IJITKM Special Issue (ICFTEM-2014) May 2014 pp. 156-161 (ISSN 0973-4414)

Image Quality Assessment Techniques

V. K. Bhola[1], T. Sharma[2],J. Bhatnagar[3]
[1]vijaymmec@gmail.com, [2]tarun2069@gmail.com,[3] jbkrishna3@gmail.com

Abstract: Image Quality assessment plays an important role in various image processing applications. It is still an active field of research. A great deal of effort has been made in recent years to develop objective image quality metrics that correlate well with perceived human quality measurement or subjective methods. Most full reference(FR) technique were derived based on pixel to pixel error such as mean square error(MSE) or peak signal to noise ratio(PSNR),structural similarity index metric(SSIM) etc. This paper reviews different techniques used for image quality assessment.
Keywords: Image quality assessment, subjective & objective method, other methods.

1. Introduction

With the development of imaging and multimedia technologies, visual information, recorded by images has become the main source for knowledge acquisition. In the process of visual information acquisition, processing, transmission, and storage, some artifacts or noise may be introduced to images which degrade the visual quality. In a typical digital imaging system, the image is captured and transformed into digital signal by the sensor .This raw digital image signal is then processed to reduce noise and compressed for storage or transmission. When the image is finally displayed on the screen to the end user, it might not be same as the original version because it has been exposed to various kinds of distortions [1]. The sources of distortion could be ranged from motion blurring, Gaussian noise, sensor inadequacy, compression, error during transmission or the combination of many factors. To improve the performance of visual information acquisition, processing, transmission, and storage systems, it is necessary to assess visual qualities of images; so that it can maintain, control and possibly enhance the quality of the image before storage or transmission. The objective of image quality assessment is to provide computational models to measure the perceptual quality of a given image. Recently, a number of techniques have been designed to evaluate the quality of images and videos. The accurate prediction of quality from an end-user perspective has received increased attention with the growing for compression and communication of digital image and video services over wired and wireless networks. Image quality methods can be categorized in two parts subjective and objective. The subjective assessment of image is done on the bases of subjective experiments [2].While objective image quality assessment methods were mainly based on some mathematical measures. The past five years have demonstrated and witnessed the tremendous and imminent demands of visual quality assessment metrics in various applications. Section II of this paper describes both of the methods of image quality assessment.

2 Image Quality Assessment Techniques

A. Subjective Methods

The evaluation of quality may be divided into two classes, subjective and objective methods. Intuitively one can say that the best judge of quality is the human himself. That is why subjective methods are said to be the most precise measures of perceptual quality and to date subjective experiments are the only widely recognized method of judging perceived quality In these experiments humans are involved who have to vote for the quality of a medium in a controlled test environment. This can be done by simply providing a distorted medium of which the quality has to be evaluated by the subject. Another way is to additionally provide a reference medium which the subject can use to determine the relative quality of the distorted medium. These different methods are specified for television sized pictures by ITU-R and are, respectively, referred to as single stimulus continuous quality evaluation and double stimulus continuous quality-scale [3]. In double-stimulus methodology, subject is presented with the source and test images before evaluating their qualities on a linear quality scale as shown in Table 1.1. In single-stimulus methodology, the subject evaluates the quality of the test images on a linear quality scale without the source as reference. The scores evaluated by multiple subjects are averaged for each test image to obtain mean opinion score and difference mean opinion score. Each image is shown to the observer which is asked to score the image on a scale from 1 to 5.Mean Opinion Score (MOS) scores are given in Table 1.1.

It is known that subjective image quality varies from one individual to another: usually, the scores given by different individuals are not identical. The observer's score depends on his general experience, on his personal appreciation and may vary according to his mood. To solve this problem, an average score is computed over all observers. This Mean Opinion Score is denoted by MOS or the Difference Mean Opinion Score. Clearly, subjective quality assessment is expensive and tedious as it has to be performed with great care in order to obtain meaningful results. Also, subjective methods are in general not applicable in environments which require real-time processing.

Table 1.1: Mean Opinion Score Classes

1	2	3	4	5
Very poor quality	Poor Quality	Good Quality	Very good quality	Excellent quality

B. Objective Methods

This is a quantitative approach where intensity of two images, reference and distorted type are used to calculate a number which indicate the image quality. The objective Image Quality Assessment (IQA) can be classified into full-reference, reduced-reference and no-reference [4]. IQA based on the availability of the reference image. The goal of objective image quality assessment models is to automatically estimate the perceptual quality of images, in a way correlated with the human appreciation. The three models of objective method on the basis of reference images are categorized as given bellow.

