Image Quality Assessment Using Gradient-weighted Structural Similarity

[1] Hongfang Li, [2,*] Shiru Zhang and [3] Yi-Ying Chang

Abstract

Digital images are subject to a wide variety of distortions during image processing application, and it is necessary to develop objective image quality metric to evaluate the degradation automatically. Images are prepared for human eyes so that the assessment result must be consistent with human visual effect.

Structure Similarity (SSIM), a well-known objective image quality assessment, is proposed by Zhou Wang. SSIM assumes that human visual perception is highly adapted to extracting structure information from a scene. Compared with PSNR or MSE，SSIM has a stronger advantage which has been proved in many different image quality assessments. However, due to the HVS characteristics of the underlying visual are neglected, SSIM has some drawbacks such as failing in blurred image measurement. In this paper, an improved SSIM method which we call it Gradient-Weighted SSIM (GWSSIM) is proposed based on the visual masking effect. GWSSIM performs in different regions of images which are weighted with different values based on their weight values. Experimental results show that GWSSIM has a better performance than both PSNR and SSIM.

Keywords: gradient, Structural Similarity, visual masking effect, image quality, HVS

1. Introduction

With the development of modern communications technology and business, and mankind into the digital age, information has been an essential element of the information society. It is well known that 75 percent of human information is obtained from images. Image information is so importance that it has been widely used in our modern life. However, in a series of processing procedures, such as image sampling, processing, storage and transmission, imperfections of processing methods, external devices and other factors will inevitably lead to image distortion or degradation, which is so-called the changed image quality perceived in human vision system. For example, in image acquisition, besides the imperfections of the imaging system, processing method, recording equipment and the transmission medium, the object movement, noise pollution and other factors will also result in image distortion and degradation and make our understanding of the objective world difficult. In image recognition, the accuracy and reliability of the collected image quality directly affects the recognition results. Again, due to the transmission errors, the impact of network latency and other unfavorable factors, all the teleconferencing and video on demand system need real-time image quality monitoring devices in order to dynamically adjust the source positioning strategy to ISP, and then meet the service quality requirements［1］. In military applications, battlefield surveillance and combat effectiveness also depends on the image or video quality obtained by the unmanned aerial vehicles and other aerial devices. Obviously, a reasonable assessment of image quality has a very important high value. Nowadays image information techniques are widely used so that the assessment of image quality becomes a fundamental problem in image processing field.

Image quality evaluation has a wide range of applications [2]: (l) It can dynamically detect and adjust the image quality. For example, in the image of the video process, we can get the image quality through the evaluation model to understand the current state of the video quality, and then reasonably adjust the video effect. (2) It can quickly evaluate the performance of image processing techniques, and can also be used for comparison of the performance of different image processing techniques.

For an image system it is necessary to have a quantitative description of image distortion degree, i.e., evaluation of image quality. Research of image quality evaluation can provide more accurate technical support ［3, 4］ for post-processing of images, and is extremely important to the image processing technology.

*Corresponding Author: Shiru Zhang
(E-mail: zsr0504@qq.com)
[1]Software Engineering Institute, Pass College of Chongqing Technology and Bussiness University, Hechuan, Chongqing 401520, P. R. China
[2]College of Communications and Information Engineering, Xi'an University of Science and Technology, No.58 Yanta Road, Yanta District, Xi'an, Shaanxi, 710054, P. R. China
[3]Department of Computer Science and Information Engineering, National Chin-Yi University of Technology, No.57, Sec.2, Chung-Shan Rd., Taiping District, Taichung City 41170, Taiwan

2. Structural Similarity of Image Quality Assessment

There are various deficiencies in HVS-based image quality evaluation methods. Zhou Wang proposed a new framework designed for the image quality - the structural similarity (Structure Similarity, SSIM) [5]. This framework derived from the human visual system is capable of highly adapting to the information hypothesis extracted from human visual sensing. Hence the change in the amount of structure information is able to reflect the approximation of human visual perception of an image distortion.

SSIM method includes three functions – luminance comparison function *l(x,y)*, contrast comparison function *c(x,y)*, and the structure comparison function *s(x,y)*. There are three important features related to SSIM:

1). Symmetry: SSIM(x,y)= SSIM(y,x)
2). Boundedness: SSIM(x,y)≤1;
3). Unique maximum: SSIM(x,y)=1 if and only if x=y (In discrete representations, $x_i=y_i$ for all i=1,2,…, N).

