Novel image-dependent quality assessment measures

Asaad Hashim

Zahir Hussain

*Edith Cowan University*

Follow this and additional works at: [https://ro.ecu.edu.au/ecuworkspost2013](https://ro.ecu.edu.au/ecuworkspost2013)

Part of the *Theory and Algorithms Commons*


This Journal Article is posted at Research Online.
NOVEL IMAGE-DEPENDENT QUALITY ASSESSMENT MEASURES

Asaad Noori Hashim and Zahir M. Hussain

1Department of Computer Science, University of Babylon, P.O. Box 4, Babylon, Iraq
2Department of Computer Science, University of Kufa, P.O. Box 21, Kufa, Najaf, and Adjunct Professor, RMIT, Vic. 3000, Australia

Received 2013-12-11; Revised 2014-01-06; Accepted 2014-04-07

ABSTRACT

The image is a 2D signal whose pixels are highly correlated in a 2D manner. Hence, using pixel by pixel error what we called previously Mean-Square Error, (MSE) is not an efficient way to compare two similar images (e.g., an original image and a compressed version of it). Due to this correlation, image comparison needs a correlative quality measure. It is clear that correlation between two signals gives an idea about the relation between samples of the two signals. Generally speaking, correlation is a measure of similarity between the two signals. An important step in image similarity was introduced by Wang and Bovik where a structural similarity measure has been designed and called SSIM. The similarity measure SSIM has been widely used. It is based on statistical similarity between the two images. However, SSIM can produce confusing results in some cases where it may give a non-trivial amount of similarity while the two images are quite different. This study proposes methods to determine a reliable similarity between any two images, similar or dissimilar, in the sense that dissimilar images have near-zero similarity measure, while similar images give near-one (maximum) similarity. The proposed methods are based on image-dependent properties, specifically the outcomes of edge detection and segmentation, in addition to the statistical properties. The proposed methods are tested under Gaussian noise, impulse noise and blur, where good results have been obtained even under low Peak Signal-to-Noise Ratios (PSNR’s).

Keywords: Image Structural Similarity, Edge Detection, Image Segmentation, Image Processing

1. INTRODUCTION

An important feature of natural images is that they are highly structured signals, meaning that the image samples exhibit strong correlation; this is more evident when samples are in spatial proximity. This 2D correlation carries important information about the structure of the objects in the image.

An objective image quality measure can have a significant role in image processing and its applications, where it can be used to monitor and adjust image quality. Also, a quality measure can be used to optimize algorithms and parameter settings of image processing systems, as to benchmark image processing algorithms. Machine evaluation of image and video quality is important for many image processing systems, for example, systems used for compression, restoration, enhancement, etc. The goal of quality assessment is to find robust techniques for objective evaluation of image quality in accord with subjective human assessment.

Over the years many researchers have contributed to the design and implementation of reference quality assessment algorithms. Wang and Bovik (2002) avoided using traditional mean-squared error methods and proposed a model for any image distortion that is dependent on a distortion in a combination of three quantities: Correlation, luminance, and contrast.

Wang et al. (2004) proposed a promising technique (SSIM) for distance covariance to measuring the structural similarity based on number of statistical measurements such as mean, standard deviation and they produced a new relation among these standards Equation 1:
\[ \rho(x, y) = \frac{(2\mu_x \mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \]  

where, \( \rho(x, y) \) is the SSIM measure between two images \( x \) and \( y \), \( \mu_x \) and \( \sigma_x^2 \) are the statistical mean and variance of pixels in image \( x \) \((\mu_y, \sigma_y^2 \text{ are defined similarly}) \) \( \sigma_{xy} \) is the statistical variance between pixels in images \( x \) and \( y \) while the constants \( C_1 \) and \( C_2 \) are defined as \( C_1 = (K_1 L)^2 \) and \( C_2 = (K_2 L)^2 \), with \( K_1 \) and \( K_2 \) are small constants and \( L = 255 \) (maximum pixel value).

