An Information-Theoretic Measure for Face Recognition: Comparison with Structural Similarity

Asmhan Flieh Hassan

PhD Candidate, Mathematics & Statistics
HuaZhong Normal University
152 Luoyu Road, Wuhan, China
Dept. of Maths, Univ. of Kufa, Najaf, Iraq

Zahir M. Hussain

Computer Science & Maths,
University of Kufa
P.O.Box 21, Kufa, Najaf, Iraq
Professor (Adjunct), Engineering,
Edith Cowan Univ., Australia

Dong Cai-lin

School of Mathematics & Statistics
HuaZhong Normal University
152 Luoyu Road, Wuhan,
China 430079

Abstract—**Automatic recognition of people faces is a challenging problem that has received significant attention from signal processing researchers in recent years. This is due to its several applications in different fields, including security and forensic analysis. Despite this attention, face recognition is still one among the most challenging problems. Up to this moment, there is no technique that provides a reliable solution to all situations. In this paper a novel technique for face recognition is presented. This technique, which is called ISSIM, is derived from our recently published information - theoretic similarity measure HSSIM, which was based on joint histogram. Face recognition with ISSIM is still based on joint histogram of a test image and a database images. Performance evaluation was performed on MATLAB using part of the well-known AT&T image database that consists of 49 face images, from which seven subjects are chosen, and for each subject seven views (poses) are chosen with different facial expressions. The goal of this paper is to present a simplified approach for face recognition that may work in real-time environments. Performance of our information - theoretic face recognition method (ISSIM) has been demonstrated experimentally and is shown to outperform the well-known, statistical-based method (SSIM).**

Keywords—Information Theoretic Similarity; Joint Histogram; Structural Similarity (SSIM); face recognition; Image Processing

I. INTRODUCTION

Face recognition is one of the fastest-growing topics within the image processing and pattern recognition. Applications of face recognition are manifold, including access control, security and video surveillance, credit card user identification, forensic analysis, entertainment, and automatic video indexing. These applications have made this topic a very popular research area in the last three decades; see [1, 2, 3] for surveys. In recent years, several approaches to face recognition have been developed. References [4, 5] give an overview of face recognition techniques.

Methods used for face recognition have improved significantly, thanks to the efforts of hundreds of researchers and funding bodies, causing its impact to spread rapidly to a variety of applications. In parallel with thousands of published papers on the topic, many useful datasets are created and used successfully for performance evaluation of various approaches.

Despite its huge literature, automatic face recognition is still a difficult problem. This difficulty is mainly due to artifacts that obscure the image features, like variations in illumination, facial expression, and head pose. Techniques that may offer efficient feature extraction with a high discrimination power and low computational complexity are crucial [6]. According to the approach of considering spectral analysis, face-recognition methods maybe divided into two main categories: visible (VIS) and infrared (IR).

Image similarity measurement could be a basic issue in each computer and human vision system, including many real-world applications. In image recognition applications, one usually has to realize the similarity between two images, i.e., a test image and a training database image. Image recognition has become a subject of great interest to researchers over the past decades because of its potential applications in several fields like Optical Character Recognition (OCR), identity authentication, human-computer interfacing, and surveillance. A variety of methods for image recognition, especially for face image recognition have been proposed [3]. In References [7-13], methods for image recognition have been classified as holistic methods, feature-based methods, and hybrid methods.

Several matching algorithms of face and object recognition systems have been designed based on image similarity measurement like SSIM [14].

This work presents a novel approach for face recognition based on information – theoretic approach rather than the statistical approach of SSIM. The new measure, which is called in this paper as ISSIM, is derived from our information-theoretic similarity measure (HSSIM) proposed in [15]. HSSIM has been modified to meet the requirements of feature extraction for the purpose of face recognition.

II. BACKGROUND

Automatic face recognition systems attempt to recognize the identity of a given face image in comparison with images that they have in their memory. In most face recognition works, this memory of a face recognizer is represented by a training set saved as a database. In this work, our training set is the AT&T well known face database [16]. Thus, the task of the face recognizer is finding most similar feature vector

among the training set to the feature vector for a given test image.

