International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 3 Issue 4, April - 2014

Image Quality Assessment Techniques: An Overview

Shruti Sonawane
Department of E&TC
TSSM's BSCOER, Pune ,
Pune University, Maharashtra, India

A. M. Deshpande
Department of E&TC
TSSM's BSCOER, Pune ,
Pune University, Maharashtra, India

***Abstract*—The image quality assessment (IQA) provides computational models to measure the perceptual quality of an image. The paper gives the comparative evaluation of different objective full reference IQA (FR-IQA) schemes such as mean squared error (MSE), peak signal to noise ratio (PSNR), structural similarity index measure (SSIM) and gradient based similarity index (GSI) can be used to evaluate quality of an image. MSE and PSNR are conventional and widely used methods. The SSIM is based on human visual system which can extract the structural changes in the image. The GSI relies on the principle that image gradients can be used to extract important visual information such as contrast and structural changes in the image with reference to a reference image. The experiments are conducted with publicly available subject-rated database i.e. Laboratory for Image and Video Engineering (LIVE) database images.**

***Index Terms*— Image quality assessment (IQA), Human visual system (HVS), Mean squared error (MSE), Peak signal to noise ratio (PSNR), Structural similarity (SSIM), Gradient based similarity.**

I. INTRODUCTION

Image quality measures play an important role in a variety of image processing applications. The goal of *image quality assessment* (IQA) is to provide quality measure which can be used to evaluate the performance of image processing systems. The IQA provides computational models to measure the perceptual quality of an image. The distortions in an image may be introduced during acquisition, transmission, compression, restoration, and processing. A large number of methods have been designed to evaluate the quality of distorted images. IQA methods can be categorized into subjective and objective methods.

Subjective methods are based on human judgment and thus they are inconvenient, time consuming and do not provide automation to the system [1]. Subjective methods are applicable where images are ultimately to be viewed by human beings, the method of quantifying visual image quality is through subjective evaluation. But due to the drawbacks, they cannot be easily performed for many scenarios, (e.g., real time systems) and are impossible to be included in automatic systems. The goal of objective image quality assessment is to design quality measures that can automatically predict perceived quality of image. The metrics used to measure image quality can be classified according to availability of an original (distortion-free) image, with which the distorted image is to be compared.

The availability of a reference image decides the classification of objective quality metrics. They are categorized as: *full-reference* (FR), *no-reference* (NR), and *reduced-reference* (RR) methods. In FR image quality assessment methods, the quality of a test image is evaluated by comparing it with a reference image which is assumed to have perfect quality. NR metrics try to evaluate the quality of an image without any reference image. In RR method, the reference image is only partially available, in the form of a set of extracted features which help to evaluate the quality of the distorted image.

To provide a compromise between FR and NR, RR methods have been designed for IQA by employing partial information of the corresponding reference.

In this paper, section II describes some widely used FR IQA methods: peak signal-to-noise ratio (PSNR) and mean square error (MSE), structural similarity approach i.e. Structural similarity index (SSIM), gradient based similarity scheme. Experimental evaluation is given in section III and a brief conclusion is given in section IV.

II. FULL REFERENCE IMAGE QUALITY ASSESSMENT METHODS

There are number of FR IQA methods available in practice. Mostly used and conventional schemes are MSE and PSNR. Some other quality measures are visual signal to noise ratio (VSNR), most apparent distortion (MAD), visual information fidelity (VIF), SSIM and gradient based similarity index.

A. MSE and PSNR

The conventional FR IQA methods i.e. *peak signal-to-noise ratio* (PSNR) and *mean square error* (MSE) calculate pixel-wise distances between a distorted image and the corresponding reference image. These methods are pixel-based, i.e. the distorted and the reference images are compared pixel-by-pixel. Given a reference image f and a test image g, both of size $M \times N$, the PSNR between f and g is defined as

$$PSNR = 10\ log_{10}\left(\frac{m^2}{MSE}\right)\ \text{[dB]}, \qquad (2)$$

where m is the maximum pixel value (e.g. 255 for 8-bit images) and

$$MSE = \left(\frac{1}{MN}\right) \sum_{i=1}^{M} \sum_{j=1}^{N} \left(f_{ij} - g_{ij}\right)^2 \quad (3)$$

The PSNR value approaches infinity as the MSE approaches zero. This shows that a higher PSNR value provides a higher image quality. These metrics are easy to calculate, they have clear physical meaning but their correlation with perceived image quality has been proven to be low. The performance of MSE is extremely poor in the sense that images with nearly identical MSE are drastically different in perceived quality [2], [3]. It is well known that MSE/PSNR does not always agree with the subjective viewing results, particularly when distortion is not additive in nature.

