GRADIENT-BASED STRUCTURAL SIMILARITY FOR IMAGE QUALITY ASSESSMENT

Guan-Hao Chen, Chun-Ling Yang, Member, IEEE, Sheng-Li Xie

School of Electronic and Information Engineering, South China University of Technology,
Guangzhou, Guangdong, 510640, China
eecghster@gmail.com, eeclyang@scut.edu.cn, 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 Gradient-based Structural Similarity (*GSSIM*). Experiment results show that *GSSIM* is more consistent with HVS than *SSIM* and *PSNR* especially for blurred images.

***Index Terms*—** Image processing, Image analysis

1. INTRODUCTION

Image quality assessment plays an important role in image processing systems. Existing image quality evaluation methods can be divided into two kinds: Subjective evaluation and objective evaluation. The HVS is the terminal of image processing systems, thus the most correct method of quantifying image quality may be subjective evaluation. In practice, however, subjective evaluation needs to organize the observers to mark the distorted images, that 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 have been also widely criticized for not correlating well with HVS for a long time.

During the last several decades, many researchers have tried to develop 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 these perceptual quality assessment models, however, are error-sensitivity approaches and follow a strategy of modifying the *MSE* measure, 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 it, the Structural Similarity (*SSIM*) is introduced to measure the distorted image's 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 failed in measuring badly blurred images. In this paper, we propose an improved quality assessment called Gradient-based Structural Similarity (*GSSIM*) 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 algorithm *GSSIM* in detail. Section 4 presents the experimental results and discussion. 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 Wang et al[6]. *SSIM* includes three parts: Luminance comparison $l(x,y)$, Contrast comparison $c(x,y)$ and Structure comparison $s(x,y)$. *SSIM* is defined as:

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

(a) (b) (c)

Fig.1 Comparison of "Cameraman" images with different types of distortions, all with MSE = 1150. (a) Original image.(b) Gaussian white noise contaminated image, MSSIM = 0.2591, MGSSIM = 0.2460. (c) Blurred image, MSSIM = 0.5114, MGSSIM = 0.1432.

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*(***x***,***y***) is, the more similar the images ***X*** and ***Y*** are.

2.2. Analysis of *SSIM*

SSIM is significantly interesting for its novel theory, low complexity 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 image contaminated with Gaussian white noise (Fig.1 b), while 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 solve this problem, the *GSSIM* is proposed in the next section.

3. GRADIENT-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, we develop an improved *SSIM* algorithm --- Gradient-based Structural Similarity (*GSSIM*), which compares the edge information between the distorted image block and the original one, and replace the contrast comparison $c(\boldsymbol{x},\boldsymbol{y})$ and structure comparison $s(\boldsymbol{x},\boldsymbol{y})$ in equation (1) with the gradient-based contrast comparison $c_g(\boldsymbol{x},\boldsymbol{y})$ and structure comparison $s_g(\boldsymbol{x},\boldsymbol{y})$ respectively.

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 operators are used to obtain the edge information due to its simplicity and efficiency.

3.1.Gradient Map

The gradient 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 gradient vector is defined as $\vec{E}_{i,j}=\{dx_{i,j},dy_{i,j}\}$, where $dx_{i,j}$ and $dy_{i,j}$ are obtained by the vertical edge mask and horizontal edge mask respectively, and the magnitude of this vector is roughly given by:

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

In keeping with tradition, we refer to the magnitude of gradient as gradient in the following. All the pixels' gradient form the image's gradient map.

The gradient map is obtained by using the two Sobel masks and equation(3). It shows the edge information of the images clearly, for the constant or slowly varying shades of gray have been eliminated. In other words, the edge and contour information of an image can be highlighted in its gradient map.