1). No Reference (NR) models

It is also called "blind models" methods, in which the QA algorithm has access only to the distorted signal and must estimate the quality of the signal without any knowledge of the 'perfect version'. Since NR methods do not require any reference information, they can be used in any application where a quality measurement is required.

2). Reduced Reference (RR) models

In this partial information regarding the 'perfect version' is available. A side-channel exists through which some information regarding the reference can made available to the QA algorithm. RR QA algorithms use this partial reference information to judge the quality of the distorted signal of the scene.

3). Full Reference (FR) model

In this method quality assessment algorithm have access to a 'perfect version' of the image or video against which it can compare a 'distorted version'. The 'perfect version' generally comes from a high-quality acquisition device, before it is distorted by, say, compression artifacts and transmission errors. There are in general two classes for objective quality assessment approach, simple statistical error metrics and human visual system feature based metrics.

A) Simple statistics error metrics:

i)MSE

It stands for the mean squared difference between the original image and distorted image. The mathematical definition for MSE [5] is:

$$MSE = \left(1/M \times N\right)\sum_{i=1}^{M}\sum_{j=1}^{N}(a_{ij} - b_{ij})^2 \qquad (1.1)$$

In Equation (1.1), a_{ij}means the pixel value at]=
position (i, j) in the original image and b_{ij} means the pixel value at the same position in the corresponding distorted image.

ii) PSNR

PSNR is a classical index defined as the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation [6]. It is given by:

$$PSNR = 10\log_{10} 255^2 / MSE$$

Where 255 is the maximal possible value the image pixels when pixels are represented using 8 bits per sample, and MSE (mean square error) is the Euclidian distance between the original and the degraded images. The major advantages of these metrics are its simplicity and mathematical tractability, but they are not correlating well with perceived quality measurement because the Human Vision System characteristics are not considered in their models. PSNR is more consistent in the presence of noise compared to the SNR. By using the CSF (contrast sensitivity function) as the weighting function, we can define weighted SNR (WSNR) as the ratio of the average weighted signal power to the average weighted noise power.

iii) Average Difference (AD)
AD is simply the average of difference between the reference signal and test image. It is given by the equation [5]:

$$\mathrm{AD} = 1/\mathrm{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}(x(i,j) - y(i,j))$$

iv) Maximum Difference (MD)
MD is the maximum of the error signal (difference between the reference signal and test image) [5]:

$$\mathrm{MD} = \mathrm{MAX}\,|x(i,j) - y(i,j)|$$

v) Mean Absolute Error (MAE)
MAE is average of absolute difference between the reference signal and test image. It is given by the equation [5]:

$$\mathrm{MAE} = 1/\mathrm{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}|x(i,j) - y(i,j)|$$

vi) Peak Mean Square Error (PMSE)
It is given by the following equation [5]:

$$\mathrm{PMSE} = \frac{1}{\mathrm{MN}} \times \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}(x(i,j) - y(i,j))^2}{(\mathrm{MAX}(x(i,j)))^2}$$

The simplest and most widely used full-reference image quality measure is the MSE and PSNR. Advantage of MSE and PSNR are that they are very fast and easy to implement. However, they simply and objectively quantify the error signal. With PSNR, greater values indicate greater image similarity, while with MSE greater values indicate lower image similarity.

B) Human Visual System (HVS) feature based metric:
i) SSIM
The structural similarity index is a method for measuring the similarity between two images [7]. The SSIM index is a full reference metric, in other words, the measuring of image quality based on an initial uncompressed or distortion-free image as reference. It compares two images using information about luminous, contrast and structure. SSIM is designed to improve on traditional methods like PSNR and MSE.SSIM metric is calculated on various windows of an image. The measure between two window x and y of common size N×N is given as follows

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

Where $\mu_x$is average of x, $\mu_y$is average of y, σ_x, $\sigma_y$are standard deviation between the original and processedimages pixels, respectively.C_1, C_2 are positive constant chosen empirically to avoid the instability of measure. SSIM is a decimal value between (-1, 1).

ii) DSSIM
This is the structural dissimilarity metric it can be derived from SSIM as follows

$$\mathrm{DSSIM}(x, y) = 1/(1-\mathrm{SSIM}(x, y))$$

iii) MSSIM
The mean of SSIM is known as mean structural similarity index metric (MSSIM) [8] and it is given as:

$$MSSIM(X,Y) = \frac{1}{M}\sum_{l=1}^{M} SSIM(xl, yl)$$

For images of very different quality which have roughly same mean square error, with respect to the original image. MSSIM gives a much better indication of image quality.

