New Iris Feature Extraction and Pattern Matching Based on Statistical Measurement

Assistant Prof Dr. Abdulamir Abdullah Kerim[1], Researcher Sarah jasim mohammed[2]

[1]Department of Computer Science,College of Science, University of Technology, Baghdad, Iraq

[2]Department of Computer Science,College of Science, University of Technology, Baghdad, Iraq

Abstract

Iris recognition system is considered one of the best biometric systems used to distinguish individuals. The iris recognition system consists of several stages and each stage must be carried out carefully, especially iris feature extraction and matching stages because it affect the accuracy of system performance and its speedup. In this paper we propose two methods for iris feature extraction from normalized iris image by using statistical measurements, the first propose method used first order statistics and the second propose method used second order statistics. The proposed methods aim to feature extraction accurately from pure iris region and isolate the eyelid and eyelash which are covers some parts of iris region. The proposed methods used weights within pattern matching measures depending on purity iris region. The proposed methods has been tested by using the iris data set (CASIA v4.0- interval), and (CASIA v1.0- interval) iris image database. Experimental results show that the first proposed method has 99.4% accuracy rate with (CASIA v4.0- interval) and 98.5% with (CASIA v1.0- interval), the second proposed method has 86.67% with (CASIA v4.0- interval).

Keywords: Iris Feature Extraction, First Order Statistical Measurement, Second Order Statistical Measurement, CASIA.

1. INTRODUCTION

Iris recognition system has become an active research since the concept of an iris recognition system was first proposed by Flom and Safir in 1987[1][2]. Iris recognition is one of the most reliable noninvasive methods of personal identification and verification owing to the stability of the iris over one's lifetime.The identification based on iris pattern has many advantages, some of these are [3]:

1-Iris is a highly protected, internal organ of the eye.
2-Iris is visible part from a distance.
3-Iris patterns possess a high degree of randomness.
4-Limited genetic penetrance.
5-Iris is stable throughout life and not affect by surgeries.

The iris is a thin circular diaphragm, which lies between the cornea and the lens of the human eye. In the iris disk, there are over 400 distinguishing characteristics that can be quantified and used to identify an individual. Although, approximately 260 of those are possible to be captured for identification, these identifiable characteristics include: contraction furrows, coronas, stripes, striations, pits, collagenous fibers, filaments, crypts (darkened areas on the iris), serpentine vasculature, rings, and freckles. Due to these unique characteristics, the iris has six times more distinct identifiable features than a fingerprint [3]. J. Daugman in 1993 developed and patented the first actual algorithms to perform iris recognition system. The Daugman system based on Integro-differential operators to detect iris inner and outer boundaries and applies a 2D version of Gabor filters in order to encode iris pattern data in the normalized polar coordinates (r, θ). After the Gabor feature extraction, a complex feature matrix is generated from the image. For each complex feature value *h*, two binary bits (h_R; h_I) are used to represent phase information at the pixel location. And the binary phase pairs from the entire image are combined into a binary feature template for pattern comparison and decision making [4]. Daugman applied the Hamming distance (HD) measures (logical exclusive OR operator) to check the dissimilarity between two strings of equal length binary templates with a threshold equal to 0.32 [5]. The Hamming distance calculates the pattern difference with a bit-by-bit comparison [6][7]. For iris templates from the same eye, their statistical independence and the Hamming distance tend to approach zero, while two different iris templates tend to have a Hamming distance of 0.5 [5]. R. Wildes in 1997 presented an Iris recognition system at Sarnoff Laboratory. Wildes used Laplacian of Gaussian filter at multiple scales to create a feature template [8]. Son et al. in 2004 used DWT, principal component analysis (PCA), linear discriminant analysis (LDA), and direct linear discriminant analysis (DLDA) for feature extraction, and experimented these features in combinations for Iris recognition. They concluded that DWT is best for feature extraction and DLDA is better to reduce the dimensionality of feature vector [9]. Chia Te Chu and Ching-Han Chen in 2012 used LPCC and LDA algorithms for iris recognition system. The advantage of feature extraction is for the dimension reduction and representation of original signal. The LPCC coefficients can improve the robust and reliability of feature vector. The basic idea of LDA is finds a linear transformation such that feature clusters are most separable after the transform. LPCC and LDA to extract the feature vector of iris images tested by using CASIA iris database [10]. Sheeba and Veluchamy in 2013 developed a technique to improve performance of iris recognition system. The feature vectors have been extracted by using Local Binary Pattern (LBP). The classification has been performed by using Learning Vector Quantization (LVQ). In matching stage used hamming distance [11].Kshamaraj et al proposed iris recognition method in 2012 using Independent Component Analysis for feature extraction. The iris region was detected using Daugman's method. In matching stage used Euclidian distance between the test iris and training iris [12].