Simulation results showed that the proposed approach (ISSIM) presents higher performance than the standard SSIM in face recognition.

In [17], modified versions of SSIM are proposed in an effort to be utilized in recognition. In [18], instead of the usual approach for applying statistics or structural methods only, the Authors proposed a methodology that integrates higher-order illustration patterns extracted by Zernike moments with a modified version of SSIM (M-SSIM). Individual measurements and metrics resulted from mixed SSIM and Zernike-based approaches provide a powerful recognition tool with great results.

There are several similarity measures that are proposed and used for varied purposes, see [19-26]. Several methods was described and applied using entropy types to handle the face recognition and edge detection problems [27-31].

The rest of this paper is organized as follows: Section III presents the mathematical foundations of statistical and information theoretic measures. Section IV describes the design of novel information - theoretic measure ISSIM to be used in face recognition. Section V shows experimental results and discusses possible modifications and improvements to the system. Section VI presents concluding remarks.

III. SIMILARITY MEASURES

A similarity measure (between two sets of data points) is defined as the distance (based on a specific norm) between various data points.

The performance of any algorithm that detects similarity is heavily dependent upon choosing a good distance function over an input data set. While similarity is an amount that reflects the strength of relationship between two database items, dissimilarity deals with the measurement of divergence between two data items.

In this work an information - theoretic measure is proposed and its performance is analyzed in comparison with a well-known statistical measure.

A. SSIM Overview: A Structural Similarity Measure

Automatic similarity finding is an important approach in many image processing systems like those used for compression, enhancement, identity check, etc. An important achievement in image similarity was the proposal of a statistical image quality measure in 2004 [14]. The measure, well-known as SSIM, has been promising in many applications. It has been widely used for image quality assessment and many algorithms of image processing systems. The technique used in SSIM is based on using statistical measurements like mean and standard deviation to find a definition for a distance function that can measure the structural similarity between a test image and a training image. The measure has been put into the form:

$$\rho(x,y) = \frac{(2\mu_x\mu_{y+}C_1)(2\sigma_{xy}+C_2)}{(\mu_x^2+\mu_y^2+C_1)(\sigma_x^2+\sigma_y^2+C_2)} \qquad (1)$$

where $\rho(x,y)$ is the SSIM measure of statistical distance (or similarity) between the image x (the test image or reference image) and (normally a corrupted version of x) image y. The statistical parameters μ_x, μ_y, σ_x^2, and σ_y^2 are the means and variances of pixel (illumination level) values in the two images x and y, while σ_{xy} represents the statistical covariance between images x and y. The constants C_1 and C_2 are given as follows: $C_1 = (K_1 L)^2$ and $C_2 = (K_2 L)^2$, with K_1 and K_2 are small constants, and $L = 255$ (the maximum pixel value). The Authors in [14] confirmed that these parameters (constants) have little effect on SSIM performance.

In [17], it has been shown that this approach gives a good level of similarity under noise free conditions, while its performance falls when noise increases. In addition, it may give a non-negligible similarity between two unrelated images. This is due to the fact that SSIM is totally dependent on the statistical features of the two images, which may have some hidden correlations. In [17] and [30], SSIM has been combined with edge detection filters (such as Canny's) to produce excellent results especially when the images are different from each other.

B. An Information-Theoretic Technique for Image Similarity

The use of information - theoretic analysis in image processing is possible if one imagines that images are 2D random variables.

The joint – histogram between two images is defined as the joint occurrence of intensity levels (pixel values) in the two images. Note that only considering a gray image is considered here, as grey version of an image contains most of the similarity information. For two images under similarity test A and B, assume that the two images have the same size and that the pixel intensity values are denoted by i and j, respectively. Note that both i and j range from 0 to $L = 255$. Now the joint histogram, denoted by $H_{ij}(A, B)$, represents the joint probability of each pixel (defined by its 2D location, as the two images have the same size) to take on the value i on the first image A and the value j on the second image B. Now every entry in the joint histogram will represent the number of times the intensity level i in any of the two images corresponds to the intensity level j in the other image.