Suppose the original image block is X, the image block to be evaluated is Y, and the definition of the structural similarity model is [2,3,6]

$$l(x,y)=\frac{2\mu_x\mu_y+C_1}{\mu_x^2+\mu_y^2+C_1} \tag{1}$$

$$c(x,y)=\frac{2\sigma_x\sigma_y+C_2}{\sigma_x^2+\sigma_y^2+C_2} \tag{2}$$

$$s(x,y)=\frac{\sigma_{xy}+C_3}{\sigma_x\sigma_y+C_3} \tag{3}$$

$$SSIM(x,y)=[l(x,y)]^{\alpha}.[c(x,y)]^{\beta}.[s(x,y)]^{\gamma} \tag{4}$$

In order to avoid the problem resulting from $\mu_x^2+\mu_y^2$ very closing to zero, the constant C_1 will be $C_1=(K_1L)^2$, where L is the bit number of one pixel in the image, K_1 is an integer and $K_1\le 1$. $C_2=(K_2L)^2$, $C_3=C_2/2$, $\alpha>0$, $\beta>0$, $\gamma>0$. It is very easy to validate that they meet the above three features.

In general $\alpha=\beta=\gamma=1$, $C_3=C_2/2$. And this time Eq. (4) will be:

$$SSIM(x,y)=\frac{(2\mu_x\mu_y+C_1)(2\sigma_{xy}+C_2)}{(\mu_x^2+\mu_y^2+C_1)(\sigma_x^2+\sigma_y^2+C_2)} \tag{5}$$

The SSIM of an image is obtained by averaging the SSIM of every sub-image, see Eq. (6)

$$MSSIM(X,Y)=\frac{1}{M}\sum_{j=1}^{M}SSIM(X_j,Y_j) \tag{6}$$

where M is the number of all the sub-images.

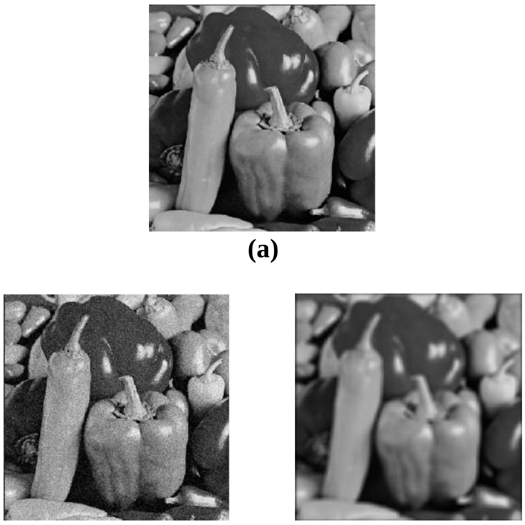
The higher value the Structural similarity is, the more similar are the two images X and Y.

PSNR is not consistent with human visual systems. We tested their performances of PSNR and SSIM, and gave the results of three 8-bit images of Lena, peppers and goldhill here. The original image and the distorted images are shown in Figures 1, 2, and 3, and the corresponding values are listed in Table 1, where (a) is original image; (b) is distorted by a Gaussian noise with 0.01 mean and 0.005 variance; (c) is blurred by mean filtering.

(a)

(b) (c)

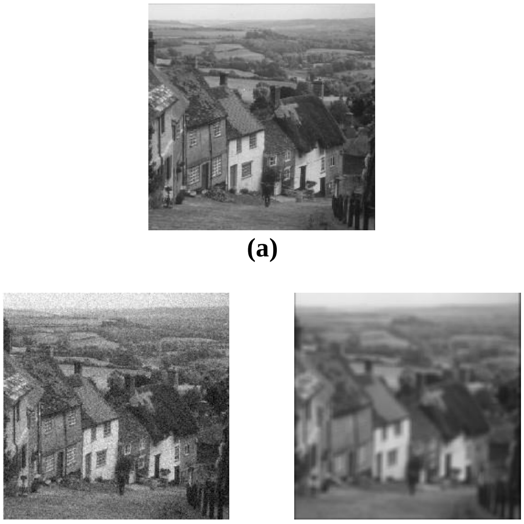
Figure 1: Lena original image and its distorted images



