconvenient, time-consuming and expensive. *PSNR* and *MSE* are still the most widely used objective metrics due to their low complexity and clear physical meaning. However they were also widely criticized for not correlating well with HVS for a long time.

During the last several decades, many researchers have tried to find a mathematic model to simulate HVS characteristics, and a great deal of effort has been made to develop new image quality assessment methods based on HVS. For example, Wen Xu and G. Hauske proposed to estimate the image quality based on segmentation error measure[1]. M. Miyahara, K. Kotani and V. R. Algazi had proposed a Picture Quality Scale (PQS) based on the characteristics of HVS and the structure and distribution of distortion[2]. In addition, other visual models based on visual interest are proposed too[3-5]. The majorities of the developed perceptual quality assessment models, however, are error-sensitivity approaches and follow a strategy of modifying the *MSE* measure so that errors are penalized in accordance with their visibility or interest.

Recently, a new philosophy for image quality measurement was proposed by Wang et al[6], based on the assumption that the HVS is highly adapted to extract structural information from the viewing field. According to this philosophy, the Structural Similarity (*SSIM*) is introduced to measure the distorted image quality, and simulation results show that it is more consistent with HVS than *PSNR* (*MSE*). In our study of SSIM, however, it was found that it fails in measuring badly blurred images. In this paper, we propose an improved quality assessment called edge-based structural similarity (*ESSIM*) based on the edge information as the most important image structure information.

The remainder of this paper is organized as follows. In section 2, the *SSIM* is simply introduced and analyzed. Section 3 describes the proposed *ESSIM* in detail. Section 4 presents the experimental results and their analysis. Finally, section 5 draws the conclusion.

2. STRUCTURAL SIMILARITY (SSIM)

2.1. Description of SSIM

Based on the assumption that the HVS is highly adapted to extract structural information from the viewing field, a new philosophy of *SSIM* for image quality measurement was proposed by Zhou Wang[6]. *SSIM* includes three parts: Luminance Comparison $l(\boldsymbol{x},\boldsymbol{y})$, Contrast Comparison $c(\boldsymbol{x},\boldsymbol{y})$ and Structure Comparison $s(\boldsymbol{x},\boldsymbol{y})$. *SSIM* is defined as:

$$SSIM(\boldsymbol{x},\boldsymbol{y}) = [l(\boldsymbol{x},\boldsymbol{y})]^{\alpha} \cdot [c(\boldsymbol{x},\boldsymbol{y})]^{\beta} \cdot [s(\boldsymbol{x},\boldsymbol{y})]^{\gamma} \tag{1}$$

(a) (b) (c)

Fig.1 Comparison of "Cameraman" images with different types of distortions, all with MSE = 1150. (a) Original image.(b) Gaussian noise contaminated image, MSSIM = 0.2591, ESSIM = 0.2510. (c) Blurred image, MSSIM = 0.5114, ESSIM = 0.1317.

The overall image quality can be evaluated by mean *SSIM* (*MSSIM*), which is defined as

$$MSSIM\ (\boldsymbol{X},\boldsymbol{Y}) = \frac{1}{M}\sum_{j=1}^{M} SSIM\ (\boldsymbol{x}_j, \boldsymbol{y}_j) \qquad (2)$$

From the definition of *SSIM*, the higher the value of $SSIM(\boldsymbol{x},\boldsymbol{y})$ is, the more similar the images $\boldsymbol{X}$ and $\boldsymbol{Y}$ are.

2.2. Analysis of *SSIM*

SSIM is significantly interesting for its novel theory and better results. However, we find *SSIM* fails in measuring badly blurred images. Let's see some results shown in Fig.1. The distortion images (b) and (c) in Fig.1 almost have the same *MSE*, but their visual quality are obviously different, the subjective quality of blurred image (Fig.1 c) is much worse than the Gaussian white noise contaminated image (Fig.1 b), but the *MSSIM* values are contrary to the perceptual quality, the blurred image has a higher *MSSIM* value than the noise contaminated image. In order to tackle this problem, the ESSIM is proposed in the next section.