This approach gives high level of similarity for noise free condition while it goes to zero when noise increase, in other words it gives similarity with two different images due to it dependent only the statistics features of images which may have some correlations. SSIM can’t reveal all image structural properties, so we need to more specific measurements that are image-dependent.

Sheikh et al. (2006) presented results of an extensive subjective quality assessment. In their study a number of distorted images were evaluated by a number of human subjects, where image quality data obtained from human quality judgments is used to evaluate several full reference image quality assessment methods. This study was the largest subjective image quality study in the literature in terms of number of images, distortion types and the number of human evaluations.

A recent improvement on SSIM is presented by Sampat et al. (2009): The Complex Wavelet SSIM (CW-SSIM). It is based on wavelet coefficients that are extracted at the same spatial locations in the same wavelet subbands of the two images under test. This approach is shown to be less sensitive to small geometric variations or distortions (such as rotations, translations and difference in scale).

Szekely et al. (2007) improved similarity testing by adding a new distance measurement called “Energy Statistics” based on the following formula: 

\[ D(\mu, v) = 2\varepsilon(d(X, Y)) - \varepsilon(d(X, X')) - \varepsilon(d(Y, Y')) \]

where, \( \varepsilon \) is the expectation and \( d(X, Y) \) is the Euclidean distance. This measure considers statistical observation and statistical potential energy. Energy statistics is a function of distance between statistical observations. This approach has a high rate of complexities and computational difficulties.

Reference Zhang et al. (2009) explains many limitations and challenges of current approach of image quality measurement. It is stated that each kind of image difference will cause a different kind of distortion in perceptual visual domain. Generally, these changes include:

- Scale, orientation, lighting and image contrast.
- Spatial distribution of texture
- Position of objects

Some kinds of distortion may highly affect the image, even if distortion is small, for example:

- Sharpness of image contours
- Other distortions or artifacts in sensitive regions like the face

Kaur et al. (2012) improved the performance of metrics like Coefficient of Correlation (CoC) and Structural Similarity Index (SSIM) for image recognition in real-time environment. Li et al. (2010) used a similarity assessment to select the images for synthesis, where a new similarity measure has been proposed using complex wavelets. This measure has been shown to be robust to small rotations and translations as well as large intensity and contrast changes.

Dan et al. (2010) proposed a novel image quality assessment technique which is based on the conventional SSIM and the discrete cosine transform (DCT). The method presents a frequency structural comparison by weighting the frequency components depending on the sensitivity of human eye.

Liu and Wang (2011) introduced a similarity measure based on edge structural similarity; while Liu et al. (2011) presented an objective fusion quality index.

Please note that the above-mentioned similarity measures are all based on statistical moments, on which we will focus in this study, while there are other moments that can also be used to test similarity (Lajevardi and Hussain, 2010a; 2010b).

Blasch et al. (2008) presented a novel approach on objective non-reference image fusion performance assessment. The proposed measure is an extension of the Universal Image Quality Index (UIQI); where its weighting factor is the similarity between blocks of pixels in the input images and the fused image.

In this study, we enhance the basic SSIM, proposed by Wang et al. (2004) and study the performance of SSIM and the proposed enhanced method under noisy conditions and blur. The enhancement is based on image segmentation and edge detection techniques to give more reliable similarity measure.

2. RATIONALE

We noticed that SSIM measure introduced by Wang et al. (2004) gives false similarity between
unrelated images; hence, it needs more image-dependent properties to be reliable. We utilized segmentation and edge properties and combined them with SSIM to get the enhanced measure mSSIM; also we tested SSIM and Mssim under disruptive conditions like Gaussian noise, impulse noise and blur.