(a)

(b) (c)

Figure 2: Peppers original image and its distorted images



(a)

(b) (c)

Figure 3: Goldhill original image and its distorted images

Table 1: Values of distorted images

Tested images		PSNR（dB）	SSIM
lena	(b)	22.9542	0.9962
	(c)	23.5463	0.8508
peppers	(b)	23.0660	0.9898
	(c)	23.2736	0.8658
goldhill	(b)	22.9647	0.9747
	(c)	23.4782	0.8526

Compared with the different images by our eyes, we find that Gaussian noise distorted images are better than that of blurred images. The PSNR values give the negative results, but the SSIM values show the consistent result. Therefore, SSIM is better than PSNR.

However, the simple linear model of SSIM is difficult to describe the complex processing of human visual into the image information. On the other hand the neglecting of HVS characteristics results in some different results with the subject ones. Again, SSIM is based on the consumptions that different regions have the same importance, which is not consistent with the practical situation. Hence SSIM is not the optimum image quality parameter. Sometimes it will not be consistent with the human visual characteristics.

3. Gradient-weighted Structural Similarity

For a seriously blurred image, SSIM evaluation results are not correct. Research has shown that human eyes are very sensitive to the edge portion of the image. It is known that the gradient value may well reflect the change in the image contrast and fine details of the texture. Therefore, we use the gradient value as the main structure to evaluate the image definition, and further propose the gradient-weighted structure distortion of image quality assessment method.

In digital image processing, the gradient magnitude reflects the texture changes among the adjacent pixels, and the gradient magnitude is relatively large in the edge portions. We first calculate the gradient magnitude values in the image, and then extract the image edge according to the gradient magnitudes. We use Sobel operator [6], which includes two operators in two different directions, i.e., the horizontal edges operator and vertical edges operator.

$$S_V = \left\{ \begin{matrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{matrix} \right\} \tag{7}$$

$$S_H = \left\{ \begin{matrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 2 & 1 \end{matrix} \right\} \tag{8}$$

The gradient structure similarity of image block x, y is defined as:

$$g(x,y)=\frac{2\sum_j\sum_i G_x(i,j)G_y(i,j)+C_3}{\sum_j\sum_i[G_x(i,j)]^2+\sum_j\sum_i[G_y(i,j)]^2+C_3} \quad (9)$$

where G_x G_y are the gradient values along horizontal and vertical directions between the original image and the distorted image respectively.

The gradient values instead of the structure comparison factor are used to calculate the gradient values by Sobel gradient operator. The weight values are calculated according to the gradient values of reference image X and the tested image Y.

Suppose the initial state is empty, take appropriate threshold *T*, and do the following steps:

Step 1: Find the magnitude g_m of the gradient value at the point (x,y).

$$g_m(x,y)=\sqrt{g_x^2(x,y)+g_y^2(x,y)} \quad (10)$$

Step 2: Find the threshold of the image gradient values.

$$T=\sqrt{\sum_{i=1}^{N}g_{m,i}^2(x,y)} \quad (11)$$

where $g_{m,i}$ is the gradient magnitude value of the ith sub-image block, and N is the total number of the sub-image in the whole image.

Step 3: Classify edge, texture and flat areas.

If $g_m < T$, it belongs to a flat area; if $g_m > T$, it belongs to an edge area or texture area.

Step 4: Define the weight function. Divide the image into 11*11 sub-images. Then a logarithmic log curve is used to modify the weight function.

$$W_j=\begin{cases}\log(G_{xj}(|G_{xj}-G_{yj}|)^2) & (G_{xj}-T)(G_{xj}-G_{yj})\geq 0\\ \log(G_{xj}(|G_{xj}-G_{yj}|)) & (G_{xj}-T)(G_{xj}-G_{yj})<0\end{cases} \quad (12)$$

Step 5: The boundary condition is generally 0≤ SSIM(x,y)≤1, but if and only if x=y then SSIM=1.

Step 6: Obtain GWSSIM.