2. METHODOLOGY

The proposed methods for iris feature extractionand pattern matching are show in figure (1):

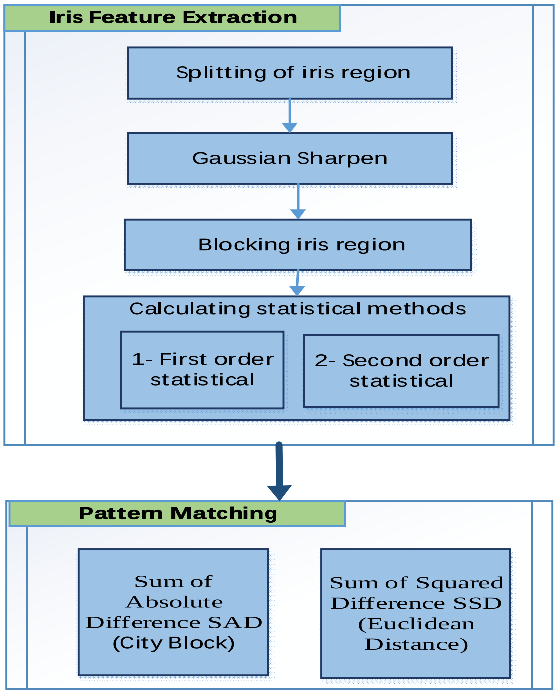
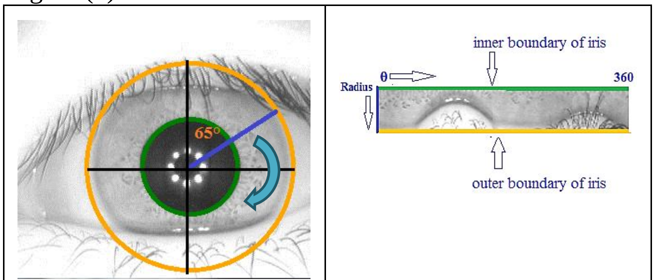


Figure 1: Block Diagram of the proposed Iris Feature Extraction and Pattern Matching.

The proposed methods applied on normalized iris image that transformed to polar coordinate using theta equal to 65° as starting point to rotation transform; that paves the way for splitting the iris images to four regions where thesecond region has eyelid, the fourth region has eyelash, and the first and third region have pure iris as shown in Figure (2).



(a) Iris region in original image

(c)Iris region in Polar Coordinate

Figure 2: Iris normalization

2.1 Iris Feature Extraction

The iris feature extraction stage aims to extract iris code from iris normalization image that can be represented as feature vector, these vectors are used in the process of pattern matching. The proposed algorithm for iris feature extraction consists of four steps (Splitting of iris region, Gaussian Sharpen, Blocking iris region and Calculating statistical methods). The proposed algorithm is described below:

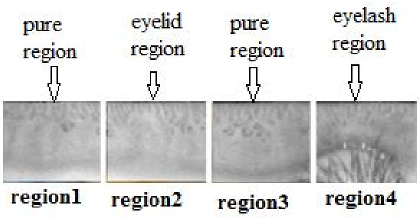
Algorithm 1:- The Proposed Iris Feature Extraction

Input: Normalizediris region image.
Output: Feature vectors of iris image.
Process:

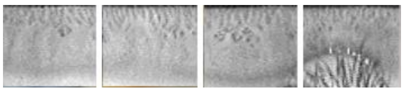
```
For i=1 to Normalized iris image.width do   // read the image
For j=1 to Normalized iris image.height do
 Begin
Step1: Split Normalized iris image into four equal image im(i,j);
Step2:Image(i,j) = call:: Gaussian Sharpen (image(i,j));
Step3:For i=1 to 4 do
            Blocking image(i,j) to 4*6 block   //each block 15*15 pixel
End
Step4:For im=1 to 4 do   // No. of iris regions
Begin
 If im= 1 || 3 then   // two pure iris regions that have weight 45%
For i=1 to 24 do    //all image blocks will compute (pure images)
Case1:
Call ::Function feature vectors (block(i)); // extract features
                        vectors using FOS and stored in DB1.
Case2:
            Create GLCM (block(i));
            Call :: Function feature vectors(GLCM(i)); // extract features
                            vectors using SOS and stored in
                            DB2.
            ElseIf im = 2   // eyelid iris region that has weight 7%
            For i=1 to 18 do   //the last 6 will removed (lower eyelid
            area)
            Call ::Function feature vectors (block(i));
              ElseIf im = 4 // eyelash iris region that has weight 3%
              For i=1 to 6 do   //the last 18 will removed (upper eyelash
                    area)
                   Call ::Function feature vectors (block(i)),
 End
Next i;
Step5:Function feature vectors (block(i))// Statistic measures
Begin
Vector1= Mean (block(i));
Vector2= Standard deviation (block(i));
Vector3= Coefficient of variation (block(i));
Vector4= Entropy (block(i));
End
Step6:End
```

The first step in the proposed iris feature extraction algorithm is region splitting, which includes spliting each normalized iris image into four equal-size regions. Region splitting will divide iris image width that is equal to 360° into four parts while iris image height will remain at the same value in all parts. This step aims to isolate upper eyelash and lower eyelid from the pure iris image so that every part will be represented in an independent region as shown in Figure (3, (b)). Splitting of iris region is invested when extracting feature vector from iris image and pattern matching stage, where feature vector is extracted only from pure iris regions without calculating the eyelash and eyelid parts which cover parts of iris region. In addition, calculate weights for each region within matching measures (45% for two pure iris image, 7% for eyelid iris image, and 3% for eyelash iris image). The second step is applying Gaussian sharpen filter to the four regions of the iris. This filter is used as enhancement technique that highlights puckers and edges and fine details in an images, as a result; the small details will accentuates as shown in Figure (3, (c)). The third step of the proposed iris feature extraction algorithm is regions blocking. Regions blocking is applied to the four regions , each region is divided into 4*6 blocks, where the region will have 24 block, each block is 15*15 pixels as shown in Figure (3, (d)). The purpose of dividing each region into a number of blocks is to the calculation in the fourth step of the proposed algorithm to extract feature vector from pure blocks and ignore blocks that include eyelid and eyelash.

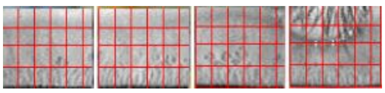
(a) Iris normalized image



(b) Split iris into four regions



(c) Gaussian Sharpen filter result



(d) Blocking of iris region

Figure 3: Iris feature extraction steps.

In the fourth step of the proposed iris feature extraction algorithm the statistical measures for each block will be calculated and its values will be stored in a specific feature vector in a database, where each vector is assigned to one statistical feature. The proposed iris feature extraction has two approaches, the first relies on first order statistical methods and the second relies on second order statistical methods.

2.1.1The First Approach: First Order Statistics

First order statistics can be used as iris texture feature extraction methods, which are based on the probability of pixel intensity values (the gray level distribution of pixel) on an image, and do not consider pixel neighborhood relationships (estimate properties of individual pixel values). The first order statistics measures that are used in the proposed algorithm are (Mean, Standard deviation, Coefficient of variation, Entropy), can be computed directly in terms of the image histogram using the equations below:

1- Mean:the mean is the average value, so it tells us something about the general brightness of the image and can be defined as follows:

$$\breve{g} = \sum_{g=0}^{L-1} gP(g) \qquad (1)$$

where P(g) is histogram probability and defined as follows:

$$P(g) = \frac{N(g)}{M}$$

where
P(g) = probability of gray level g in image,
N(g) = number of pixel with gray level g in image,
M = total number of pixel in image.