B. Structural Similarity Based Image Quality Assessment

The structural similarity approach is based on the assumption that the human visual system is highly adapted to extract structural information from the viewing field. It follows that a measure of structural information change can provide a good approximation to perceived image distortion. It considers image degradations as perceived changes in structural information [4]. The Structural similarity index (SSIM) is based on comparing the structures of the reference and the distorted images. The structural information in an image can be defined as those attributes that represent the structure of objects in the scene, independent of the average luminance and contrast.

The system separates the task of similarity measurement into three comparisons: luminance, contrast and structure. First, the luminance of each signal is compared. Assuming discrete signals, this is estimated as the mean intensity:

$$\mu_x = \frac{1}{N} \sum_{i=1}^{N} x_i \quad (4)$$

The luminance comparison function *l(x, y)* is then a function of μ_x and μ_y. The standard deviation (the square root of variance) can be used as an estimate of the signal contrast. An unbiased estimate in discrete form is given by

$$\sigma_x = \left(\frac{1}{N-1} \sum_{i=1}^{N} \left(x_i - \mu_x\right)^2\right)^{\frac{1}{2}} \quad (5)$$

The contrast comparison *c(x,y)* is then the comparison of σ_x and σ_y. Third, the signal is normalized (divided) by its own standard deviation, so that the two signals being compared have unit standard deviation. The structure comparison *s(x,y)* is conducted on these normalized signals $(x - \mu_x)/\sigma_x$ and $(y - \mu_y)/\sigma_y$. Finally, the three components are combined to yield an overall similarity measure:

$$S(x, y) = f(l(x, y), c(x, y), s(x, y)) \quad (6)$$

For luminance comparison following equation is used,

$$l(x, y) = \frac{2\mu_x \mu_y + C_1}{{\mu_x}^2 + {\mu_y}^2 + C_1} \quad (7)$$

where the constant C_1 is included to avoid instability when ${\mu_x}^2 + {\mu_y}^2$ is very close to zero. Specifically, $C1 = (K_1 L)^2$, where L is the dynamic range of the pixel values (255 for 8-bit grayscale images), and K_1 << 1 is a small constant. The contrast comparison function takes a similar form:

$$c(x, y) = \frac{2\sigma_x \sigma_y + C_2}{{\sigma_x}^2 + {\sigma_y}^2 + C_2} \quad (8)$$

where, $C_2 = (K_2 L)^2$ and $K_2 << 1$

The structure comparison function is as follows:

$$s(x, y) = \frac{\sigma_{xy} + C_3}{(\sigma_x \sigma_y + C_3)} \quad (9)$$

Structural Similarity (SSIM) index between signals x and y is,

$$\mathrm{SSIM}(x, y) = [l(x, y)]^{\alpha}.[c(x, y)]^{\beta}.[s(x, y)]^{\gamma} \quad (10)$$

where $\alpha > 0$, $\beta > 0$ and $\gamma > 0$ are parameters used to adjust the relative importance of the three components. This results in a specific form of the SSIM index:

$$\mathrm{SSIM}(x, y) = \frac{\left(2\mu_x \mu_y + C_1\right)\left(2\sigma_{xy} + C_2\right)}{\left({\mu_x}^2 + {\mu_y}^2 + C1\right)\left({\sigma_x}^2 + {\sigma_y}^2 + C_2\right)} \quad (11)$$

The SSIM indices are calculated within the sliding window, which moves pixel-by-pixel from the top-left to the bottom-right corner of the image. This results in a SSIM index map of an image, which is also considered as the quality map of the distorted image being evaluated. The overall quality value is defined as the average of the quality map – the mean SSIM (MSSIM) index.

