**An Efficient InformationTheoretic Measure for Image Similarity**

**Farah M. Altufaili1, Zahir M. Hussain2 and Hind Rustum Mohammed3**

1 Department of Computer Science, University of Kufa, Najaf, Iraq,

*farahm.altufaily@student.uokufa.edu.iq*

2 Department of Computer Science, University of Kufa, Najaf, Iraq,

School of Engineering, Edith Cowan University, Australia

*Zahir.hussain@uokufa.edu.iq*

3 Department of Computer Science, University of Kufa, Najaf, Iraq,

*hindrustum.shaaban@uokufa.edu.iq*

**ABSTRACT**

A new image information theoretic image similarity measure is proposed for grey scale image. The proposed measure is based on equation that used SSIM and HSSIM that base on joint histogram. The new measure that called SITS outperforms existing measures in detecting image similarity at low PSNR under Gaussian noise. The new measure SITS can detect similarity with confidence at least 20% increase than existing measures that used joint histogram alone, especially at low PSNR.

***Keywords:*** Information Theoretic (IT), Joint Histogram, Structural Similarity (SSSIM), Image Processing.

**INTRODUCTION**

Similarity measure can be defined as the distance between two different points. Today image similarity measures become an essential point in real world application. It can be used for various image processing applications such as dynamic monitor and adjust image quality, enhancement, compression, restoration, etc. [1].

Earlier and simplest method to measure the image similarity between two input images or image patches is accomplish by using the Mean-Squared Error (MSE). It is famous because of it’s mathematical simplicity and it easy to optimize [[2].](#Liu_Lin_2012)

Since human visual system perception is highly adapted for extracting structural information from a scene. In 2004, Wang and Bovik propose a fundamental alternative similarity measure that base on image structure, it used a combination of three quantities: luminance, contrast and correlation [1].

Image similarity measure methods can be classified into two main directions: statistical base and information theoretic IT base [[3](#Asmahan_2014_1)].

**STATISTICAL BASE**

In 2004, Wang and Bovik proposed a measure called (SSIM) that used statistical parameter such as mean, variance and standard deviation as following Equation 1 [1], [4]:

$ρ\left(x,y\right)=$ $\frac{(2μ\_{x}μ\_{y}+ C\_{1})(2σ\_{xy}+ C\_{2})}{(μ\_{x}^{2}+ μ\_{y}^{2}+ C\_{1})(σ\_{x}^{2}+ σ\_{y}^{2}+ C\_{2})} (1)$

where, $ρ\left(x,y\right)$ is SSIM between original (reference) image $x$ and a noisy version $y$,$ μ\_{x}$ and $μ\_{y}$ are the statistical mean of $x$ and $y$ respectively, $σ\_{x}^{2} $and $σ\_{y}^{2}$ are the variance of $x$ and $y$ respectivel. $σ\_{xy}$ the covariance of x and y finally 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).

In 2011, L. Zhang, L. Zhang,, X. Mou and D. Zhang proposed feature similarity (FSIM) index for full reference image quality assessment. The paper base on the idea of human visual system understand image mainly according to its low-level features [5]

**INFORMATION-THEORETIC BASE**

Information-theoretic base also employed to compute image similarity and since images is 2D signals mutual information can be used to find similarity.

In IT the similarity between $x$ and $y$ (where $x$ and $y$ are images) is corresponding differences between them. The more differences they have, the less similar they are [6].

In 2013, D. Mistry, A. Banerjee and A. Tatu propose a new measure that base on joint entropy (joint histogram) [7].

In 2014, A. F. Hassan, D. Cai-lin and Z. M. Hussain proposed a new measure called HSSIM that base on joint histogram. HSSIM outperforms statistical similarity SSIM it is have the ability to detect the similarity under significant noise (low PSNR). It gave an average difference about 20dB.HSSIM Proposed as in Equation 2:

$$E\left(x,y\right)=\sqrt{\frac{\sum\_{i}^{}\sum\_{j}^{}\left[(H\_{ij}-H\_{ji})\frac{1}{h\_{i}-C}\right]}{2L^{2}} } (2)$$

where, $H$ is the joint histogram of $x$ and $y$, $h\_{i}$ is a reference image histogram and $C$ is very small positive constant to avoid division by zero.