3. EDGE-BASED STRUCTURAL SIMILARITY

Many researchers' studying results find that human eye is very sensitive to the edge and contour information of an image, that is, the edge and contour information may be the most important information of an image's structure for human to 'capture' the scene. Based on this thought, we propose an improved *SSIM* algorithm---Edge-based structural similarity (*ESSIM*), which compares the edge information between the distorted image block and the original one, and replace the structure comparison $s(\boldsymbol{x},\boldsymbol{y})$ in equation (1) by the edge-based structure comparison $e(\boldsymbol{x},\boldsymbol{y})$.

There are a number of ways to get the edge information, such as the simple edge detection algorithm, and the local gradients, etc. In this paper, the Sobel operator is used to obtain the edge information due to its simplicity and efficiency.

3.1.Edge Map

The edge map of an image is generated by using the Sobel operator. Fig. 2 shows the two 3×3 masks of Sobel operators used in this paper.

-1	0	+1
-2	0	+2
-1	0	+1

Vertical edge mask

−1	−2	−1
0	0	0
+1	+2	+1

Horizontal edge mask

Fig.2 Sobel operator masks

For each pixel $p_{i,j}$, its edge vector is defined as $\vec{D}_{i,j} = \{dx_{i,j}, dy_{i,j}\}$, where $dx_{i,j}$ and $dy_{i,j}$ are obtained by the vertical edge mask and horizontal edge mask respectively. The edge vector can be also represented as its amplitude and direction, the amplitude can be roughly estimated by

$$Amp_{i,j} = \left|dx_{i,j}\right| + \left|dy_{i,j}\right| \qquad (3)$$

The angle representing the pixel's edge direction is decided by

$$Ang_{i,j} = \frac{180^0}{\pi} \times \arctan(\frac{dy_{i,j}}{dx_{i,j}}) \qquad (4)$$

Each pixel in the image has an edge vector containing its edge amplitude and direction, and all the pixels' edge vectors form the image's edge map. It must be noted that in the actual implementation of the algorithm, equation (4) is not necessary. In this paper, the $dx_{i,j}/dy_{i,j}$ is used to express the pixel's edge direction.

3.2. Edge Direction Vector (Histogram)

The edge direction histogram[7] is used to compare the edge information between the distorted image blocks and the reference ones. In this paper, the continual direction (0^o~180^o) is divided into 8 discrete directions which are shown in Fig.3 and are not the same as reference [7].

For each image block, its edge direction histogram can be obtained by the following steps:

Step1. Calculate each pixel's edge amplitude and direction by equations (3) and (4) respectively.

Step2. Quantify each pixel's direction as one of the 8 discrete directions (see Fig.3).

Step3. Sum up all the pixels' edge amplitudes with the same direction in the block.

Now edge direction histogram (vector) is obtained. An example of a 16×16 image block and its edge direction histogram are showed in fig.4.

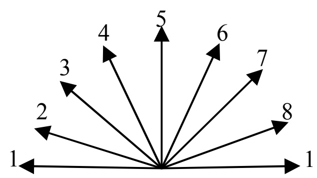


Fig.3 The 8 discrete directions

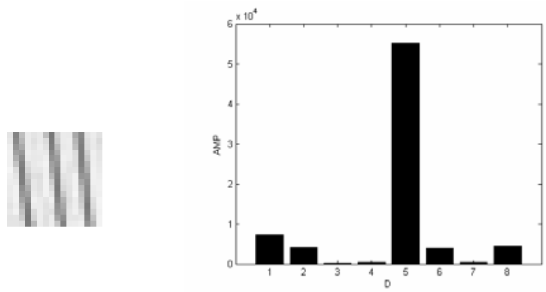


Fig.4 A 16 × 16 image block and its edge direction histogram, D represents the directions and AMP respresents the amplitude.