The SSIM provides the distortion/similarity map in the pixel domain [4]. The SSIM is widely accepted due to its reasonably good evaluation accuracy, pixel wise quality measurement, and simple mathematical formulation, which facilitates analysis and optimization. But it is less effective for badly blurred images since it underestimates the effect of edge damage and treats every region in an image equally.

C. Gradient-based Similarity Approach

As edge information and differentiated distortion at the edges are important for perceptual quality gauging, new IQA scheme based on the edge/gradient similarity has been recently proposed [6], [7]. Similar to the SSIM, this scheme considers luminance and contrast–structural changes.

The gradient similarity is defined as:

$$g(x, y) = \frac{2g_x g_y + C_4}{{g_x}^2 + {g_y}^2 + C_4} \quad (12)$$

where g_x and g_y are the gradient values for the central pixel of image blocks x and y, respectively, and C_4 is the small constant to avoid the denominator being zero. $g(x, y)$ is the gradient similarity between two images and its value lies in the range [0, 1]. Fig.1 shows gradient-based similarity measure scheme.

Gradient value g_x is calculated as the maximum weighted average of difference for the block (same for g_y):

$$g_x = \max_{k=1,2,3,4}\{mean_2(|x.M_k|)\} \quad (13)$$

with, $M_k (k = 1,2,3,4)$ as shown in Fig.1, where the weighting coefficient decreases as the distance from the central pixel increases, and mean2 is the mean value for a matrix. $g(x, y)$can be interpreted either as a blockwise version (i.e., gradient similarity for image blocks *x* and *y*) or a pixel wise version (i.e., the gradient similarity for the central pixels of image blocks *x* and *y*) i.e.

ISSN: 2278-0181
Vol. 3 Issue 4, April - 2014

$$g(x_c, y_c) = g(x, y) \qquad (14)$$

where x_c and y_c are the central pixels of image blocks x and y respectively. The formulation for $g(x, y)$ is able to measure both image contrast (the degree of signal variation)

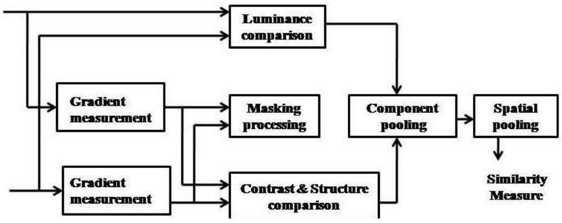


Fig.1: Block diagram of Gradient-based similarity scheme

change and image structure (structure of objects in the scene) change since the gradient value (i.e., g_x and g_y) is a contrast-and-structure variant feature [6]. Although both standard deviation and gradient can be used to measure the contrast, their difference is as follows. For given a group of pixel values, the standard deviation for the pixels is a constant, no matter how these pixels are positioned (i.e., independent of pixel positioning), whereas the gradient values for the same group of pixels change according to the positioning of these pixels. $g(x, y)$ can be rewritten as,

$$g(x, y) = \frac{2(1-R)+K}{1+(1-R)^2+K} \qquad (15)$$

with the masked gradient change defined as

$$R = \frac{|g_x - g_y|}{\max(g_x, g_y)} \qquad (16)$$

Where, $= \frac{C_4}{\max(g_x, g_y)^2}$. The value of gradient change R lies in the range of [0, 1]. The $g(x, y)$ is less sensitive to the case of higher masking contrast than that of lower masking contrast, and this is consistent with the contrast masking of the HVS for high masking contrast[3]. For this approach to be matched better with the masking effect and visibility threshold, $g(x, y)$ is further modified as

$$g(x, y) = \frac{2(1-R)+K'/\max(g_x, g_y)}{1+(1-R)^2+K'/\max(g_x, g_y)} \qquad (17)$$

Where, $K = K'/\max(g_x, g_y)$ and K' is is a positive constant and called as a masking parameter. The gradient based quality index (GBQI) can be calculated as

$$q = \big(1 - W(g, e)\big).g + W(g, e).e \qquad (18)$$

Where

$$e(x, y) = 1 - \left(\frac{x_{i-}y_i}{L}\right)^2 \qquad (19)$$

is the luminance similarity between two images and L is the dynamic range of the pixel values (255 for 8-bit grayscale images) and is the luminance similarity.

$$W(g, e) = p.g \qquad (20)$$

and where is a positive weighting parameter.Its value is taken as 0.1 in this paper [6].