**METHODOLOGY**

Image similarity measure methods can based on different bases, but in general it’s classify into two main directions: statistical base and information-theoretic base. The proposed measure is combination of both bases.

**STATISTICAL MEASURES**

Gray scale image is a two dimensional array that is highly structured and by compute it’s statistical measures such as mean, variance and standard deviation valuable information can be obtained. That kind of information can be used to compute image similarity [1].

SSIM that proposed by Wang and Bovik in 2004 is a standard image similarity measure that base on these statistical measures.

**JOINT HISTOGRAM**

The joint histogram $H$ of a pair of image can be define as a function of two variables$ H(x,y)$, $x$ the first image gray level intensity and $y$ the second image gray level intensity. The value of $H$ at coordinate $(x,y)$ can be defined as the number of corresponding pairs having gray-level $x$ and $y$ in the first and second images respectively. Joint histogram is a multidimensional histogram that base on a set of local pixel features [[8](#Pass_G)]. Many researchers are employing joint histogram such as [3] and [8] - [10].

A. F. Hassan, D. Cai-lin and Z. M. Hussain proposed HSSIM that base on joint histogram as Equation 3,4 and 5 [3]:

$H\_{ij}\left(x,x\right)=H\_{ji}\left(x,x\right) (3)$

$H\_{i+1, j+1}=H\_{i+1, j+1}+1 (4)$

where, $i=x\left(m,n\right) $and$ j=y\left(m,n\right)$, for $m$ columns and $n$ rows.

$$E\left(x,y\right)=\sqrt{\frac{\sum\_{i}^{}\sum\_{j}^{}\left[(H\_{ij}-H\_{ji})\frac{1}{h\_{i}-C}\right]}{2L^{2}} } (5)$$

where, $h\_{i}$ is a reference image histogram and $C$ is very small positive constant to avoid division by zero. Note that:

$$E(x,y)\geq 0$$

the above value can be normalize by using respect maximal error estimated value $E\_{\infty }\left(x,y\right)$ in significant noise (very low PSNR) as in Equation 6 and 7:

$$e\left(x,y\right)=\frac{E\left(x,y\right)}{E\_{\infty }\left(x,y\right)} \left(6\right)$$

$λ\left(x,y\right)=1- e\left(x,y\right) (7)$

**THE PROPOSED IT MEASURE**

SSIM is designed base on statistical approach that failed under significant noise (low PSNR). The new measure is used combination of SSIM and HSSIM as a new image dependent measure as in Equation 8:

$τ\left(x,y\right)=\sqrt{ρ\left(x,y\right)C+λ\left(x,y\right)\left(1-C\right)} (8)$

note that:

$$τ(x,y)\geq 0$$

The above value can be normalize by using respect minimal error estimated value $τ\_{-\infty }\left(x,y\right)$ in low noise (very high PSNR) as in Equation 9:

$T\left(x,y\right)=\frac{τ(x,y)}{τ\_{-\infty }\left(x,y\right)} (9)$

The process of normalization will ensure that:

$$0\leq T\left(x,y\right)\leq 1$$

**THE TEST ENVIRONMENT**

The proposed measure has been tested under Gaussian noise and simulate with MATLAB. Gaussian noise is the most popular noise source in image processing system. The tested images are distinct in types such as human face, geometric shape and landscape with different sizes and formats (such as PNG, JPG, TIF, etc.).

**RESULTS AND DISCUSSION**

The new measure simulated using MATLAB, Note that $0\leq ρ\left(x,y\right)\leq 1$ and $0\leq λ\left(x,y\right)\leq 1$ so is$ 0\leq τ(x,y)\leq 1$. We have 1 for completely similar images and 0 for completely different images. The proposed measure is compared with SSIM and HSSIM [1] and [3] as in Figures 1 to 3.