3.2 Edge comparison

We set $\boldsymbol{X}'$ and $\boldsymbol{Y}'$ represent the gradient map of the original image and the distorted one respectively, and let $\boldsymbol{x}'$ and $\boldsymbol{y}'$ be the block vectors of $\boldsymbol{X}'$ and $\boldsymbol{Y}'$, then gradient-based contrast comparison $c_g(\boldsymbol{x},\boldsymbol{y})$ and structure comparison $s_g(\boldsymbol{x},\boldsymbol{y})$ can be described by:

$$c_g(\boldsymbol{x},\boldsymbol{y}) = \frac{2\sigma_{x'}\sigma_{y'} + C_2}{\sigma^2_{x'} + \sigma^2_{y'} + C_2} \qquad (4)$$

$$s_g(\boldsymbol{x},\boldsymbol{y}) = \frac{\sigma_{x'y'} + C_3}{\sigma_{x'}\sigma_{y'} + C_3} \qquad (5)$$

where $\sigma_{x'}$ and $\sigma_{y'}$ are the standard deviation of vector $\boldsymbol{x}'$ and $\boldsymbol{y}'$ respectively, $\sigma_{x'y'}$ is the covariance of vector $\boldsymbol{x}'$ and $\boldsymbol{y}'$, and C_2, C_3 are small constants to avoid the denominator being zero. And the Gradient-based Structural Similarity(*GSSIM*) is described as follows:

$$GSSIM(\boldsymbol{x},\boldsymbol{y}) = [l(\boldsymbol{x},\boldsymbol{y})]^{\alpha} \cdot [c_g(\boldsymbol{x},\boldsymbol{y})]^{\beta} \cdot [s_g(\boldsymbol{x},\boldsymbol{y})]^{\gamma} \qquad (6)$$

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

$$MGSSIM\ (\boldsymbol{X},\boldsymbol{Y}) = \frac{1}{M}\sum_{j=1}^{M} GSSIM\ (\boldsymbol{x}_j, \boldsymbol{y}_j) \qquad (7)$$

4. SIMULATION RESULTS AND DISCUSSION

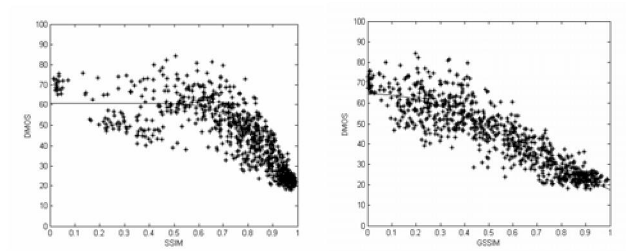
4.1. Simulation Details

The performance of the proposed *GSSIM* 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 779 distorted images were used in our simulations, including five types: JPEG, JPEG2000, white noise, Gaussian blur, and Fastfading (Transmission errors in the JPEG2000 bit stream using a fast-fading Rayleith channel model).

We compare the performance of the proposed *GSSIM* against *PSNR* and *SSIM*. As for *GSSIM*, each image is partitioned into overlapped 8×8 blocks, for SSIM, the image is partitioned into overlapped 11×11 blocks which is used in reference[6]. Then the *GSSIM* and *SSIM* are calculated for each image block and the *MGSSIM* and *MSSIM* are obtained by equation (7) and equation (2), respectively. The constants C_2 and C_3 are the same values as that used in reference[6].

4.2. Discussion

The simulation results for all the distorted images are shown in Fig.3 and Table 1. Fig.3 shows the scatter plots of Difference Mean Opinion Score (DMOS) versus *SSIM* and *GSSIM*. It is clear that the proposed *GSSIM* is consistent with the subjective scores much better than *SSIM*.

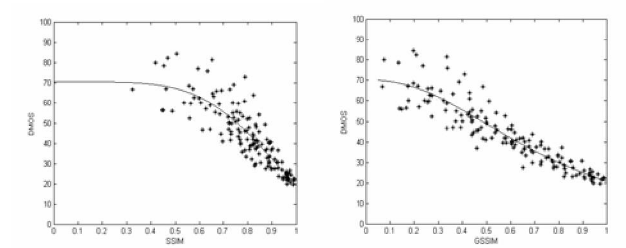


(a) (b)

Fig.3 Scatter plots of DMOS versus model prediction for JPEG2000, JPEG, Gaussian blur, white noise, and Fastfading distorted images. (a) SSIM and (b) GSSIM

Table 1. Performance comparison of image quality assessment models (PSNR, SSIM, and GSSIM) for JPEG2000, JPEG, Gaussian blur, white noise, and Fastfading distorted images.