3.3 Edge comparison

Let D_x and D_y represent the original image block edge direction vector (histogram) and the distorted one respectively, then the edge comparison e(x,y) can be obtained by calculating the correlation coefficient of D_x and D_y, that is:

$$e(\boldsymbol{x},\boldsymbol{y}) = \frac{\sigma'_{xy} + C_3}{\sigma'_x \sigma'_y + C_3} \tag{5}$$

where σ'_x and σ'_y are the standard deviation of vector D_x and D_y respectively, σ'_{xy} is the covariance of vector D_x and D_y, and C_3 is a small constant to avoid the denominator being zero. And the edge-based structural similarity(*ESSIM*) is described as follows:

$$ESSIM(\boldsymbol{x},\boldsymbol{y}) = [l(\boldsymbol{x},\boldsymbol{y})]^{\alpha} \cdot [c(\boldsymbol{x},\boldsymbol{y})]^{\beta} \cdot [e(\boldsymbol{x},\boldsymbol{y})]^{\gamma} \tag{6}$$

The overall image structure similarity is calculated as the mean of all the subimages *ESSIM*.

$$MESSIM\ (X,Y) = \frac{1}{M}\sum_{j=1}^{M} \mathrm{E}SSIM\ (x_j, y_j) \tag{7}$$

4. SIMULATION RESULTS

4.1. Simulation Details

The proposed ESSIM's performance was evaluated based on the Live Image Quality Assess Database Release2 of the Laboratory for Image & Video Engineering in the University of Texas at Austin. A total of 489 distorted images were used in our simulations, including 169 JPEG2000 compressed images, 175 JPEG compressed images, and 145 Gaussian blurred images.

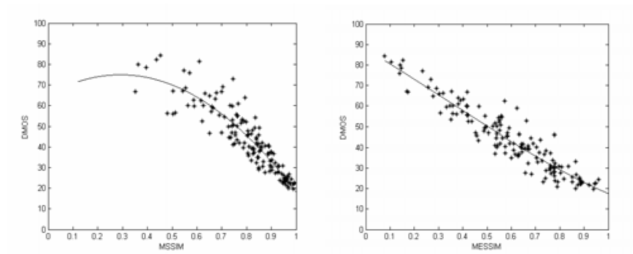
We compare the performance of the proposed *ESSIM* against *PSNR* and *SSIM*. For *ESSIM* and *SSIM*, each image is partioned into non-overlapping 8×8 blocks, the *ESSIM* and *SSIM* are calculated for each image block and then the *MESSIM* and *MSSIM* are obtained by equation (2) and equation (7), respectively. We choose the constant C_3 with the same value as that used in reference[6].

4.2. Discussion

The simulation results for 145 Gaussian blur distorted images are shown in Fig.5 and Table 1. Fig.5 shows the scatter plots of Difference Mean Opinion Score (DMOS) versus *MSSIM* and *MESSIM*. It is clear that the proposed *ESSIM* consistent with the subjective scores much better than *SSIM*.

Table 1 shows the quantitative measures of the performance of *ESSIM*, *SSIM* and *PSNR*, and five metrics[5] are used to measure these three objective models. The correlation coefficiences (CC) before and after non-linear regression means the correlation degree between each model and DMOS, they provide the prediction accuracy evaluation, and the large CC value means the better accuracy. The mean absolute error (MAE), root mean squared error (RMS) and outlier ratio (OR) after non-linear regression are measures of prediction consistency, and perform the different way from CC, small value means the better performance. We can see that *MESSIM* is better than *MSSIM* and *PSNR* in all the criteria.

The primary reason of performance improvement in *ESSIM* is that it pays more attention to the edges and details in images, which represents the higher layer image structure information.

