ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

Contemporary Full Reference Image Quality Assessment Metrics

Parineeta Sachin Gengaje

Final Year B.E Student, Dept. of E&TC Cummins College of Engineering, Karvenagar, Pune, India

ABSTRACT: Image quality assessment plays a vital role in design and evaluation of many image processing algorithms and applications. Objective methods for assessing full reference perceptual image quality traditionally attempts to quantify the visibility of errors (difference) between an original (reference) image and test (distorted) image. Over the years, mean square error (MSE) and peak signal to noise ratio (PSNR) are used as popular metrics to assess image fidelity. However it is now proved that MSE & PSNR exhibits weak performance. They have been widely criticized for serious shortcomings, especially when dealing with perceptually quality of images. New approaches towards fidelity based image quality assessment includes Human Vision Modeling based Metrics, Information Theoretic Metrics and Structural Similarity Metrics. This paper presents a brief survey of first two approaches and discusses structural similarity approach in detail. Structural information in an image is defined as those attributes that represent the structure of the object in the scene, independent of the average luminance and contrast. The structural similarity approach incorporates image structures along with perceptual modeling in calculating image fidelity values.

KEYWORDS: Mean Square Error, Peak Signal to Noise Ratio, Full Reference Image Assessment, Structural Similarity

I. INTRODUCTION

Recent advances in digital image technology, computational speed, storage capacity and networking have resulted into huge proliferation of digital images and videos. Digital images are subjected to wide variety of distortions during acquisition, processing, compression, storage, transmission and reproduction, any one of which may result in a degradation of visual quality of the image. Image quality can be accessed through ‘subjective’ and ‘objective’ metrics. Subjective assessment (assessment by human observer) is usually inconvenient, time consuming and expensive. Hence many researchers are now working on ‘objective quality assessment’ of images predicting perceived image quality by human observer.

For more than 50 years, the mean-squared error (MSE) along with the related quantity of peak signal-to-noise ratio (PSNR) has been the dominant full reference quality performance metric in the field of image processing. MSE and PSNR are popular because they are simple to calculate, have clear physical meanings, and are mathematically convenient in the context of optimization. However in many of image processing applications the MSE and PSNR exhibits weak performance and have been widely criticized for serious shortcomings, especially when dealing with perceptually important signals such as watermarked images. [1]. Owing to the poor performance of the MSE as a visual metric, interesting alternatives are arising in the image processing field. This paper presents a survey new image quality assessment metrics for image fidelity applications like image watermarking.

Section II of this paper discusses the conventional image quality assessment (QA) metrics namely MSE and PSNR and their limitations, section III presents a general view of new QA approaches. The structural similarity index (SSIM) along with its rationale and necessary mathematical background is presented in section IV. Applications and variations of SSIM are discussed in section V and this paper is concluded in Section VI.

II. CONVENTIONAL METRICS – MSE AND PSNR

We begin with a discussion of the MSE as an image fidelity measure. The goal of an image fidelity measure is to compare two images by providing a quantitative score that describes the degree of similarity/fidelity or, conversely, the level of error/distortion between them. Usually, it is assumed that one of the images is a pristine original, while the other is distorted or contaminated by errors.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

Suppose that $x = \{x_i | i = 1,2, \dots, N\}$ and $y = \{y_i | i = 1,2, \dots, N\}$ are two finite-pixel images, where N is the number of pixels in each image and x_i and y_i are the values of the *ith* pixel in x and y, respectively. The MSE between the signals is

$$MSE\ (x, y) = \frac{1}{N} \sum_{i=1}^{N} (x_i - y_i)^2 \tag{2.1}$$

In the MSE, we often refer to the error signal $e_i = x_i - y_i$

In the literature of image processing, MSE is often converted into a peak signal-to-noise ratio (PSNR) measure as