In [15], a similarity error estimate between symmetrically – located entries of the joint histogram (around y=x relation) has been designed as follows:

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

where H_{ij} and H_{ji} are symmetrically- located entries of the joint histogram between two images, and h_i is the normal image histogram of the reference image histogram. The constant c is a very small positive constant, inserted principally to avoid division by zero, and $L = 255$ is the maximum pixel value.

It is easy to show the above relation is non-negative, that is:

$$E(x,y) \geq 0 \qquad (3)$$

To be used as a similarity measure and compared with SSIM, a normalization process is necessary to ensure that the range of the measure is kept inside [0,1]. In [15], maximal error estimate $E_\infty(x,y)$ was used for normalization as follows:

$$e_o(x,y) = \frac{E(x,y)}{E_\infty(x,y)} \qquad (4)$$

The normalization process will ensure that:

$$0 \leq e_o(x,y) \leq 1 \qquad (5)$$

Based on the above error estimate an information - theoretic similarity measure (which was named as HSSIM) was proposed as follows:

$$\lambda_o(x,y) = 1 - e_o(x,y) \qquad (6)$$

where:

$$0 \leq \lambda_o(x,y) \leq 1 \qquad (7)$$

If one refers to the well-known SSIM by $\rho(x,y)$, then a similar inequality between 0 and 1 is obtained as follows:

$$0 \leq \rho(x,y) \leq 1 \qquad (8)$$

Note that $\rho(x,y) = 1$ for totally similar (identical) images, while $\rho(x,y) = 0$ for totally dissimilar images. In practical applications for face and image similarity, SSIM and HSSIM adopt values above zero and less than one.

In this work the intention is to design a measure for face recognition based on HSSIM. However, a change in HSSIM structure is necessary as follows.

IV. A FACE RECOGNITION APPROACH BASED ON JOINT HISTOGRAM

In [18], SSIM was utilized for face recognition between a reference face image x and images $\{y_{pr}\}$ for a number P of people, each person (whose database number is p such that $p \leq P$) has R face images with different poses (each pose is indexed by the variable r); hence a total of $N = R \cdot P$ face images in the database. One of the poses is chosen with a number $r = a$ that corresponds to the person p to serve as a reference image. Then two tests are performed: the first keeps the ath pose in the database then starts recognition. This is an easy task, but helps in testing and comparing the performances of the proposed measures. The second test removes the ath pose and starts recognition process, which is a hard task since different poses may be considered as different images.

This section describes a face recognition algorithm using the information - theoretic error estimate based on joint in (2) above.

Let us define a new error measure as follows. After running the test, define the maximum information – theoretic error for the whole test as E_m as follows:

$$E_m = \max_{p,r}\{E(x, y_{pr})\} \qquad (9)$$

Now define the information – theoretic error measure $e(x,y)$ as follows:

$$e(x,y) = \frac{E(x,y)}{E_m} \qquad (10)$$

This *database-dependent* (or, more accurately, *test-dependent*) normalization process will ensure that:

$$0 \leq e(x,y) \leq 1 \qquad (11)$$

After normalization, an information - theoretic similarity measure for face recognition (which will be called henceforth as ISSIM) can be proposed as follows:

$$\lambda(x,y) = 1 - e(x,y) \qquad (12)$$

where it is evident that:

$$0 \leq \lambda(x,y) \leq 1 \qquad (13)$$

Hence, the new measure is well-defined.

For a comparison between a reference face x and a person p in the database, ISSIM is defined as follows:

$$\lambda(x,p) = \max_r\{\lambda(x, y_{p,r})\} \qquad (14)$$

Here, the value of $\lambda(x,p)$ will indicate how much confidence one can put in the recognition process. Of course, a value of $\lambda(x,p) = 1$ is obtained only in the *ideal case* where the reference face pose exists in the database. Recognition is based on the following search:

$$p_i = \arg[\max_p\{\lambda(x,p)\}] \qquad (15)$$

For SSIM-based statistical face recognition, one can define a similar process to find the best match as follows:

$$\rho(x,p) = \max_r\{\rho(x, y_{p,r})\} \qquad (16)$$

with the recognition decision based on the following search:

$$p_s = \arg[\max_p\{\rho(x,p)\}] \qquad (17)$$

Recognition Confidence

In most cases SSIM and ISSIM agree such that:

$p_i = p_s$.