Since Gaussian noise is the most popular image noise comparison is performed under result it and show that proposed measure gave great throughput. It notice that the proposed measure is gave a better result when working on human face images as shown in Table 1:

Table 1: compere the proposed measure with SSIM and HSSIM (when PSNR=-20)

|  |  |
| --- | --- |
| **Images** | **Similarity measures** |
| ***SSIM*** | ***HSSIM*** | ***Proposed measure*** |
| Image1 (landscape) | 0.0056 | 0.0389 | 0.1984 |
| Image2 (human face) | 0.0044 | 0.0395 | 0.1987 |
| Image3 (geometric shape) | 0.0039 | 0.0298 | 0.1730 |



(a)



(b)



(c)

Fig. 1. Performance comparison of SSIM, HSSIM and SITS using similar images (landscape image) under Gaussian noise. (a): tested images: Original Image, Noisy Image, PSNR (dB) = 28, (b): Performance comparison of SSIM, HSSIM and SITS, (c): Histograms Original and noisy, PSNR (dB) =28.



(a)



(b)



(c)

Fig. 2. Performance comparison of SSIM, HSSIM and SITS using similar images (Human face) under Gaussian noise. (a): tested images: Original Image, Noisy Image, PSNR (dB) = 28, (b): Performance comparison of SSIM, HSSIM and SITS, (c): Histograms Original and noisy, PSNR (dB) =28.



(a)



(b)



(c)

Fig. 3. Performance comparison of SSIM, HSSIM and SITS using similar images (Geometric shape) under Gaussian noise. (a): tested images: Original Image, Noisy Image, PSNR (dB) = 28, (b): Performance comparison of SSIM, HSSIM and SITS, (c): Histograms Original and noisy, PSNR (dB) =28.

**CONCLUSION**

A new image similarity measure is proposed and tested versus structural similarity (SSIM) and information theoretic (HSSIM) measures under Gaussian noise. It’s have been notice that the new measure SITS give a better performance under significant noise (low PSNR) and can increase the confidence with at least 20% than existing measure and at least 80% than SSIM.

**REFERENCES**

1. Z. Wang, A.C. Bovik, H.R. Sheikh and E.P. Simoncelli, “Image quality assessment: From error visibility to structural similarity”, IEEE Trans. Image Process., 2004.
2. Liu and W. Lin, “Image Quality Assessment Based on Gradient Similarity”, IEEE transactions on image processing, 2012
	1. F. Hassan, D. Cai-lin and Z. M. Hussain, “An Information-Theoretic Image Quality Measure: Comparison with Statistical Similarity,” Journal of Computer Science, 2014.
3. Z. Wang and A.C. Bovik, *Modren image quality assessment*, Morgan & Claypool Publishers, 2006.
4. L. Zhang, L. Zhang,, X. Mou and D. Zhang, “FSIM: A Feature Similarity Index for Image Quality Assessment”, IEEE Transactions On Image Processing, 2011.
5. D. Lin, “An Information-Theoretic Definition of Similarity”, ICML '98 Proceedings of the Fifteenth International Conference on Machine Learning, 1998.
6. D. Mistry, A. Banerjee and A. Tatu, “Image Similarity based on Joint Entropy (Joint Histogram)”, International Conference on Advances in Engineering and Technology (ICAET – 2013), 2013.
7. G. Pass and R. Zabih, “Comparing images using joint histograms”, Multimed. System-Springer, 1999.
8. X. Lu, S. Zhang, H. Su and Y. Chen, “Mutual information-based multimodal image registration using a novel joint histogram estimation”, Computerized Medical Imaging and Graphics, 2009.
9. 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.
	1. N. Hashim and Z. M. Hussain, “Novel image dependent quality assessment measures,” Journal of Computer Science, 2014.
	2. A. Goshtasby, “Image Registration: Principles, Tools and Methods”, Springer-Verlag London Limited, 2012.
	3. George and S. J. Livingston, “A survey on full reference image quality assessment algorithms”, International Journal of Research in Engineering and Technology, 2013.
10. R. Soundararajan and A. C. Bovik, “Survey of information theory in visual quality assessment”, Springer-Verlag London, 2012.
11. Z. Wang, A. C. Bovik and L. Lu, “Why is image quality assessment so difficult”, Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 2002.
	1. F. Hassan, Z. M. Hussain and D. Cai-lin, “An Information-Theoretic Measure for Face Recognition: Comparison with Structural Similarity”, International Journal of Advanced Research in Artificial Intelligence, 2014.