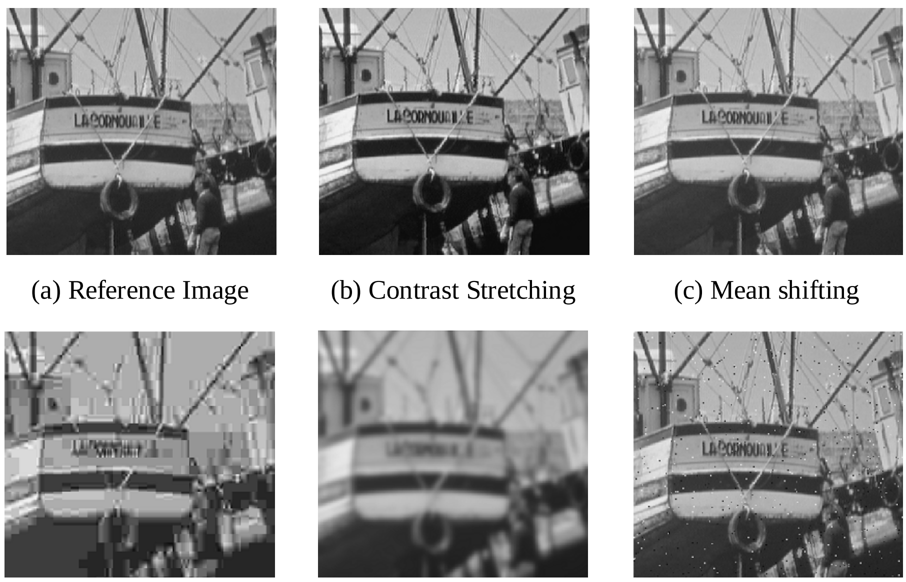
$$PSNR = 10 log_{10} \frac{L^2}{MSE} \tag{2.2}$$

where L is the dynamic range of allowable image pixel intensities. For example, for images that have allocations of *8 bits/pixel* of gray-scale, $L = 2^8 - 1 = 255$. The PSNR is useful if images having different dynamic ranges are being compared, but otherwise contains no new information relative to the MSE.

The MSE (and hence PSNR) is a popular image fidelity assessment metric since long because it is simple, it has a clear physical meaning, it is an excellent metric in the context of optimization and it is a desirable measure in the statistics and estimation framework. Finally, the MSE is widely used simply because it is a convention. But unfortunately MSE and PSNR are not very well matched to perceived visual quality [2, 3]. MSE (and hence PSNR) suffers with below limitations when used to assess image fidelity –

1. MSE do not consider temporal or spatial relationships between the samples of the original image and test image. In other words, if the original and distorted images are randomly re-ordered in the same way, then the MSE between them will be unchanged.
2. Image fidelity is independent of any relationship between the original image and the error signal. For a given error signal, the MSE remains unchanged, regardless of which original signal it is added to
3. MSE fails to incorporate knowledge about source, channel and receiver in an information communication framework

Fig. 2.1 (a) below shows an original image 'Boat' with figures (b) to (f) showing images with different distortions [3]. However each of the distorted images, which are visually very different from others, yields nearly identical MSE of 210 with the original image. This example clearly reveals limitations of MSE as fidelity metric.



(a) Reference Image (b) Contrast Stretching (c) Mean shifting

(d) JPEG Compression (e) Blurring (f) Salt & Pepper Noise

Fig 2.1 – Original Boat image (a) and its distorted versions (b) to (f)

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

III. What Are New Metrics

Following the learning out of communication theory, a good signal fidelity measure would need to be able to effectively and efficiently make use of knowledge about the transmitter, channel, and receiver. Depending on the application field and the type of the signals being considered, the nature of this knowledge might vary considerably; it is unlikely that there is a universal signal fidelity measure that works in all situations. Based on these rationales there are three types of alternate or new approaches towards fidelity based image quality assessment –

1. Human Vision Modeling based Metrics (HVM)
2. Structural Similarity Metrics (SSIM)
3. Information Theoretic Metrics

The most obvious type of information to incorporate into an image QA metric can be receiver information. HVM based metrics utilizes mathematical models of certain stages of human visual processing systems These metrics are designed by measuring thresholds of visibility of signals and noise in test images considering response of human to average brightness, contrast, spatial frequencies, orientation etc. Figure 3.1 shows a general framework for HVM based QA metrics utilizing error sensitivity. [3]