However, the confidence in recognition varies significantly. By "confidence in recognition" it is meant that there is a "good" distance between the peaks defined in (15) and (17) and the next-in-height peaks who can confuse the decision when two peaks (using the same measure) are nearby in magnitude.

Now define the confidence in recognizing the test image x using a measure μ. First, find the second peak (maximum) in the above measure $\{\mu(x,p)|p = 1, \cdots, P\}$. Then, find the difference between the maximum and the second maximum for each measure, which is called here the MM $-$ difference, denoted by $\partial(\mu,x)$, which can be $\partial(\rho,x)$ or $\partial(\lambda,x)$. The resulting quantity is the MM-difference for that measure.

Confidence of recognition of reference image x using a similarity measure μ is defined as:

$$\partial(\mu,x) = \max_p\{\lambda(x,p)\} - \max2_p\{\lambda(x,p)\} \qquad (18)$$

where $\max2_p\{\lambda(x,p)\}$ is the second maximum of the curve of similarity between the reference image x and other persons in database (note that all poses for each person are exhausted in $\mu(x,p)$ as per (14)). ISSIM outperforms SSIM as one has for most cases:

$$\partial(\lambda, x) > \partial(\mu, x) \quad \text{for most } x \qquad (19)$$

If ∂ is extended to other peaks then the average of ISSIM distance is always larger:

$$\text{mean } \partial(\lambda, x) > \text{mean } \partial(\mu, x) \quad \forall \, x \qquad (20)$$

Hence, ISSIM gives a clearer decision in most cases. In easier cases, this difference can reach up to 0.5, equivalent to half the value of the total measure, which is 1. Of course, one can get even clearer decisions when the reference image is included in the database, although this case is not realistic.

V. EXPERIMENTAL RESULTS AND PERFORMANCE

A. Image Database

A well-organized face image database was created by AT&T [16] and has been widely used for testing face recognition systems. This database is divided into two subgroups, for separate training and testing purposes. Throughout training, 49 images were used, containing seven subjects and every subject having 7 images with different facial expressions. A sample is shown in Figure (1) below. Note that in our work, the dimensions (size) of every face image are modified to be 92×92 pixels.
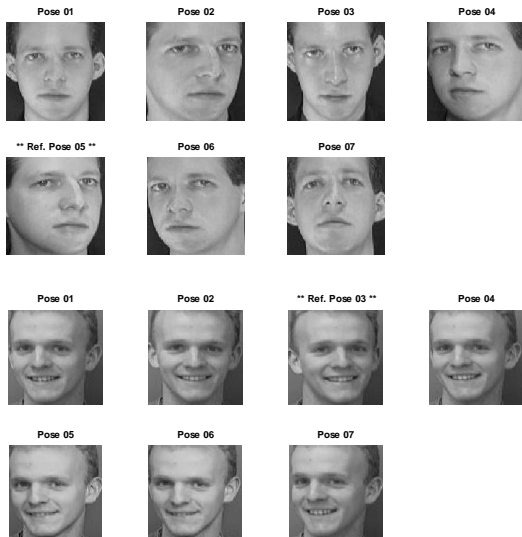

Fig. 1. Examples of the training and testing images taken from the well-known AT&T face image database

B. Testing

The face recognition system presented in this paper was developed, trained, and tested using MATLAB. The performance of the proposed information – theoretic algorithm ISSIM is tested and then compared with the well-known structural similarity measure SSIM. The results illustrate the efficacy of ISSIM for face recognition. The algorithm has been tested using two cases as follows:

Case 1- Test Image Included in the Database

If the test image (the reference pose) represents an image that belongs to the training dataset, then the program utilizes all poses related to the person under test.