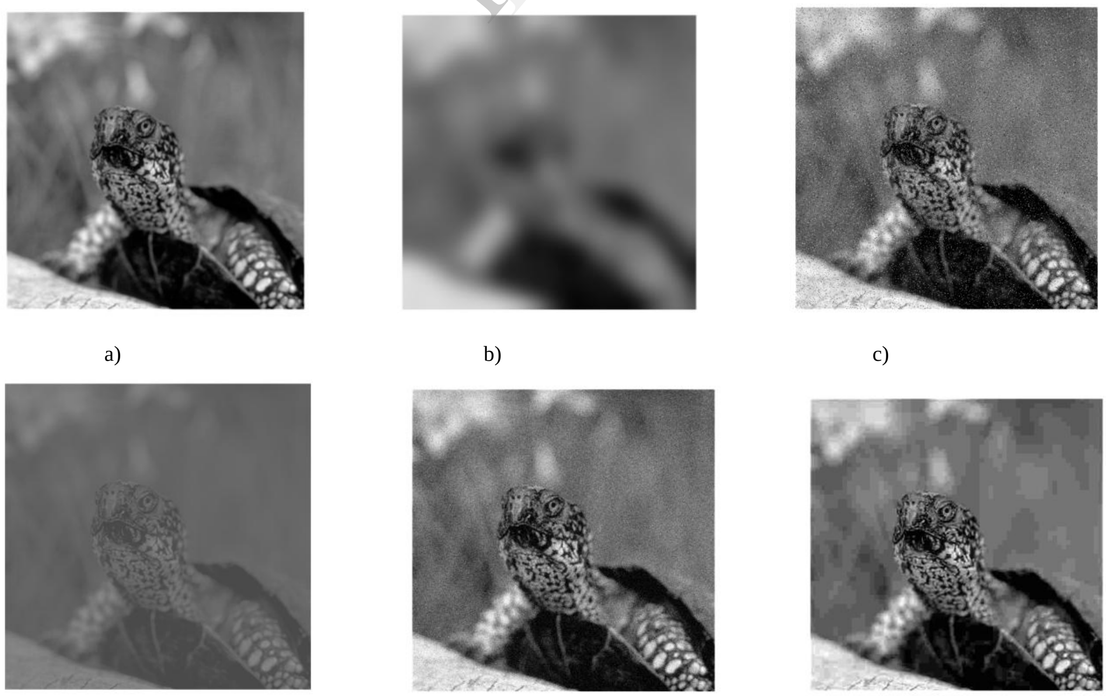


Fig.2: a) Original image, b) Gaussian Blurred image (PSNR=30.5166 , MSE=57.7327, SSIM =0.6933 , QI=0.1928) c) Salt & Pepper noise contaminated image (PSNR=35.7266, MSE=17.3948, SSIM =0.2613, QI=0.8023), d) Contrast changed image (PSNR=27.5762 , MSE= 113.6218 , SSIM = 0.7672 ,

MSSIM = 0.4189, QI1=0.7090), e) Gaussian noise contaminated image with (PSNR =30.1032 , MSE= 63.4984, SSIM =0.3422, QI1 = 0.8162), f) JPEG compressed image (PSNR = 35.4025 , MSE= 18.7425 , SSIM =0.8696 , QI1 = 0.9459)

III. EXPERIMENTAL EVALUATION

Experimental evaluation is carried out on Laboratory for Image and Video Engineering (LIVE) database images under various degradations [8]. In Fig. 2, (a) shows original image and other images are the distorted images, each having different type of distortion. The measures i.e. PSNR, MSE, SSIM, MS-SSIM, gradient similarity index (GSI) and gradient based quality index (GBQI) for different types of distorted images with reference to a reference image have been calculated. In the Fig.2 (c) and (f), distortions are different but PSNR and MSE measures are nearly same. So PSNR and MSE gives approximately same value for different types of distortions. The SSIM works well with other distortions but it is less effective for blurred images. . As shown in Tables I-V the gradient similarity approach can measure the relative quality loss (mainly the quality loss around edge and that in the non edge regions) better than the SSIM.

