A Universal Image Quality Index

Zhou Wang, *Student Member, IEEE*, and Alan C. Bovik, *Fellow, IEEE*

Abstract— **We propose a new universal objective image quality index, which is easy to calculate and applicable to various image processing applications. Instead of using traditional error summation methods, the proposed index is designed by modeling any image distortion as a combination of three factors: loss of correlation, luminance distortion, and contrast distortion. Although the new index is mathematically defined and no human visual system model is explicitly employed, our experiments on various image distortion types indicate that it performs significantly better than the widely used distortion metric mean squared error. Demonstrative images and an efficient MATLAB implementation of the algorithm are available online at *http://anchovy.ece.utexas.edu/~zwang/research/quality_index/demo.html***

Keywords—**image quality measurement, human visual system (HVS), mean squared error (MSE)**

I. Introduction

Objective image quality measures play important roles in various image processing applications. There are basically two classes of objective quality or distortion assessment approaches. The first are mathematically defined measures such as the widely used mean squared error (MSE), peak signal to noise ratio (PSNR), root mean squared error (RMSE), mean absolute error (MAE), and signal to noise ratio (SNR). The second class of measurement methods consider human visual system (HVS) characteristics in an attempt to incorporate perceptual quality measures [1],[2]. Unfortunately, none of these complicated objective metrics in the literature has shown any clear advantage over simple mathematical measures such as RMSE and PSNR under strict testing conditions and different image distortion environments [3],[4],[5].

Mathematically defined measures are still attractive because of two reasons. First, they are easy to calculate and usually have low computational complexity. Second, they are independent of viewing conditions and individual observers. Although it is believed that the viewing conditions play important roles in human perception of image quality, they are, in most cases, not fixed and specific data is generally unavailable to the image analysis system. If there are N different viewing conditions, a viewing condition-dependent method will generate N different measurement results that are inconvenient to use. In addition, it becomes the *user*'s responsibilities to measure the viewing conditions and to calculate and input the condition parameters to the measurement systems. By contrast, a viewing condition-independent measure delivers a single quality value that gives a general idea of how good the image is.

In this paper, we propose a mathematically defined universal image quality index. By "universal", we mean that the quality measurement approach does not depend on the images being tested, the viewing conditions or the individual observers. More importantly, it must be applicable to various image processing applications and provide meaningful comparison across different types of image distortions. Currently, the PSNR and MSE are still employed "universally", regardless of their questionable performance. This work attempts to develop a new index to replace their roles.

The authors are with Laboratory for Image and Video Engineering (LIVE), The Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX 78712-1084 USA. E-mail: zhouwang@ieee.org, bovik@ece.utexas.edu.

II. Definition of the New Quality Index

Let $\mathbf{x} = \{x_i | i = 1, 2, \cdots, N\}$ and $\mathbf{y} = \{y_i | i = 1, 2, \cdots, N\}$ be the original and the test image signals, respectively. The proposed quality index is defined as

$$Q = \frac{4\,\sigma_{xy}\,\bar{x}\,\bar{y}}{(\sigma_x^2 + \sigma_y^2)\,[(\bar{x})^2 + (\bar{y})^2]}\,, \qquad (1)$$

where

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i\,, \qquad \bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i\,,$$

$$\sigma_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2\,, \quad \sigma_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \bar{y})^2\,,$$

$$\sigma_{xy} = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})\,.$$

The dynamic range of Q is $[-1, 1]$. The best value 1 is achieved if and only if $y_i = x_i$ for all $i = 1, 2, \cdots, N$. The lowest value of -1 occurs when $y_i = 2\bar{x} - x_i$ for all $i = 1, 2, \cdots, N$. This quality index models any distortion as a combination of three different factors: loss of correlation, luminance distortion, and contrast distortion. In order to understand this, we rewrite the definition of Q as a product of three components:

$$Q = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \cdot \frac{2\,\bar{x}\,\bar{y}}{(\bar{x})^2 + (\bar{y})^2} \cdot \frac{2\,\sigma_x \sigma_y}{\sigma_x^2 + \sigma_y^2}\,. \qquad (2)$$

The first component is the correlation coefficient between $\mathbf{x}$ and $\mathbf{y}$, which measures the degree of linear correlation between $\mathbf{x}$ and $\mathbf{y}$, and its dynamic range is $[-1, 1]$. The best value 1 is obtained when $y_i = a x_i + b$ for all $i = 1, 2, \cdots, N$, where a and b are constants and $a > 0$. Even if $\mathbf{x}$ and $\mathbf{y}$ are linearly related, there still might be relative distortions between them, which are evaluated in the second and third components. The second component, with a value range of

$[0,1]$, measures how close the mean luminance is between $\mathbf{x}$ and $\mathbf{y}$. It equals 1 if and only if $\bar{x} = \bar{y}$. σ_x and σ_y can be viewed as estimate of the contrast of $\mathbf{x}$ and $\mathbf{y}$, so the third component measures how similar the contrasts of the images are. Its range of values is also $[0,1]$, where the best value 1 is achieved if and only if $\sigma_x = \sigma_y$.