In this testing case the recognition is high for both SSIM and ISSIM. Although less realistic than the next case, it can be used to compare the power of the proposed ISSIM as compared to SSIM in performance for face recognition. ISSIM outperforms SSIM by far.

Case 2- Test Image Excluded from the Database

If the test image is not included in the data set, the recognition is still very good and ISSIM still performs better than SSIM in recognizing the face despite the difference in poses.

To compare performance, the distance between he first and second maxima of the similarity curve are utilized for each recognition method. The more this distance the more confident is the decision of recognizing the person.

The Figures below show recognition results and performance analysis of the proposed ISSIM.

C. Discussion

Figures (2) and (3) present poses for two persons from the database, used for testing. The reference pose, considered as test image, is indicated.

A pose with no significant angle is chosen for the purpose of testing and comparison, although eye-glasses are considered in the second test (see Figure (3)), which increases the difficulty of recognition.

Figures (4) - (7) show results of similarity test between the reference image and other people in the database, where, for each person, the maximum similarity among all of his / her poses is considered.

Note that, for almost all cases, the proposed ISSIM keeps larger difference between the test image and other persons.

Figures (4) - (5) keeps the test pose as part of the database, giving an ideal maximum of 1 for both approaches.

Figures (6)-(7) are more realistic as the test excludes the image under test from the database. In this case the maximum similarity is less than 1 for both approaches, but ISSIM gives much better results.

Figures (8)-(9) show the performance measure for both approaches, where the proposed ISSIM keeps larger difference than SSIM. Hence, according to (18) and (19), better confidence in the recognition process is obtained.
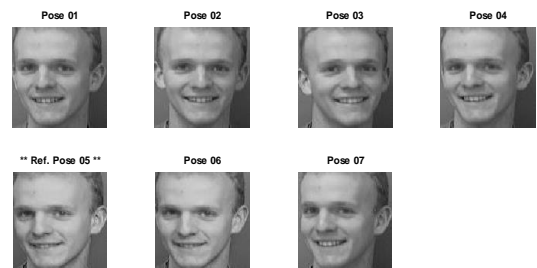

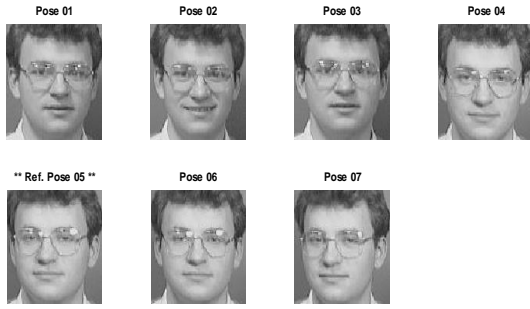Fig. 2. Poses for Person no.5 in database. Reference Pose is indicated

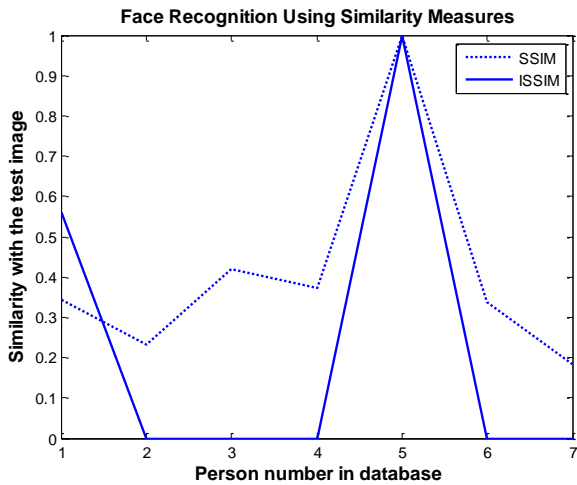Fig. 3. Poses for Person no.6 in database. Reference pose is indicated