2- Standard deviation:which is also known as the square root of the variance, tell us something about the contrast, The equation for standard deviation is as follows:

$$\sigma_g = \sqrt{\sum_{g=0}^{L-1} (g-\breve{g})^2 P(g)} \qquad (2)$$

3- Coefficient of variation: tells us something about how gray levels which are variant. It is calculated by dividing the standard deviation by the mean value. The coefficient of variation indicates the relative amount of variability in distribution. The equation for variation is as follows:

-

$$C_v = \frac{\sigma_g}{\breve{g}} \qquad (3)$$

4- Entropy:it is used as a measure to the randomness of the data. It can also tell us how many bits do we need to code the image data. The equation for entropy is as follows:

$$En = -\sum_{g=0}^{L-1} P(g) log_2 [P(g)] \qquad (4)$$

After extracting the required features they are stored into a database that contains for each image four region (reg1,reg 2,reg 3,reg 4), each region has 24 blocks and each block has four feature vectors (Mean, Standard deviation, Coefficient of variation, Entropy).

2.1.2The Second Approach: Second Order Statistics

In this proposal the second order statistics use the gray level co-occurrence matrix (GLCM) to calculate various features from iris region images and store them as a frequency matrix.In the proposed algorithm GLCM is applied to blocking of iris images for two iris regions (the first one and the third) which represent the pure regions. In the beginning, for each block created GLCM, its size ranges between small number and large number of grayscale levels in a block, so the row and column of GLCM will be (small number, large number). The displacement h=1 and the direction is (0°, 45°, 90°, 135°). Calculating the GLCM for all blocks will be explained as shown in Figure (4).

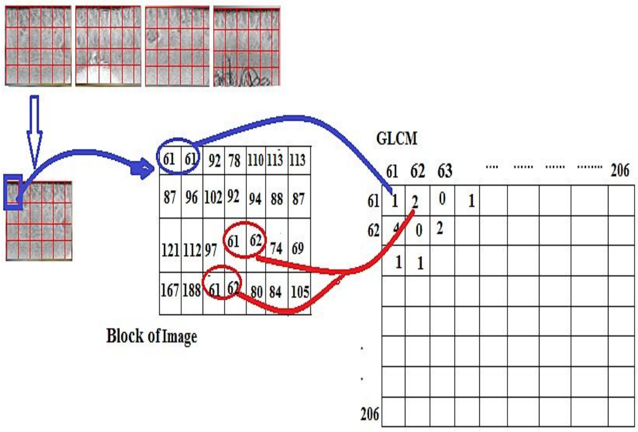


Figure 4: The GLCM for each iris region block.

After calculating the co-occurrence matrices, the second statistical measures (Mean, Standard deviation, Coefficient of Variation, Entropy) will be computed from it, and those measures will be stored as feature vectors in database.

2.2 Pattern Matching

After extracting the features and storing them as templates (numerical vectors) in database, a matching measure is used to find the similarity between the two irises; one is the tested iris template and the other is the template stored in the database. In order to determine whether the two iris templates match (i.e. extracted from the same eye), distance measures have to be employed to measure the closeness of that match. The measures used for the proposed approach are Euclidean distance (Sum of Squared Difference SSD) and city block distance (Sum of Absolute Difference SAD).

2.2.1 Euclidean Distance (Sum of Squared Difference SSD)

Euclidean distance (ED) or (Sum of Squared Difference SSD) is a distance measure that can be used in a biometric system. It defines the closeness of match between two iris feature templates. The Euclidean distance is defined as:

$$ED = \sum_{i=1}^{n} (ai - bi)^2 \qquad (5)$$

where:

(n) = the number of digits in template,

(a_i) = tested iris template,

(b_i) = iris template stored in database.