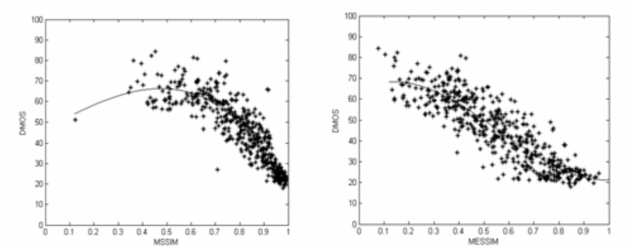


(a) (b)

Fig.5 Scatter plots of DMOS versus model prediction for Gaussian blur distorted images. (a) MSSIM and (b) MESSIM

Table 1. Performance comparison of image quality assessment models (PSNR, MSSIM, and the ESSIM) on Gaussian blurr distorted images

Model	CC	Non-linear Regression			
		CC	MAE	RMS	OR
PSNR	0.774	0.784	7.695	9.757	0.055
MSSIM	0.878	0.914	4.935	6.373	0.048
ESSIM	0.939	0.940	4.122	5.371	0.041



(a) (b)

Fig.6 Scatter plots of DMOS versus model prediction for JPEG2000, JPEG, Gaussian blur distorted images. (a) MSSIM and (b) MESSIM

Table 2. Performance comparison of image quality assessment models (PSNR, MSSIM, and MESSIM) for JPEG2000, JPEG and Gaussian blur distorted images

Model	CC	Non-linear Regression			
		CC	MAE	RMS	OR
PSNR	0.787	0.804	7.545	9.526	0.048
MSSIM	0.840	0.906	5.098	6.787	0.049
ESSIM	0.873	0.978	5.966	7.661	0.049

In addition to the performance comparison on the Gaussian blurred images, we test the JPEG2000 and the JPEG too. The scatter plots of the total 489 distorted images are shown in Fig.6, and the corresponding quantitative measures are listed in Table 2. *ESSIM* has better prediction accuracy with the DMOS than *PSNR* and *SSIM* (larger CCs), but its prediction consistency is a little worse than *SSIM* and better than *PSNR*.

We will study the image blocks properties deeply and improve our method in the future work.

5. CONCLUSIONS

In this paper we propose an edge-based structural similarity (*ESSIM*) for image quality assessment, which follows the HVS's characteristic that human eye is very sensitive to the edge and contour information of an image, and the edge and contour information is the most important strutural information for images. This may be the primary reason that our proposed *ESSIM* has better performance than *PSNR* and *SSIM*, especially as for the Gaussian blurred images.

6. ACKNOWLEDGMENTS

The work described in this paper was substantially supported by research projects from National Natural Science Foundation of China. [Project No. 60402015, No.60325310]

7. REFERENCES

[1] Wen Xu and G.Hauske, "Picture Quality Evaluation Based on Error Segmentation",SPIE vol.2308,pp.1454-1465,1994

[2] M.Miyahara,K.Kotani and V.Algazi,"Objective Picture Quality Scale (PQS) for Image Coding", Tech.Rep.CIPIC, University of California,Davis,1996

[3] Claudio M privitera, Lawrence W.stark, "Algorithms for Defining Visual Regions of Interest Comparison with Eye Fixation," IEEE trans on PAMI, Vol.22, No.9, pp.970-980,2000

[4] Wang Zhou, Bovik.A.C. "Wavelet-based foveated image quality measurement for region of interest image coding" Image Processing. Proceedings. Vol.2. pp.89 – 92, 2001

[5] Wang Kong-qiao,Shen Lan-sun,Xing Xin."A Quality Assessment Method of Image Based on Visual Interests". Journal of Image and Graphics. pp.300-303,2000

[6] Wang Zhou, Alan C. Bovik, Hamid R. Sheikh, and Eero P. Simoncelli, "Image Quality Assessment: From Error Measurement to Structural Similarity", IEEE Transactions on Image Processing, Vol. 13, No.4, pp600-613, April 2004.

[7] Pan Feng, Lin Xiao, and Susanto Rahardja, "Fast mode decision algorithm for intraprediction in H.264/AVC video coding " IEEE Trans.Circuits and Systems for Video Tech, pp813-822, Jul.2005.