International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 22 (2017) pp. 12755-12761
© Research India Publications. http://www.ripublication.com

A Joint Histogram - 2D Correlation Measure for Incomplete Image Similarity

Nisreen Ryadh Hamza
MSc Candidate, Faculty of Computer Science & Mathematics, University of Kufa, Iraq.
ORCID: 0000-0003-0607-3529

Hind Rustum Mohammed
Professor, Faculty of Computer Science and Mathematics, University of Kufa, Najaf, Iraq.

Zahir M. Hussain
Professor, Faculty of Computer Science & Mathematics, University of Kufa, Najaf, Iraq.

ORCID: 0000-0002-1707-5485

Abstract

A hybrid similarity measure is proposed for evaluate the similarity between gray images. The well-known Structural Similarity Index Measure has been designed using a statistical method that fails under high noise .The proposed scale , referred by JhCorr2, uses a mixture of two parts: the first part is based on a two-dimension correlation , while the second part is information - theoretic that using concept of joint histogram with original histogram . The new measure shows the features of information - theoretic approaches and statistical approaches. The proposed similarity measure is robust under conditions of noise and incomplete information. The new measure superior the classical structure similarity and correlation in detecting image similarity at low PSNR under Gaussian, Impulse and multiplicative noise. A face image database AT&T (The image size is 92x112 pixels, with 256 grey levels per pixel) and the well-known image similarity techniques SSIM, corr2 and Jhcorr2 are considered. Peak signal to noise ratio (PSNR), window length in pixel and maximum burst length in pixel were used in this test.

Keywords: Joint histogram, statistical similarity, information-theoretic similarity, image similarity, Gaussian noise, impulse noise, multiplicative noise, 2D correlation.

INTRODUCTION

In image processing, applications that require comparing two images according to their content, image matching is an essential component in this process. One of the most important examples is the image database retrieval systems [1]. Image similarity has become in the recent years a basic point in image processing applications like monitoring, image compression, restoration, and many other application.

Image similarity can be defined as the difference between two images, and image similarity measure is a numerical difference between two different images under comparison. Similarity techniques can be classified into: the statistical techniques and the information theoretical techniques [2]. An easy way to measure the similarity between two images is calculate the Mean-Squared Error (MSE) [3, 4, 5]. A significant objective measure is Structure Similarity Index Measure was proposed by (Wang and Bovik), 2004 [1,6].

Image discrimination has become an interesting subject over the past decennium because of its implementation in many fields such safety, identity authentication and video monitoring. Different ways for image discrimination, especially for Face recognition, have been proposed [7]. Many algorithms for recognition have recently been prepared according to similarity measure between two images [1].

In this work we propose a hybrid similarity measure that combines features of a two-dimension correlation and the information - theoretic features represented by a joint histogram with original histogram. The proposed measure is shown outperform structure similarity in detecting similarity under Gaussian, impulse and multiplicative noise.

The paper is organized as follows: Section 2 deals with Similarity techniques. Section 3 presents the incomplete images .Section 4 presents the proposed measure. In Section 5, environment test. Section 6, experimental results. The conclusions of this study are given in Section 7.

SIMILARITY TECHNIQUES

Similarity techniques can be classified into: statistical techniques and information theoretical techniques**.**

A. Statistical Techniques:

A valuable information can be obtained from image by compute statistical measurements such as mean, variance and standard deviation. This information can be used to compute image similarity [1]. Mean is average of a range of values or quantities, computed by dividing the total of all values by the number of values [2]

$$\mu_A = \frac{1}{M}\sum_{i=1}^{M} A_i \quad (1)$$

where M is the number of values is, A_i is the single value in the data set A and $p(A_i)$ is the probability of A_i. When $\{A_i\}$ is probabilistic with $p(A_i)$ is the probability of A_i, then

$$\mu_A = \sum_{i=1}^{M} A_i \, p(A_i).$$

Variance and standard deviation are measures of difference of a set of data values. The standard deviation is equal to the square root of the variance [9].

$$\sigma_A = \sqrt{\frac{1}{M-1}\sum_{i=1}^{M}(A_i - \mu_A)^2} \quad (2)$$

where A represent a set of data values, M is the number of values and μ_A is the mean value of A.