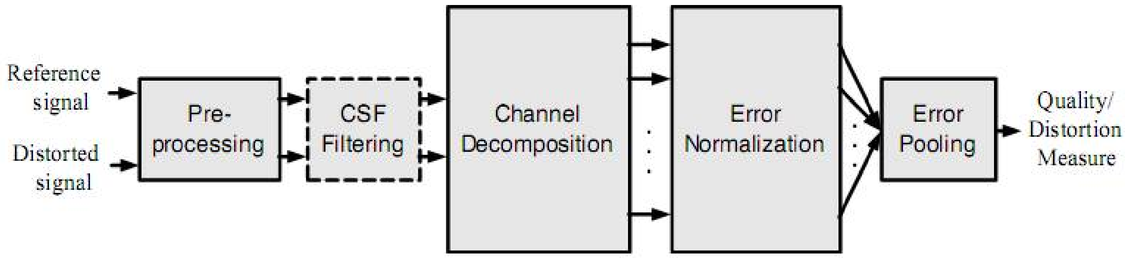


Fig 3.1 – A QA system based on error sensitivity

Most of such models are general purpose, in the sense that they do not assume any specific distortion type. They are intended to be flexible enough to be used in a variety of different applications. There are also many methods that are designed for specific applications. For example, many image fidelity measurement methods have been developed specifically for block-discrete cosine transfer (DCT) and wavelet-based image compression. Some of the general purpose HVM based QA metrics includes - Visible Differences Predictor, Sarnoff JND Vision Model, Teo and Heeger Model [1]. However HVM based models suffers with limitations of quality definition problem, natural image complexity problem, decorrelation problem and cognitive interaction problem.[1]

In information theoretic approach, the image quality can be viewed as the information fidelity problem rather that signal (image) fidelity problem [4]. The image source communicates through a channel that limits the amount of information that can flow through it, thereby introducing distortion. Output of image source is original image which is distorted through channel to become test image. Figure 3.2 shows this scenario.

Fig 3.2 – Information theoretic framework for image QA

Information fidelity criteria attempt to relate visual quality to the amount of information shared between original image and test image. This shared information can be quantified by commonly used measure – Mutual Information. Information theoretic approach requires accurate modeling of the image source, channel and receiver. Source modeling involves statistical modeling of the natural images (real world images). Statistical properties of natural images have been studied extensively to empirically aid for source modeling. The distortion model is a simple signal attenuation and additive Gaussian noise model in each subband of the frequency (eg- Wavelet subbands).

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

The HVS based (perceptual) models attempts to use knowledge about the receiver (human visual system) but do not use any knowledge about the transmitter (source of the images) and hence only accounts for non structural distortions. A better structural similarity approach exploits knowledge about natural image source (the transmitter) and of the human visual systems (the receiver). Section 4 describes structural similarity approach for image QA in detail.

IV. STRUCTURAL APPROACH FOR IMAGE QA- STRUCTURAL SIMILARITY INDEX

Natural (real world) images are highly structured. Their pixels exhibit strong dependencies, especially when they are spatially proximate, and these dependencies carry important information about the structure of the objects in the visual scene. Therefore, at least for image fidelity measurement, the retention/degradation of image structure should be an important ingredient. The structural similarity approach incorporates image structures along with perceptual modeling in calculating image fidelity values.

The structural approach for QA is based on assumption that the human visual system is highly adapted to extract structural information from the viewing field and therefore a measurement of structural similarity should provide a good approximation to perceptual image quality. Depending on how structural information and structural distortion are defined, there may be different ways to develop image QA metrics. Structural Similarity Index (SSIM) is one such measure.