By using gradient information instead of the structure comparison factor in SSIM, we can obtain the GWSSIM value defined as Eq. (13).

$$GWSSIM(x,y)=\frac{\sum_{j=1}^{N}(W_j(x_j,y_j)\cdot SSIM(x_j,y_j))}{\sum_{j=1}^{N}W_j(x_j,y_j)} \quad (13)$$

Because the value of the structure comparison factor s (x, y) in SSIM is in the range [-1, 1], which does not meet the boundary condition, negative values may occur and lead to an unreasonable result. In order to overcome this drawback, *s(x,y)* will be:

$$s(x,y)=\frac{\sigma_{xy}+\sigma_x\sigma_y+c_3}{2\sigma_x\sigma_y+c_3} \quad (14)$$

The principle of the proposed GWSSIM algorithm is shown in Figure 4, and the detail computing procedure is shown in Figure 5.

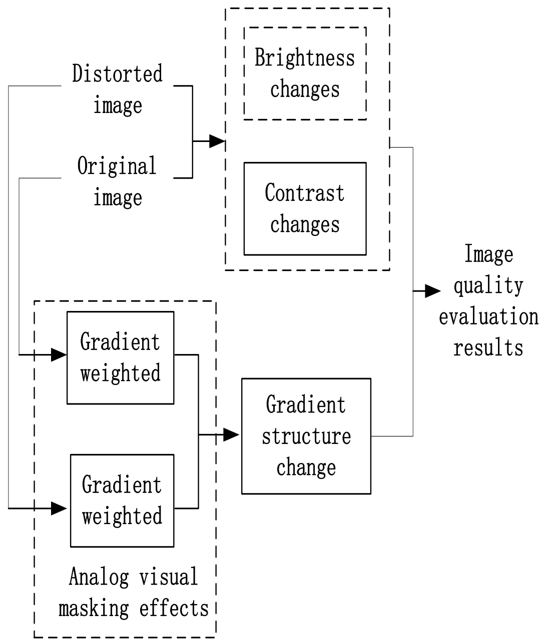


Figure 4: Principle of GWSSIM algorithm

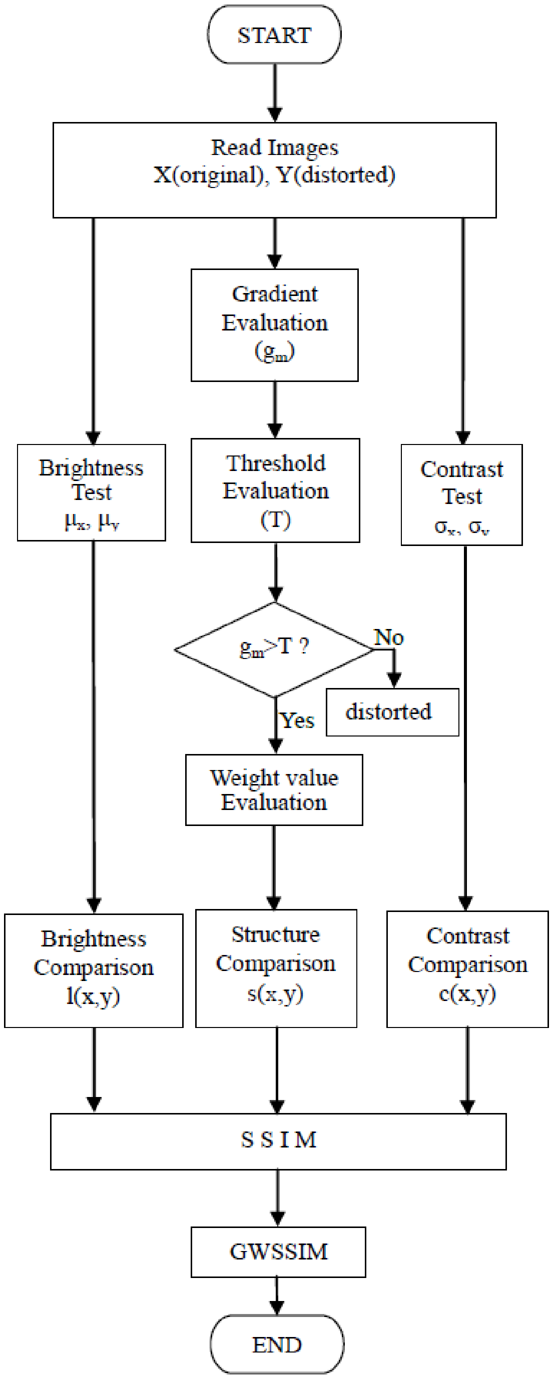


Figure 5: Flow chat of GWSSIM algorithm

4. Experiment and Analysis

SSIM evaluation method is not consistent with the actual situation in terms of blurred images. We select LIVE Database2 [7] as our image library and focus on the Gaussian blur images.