3. THE PROPOSED MEASURES

The design of SSIM was based on image statistical properties, Wang et al. (2004), hence the non-zero SSIM measure $\rho(x, y)$ between unrelated images $x$ and $y$. We noticed that even straightforward segmentation (of the two images $x, y$ into $K$-pairs of corresponding sub-images $x_i, y_i$, $i = 1, 2, \ldots, K$) can substantially reduce the chance of statistical similarity between all available segments, therefore we propose the following image dependent measure Equation 2:

$$\zeta(x, y) = \prod_{i=1}^{K} \rho(x_i, y_i)$$  \hspace{1cm} (2)

Similarly, the inclusion of edge effects into SSIM will highly reduce the chance of statistical similarity, hence we propose the following image-dependent measure Equation 3:

$$\eta(x, y) = R(x, y) \rho(x, y)$$  \hspace{1cm} (3)

Noting that $R(x, y)$ is the 2D edge correlation coefficient defined as Equation 4:

$$R(x, y) = \frac{\sum \sum (g_{x_n} - g_x)(h_{y_n} - h_y)}{\sqrt{\sum \sum (g_{x_n} - g_x)^2\sum \sum (h_{y_n} - h_y)^2}}$$  \hspace{1cm} (4)

where, g and h are the new images resulting from applying an edge detection technique to the test images x and y, respectively, while $g_x$ and $h_y$ are their global means.

4. THE TEST ENVIRONMENT

The proposed SSIM measures have been tested under Gaussian noise and blur. Impulse noise, $e = \{e(i,j)\}$, which is a source of noise in many image processing systems, has also been considered. The arrival time of this noise process at an instant $k$ is formulated as a Poisson process $b_k$ with parameter $\lambda$, while the amplitude of any noisy sample is formulated as a Gaussian process $g_k$ with zero mean and variance of $\sigma^2$. The overall impulsive noise process $i_k$ is given by Al-Mawali et al. (2010) Equation 5:

$$i_k = b_k g_k$$  \hspace{1cm} (5)

If the random variable that represents the time count of arrival (since the last impulse) is $T$, then the probability of arriving $m$ samples after the previous impulse, $p(m)$, will be Equation 6:

$$p(k) = p(T = k) = \exp(-\lambda) (\lambda^k / k!); \quad k = 0, 1, 2, \ldots$$  \hspace{1cm} (6)

Noting that Equation 7:

$$\epsilon(T) = \text{var}(T) = \lambda$$  \hspace{1cm} (7)

The power of the Gaussian amplitude $\sigma^2$ will contribute a total noise power of Equation 8:

$$n_p = \sigma^2 / \lambda$$  \hspace{1cm} (8)

Hence, we define $r$, the Peak Signal to Noise Ratio (PSNR), as follows Equation 9:

$$r = \frac{L^2}{n_p} = \frac{\lambda L^2}{\sigma^2}$$  \hspace{1cm} (9)

5. RESULTS

The proposed measures as well as SSIM have been simulated using MATLAB. Note that $0 \leq \rho(x, y) \leq 1$, so are $\zeta(x, y)$ and $\eta(x, y)$. For completely similar images we have $\rho(x, y) = 1$; while for totally different images we have $\rho(x, y) = 0$. It is better to calculate similarity measures locally not globally; hence, an $M \times M$ window ($M = 11$) is used with a standard deviation of 1.5, Wang et al. (2004). The constants $C_1 = (K_1L)^2$ and $C_2 = (K_2L)^2$ ($K_1$ and $K_2$ being small constants, $L = 255$) where chosen as $K_1 = 0.01$ and $K_2 = 0.03$, Wang et al. (2004). Note that the performance of SSIM is insensitive to these constants, Wang et al. (2004).

5.1. Performance under Gaussian Noise

First we implemented the Segmentation-based Measure (mSSIM) as per Equation 2 and tested its performance when the other image is corrupted with Gaussian noise. Peak Signal to Noise Ratio (PSNR) was used in this test as follows:

$$\text{PSNR} = \frac{L^2}{p_n}$$
where, $p_n$ is the Gaussian noise variance (power). The result of using mSSIM for two similar images is shown in Fig. 1a, with performance of mSSIM as compared to SSIM (represented by Equation 1) is shown in Fig. 1b; while the result of comparing two dissimilar images is shown in Fig. 2a and b, respectively. We used the images "woman" and "moon" from MATLAB.