III. Application to Images

Image signals are generally non-stationary while image quality is often also space variant, although in practice it is usually desired to evaluate an entire image using a single overall quality value. Therefore, it is more appropriate to measure statistical features locally and then combine them together. We apply our quality measurement method to local regions using a sliding window approach. Starting from the top-left corner of the image, a sliding window of size $B \times B$ moves pixel by pixel horizontally and vertically through all the rows and columns of the image until the bottom-right corner is reached. At the j-th step, the local quality index Q_j is computed within the sliding window. If there are a total of M steps, then the overall quality index is given by

$$Q = \frac{1}{M} \sum_{j=1}^{M} Q_j \, . \tag{3}$$

We use images with different types of distortions to test the new quality index and compare the results with the MSE and with subjective evaluations. The test images are distorted by a wide variety of corruptions: impulsive salt-pepper noise, additive Gaussian noise, multiplicative speckle noise, mean shift, contrast stretching, blurring, and JPEG compression. Some sample images are shown in Fig. 1 and Fig. 2, where we tuned all the distortions to yield the same MSE relative to the original image, except for the JPEG compressed image, which has a slightly smaller MSE. The overall quality index of each image is also calculated, where the sliding window size is $B = 8$. In the subjective experiments, we showed the original and the 7 distorted images to 22 subjects (9 of them have been working in the area of image processing and the rest are naive) and asked them to compare each distorted image with the original one and rank them according to quality. The average ranking values, together with the MSE and the proposed quality measurement results, are given in Table I. In this experiment, the performance of MSE is extremely poor in the sense that images with nearly identical MSE are drastically different in perceived quality. By contrast, the new quality index exhibits very consistent correlation with the subjective measures. In fact, the ranks given by the new quality index are the same as the mean subjective ranks! It can be observed that the rank of Fig. 2(b) is almost tied with Fig. 2(a). This is no surprise because contrast stretching as Fig. 2(b) is often an image enhancement process, which is supposed to increase the visual quality of the original image. Actually, many subjects regarded it as better than the original image. However, if we assume that the original image is the perfect one (as our quality measurement method does), then it is fair to give Fig. 2(a) a higher quality value. Only a few images are presented in this paper due to space limit. More demonstrative images and an efficient MATLAB implementation of the proposed algorithm are available online at *http://anchovy.ece.utexas.edu/~zwang/research/quality_index/demo.html*, where we try to demonstrate the new index's "universal" property by comparing the quality measurement results of different images with different distortion types and layers.

IV. Conclusion and Discussion

A new universal image quality index was proposed. Our experimental results indicate that it outperforms the MSE significantly under different types of image distortions. It is perhaps surprising that such a simple mathematically defined quality index performs so well without any HVS model explicitly employed. We think the success is due to its strong ability in measuring structural distortion occurred during the image degradation processes. This is a clear distinction with MSE, which is sensitive to the energy of errors, instead of structural distortions. In the future, more extensive experiments are needed to fully validate the new index.

There is no doubt that more precise modeling of the HVS is always advantageous in the design of image quality metrics. However, without a well-defined mathematical framework, the efforts in HVS modeling will not result in a successful quality measure. For example, error summation in the form of the Minkowski metric

$$Err = \left(\sum_k |s_k - s'_k|^\beta\right)^{1/\beta} \tag{4}$$

or its equivalent has been widely adopted by most previous image and video quality assessment models, where β is a constant typically with a value between 1 and 4, and s_k and s'_k are the corresponding image components (in various formats such as pixel value, weighted pixel value, weighted DCT coefficient and weighted wavelet coefficient) of the original and the test images, respectively. We maintain that this is not an appropriate mathematical form for image quality evaluation, since image differencing does not adequately capture an estimate of the correlation between s_k and s'_k, which our evidence implies is very important. We believe that the basic idea introduced in this paper is a promising starting point for the future development of more successful image and video quality assessment methods.

References

[1] T. N. Pappas and R. J. Safranek, "Perceptual criteria for image quality evaluation," in *Handbook of Image and Video Processing* (A. C. Bovik, ed.), Academic Press, May 2000.

[2] "Special issue on image and video quality metrics," *Signal Processing*, vol. 70, Nov. 1998.

[3] J.-B. Martens and L. Meesters, "Image dissimilarity," *Signal Processing*, vol. 70, pp. 155–176, Nov. 1998.

[4] VQEG, "Final report from the video quality experts group on the validation of objective models of video quality assessment," *http://www.vqeg.org/*, Mar. 2000.

[5] A. M. Eskicioglu and P. S. Fisher, "Image quality measures and their performance," *IEEE Trans. Communications*, vol. 43, pp. 2959–2965, Dec. 1995.

TABLE I
QUALITY MEASUREMENT OF "LENA" IMAGE WITH DIFFERENT TYPES OF DISTORTIONS.

Image	Distortion Type	Mean Subjective Rank	MSE	Q
Fig. 2(a)	Mean Shift	1.59	225	0.9894
Fig. 2(b)	Contrast Stretching	1.64	225	0.9372
Fig. 1(b)	Impulsive Salt-Pepper Noise	3.32	225	0.6494
Fig. 1(d)	Multiplicative Speckle Noise	4.18	225	0.4408
Fig. 1(c)	Additive Gaussian Noise	4.27	225	0.3891
Fig. 2(c)	Blurring	6.32	225	0.3461
Fig. 2(d)	JPEG Compression	6.68	215	0.2876

Fig. 1. Evaluation of "Lena" images contaminated by impulsive salt-pepper, additive Gaussian, and multiplicative speckle noises. (a) Original "Lena" image, 512×512, 8bits/pixel; (b) Salt-pepper noise contaminated image, MSE = 225, Q = 0.6494; (c) Gaussian noise contaminated image, MSE = 225, Q = 0.3891; (d) Speckle noise contaminated image, MSE = 225, Q = 0.4408.

(a) (b)

(c) (d)

Fig. 2. Evaluation of "Lena" images distorted by mean shift, contrast stretching, blurring, and JPEG compression. (a) Mean shifted image, MSE = 225, Q = 0.9894; (b) Contrast stretched image, MSE = 225, Q = 0.9372; (c) Blurred image, MSE = 225, Q = 0.3461; (d) JPEG compressed image, MSE = 215, Q = 0.2876.