Structural Similarity Measure

The measure proposed by Wang and Bovik. (2004) which was called SSIM, used distance function to measure the likeness depended on statistical feature Equ.1 shows the measure:

$$\mathrm{ssim}(a,b) = \frac{(2\mu_a + co_1)(2\sigma_{ab} + co_2)}{(\mu_a^2 + \mu_b^2 + co_1)(\sigma_a^2 + \sigma_b^2 + co_2)} \quad (3)$$

where μ_a, μ_b represents the means and $\sigma^2{}_a$, $\sigma^2{}_b$ represents the variance of a and b, σ_{ab} is the covariance of a and b, and co_1 and co_2 are constants inserted to avoid division by zero, and are defined as $co_1 = (T_1L)^2$ and $co_1 = (T_2L)^2$, T_1 and T_2 are small constants and L (maximum pixel value) [1].

B. Information - Theoretic Techniques:

Information-theoretic technique is the similarity measure for images. Information-theoretic technique is aimed to find the similarity between images according to their content (intensity values) [10]. In 2014, A. F. Hassan, D. Cai-lin and Z. M. Hussain proposed a new measure that based on joint histogram. The measure outperforms statistical similarity of SSIM; it has the ability to detect similarity under significant noise (low PSNR) [2].

HISTOGRAM AND JOINT HISTOGRAM

The histogram shows how levels of brightness are occupied in an image. This levels divide into a series of intervals— then count how many values fall into each interval. For example, if an image pixel is 8-bit, then the brightness ranges from zero to 255 [13]. The joint histogram JH of a pair of gray images can be defined as a function of two variables JH (a,b), a is the first gray image and b is the second gray image. The value of JH at coordinates (i,j) can be defined as the number of corresponding pairs containing gray level i and j in the first and second images respectively. Many researchers had been utilize joint histogram such as [11, 12].

Incomplete Image

Create the incomplete image based on:

1. Window length

Missing area of image is square W-by-W; for example window length=10 pixel, the missing area = 10*10.

2. Burst noise

Burst Noise is a type of internal electronic noise (undesirable electrical energy) that produced inside the communication system or in the receiver. It occurs because of imperfections in semiconductors material and heavy ion implants. Burst noise consists of sudden step-like transitions between two or more current levels, as high as several hundred microvolts, at random and unpredictable times lasts for several milli-seconds [14]. Effect Burst noise on image represent by strings of pixel errors, each with random length and waiting times Wt between bursts are random (distribution of burst length and waiting times is Uniform [15], Poisson [16] or Rayleigh [17]). Each string is set of 0's.

The Proposed Measure (Joint istogram -2D Correlation)

The proposed measure uses a combination of correlation 2D and Histogram Similarity. It takes advantages of both statistical features and information theoretic features. An image dependent measure with better result than statistical measure and information theoretic measure individually is proposed as follows:

$$\rho(x,y) = \sqrt{R(x,y)k + r(x,y)(1-k)} \quad (4)$$

where $R(x,y)$ is the correlation 2D measure between image x and y and it is given as follows:

$$R(x,y) = \frac{\sum_i\sum_j (x(i,j)-\bar{x})(y(i,j)-\bar{y})}{\sqrt{(\sum_i\sum_j (x(i,j)-\bar{x})^2)(\sum_i\sum_j (y(i,j)-\bar{y})^2)}} \quad (5)$$

and $r(x,y)$ is the information - theoretic measure between image x and y. The researcher had been utilized the joint histogram and then combined it with the original histogram as follows as:

$$Q(x,y) = \sqrt{\frac{\sum_i\sum_j \left[(JH_{ij} - JH_{ji})\frac{1}{h_i + b}\right]^2}{2L^2}} \quad (6)$$

where, JH is the Joint Histogram, h_i is the original reference image histogram and b is a small positive constant to avoid

division by zero. Note: $Q(x,y) \geq 0$.
The above value can be normalized by using the maximal error estimate value $Q_\infty(x,y)$ in significant noise as follows:

$$z(x,y) = \frac{Q(x,y)}{Q_\infty(x,y)} \qquad (7)$$

$$r(x,y) = 1 - z(x,y) \qquad (8)$$

where $0 \leq r(x,y) \leq 1$. Note that k is a small positive constant: $0 < k < 1$.