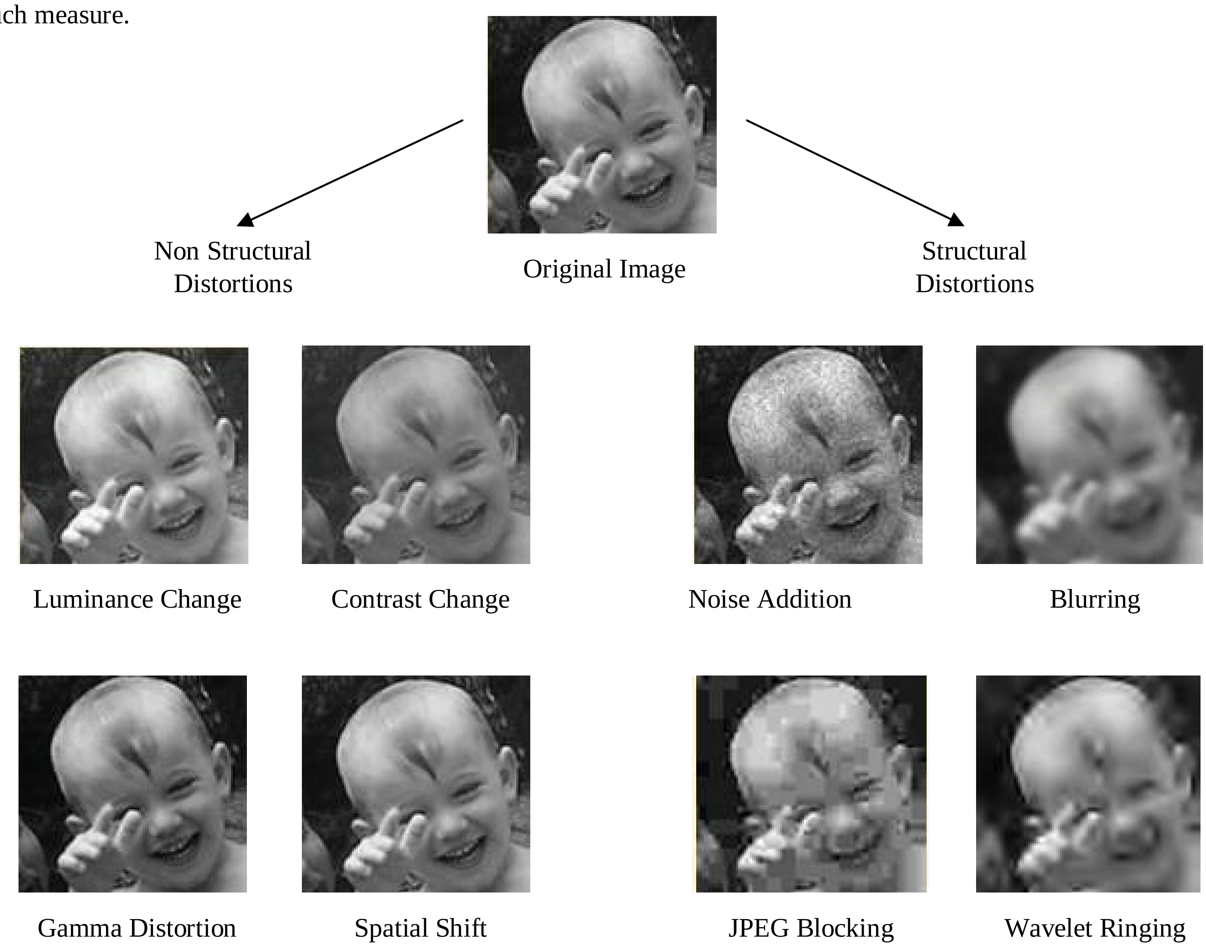


Fig 4.1 – Structural and Non structural distortions

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

The principle philosophy underlying the original SSIM approach is that the human visual system is highly accustomed to structural information in the image and this image structures shall be preserved in spite of attacks or degradation of image. Equivalently, a QA metric may seek to measure structural distortion to achieve image fidelity measurement. Figure 4.1 borrowed from [2] helps illustrate the distinction between structural and nonstructural distortions. In the figure, the nonstructural distortions (a change of luminance or brightness, a change of contrast, Gamma distortion, and a spatial shift) are caused by ambient environmental or instrumental conditions occurring during image acquisition and display. These distortions do not change the structures of images of the objects in the visual scene. However, other distortions (additive noise and blur and lossy compression) significantly distort the structures of images of the objects. If we view the human visual system as an ideal information extractor that seeks to identify and recognize objects in the visual scene, then it must be highly sensitive to the structural distortions and automatically compensates for the nonstructural distortions. Consequently, an effective objective signal fidelity measure should simulate this functionality. The luminance of the surface of the object being observed is the product of the illumination and the reflectance, but the structures of the objects in the scene are independent of the illumination. Hence to explore structural information in the image, influence of the illumination must be separated out. Structural information in an image is defined as those attributes that represent the structure of the object in the scene, independent of the average luminance and contrast. Since, the luminance and contrast may vary across a scene; local illumination and contrast are considered.