Other Methods
i) A No Reference Image Quality Assessment by using a general regression neural network (GRNN): The general regression neural network is a powerful regression tool that has a dynamic network structure [9], [10]. It is based on established statistical principles, and asymptotically converges with an increasing number of samples to the optimal regression surface [11]. GRNN has been observed to yield better results than the back-propagation network or RBF (radial basis function) network in terms of prediction performance.

The GRNN was implemented using the MATLAB function newgrnn. The four perceptually motivated features can be used as inputs to the GRNN: 1) The mean value of the phase congruency image of distorted image (MPC), 2) The entropy of the phase congruency image of distorted image (EPC), 3) The entropy of the distorted image (EDIS), and 4) The mean value of the gradient magnitude of the distorted image (MGDIS).

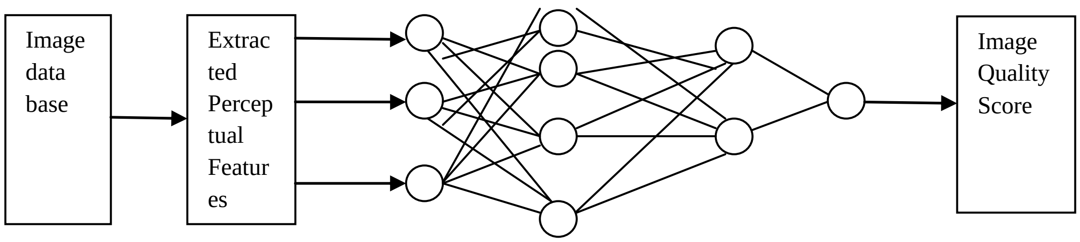

Fig c.1 Schematic diagram of GRNN for assessing image quality

ii) Image Quality Assessment based on a degradation model: The degradation can be modelled to develop efficient methods for minimizing the visual impact of degradation. A degraded image can be modelled as an original image which has been subject to two independent sources of degradation, linear frequency distortion and additive noise injection. This model is commonly used in image restoration. Based on the model, methods to measure the quality of images and demonstrate how one may use the quality measures in quantifying the performance of image restoration algorithms can be developed. The distortion (relative to the original image) can be modelled as linear and spatially invariant. The noise is modelled as spatially varying additive noise. A degraded image is referred as an image degraded by the two-source degradation model. The degradation in the restoredimage can be quantified as compared with the original, uncorrupted image. Two complementary quality measures that separately measure the impact of frequency distortion and noise injection on the human visual system (HVS) are used. This decoupled approach allows a designer to explore the fundamental tradeoffs between distortion and noise to improve restoration algorithms, which is not possible with a scalar-valued quality measure. SNR measures, such as peak SNR (PSNR), assume that distortion is only caused by additive signal-independent noise. As a consequence, noise measures applied directly to a restored image and its original do not measure visual quality. Two measures of degradation: a) distortion measure (DM), b) noise quality measure (NQM), based on the observation that the psychovisual effects of filtering and noise are separate. Instead of computing a residual image, a restored image by passing the original image through the restoration algorithm using the same parameters as were used while restoring a degraded image is computed. The DM is computed in three steps. First, find the frequency distortion in the restored image by comparing the restored and the model restored images. Second, compute the deviation of this frequency distortion from an allpass response of unity gain (no distortion). Finally, weight the deviation by a low pass CSF (contrast sensitivity function) and integrate over the visible frequencies. The NQM is computed in two steps. First, process the original image and the modelled restored image separately through a contrast pyramid. The contrast pyramid, which is based on Peli's work [12], computes the contrast in an image at every pixel and at spatial frequencies separated by an octave, and models the following nonlinear spatially varying visual effects:

1) Variation in contrast sensitivity with distance, image dimensions, and spatial frequency
2) Variation in the local luminance mean
3) Contrast interaction between spatial frequencies
4) Contrast masking effects.

Second, form the NQM by computing the SNR of the restored degraded image with respect to the model restored image.

iii) A robust quality metric for colour image quality assessment: This is a full reference model based on human visual system properties Two main stages are used:
a) First one in order to compute visual representation of image.
b) Second in order to pool errors between visual representations of image.

This metric does not use any a priori knowledge of the type of degradations introduced by any image processing.