Fig. 4. Face recognition using the proposed ISSIM and the well-known SSIM with a test image as a pose belonging to person no. 5. Reference pose (as indicated in Figure (2)) is included in the database during the test
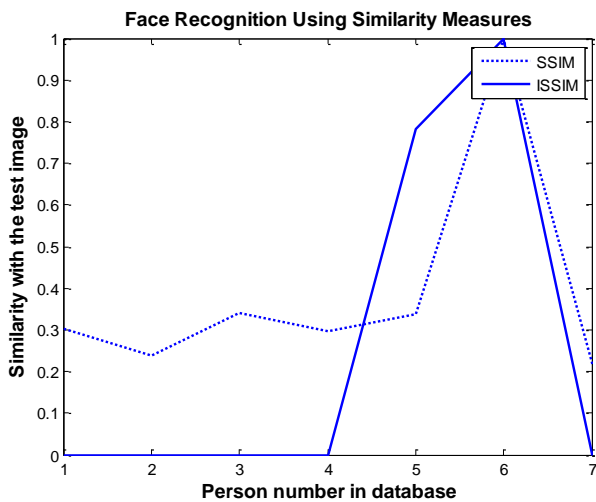

Fig. 5. Face recognition using the proposed ISSIM and SSIM with a test image as a pose belonging to person no. 6. Reference pose (as indicated in Figure (3)) is included in the database during the test
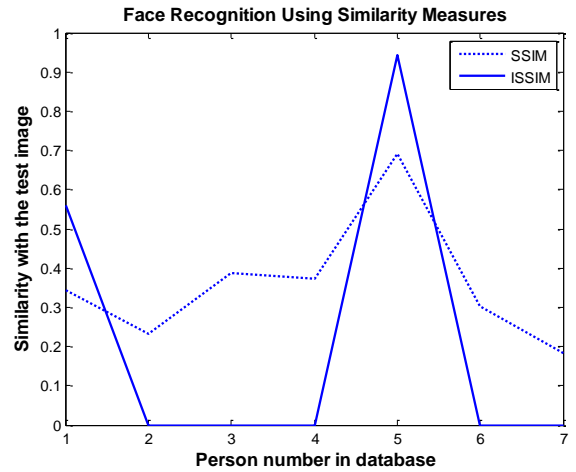

Fig. 6. Face recognition using the proposed ISSIM and SSIM with a test image as a pose belonging to person no.5. Reference pose (as indicated in Figure (2)) is excluded from the database in the test
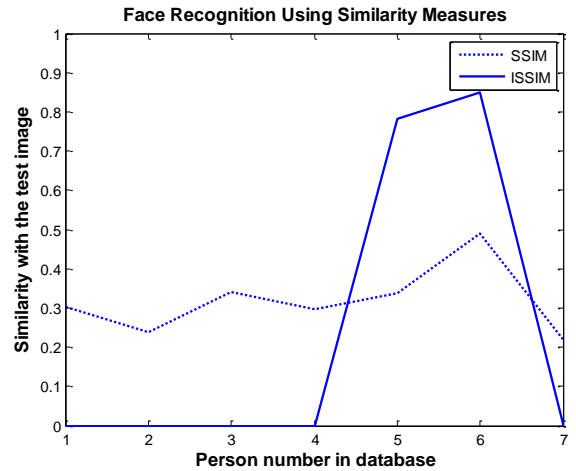

Fig. 7. Face recognition using the proposed ISSIM and SSIM with a test image as a pose belonging to person no. 6. Reference pose (as indicated in Figure (3)) is excluded from the database during the test
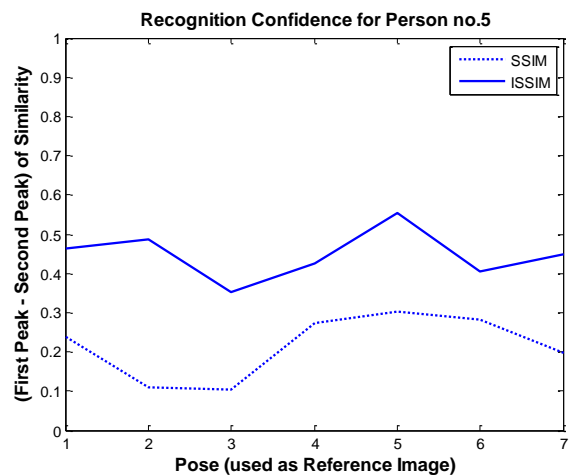