All the experiments in this paper were conducted under the Matlab2010a environment. Many images are used to verify the proposed algorithm. Due to the limited space here, we only list the experimental results of Lena.

We converted the RGB images into gray-scale images before computing. The SSIM evaluation results are obtained by running Matlab programs provided by Zhou Wang.

Since SSIM is not sensitive to its parameters, we can set them simply as $\alpha=\beta=\gamma=1$. In principle, α, β, and γ should be different values according to the different distortion components, but generally it can be taken as $\alpha=\beta=\gamma=1$.

In order to verify the accuracy of the above method, we focus on the LIVE image database and take the image "lena" as an example. The degraded images obtained through MATLAB are given in Figure 6, where (a) is the original image; (b) is the image after adding a Gaussian noise of zero-mean and 0.02 variance; (c) is the image after adding 6% salt and pepper noise; (d) is the image after adding a multiplicative noise with variance of 0.02; (e) is the image after Gaussian blurred. The values of PSNR, SSIM and GWSSIM are shown in Table 2.

From Figure 6 and Table 2, we can see different distortion images and their quality indexes of PSNR, SSIM and GWSSIM. In terms of the human eyes, the image of adding salt & pepper noise is better than Gaussian blur images, and the image of adding multiplicative noise is slightly better than Gaussian noise images. But the corresponding GWSSIM values of salt & pepper noise and Gaussian blur are 0.7222 and 0.6388 respectively, while SSIM values of salt & pepper noise, Gaussian blur are 0.6644, 0.7684 respectively. SSIM evaluation shows that the fuzzy image evaluation is ineffective, and our GWSSIM is better than SSIM when blurred images are tested.

Table 2: Parameters involved in the experiments

K_1	K_2	L	α	β	γ
0.01	0.03	255	1	1	1

(a) Original image

(b) Gaussian noise

(c) Add salt & speckle noise

(d) Add multiplicative noise

(e)Gaussian blur

Figure 6: Different distorted images of Lena

Table 3: Image quality index of the distorted images of lena corresponding to Figure 6

quality Index	(b)	(c)	(d)	(e)
PSNR	19.985	21.845	22.838	24.957
SSIM	0.4080	0.6644	0.5717	0.7684
GWSSIM	0.3684	0.7222	0.4865	0.6388

In order to further illustrate the performance superiority of GWSSIM algorithm, blurred images with different parameters, as shown in Figure 7, were obtained by MATLAB. The corresponding quality indexes of PSNR, SSIM and GWSSIM are shown in Table 3.

(a) Original image

(b) Blurred image ($\sigma = 0.1$)

(c) Blurred image ($\sigma = 0.3$)

(d) Blurred image ($\sigma = 0.5$)

(e) Blurred blur ($\sigma = 50$)

Figure 7: Blurred images with different parameters

Table 4: Image quality index of the blurred images corresponding to Figure 7

Parameter σ	0.1	0.3	0.5	50
PSNR	Inf	63.3338	35.1563	24.2499
SSIM	1	1.0000	0.9755	0.7349
GSSIM	1	0.9994	0.9311	0.5942

From Table 4, for σ=0.1 and σ=0.3, SSIM value is the same, but GWSSIM has two different values. Obviously the proposed GWSSIM method is better than that of SSIM for the blurred images.

The objective quality assessment values are computed for same images by using different models. We compare the results with that of subjective perception of quality value provided by the database (DMOS). The comparison data are shown in Table 5.