The main ideas of SSIM were introduced in [5], and more formally distilled in [6] and [7]. The basic form of SSIM is very easy to understand. Suppose that *x* and *y* is local image patches taken from the same location of two images that are being compared. We may consider one of the image patches to have perfect image quality (original image patch), then SSIM can serve as a quantitative measurement of the quality of the second image patch (test image patch). The local SSIM index measures the similarities of three elements of the image patches: the similarity of the local patch luminance (brightness values), the similarity of the local patch contrasts, and the similarity of the local patch structures. These local similarities are expressed using simple, easily computed statistics, and combined together to form local SSIM [3].

Below is the step by step procedure to calculate SSIM index-

1. First luminance of two image patches is compared. As summing discrete image signal, this is estimated as mean intensity

   $$\mu_x = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{4.1}$$

   where *N* are the number of pixels in each image patches *x* and *y*. The luminance comparison function *l(x,y)* is then function of μ_x *and* μ_y

2. Mean intensity of the respective signals *x* and *y* are removed from the signals *x* and *y*
3. Standard deviation of the signal is calculated as an estimate of the signal contrast

   $$\sigma_x = \frac{1}{N-1}\left\{\sum_{i=1}^{N}(x_i - \mu_x)^2\right\}^{\frac{1}{2}} \tag{4.2}$$

4. The signals are normalized (divided) by its own standard deviation so that the two signals being compared have unit standard deviation. The structured comparison *s(x,y)* is then carried out on these normalized signals $(x-\mu_x)/\sigma_x$ and $(y-\mu_y)/\sigma_y$. The correlation (inner product) between these is a simple and effective measure to quantify the structural similarity.

5. The structural comparison is then defined as

   $$s(x,y) = \frac{\sigma_{xy}+C_3}{\sigma_x\,\sigma_y+C_3} \tag{4.3}$$

   where σ_{xy} can be estimated as

   $$\sigma_{xy} = \frac{1}{N-1}\sum_{i=1}^{N}(x_i-\mu_x)\,(y_i-\mu_y) \tag{4.4}$$

   Constant C_3 is introduced to avoid instability when $\sigma_x\,\sigma_y$ is very close to zero

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

6. For luminance comparison, we define
   $l(x, y) = \frac{2\,\mu_x\mu_y+C_1}{\mu_x^2+\mu_y^2+C_1}$ (4.5)
   where constant C_1 is introduced to avoid instability when $\mu_x^2 + \mu_y^2$ is very close to zero.
7. The contrast comparison function takes a similar form
   $c(x, y) = \frac{2\,\sigma_x\sigma_y+C_2}{\sigma_x^2+\sigma_y^2+C_2}$ (4.6)
   where constant C_2 is introduced to avoid instability when $\sigma_x^2+\sigma_y^2$ is very close to zero
8. Three comparisons of 4.3, 4.5 and 4.6 are combined to result into SSIM index between image patches x and y

   $SSIM(x, y) = [i(x, y)]^{\alpha}.[c(x, y)]^{B}.[s(x, y)]^{\gamma}$ (4.7)
   where $\alpha > 0, \beta > 0, \gamma > 0$ are parameters used to adjust the relative importance of the three components in final SSIM. For $\alpha = \beta = \gamma = 1$