Fig. 8. Recognition confidence using the proposed ISSIM and SSIM with a test image as a pose belonging to person no. 5. Reference pose x (as in Figure

(2)) is excluded from the database during the test. It is clear that ISSIM $\lambda(x,p)$ gives more confidence than SSIM $\mu(x,p)$ as $\partial(\lambda,x) > \partial(\mu,x)$
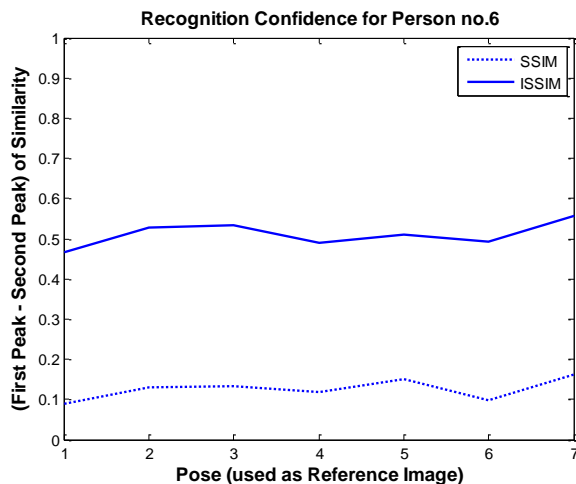

Fig. 9. Recognition confidence using the proposed ISSIM and SSIM with a test image as a pose belonging to person no. 6. Reference pose x (as in Figure (3)) is excluded from the database during the test. It is clear that ISSIM $\lambda(x,p)$ gives more confidence than SSIM $\mu(x,p)$ as $\partial(\lambda,x) > \partial(\mu,x)$.

VI. CONCLUSIONS

This paper presents a novel face recognition technique (called ISSIM) that uses information - theoretic similarity measure based on joint histogram. This measure utilizes the fact that similarity between two images would result in more symmetry along y=x curve in the joint histogram. Divergence from this symmetry is considered as a measure of dissimilarity.

The system was simulated numerically using on MATLAB using part of AT&T image database that consists of 49 face images, containing seven subjects with every subject having seven poses with different facial expressions. Results showed that ISSIM outperforms the well-known SSIM when used in face recognition.

The main performance measure that was considered here (and used for comparison purposes) is the distance between the maximum similarity found by the specific approach and the second maximum, which can be more confusing in the recognition task if it takes smaller values. As no more complications are used in the proposed system, it is evident that this system is well-suited for low-cost, real-time hardware or software implementation.

In this work, global face analysis is applied, where the whole image is treated at once. While good results are obtained under a standard database, difficulties may arise in practice. Local analysis for face images proved to play a significant role in improving face recognition.

The Authors intend to pursue this point in future works and extend their previous studies on local analysis to improve the performance of the above defined measure.

REFERENCES

[1] W. A. Barrett, "A survey of face recognition algorithms and testing results," *Systems and Computers*, 1998.

[2] T. Fromherz, P. Stucki, M. Bichsel, *A Survey of Face Recognition*, MML Technical Report, 1997.

[3] W. Zhao, R. Chellappa, P. J. Phillips, A. Rosenfeld, "Face recognition: a literature survey," *ACM Computing Surveys*, 35(4), 2003.

[4] R. Chellappa, C. L. Wilson, and S. A. Sirohey, "Human and machine recognition of faces: a survey," *Proceedings of IEEE*, 83(5), 1995.

[5] S. Lawrence, C. L. Giles, A. C. Tsoi, and A. D. Back, "Face recognition: a convolutional neural network approach," *IEEE Transactions on Neural Networks*, 8(1), 1997.

[6] A. K. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition," *IEEE Trans. Circuits Systems for Video Technology*, 14(1), 2004.

[7] M.Turk, A.Pentland, "Eigenfaces for recognition," Journal of Cognitive Neuroscience, 3(1), 1991.