Algorithm:
Input:
Images x and y, which are the reference image and the incomplete, k is small constant and L = 255 which represents the maximum pixel value.
Output:
Similarity, a number ranging between 0 and 1.
Step 1: Transform the images into double precision.
Step 2: Set R

$$R(x,y) = \frac{\sum_i \sum_j (x(i,j)-\bar{x})(y(i,j)-\bar{y})}{\sqrt{(\sum_i \sum_j (x(i,j)-\bar{x})^2)(\sum_i \sum_j (y(i,j)-\bar{y})^2)}}.$$

Step 3: Set Q:

$$Q(x,y) = \sqrt{\frac{\sum_i \sum_j \left[(JH_{ij} - JH_{ji}) \frac{1}{h_i + c}\right]^2}{2L^2}}$$

Step 4: Set $Q_\infty = Q(x,y)$ when noise is maximum.
Step 5: normalization: $z = Q/\,Q_\infty$.
Step 6: Set $r = 1 - z$.
Step 7: Compute $\rho(x,y)$ =JhCorr2(x,y)
$\rho(x,y) = \sqrt{R(x,y) \times k + r(x,y) \times (1-k)}$.
End of Algorithm

TEST ENVIRONMENT

Four types of noise have been considered in simulation and testing: Gaussian noise, multiplicative noise which are the most popular noise types that are encountered in signal processing systems, impulse noise which is one of the most popular noise types in image processing systems, and burst noise which is electronic noise .

To test the performance of the proposed measure, a human face images have been considered (from AT&T database, [8]).

RESULTS AND DISCUSSION

The proposed measure has been tested and simulated using MATLAB.

A. Performance vs. window length

Performance of similarity measures has been tested according to window length of missing information. When the window length increases, Jhcoor2 measure can detect similarity better than SSIM. Fig 1 shows the test images and performance of similarity vs. window length.

Table 1: Comparison of similarity measures for same images vs. window length

Measure / window length	SSIM	Corr2	JHCORR2
20	0.9520	0.9771	0.9954
40	0.8174	0.7540	0.9495
60	0.6088	0.4620	0.8859
80	0.3269	0.1972	0.8239

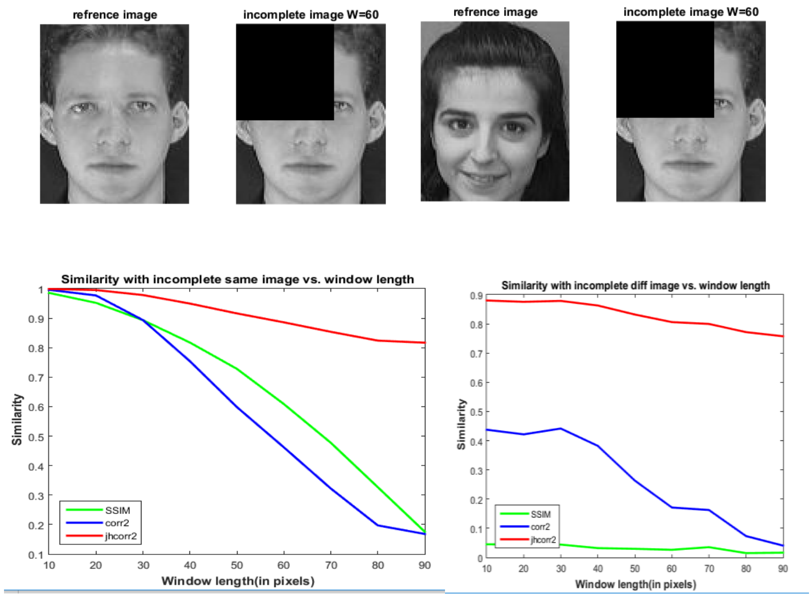

Figure 1: Performance of SSIM, corr2 and JhCorr2 using similar images and different images vs. window length

B. Performance under Gaussian Noise:

The proposed measure has been tested under Gaussian noise. Results are shown in Figures 2. Table 2 shows a comparison between the SSIM, corr2, and JhCorr2 for face images. The proposed measure gives larger similarity under Gaussian noise (when low PSNR).