Model	Non-linear Regression			
	CC	MAE	RMS	OR
PSNR	0.815	7.475	9.328	0.044
SSIM	0.864	6.269	8.113	0.053
GSSIM(proposed)	0.893	5.538	7.260	0.060



(a) (b)

Fig.4 Scatter plots of DMOS versus model prediction for Gaussian blur distorted images. (a) SSIM and (b) GSSIM

Table 2. Performance comparison of image quality assessment models (PSNR, SSIM, and GSSIM) on Gaussian blur distorted images

Model	Non-linear Regression			
	CC	MAE	RMS	OR
PSNR	0.783	7.756	9.777	0.048
SSIM	0.874	5.759	7.639	0.041
GSSIM(proposed)	0.917	4.541	6.284	0.055

Table 1 shows the quantitative measures of the performance of *GSSIM*, *SSIM* and *PSNR*. According to the video quality experts group (VQEG) Phase I FR-TV test [7], four metrics are used to measure these objective models. The correlation coefficiences (CC) after non-linear regression means the correlation degree between each model and DMOS, and the larger CC value means the better accuracy. The mean absolute error (MAE), root mean squared error (RMS) and outlier ratio (OR) are measures of prediction consistency, smaller value means better performance. We can see that *GSSIM* is better than *SSIM* and *PSNR* in all the criteria except OR.

In addition to the performance comparison on the whole database, the results only for the blurred images are given in Fig.4 and Table 2 in order to further show the performance of GSSIM. It is clear that the proposed G*SSIM* are in closer agreement with DMOSs than *PSNR* and *SSIM*.

The *GSSIM* of the distorted images in Fig.1 are given too (See Fig.1). Compare the *GSSIM* values with the corresponding *SSIM* values, we can see that the *GSSIM* is more rational than *SSIM*. Furthermore, the original image "Caps" and its two blurred images from the database with subjective scores(DMOS) are shown in Fig.5, and their *GSSIM* and *SSIM* are given, we can sum up the same conclusion that the proposed *GSSIM* is more consistent with HVS.

The primary reason of performance improvement in *GSSIM* is that it pays more attention to the edges and details of images, which represents the most important structural information of images.

5. CONCLUSIONS

In this paper we propose the Gradient-based Structural Similarity (*GSSIM*) 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 structural information of images. This may be the primary reason that the proposed *GSSIM* has better performance than *PSNR* and *SSIM*, especially as for the Gaussian blurred distorted 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 and 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.

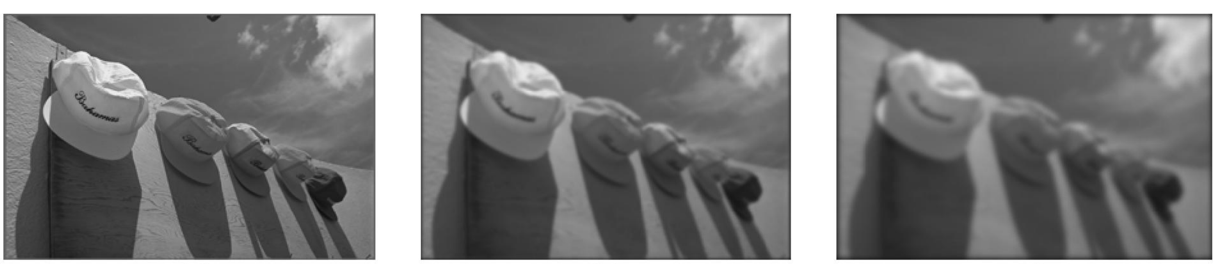
[4] Zhou Wang and Bovik.A.C. "Wavelet-based foveated image quality measurement for region of interest image coding" *ICIP2001 Proceedings*. Vol.2. pp.89 – 92, 2001.

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

[6] Zhou Wang, 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] VQEG. (2000, Mar.) Final Report From the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment. [Online] Available: http://www.vqeg.org/

[8] Rafael C. Gonzalez and Richard E. Woods, *Digital Image Processing (Second Edition),* Publishing House of Electronics Industry, 2002.



(a) (b) (c)

Fig.5 Sample Gaussian blur distorted image "caps" of different quality levels. (a) original image (b) DMOS = 60.8332, MSSIM = 0.7824, MGSSIM = 0.4310 (c) DMOS = 69.1550, MSSIM = 0.7436, MGSSIM = 0.3465.