[8] W. Zhao, R. Chellappa, A. Krishnaswamy, "discriminant analysis of principal components for face recognition," *Proceedings of the 3rd IEEE International Conference on face and Gesture Recognition*, FG'98, Nara, Japan, 1998.

[9] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: recognition using class specific linear projection," IEEE Transactions on Pattern Analysis & Machine Intelligence (PAMI), Vol.19, 1997.

[10] T. Kanade, *Computer Recognition of Human Faces*, Birkhauser, Basel, Switzerland, and Stuttgart, Germany, 1977.

[11] A. V. Nefian, M. H. Hayes, "Hidden Markov models for face recognition," *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, Vol.5, 1998.

[12] L. Wiskott, J. M. Fellous, N. Kruger, C. Malsburg, "Face recognition by elastic bunch graph matching," *IEEE Transactions on Pattern Analysis & Machine Intelligence (PAMI)*, Vol.19, 1997.

[13] J. Huang, B. Heisele, V. Blanz, "Component-based face recognition with 3d morphable models," *Proceedings of the 4th International Conference on Audio and Video Based Biometric Person Authentication (AVBPA)*, Guildford, UK, 2003.

[14] 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.*, 13 (4), 2004.

[15] A. F. Hassen, Z. M. Hussain, D. Cai-Lin, "An information - theoretic image quality measure: comparison with statistical similarity," *Journal of Computer Science*, in press, 2014.

[16] AT&T Laboratories, *The Database of Faces,* Cambridge [online], ©2002 [accessed 10/09/2014]. Available from: http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html

[17] A. N. Hashim and Z. M. Hussain, "Novel image dependent quality assessment measures," *Journal of Computer Science*, 10 (9): 1548-1560, 2014.

[18] A. N. Hashim and Z. M. Hussain, "Local and semi-global feature-correlative techniques for face recognition," *International Journal of Advanced Computer Science and Applications*, 5(8), 2014.

[19] F. G. Ashby, "towards a unified theory of similarity and recognition," *Psychological Review*, 95(1), 1988.

[20] R. Jain, R., S. N. J. Murthy, P. L-J., Chen, S. Chatterjee, "Similarity measures for image databases," *IEEE Int. Conference on Fuzzy Systems*, V.3, 1995.

[21] S. Santini, R. Jain, "Similarity is a geometer," *Multimedia Tools and Applications*, 5(3), 1997.

[22] S. Santini, R. Jain, "Similarity measures," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(9), 1999.

[23] J. Williams, N. Steele, N. Deference, "Distance and similarity ASA basis for fuzzy decision support based on prototypical decision classes," *Fuzzy Sets and Systems,* V.131, 2002.

[24] A. K. Patidar, J. Agrawal, N. Mishra, "Analysis of different similarity measure functions and their impacts on shared nearest neighbor clustering approach," *International Journal of Computer Applications* , 40(16), 2012.

[25] A. Huang, "Similarity measures for text document clustering," *NZCSRSC 2008*, Christchurch, New Zealand, April 2008.

[26] K. Taghva and R. Veni, "Effects of Similarity Metrics on Document Clustering," *Seventh International Conference on Information Technology*, 2010.

[27] Y. D. Zhang and L. N. Wu, "Pattern recognition via PCNN and Tsallis entropy," *Sensors Journal*, 8, 2008.

[28] B. H. Shekar, M. S. Kumari, L. M. Mestetskiy, N. F. Dyshkant, "Face recognition using kernel entropy component analysis," *Neurocomputing Journal*, 74(6), 2011.

[29] M. A. El-Sayed, S. Abdel-Khalek, E. Abdel-Aziz, "Study of efficient technique based on 2D Tsallis entropy for image thresholding," *International Journal on Computer Science And Engineering (IJCSE),* 3(9), 2011.

[30] M. A. El-Sayed, "A new algorithm based entropic threshold for edge detection in images," *International Journal of Computer Science*, 8(1), 2011.

[31] M. A. El-Sayed, T. Abd-El Hafeez, "New edge detection technique based on the Shannon entropy in Gray level images," *International Journal on Computer Science and Engineering (IJCSE),* 3(6), 2011.