Table 2: Comparison of similarity measures for same images under Gaussian noise

Measure PSNR	SSIM	Corr2	JHCORR2
-50	0.0017	0.0061	0.0316
-20	0.0025	0.0251	0.2542
0	0.0100	0.1228	0.4407
30	0.5022	0.7828	0.9870
50	0.8702	0.7925	0.9897

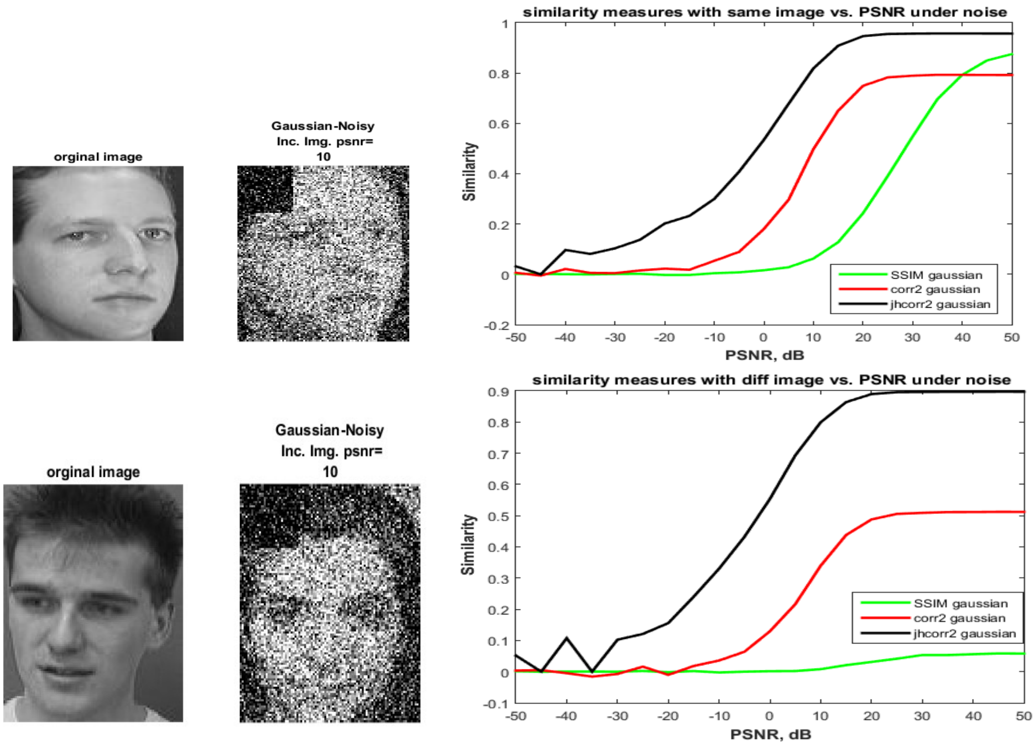

Figure 2: Performance comparison of SSIM, corr2, and JhCorr2 under Gaussian noise and window length 30.

C. Performance under multiplicative Noise:

Second test was performed under multiplicative noise. Results are shown in Table 3 and Figure 3. It can be note that the proposed measure gives better under multiplicative noise (when low PSNR).

Table 3: Comparison of similarity measures using same images under multiplicative noise

Measure PSNR	SSIM	corr2	JHCORR2
8.7162	0.0435	0.4090	0.6980
20.6930	0.3129	0.8248	0.9622
29.1123	0.5622	0.8695	0.9736
48.3285	0.8732	0.8774	0.9752