TABLE I. RESULTS FOR BLURRED IMAGES

Quality Measure	Test Image 1	Test Image 2	Test Image 3	Test Image 4
PSNR	43.8663	37.0868	35.0864	34.2307
MSE	2.6696	12.7174	20.1575	24.5476
SSIM	0.7338	0.3352	0.2335	0.1972
GSI	0.8199	0.8058	0.8019	0.7996
GBQI	0.8496	0.8336	0.8292	0.8266

TABLE II: RESULTS FOR DIFFERENT CONTRAST STRETCHED IMAGES

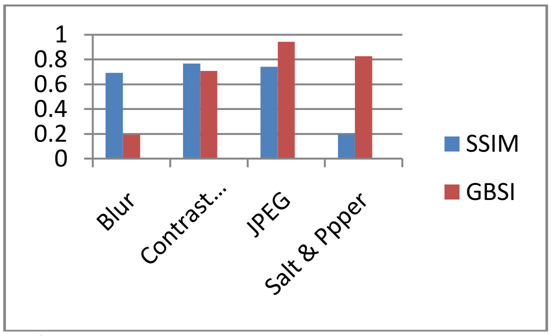
Quality Measure	Test Image 1	Test Image 2	Test Image 3	Test Image 4
PSNR	33.3841	30.2515	28.1930	27.5762
MSE	29.8312	61.3665	98.5775	113.6218
SSIM	0.9416	0.9109	0.8302	0.7672
GSI	0.9599	0.9901	0.9343	0.7443
GBQI	0.9118	0.9317	0.8763	0.7090

TABLE III: RESULTS FOR JPEG COMPRESSED IMAGES

Quality Measure	Test Image 1	Test Image 2	Test Image 3	Test Image 4
PSNR	39.5710	35.4025	32.1332	30.9283
MSE	7.1776	18.7425	39.7884	52.5112
SSIM	0.9477	0.8696	0.7904	0.7413
GSI	0.9758	0.9982	0.9975	0.9999
GBQI	0.9292	0.9459	0.9479	0.9440

TABLE V: RESULTS FOR SALT AND PEPPER NOISE CONTAMINATED IMAGES

Quality Measure	Test Image 1	Test Image 2	Test Image 3	Test Image 4
PSNR	43.8663	37.0868	35.0864	34.2307
MSE	2.6696	12.7174	20.1575	24.5476
SSIM	0.7338	0.3352	0.2335	0.1972
GSI	0.8199	0.8058	0.8019	0.7996
GBQI	0.8496	0.8336	0.8292	0.8266



CONCLUSION

The two distorted image signals with the same amount of error energy may have very different structure of errors, and hence different perceptual quality. The conventional methods PSNR and MSE do not always agree with the subjective viewing results in case of additive distortion. The SSIM gives good evaluation accuracy and simple mathematical formulation. But in case of blurred images, the SSIM neglects the effect of edge damage and treats every region equally in an image. The gradient based similarity can measure the relative quality loss, mainly the quality loss around edge and stands better than the SSIM.

REFERENCES

[1] Methodology for the Subjective Assessment of the Quality of Television Pictures, Recommendation ITU-R Rec. BT. 500- 11.

[2] 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, pp. 3313–3316, May 2002.

[3] Alain Horé, Djemel Ziou, "Image quality metrics: PSNR vs. SSIM", 2010 International Conference on Pattern Recognition.

[4] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, April 2004.

[5] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multi-scale structural similarity for image quality assessment," in *Proc. IEEE Asilomar Conf. Signals, Syst., Comput.*,pp. 1398–1402. Nov. 2003

[6] Anmin Liu, Weisi Lin and Manish Narwaria, "Image Quality Assessment Based on Gradient Similarity", *IEEE Trans. Image Process.*, vol. 21, no. 4, pp.1500–1512, April 2012.

[7] G. Chen, C. Yang, and S. Xie, "Edge-based structural similarity for image quality assessment," in *Proc. Int. Conf. Acoust., Speech, Signal Process.*, 2006.

[8] H. R. Sheikh, Z. Wang, A. C. Bovik, and L. K. Cormack, "Image and video quality assessment research at LIVE," http://live.ece.utexas.edu/research/quality/.