   $SSIM(x, y) = \left(\frac{2\,\mu_x\mu_y+C_1}{\mu_x^2+\mu_y^2+C_1}\right)\left(\frac{2\,\sigma_x\sigma_y+C_2}{\sigma_x^2+\sigma_y^2+C_2}\right)\left(\frac{\sigma_{xy}+C_3}{\sigma_x\,\sigma_y+C_3}\right)$ (4.8)
9. For image quality assessment, the SSIM index is applied locally rather than globally. This is because image features are highly non stationary. Additionally, using local windows provides quality map of an image, as apposed to a single index for an entire image, thus providing valuable information about local quality.
   The quantities $\mu_x, \sigma_x, \mu_y, \sigma_y$ and σ_{xy} are computed in a local sliding window that moved pixel by pixel over entire image. To avoid blocking artifacts, the resulting values are weighted using a circular symmetric *11X11* Gaussian function. The weighing function W =$\{w_i|1,2, \dots, N\}$, with standard deviation 1.5 samples normalized to unit sum $\sum_{i=1}^{N} w_i = 1$.
   The estimates of the local statistics are then modified as
   $\mu_x = \frac{1}{N}\sum_{i=1}^{N} w_i x_i$ (4.9)
   $\sigma_x = \frac{1}{N-1}\left\{\sum_{i=1}^{N} w_i (x_i - \mu_x)^2\right\}^{\frac{1}{2}}$ (4.10)
   $\sigma_{xy} = \frac{1}{N-1}\sum_{i=1}^{N} w_i (x_i - \mu_x)(y_i - \mu_y)$ (4.11)

10. As a final step, the overall quality of an image can be defined by a single QA index named Mean SSIM (MSSIM).
    Then, this Mean SSIM between two images *X* and *Y* can be evaluated as,
    $MSSIM(X, Y) = \frac{1}{M}\sum_{j=1}^{M} SSIM(x_j y_j)$ (4.12)
    where *X* and *Y* are reference and test images, respectively; x_j and y_j are the image contents at j^{th} local window and *M* is the number of local windows in the image.

Figure 4.2 shows the block diagram for calculation of SSIM [3]. The SSIM index is symmetric: $S(x, y) = S(y, x)$, so that two image patches being compared give the same index value regardless of their ordering. It is also bounded: $-1 < S(x, y) \leq 1$, achieving maximum value $S(x, y) = 1$ if and only if $x = y$.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

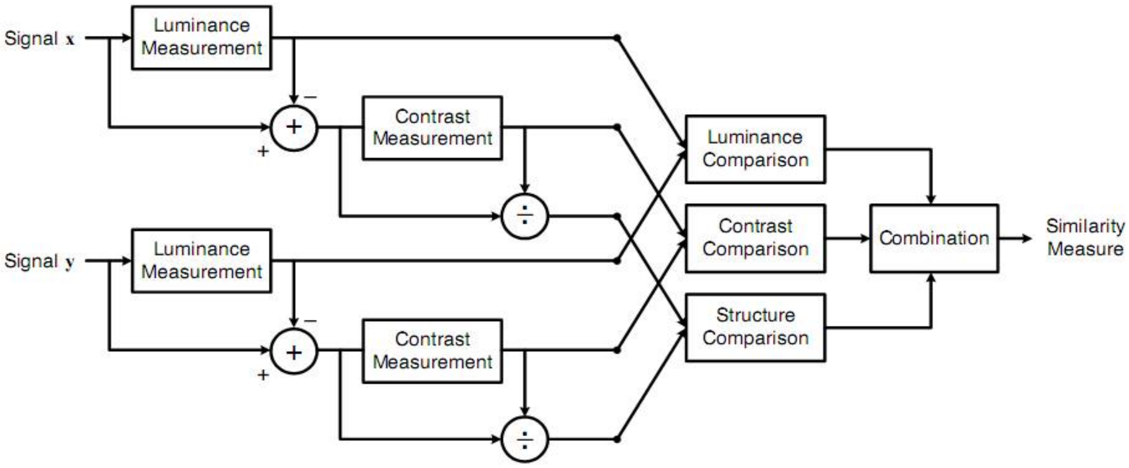


Fig 4.2 – Block Diagram for calculation of SSIM index

Despite its simplicity, the SSIM index performs remarkably well across a wide variety of image and distortion types as has been shown in intensive human studies [8]. By example, Figure 4.3 reproduced from [2], shows the SSIM scores of images having near identical MSE values. Without much effort, it can be seen that the SSIM scores are much more consistent than the MSE scores relative to visual perception. Luminance-shifting and contrast-stretching, which generally does not degrade image structure, lead to very high SSIM values, while noise contamination and excessive JPEG-compression lead to low SSIM values.