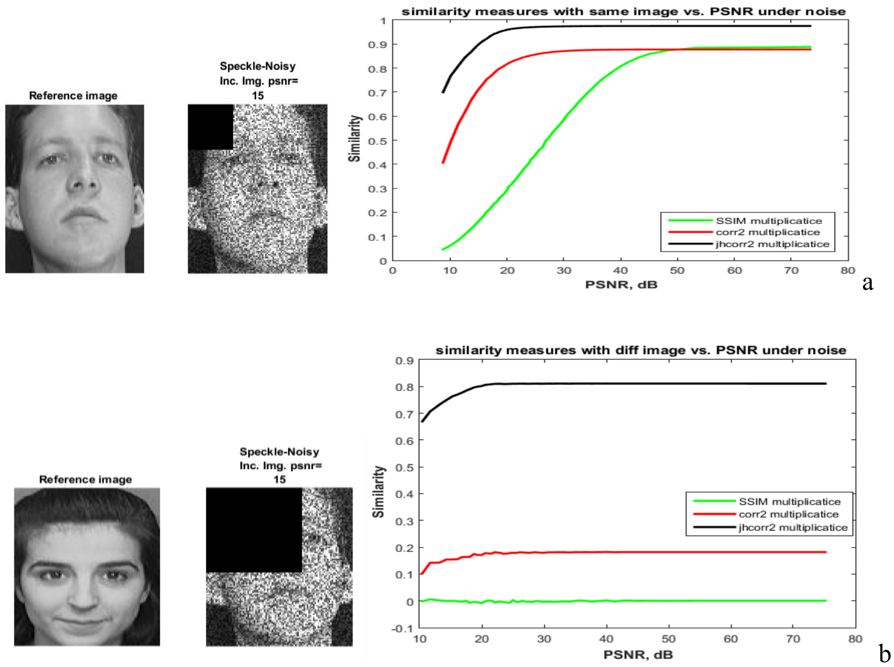

Figure 3: Performance comparison of SSIM and JhCorr2 using face images under multiplicative noise ((a) window length 30 (b) window length 60).

D. Performance under Impulsive Noise:

Third test was performed under impulsive noise. Results are shown in Table 4 and Figure 4. It can be seen that the proposed measure gives better results under impulsive noise for human face images.

Table 4: Comparison of similarity measures using same images under multiplicative noise.

Measure / PSNR	SSIM	corr2	JHCORR2
5.0680	0.0011	0.0201	0.0856
15.2293	0.2602	0.7485	0.9176
28.3971	0.8030	0.9372	0.9861
51.2357	0.8975	0.9493	0.9898

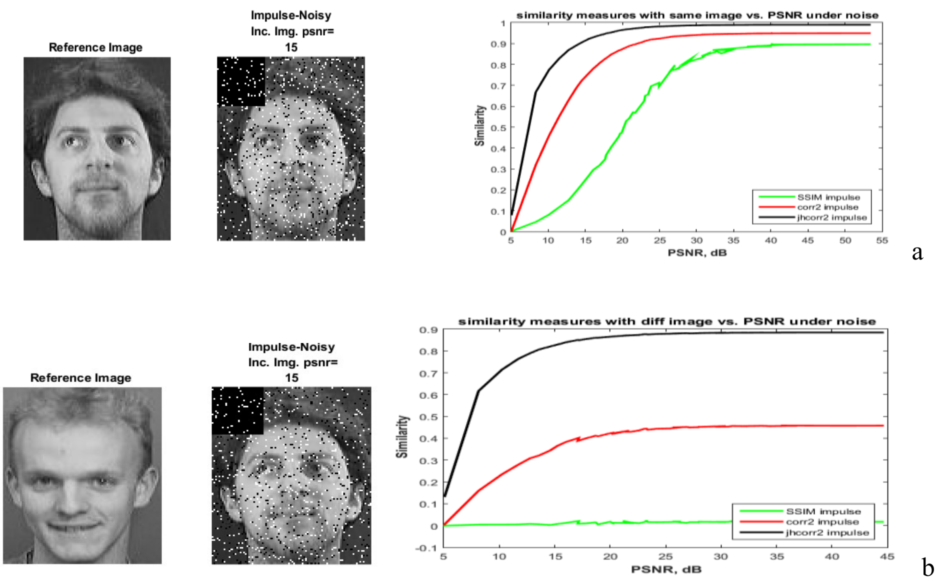

Figure 4: Performance of SSIM and JhCorr2 using face images under multiplicative noise (window length 30).