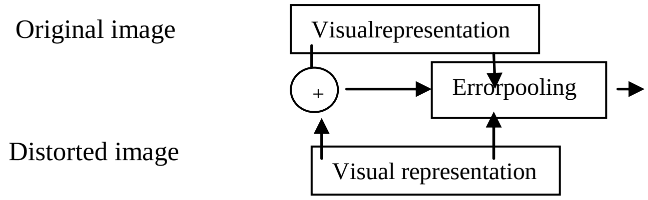

Fig .2 structure of the criteria

Output of error pooling block gives the measure of quality of image. Here visual representation is highly based on results provided by psychophysics experiments on colour perception and on masking effect.

iv) Image Quality Evaluation method based on digital watermarking: Digital watermarking based method can estimate the quality of an image in terms of the classical objective metrics PSNR, Weighted PSNR, Watson just noticeable difference (JND)[13][14] without need of original image.

In this method a watermark is embedded into the discrete wavelet transform domain of original images using a quantization method. If it is considered that different images have different frequency distributions, then vulnerability of the watermark for the image is adjusted using automatic control. After auto adjustment, the degradation of the extracted watermark can be used to estimate image quality in terms of other classical metrics with high accuracy.

So calculated PSNR, wPSNR (weighted PSNR), are compared with those calculated using watermarking based approach.

3. Conclusion

In the field of image processing, image quality assessment is a fundamental and challenging problem with many interests in a variety of applications, such as dynamic monitoring and adjusting image quality, optimizing algorithms and parameter settings of image processing systems, and benchmarking image processing system and algorithms[15][16].So full reference(FR) methods like structural similarity index metric(SSIM),mean structural similarity index metric(MSSIM) are more efficient because some mathematical formula like peak signal to noise ratio(PSNR), mean square error(MSE) become unstable if image has a significant amount of degradation.

4.References

[1] Special issue of Signal Processing on Image quality assessment, vol. 70, 1998.

[2] Z. Wang and A. C. Bovik. Modern Image Quality Assessment. Morgan and ClaypoolPublishing Company, New York, 2006.

[3] Horé, A., Ziou, D., “Image Quality Metric: PSNR Vs SSIM” IEEE conference on Pattern recognition, pp 2366-2369,2010.

[4] Combining and Selecting Indicators for Image Quality Assessment Proceedings of the ITI 2009 31st Int. Conf. on Information Technology Interfaces, June 22-25, 2009, Cavtat, Croatia.

[5] Ahmet M. Eskicioglu, Paul S. Fisher, “Image Quality Measures and Their Performance” IEEE Transactions on Communication, Vol. 43, No. 12, pp. 2959-2965, December1995.

[6] ZhizhongZhe, Hong Ren Wu, “ A New Way of Pooling: Starting from An Image Quality Measure, IEEE proceedings on signal processing”,pp: 1080 - 1083 vol.2,2004

[7] Z. Wang and A. C. Bovik.Modern Image Quality Assessment. Morgan and ClaypoolPublishing Company, New York, 2006.

[8] Wang Z. , Bovik A. C., “A universal image quality index,” IEEE Processing Letters, vol. 9, pp. 81–84, Mar. 2002.

[9] D. F. Specht, “A general regression neural network,” IEEE Trans. Neural Netw., vol. 2, no. 6, pp. 568–576, Nov. 1991.

[10] S. Chartier, M. Boukadoum, and M. Amiri, “BAM learning of nonlinearly separable task by using an asymmetrical output function and reinforcement learning,” IEEE Trans. Neural Netw., vol. 20, no. 8, pp. 1281–1292, Aug. 2009.

[11] D. Tomandl and A. Schober, “A modified general regression neural network (MGRNN) with new, efficient training algorithms as a robust ‘black box’-tool for data analysis,” Neural Netw., vol. 14, no. 8, pp. 1023–1034, Oct. 2001.

[12] E. Peli, “Contrast in complex images,” J. Opt. Soc. Amer. A, vol. 7, pp. 2032–2039, Oct. 1990.

[13] I. Cox, M. Miller, and J. Bloom, Digital Watermarking. New York: Morgan Kaufmann, 2002.

[14] Z. Wang, H. R. Sheikh, and A. C. Bovik, “Objective video quality assessment,” in Handbook of Video Databases: Design and Applications. Boca Raton, FL: CRC Press, 2003, pp. 1041–1078.

[15] Z. Wang, A. C. Bovik, L. Lu, “Why is Image Quality Assessment so Difficult?”, in Proc. of Int. Conf. on Acoustics,Speech and Signal Proc, Orlando, Florida, USA, pp: IV-3313-IV3316, May 2002.

[16] R. F. Zampolo, R. Seara, “A Measure for Perceptual Image Quality Assesment”, in Proc. of Int. Conf. on ImageProc.,Barcelona,Spain,pp:433-436,Sept.2003