Table 5: Performance comparison of different image quality evaluation models for Gaussian blurred image library

Quality index	CC	MAE	RMSE	OR
PSNR	0.7834	7.7391	9.7623	0.0490
SSIM	0.8740	5.7753	7.6386	0.0421
GWSSIM	0.9182	4.5312	6.2253	0.0710

First, as we know, the larger CC indicates the better correlation of predicted quality values with DMOS. From Table 5 we find that the CC value of GWSSIM is the largest one. Second, the smaller value of MAE, RMSE and OR indicates the smaller prediction error of the model. From Table 5 we find we have the smallest values of MAE and RMSE.

From all above experimental results on Gaussian blur distorted image library, we find that the proposed GWSSIM is better than conditional PSNR and SSIM. The significant improvement comes from the sufficient consideration of human visual characteristics.

5. Conclusions

Based on the SSIM, we propose a new image quality evaluation method - GWSSIM. In principle, GWSSIM overcome the negative values phenomenon in SSIM algorithm. Due to taking gradient value as the structural information, GWSSIM highly agrees with the human visual perception characteristics, so that it can better descript the image quality, especially for blurred images, and of course for other types of degradation images. Again, computer simulation experiments also show that the proposed GWSSIM model outperforms SSIM and PSNR model.

GWSSIM model only uses the magnitude of the gradient, so the gradient direction will affect the image quality evaluation results. Therefore, how to fully use gradient values in evaluating image quality is our next research goal in the near future.

References

[1]. Lou Bin, Shen Hai-bin, Zhao Wu-feng and Yan Xiao-lang, “No-reference image quality assessment based on statistical model of natural image”, Journal of Zhejiang University (Engineering Science), Vol. 44, No. 2, pp.248-252, 2010. (In Chinese)

[2]. Ye sheng-nan, Su Kai-na, Xiao Chuang-bai, Duan Juan, “Image quality assessment based on structural information extraction”, ACTA electronica SINICA, Vol. 36, No. 5, pp.856-861, 2008 (In Chinese)

[3]. Zhou Wang, Alan C.Bovik, Hamid R. Sheikh, Eero P. Simoncelli. "Image quality qssessment: From error visibility to structual similarity", IEEE Transactions on Image Processing, Vol.13, No.4, pp.600-612. 2004.

[4]. Vittoria Bruni, Domenico Vitulano, Zhou Wang, “Special issue on human vision and information theory”, Signal, Image and Video Processing, Vol. 7, No.3, pp389-390, 2013.

[5]. Zhou Wang, Ligang Lu, Alan C. Bovik, “Video quality assessment based on structural distortion measurement”, Signal Processing: Image Communication, Vol. 19, No. 2, pp. 121- 132, 2004.
[6]. Duan ying-ying, Ma Jian-hua, Chen Wu-Fan, et al. “Improved SSIM medical image quality assessment”, Computer Engineering and Applications, Vol. 46, No. 2, pp145-149, 2010. (In Chinese)
[7]. 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.

Hongfang Li was born in 1985, Baoji, Shaanxi, China. She received her Master’s degree in Xi’an University of Science and Technology in 2012. Her research direction is digital image processing.

Shiru Zhang (used name Minrui Zhang) was born in Xian, China, on July, 1965. She received her Ph.D. degree in signal and information processing from Xidian University in China in 2005. She is with Xi’an University of Science and Technology in China, and now is a professor there and a visiting professor of National Chin-Yi University of Technology in Taiwan. She was a visiting scholar in UIUC, Illinois, USA. from May 2008 to Nov. 2009. She is a member of IEEE, a senior member of China Institute of Communications, a member of China Society of Image and Graphics, a vice general secretary of Shaanxi Signal Processing Society. Her current research interests include information hiding and digital image processing.

Yi-Ying Chang received his Bachelor’s degree in Electronic Engineering (1982), and his Master’s degree in Computer Information Science (1997) from the Department of Computer Information & Science Engineering in the Knowledge System Institute, Skokie, Illinois, USA. He received his Ph.D. in Electrical Engineering degree (2011) from the Electrical Engineering Department in National Cheng Kung University, Tainan, Taiwan. His research interests include image enhancement and segmentation. He joined the National Chin-Yi University of Technology in 1982, and he is an associate professor in the Department of Computer Science and Information Engineering at the National Chin-Yi University of Technology.