International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 22 (2017) pp. 12755-12761

E. Performance under Burst Noise:

The length of the burst (the error string) base on the maximum burst length, where generated the numbers of random according to maximum burst length. There are three cases for waiting time and burst length are uniform or Poisson or Rayleigh. Observe that a Jhcorr2 measure, gives a similarity despite maximum burst length is large. Performance of similarity measures is tested according to maximum burst length as shown in Figure 5.

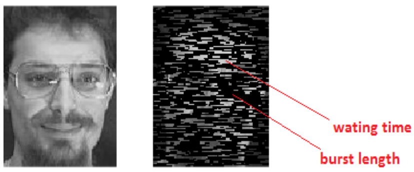

Table 5: Comparison of similarity measures burst noise (burst length uniform).

Measure / maximum burst length.	JHCORR2 Wt uniform	JHCORR2 Wt Poisson	JHCORR2 Wt Rayleigh
7	0.9019	0.8875	0.8925
17	0.8366	0.8173	0.8136
27	0.8048	0.7834	0.7860

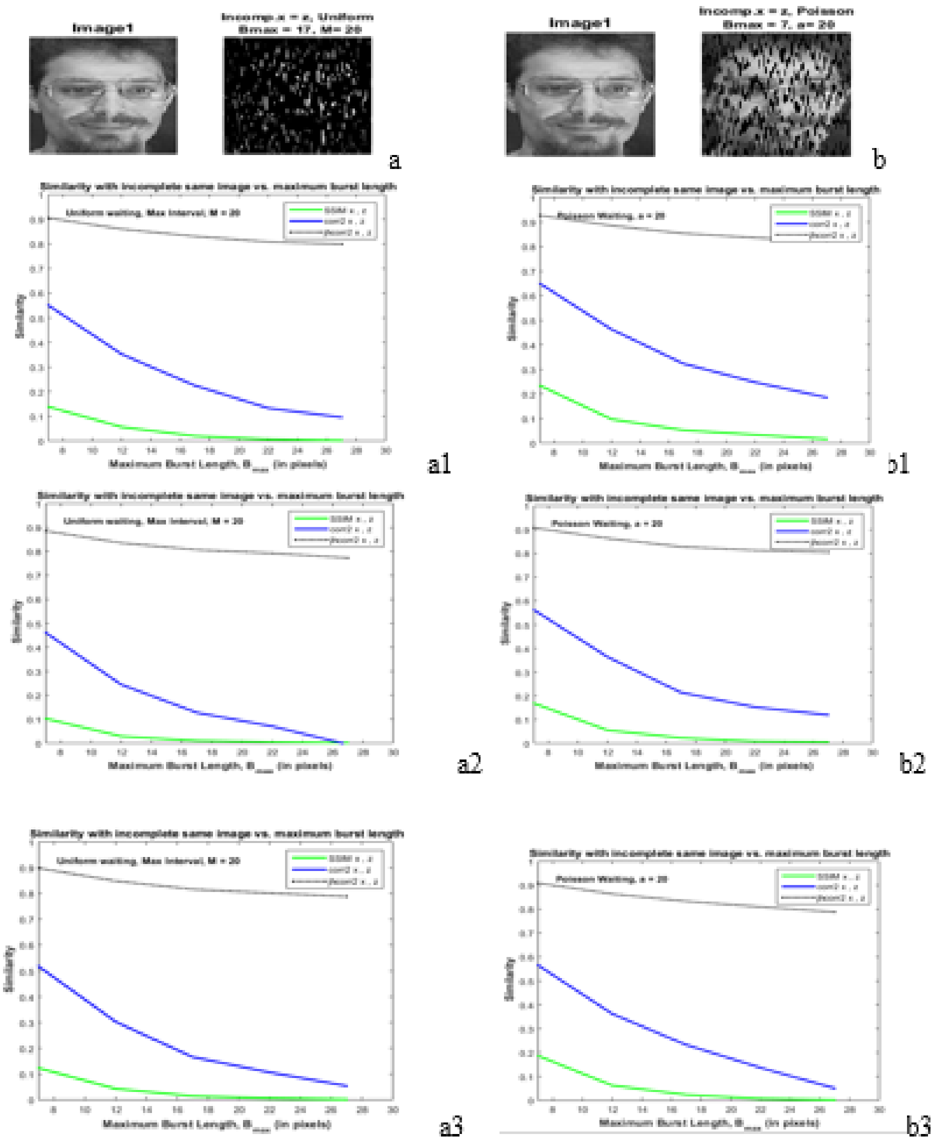

Figure 5: Performance comparison of SSIM and JhCorr2 using face images under burst noise ((a) burst length is uniform (b) burst length is Poisson ,waiting time (a1,b1)uniform ,(a2,b2) Poisson and (a3,b3) Rayleigh) .

CONCLUSION

1. A hybrid similarity measure, called joint histogram -two dimension correlation (JhCorr2), has been proposed.
2. Design of the proposed measure is based on a combination of information theoretic features and statistical features. A joint histogram with original histogram have been used as information - theoretic tool, and 2D correlation has been used as a statistical tool
3. The proposed measure has been tested versus the well-known structural similarity (SSIM) under non ideal conditions of noise (Gaussian, impulse and multiplicative noise) and incomplete information.
4. The new measure gave better performance (more similarity) than the structure similarity under non ideal conditions (80% when the window length is large, 25% when the noise is very high PSNR -20 and 80% when Maximum burst length is large).

REFERENCES

[1] Z. Wang, A. Bovik, H. Sheikh, E. Simoncelli, Image quality assessment: From error visibility to structural similarity, IEEE Trans. Image Process. 13 (4) (2004).

[2] A. F. Hassan, D. Cai-lin, Z. M. Hussain, An information-theoretic image quality measure: comparison with statistical similarity, Journal of Computer Science 10 (11) (2014).