**Fig. 1.** Performance of SSIM and mSSIM using similar images under Gaussian noise. (a) Above: The test images. (b) Below: Performance comparison between SSIM and mSSIM
Secondly, we implemented the Edge-based Measure (eSSIM) as per Equation 3 and tested its performance under Gaussian noise. Canny method has been utilized for edge detection, Canny (1986); though other methods can also be used. The results are shown in Fig. 3 and 4, with performance of eSSIM compared to that of SSIM (represented by Equation 1). In case of dissimilar images, a clearer comparison can be viewed using logarithmic scale as shown in Fig. 4.

Fig. 2. Performance of SSIM and mSSIM using dissimilar images under Gaussian noise. (a) Above: The test images. (b) Below: Performance comparison between SSIM and mSSIM.
5.2. Performance Under Blur

The proposed methods have also been tested under blur. We simulated blur effect as spatial windowing (convolution) with a 2D averager, with window length $W$. Figure 5 and 6 show the performance of eSSIM as compared to that of SSIM under blur for different window lengths.
Fig. 4. Performance of SSIM and eSSIM using dissimilar images under Gaussian noise (a) Above: The test images (b) Below: Performance comparison between SSIM and eSSIM. Logarithmic scale is used.
Fig. 5. Performance of SSIM and eSSIM using similar images under blur (a) Above: The test images (b) Below: Performance comparison between SSIM and eSSIM
Fig. 6. Performance of SSIM and eSSIM using dissimilar images under blur (a) Above: The test images (b) Below: Performance comparison between SSIM and eSSIM (using logarithmic scale)
Fig. 7. Performance of SSIM and eSSIM using dissimilar images under impulse noise with low arrival rate $\lambda = 50$ (a) Above: The test images (b) Below: Performance comparison between SSIM and eSSIM.
Fig. 8. Performance of SSIM and eSSIM using dissimilar images under impulse noise with high arrival rate $\lambda = 10$ (a) Above: The test images (b) Below: Performance comparison between SSIM and eSSIM
5.3. Performance under Impulse Noise

Impulse noise has been simulated as per Equation 5-9. Performance of SSIM and the proposed measures have been compared under impulse noise as shown in Fig. 7 and 8 for different values of Poisson parameter $\lambda$.

6. DISCUSSION

The conventional SSIM, published by Wang and Bovik (2002), outperforms mSSIM or eSSIM in discovering similarity between similar images, where it gives higher correlation coefficient at similar SNR and blur. Hence, SSIM outperforms the proposed measures in case of comparing two similar images, where it gives reasonable similarity at lower PSNR’s than those thresholds of our proposed measures. The reason is that similarity is diluted by using edges or segmentation, which are the bases of our approach. However, SSIM can be misleading for dissimilar images, where mSSIM and eSSIM give almost zero correlation between un-related images.

7. CONCLUSION

Two new image-dependent quality assessment measures have been proposed and tested versus structural Similarity Measure (SSIM) under noise (Gaussian and impulse) and blur. It is shown that the proposed measures can rid SSIM from the disadvantage of giving non-zero correlation between dissimilar images, while SSIM still outperforms the proposed measures in case of comparing two similar images, where it gives reasonable similarity at lower Peak Signal-to-Noise Ratios (PSNR’s) than those thresholds of our proposed measures. Little are the works that utilized the capabilities of SSIM for face recognition. As a future direction, we are currently working on using SSIM as a tool for face recognition, where initial results are promising. Also, an extension towards facial expression recognition as per (Lajevardi and Hussain, 2012; 2009) is under consideration.

8. REFERENCES