V. APPLICATIONS AND FUTURE OF SSIM

SSIM has been used for evaluating image processing results in a rapidly increasing number of exciting applications. Some of them includes - image fusion, image compression, image watermarking, chromatic image quality, retinal and wearable displays, video hashing, wireless video streaming, visual surveillance, radar imaging, digital camera design, infrared imaging, MRI imaging, chromosome imaging, remote sensing, target recognition. [2]

An exciting consideration is the possibility of numerous extended applications beyond image processing, since the SSIM index does not rely on specific image or visual models. The generic definition of SSIM suggests that it should find broad applicability.

A drawback of the basic SSIM index is its sensitivity to relative translations, scaling and rotations of images, as seen in Figure fig. 4.3 (h)–(l). This is undesirable and contradictory to the philosophy of structural similarity, since small geometric distortions are nonstructural. To handle such situations, a wavelet domain version of SSIM, called the complex wavelet SSIM (CW-SSIM) index was developed [9]. The CW-SSIM index is also inspired by the fact that local phase contains more structural information than magnitude in natural images [10], while rigid translations of image structures leads to consistent phase shifts.

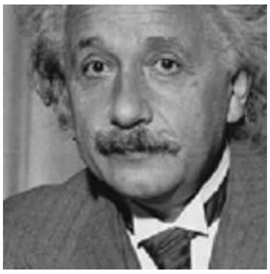
The SSIM index can also be applied to videos. SSIM has been deployed in this manner with rather good results [11], and is now deployed as a basic video quality assessment tool in popular public-domain software such as the Moscow State University video quality measurement tool [12] and the award-winning freeware H.264 codec x.264 [13]. A MATLAB code for SSIM implementation is available at [14].

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

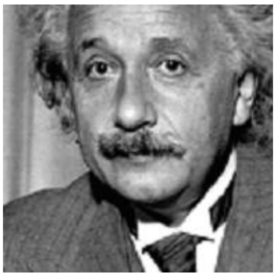
(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016