[3] P. Premaratne, M. Premaratne, New structural similarity measure for image comparison, ProceedingsoftheInternationalConferenceonEmerging IntelligentComputing Technology and Applications, ICIC 2012, Huangshan, China (2012).

[4] A. M. Eskicioglu, Quality measurement for monochrome compressed images in the past 25 years, IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP2000) (2000).

[5] B. Girod, Psychovisual aspects of image processing: What's wrong with mean squared error?, Proceedings of the Seventh Workshop on Multidimensional Signal Processing (1991).

[6] J. Goldberger, S. Gordon, H. Greenspan, An efficient image similarity measure based on approximations of kl-divergence between two Gaussian mixtures, IEEE International Conference on Computer Vision (ICCV) (2003).

[7] W. Zhao, R. Chellappa, P. J. Phillips, A. Rosenfeld, Face recognition: a literature survey, ACM Computing Surveys (2003).

[8] AT & T Laboratories Cambridge, AT & T face database (formerly 'the ORL Database of Faces'), http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html (2002).

[9] M. Bland and D. G. Altman, "Statistics notes: measurement error," BMJ volume 312, 1996.

[10] R. Soundararajan and A. C. Bovik, "Survey of information theory in visual quality assessment," Springer-Verlag London, 2013.

[11] G . Pass and R. Zabih, "Comparing images using joint histograms," Multimed. System-Springer, 1999.

[12] H. Chen and P. K. Varshne, "Mutual Information-Based CT-MR Brain Image Registration Using Generalized Partial Volume Joint Histogram Estimation," IEEE Transactions on Medical Imaging, 2003.

[13] M. S. Nixon and A. S. A. Nixon, in Feature Extraction and Image Processing, 2Newnes, 2002.

[14] Alicja Konczakowska, Bogdan M. Wilamowski" Noise in Semiconductor Devices" (2010).

[15] Charles Boncelet, "Image Noise Models". In Alan C. Bovik. Handbook of Image and Video Processing. Academic Press, 2005.

[16] Frank A. Haight (1967). Handbook of the Poisson Distribution. New York: John Wiley & Sons.

[17] Hogg, R. V., and Craig, A. T., Introduction to Mathematical Statistics, Macmillan Publishing Co., Inc., New York, (1978).