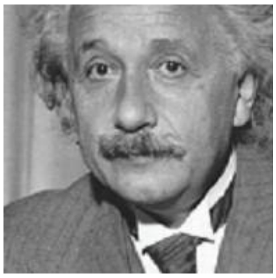
(a) Reference Image

MSE=0, SSIM=1



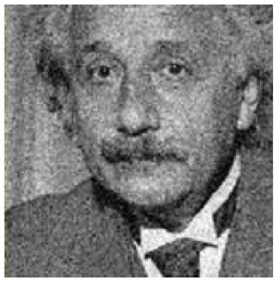
(b) Mean Contrast Stretch

MSE=306, SSIM=0.928



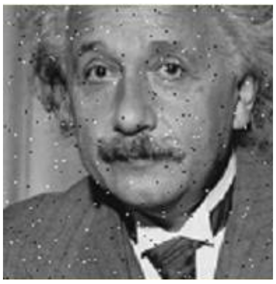
(a) Luminance Shift

MSE=309, SSIM=0.987



(d) Gaussian Noise

MSE=309, SSIM=0.576

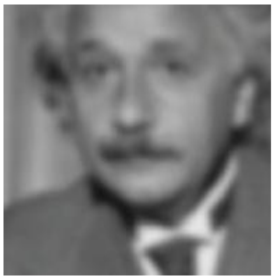


(e) Impulsive Noise

MSE=313, SSIM=0.730

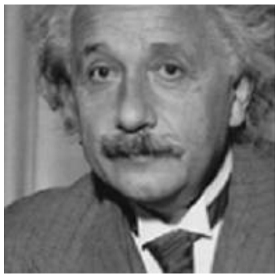
(f) JPEG Compression

MSE=309, SSIM=0.580



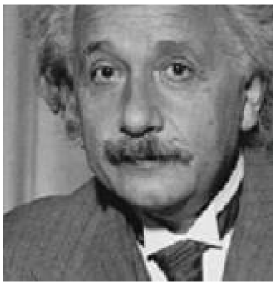
(g) Blurring

MSE=308, SSIM=0.641



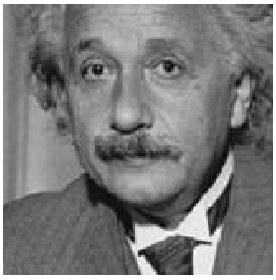
(h) Spatial Scaling

MSE=694, SSIM=0.505



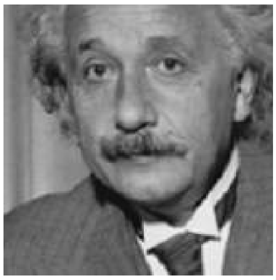
(i) Spatial Shift - Right

MSE=871, SSIM=0.404



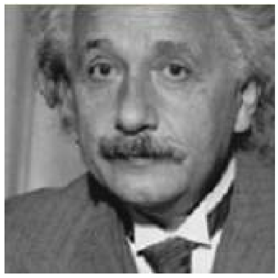
(j) Spatial Shift - Left

MSE=873, SSIM=0.399



(k) Rotation - Anticlock

MSE=590, SSIM=0.549



(l) Rotation - Clock

MSE=577, SSIM=0.551

Fig 4.3 - Comparison of image fidelity measures with different types of distortions.

One of the interesting directions for future research work is use of SSIM index as objective functions for image optimization problems like restoration, quantization and denoising. Another path is to use of SSIM as a measure of fidelity for non image/video applications like audio or speech.

VI. CONCLUSION

This paper first discusses MSE and PSNR's limitations when used as full reference quality assessment metrics for images. The new generation metrics are focused on Human Vision Modelling, Structural Similarity and Information Theoretic approach. SSIM has proved to be the most promising metric out of these. SSIM attempts to

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

measure the closeness between original image and test (distorted) image by measuring the amount of structural distortion present in the test image. Results of SSIM are found to be encouraging and it shall be used as an image fidelity measure in applications where perceptual criteria might be relevant.

REFERENCES

1. Editor- Al Bovik, "The Essential Guide to Image Processing", Second Edition, Elsevier
2. Zhou Wang and Alan C. Bovik, "Mean Squared Error: Love it or leave it", IEEE Signal Processing Magazine, January 2009
3. Zhou Wang and Alan C. Bovik, "Image quality assessments: From error visibility to structural similarity", IEEE Transaction on Image Processing, Vol 13, No 04, April 2004
4. Hamid Rahim Sheikh, Alan C. Bovik, Gustavo de Veciana, "An information fidelity criterion for image quality assessment using natural scene statistics", IEEE Transaction on Image Processing, Vol 14, No 12, December 2005
5. Z.Wang and A.C.Bovik, "A universal image quality index", IEEE Signal Processing Lett., vol. 9, no. 3, Mar.2002
6. Z.Wang and E.P.Simoncelli, "Maximum differentiation (MAD) competition: A methodology for comparing computational models of perceptual quantities," J. Vis., vol. 8, no. 12, Sept. 2008.
7. Z.Wangand A.C.Bovik, "Modern Image Quality Assessment. San Rafael, CA: Morgan & Claypool", 2006.
8. H.R.Sheikh, M.F.Sabir, and A.C.Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," IEEE Transaction on Image Processing, vol. 15, no. 11, Nov.2006
9. Z.Wangand E.P.Simoncelli, "Translation insensitive image similarity in complex wavelet domain," in Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing, Mar.2005
10. T.S.Huang, J.W.Burdett, and A.G.Deczky, "The importance of phase in image processing filters," IEEE Trans. Acoust., Speech Signal Process., vol. 23, no. 6, Dec.1975
11. Z.Wang, L.Lu and A.C.Bovik, "Video quality assessment based on structural distortion measurement," Signal Process. Image Commun., vol. 19, no. 2, Feb.2004.
12. www.compression.ru/video/quality_measure/
13. www.videolan.org/developers/x264.html
14. Z.Wang. The SSIM index for image quality assessment. [Online].Available: http://www.cns.nyu.edu/lcv/ssim

BIOGRAPHY

Ms. Parineeta Sachin Gengaje is a Final Year student for Electronics and Telecommunication Engineering UG Program at Cummins College of Engineering for Women, Pune. Her area of interest includes Image Processing and Embedded Systems.