2.2.2 City Block (Sum of Absolute Difference SAD)

Another distance measure, called the city block or (Sum of Absolute Difference SAD), can be used to measure the degree of similarity between two iris feature templates, and can be defined as follows:

$$SAD = \sum_{i=1}^{n} |\mathrm{ai} - \mathrm{bi}| \qquad (6)$$

The proposed algorithm for pattern matching uses weights within the matching measures. The idea of this algorithm is allocate weights for each region of iris regions. The value of weights depends on the purity of the four regions and as shown in Figure (5):

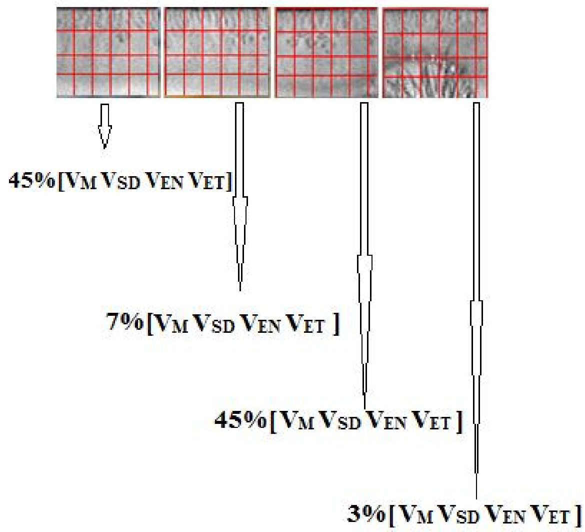


Figure 5: The weights distribution in the calculation of matching measurement.

Those weights have been determined after applying a number of tests to the experimental data and results show that those weights give the most efficient performance. The pure regions (without eyelash and eyelid) have been allocated the weight 45%; the region that has eyelash has been allocated the weight 3% and the region that has eyelid has been allocated the weight 7% because it's considered as damaged regions and badly affects the matching results. The proposed algorithm of weights in pattern matching is described below.

Algorithm 2:- The Proposed weights in pattern matching measures

Input: Feature vectors of iris image. //the tested iris.

Output: Feature vectors of iris image from DB that match tested iris.

Process:

For i=0 to N do // N is no. of records in the DB

For im=1 to 4 do // four iris region

Begin

Step1: If im = 1 || 3 then // pure iris region

Case 1: [45 % × $\sum | a_{im} - b_i |$] // SAD measure

Case 2: [45 % × $\sum (a_{im} - b_i)^2$] //SSD measure

Step2: If im = 2 then // eyelid iris region

Case 1: [7 % × $\sum | a_{im} - b_i |$] // SAD measure

Case 2: [7 % × $\sum (a_{im} - b_i)^2$] //SSD measure

Step3: If im = 4 then // eyelash iris region

Case 1: [3 % × $\sum | a_{im} - b_i |$] //SAD measure

Case 2: [3 % × $\sum (a_{im} - b_i)^2$] //SSD measure

Step4: Case 1: $SAD_{sample} = SAD_{(im)}$ //Summation four regions

Case 2: $SSD_{sample} = SSD_{(im)}$

Step5: Case1: If ($SAD_{sample(i)} < SAD_{(i)}$)then //Select the matched iris

$SAD_{Final} = SAD_{sample(i)}$

Case2: If ($SSD_{sample(i)} < SSD_{(i)}$)then // Select the matched iris

$SSD_{Final} = SSD_{sample(i)}$

Next im;

Next i;

End

Step6: End

To make a decision (to accept/ reject a person) the tested image is processed by the operations mentioned in the system; i.e. extract its feature vector and compare it with all of feature vectors in database.

3.EXPERIMENTAL RESSULTS

The proposed iris feature extraction and pattern matching methods tested on the iris data set (CASIA v4.0- interval), and tested on (CASIA v**1**.0- interval) iris image database.The proposed methods has implemented in c # language and Microsoft access office 2013 run under Microsoft windows 7 Home Premium Operating System Version 6.1.7601 Service Pack 1 Build 7601.Experimental results are shown in Table (1) and described as follows:

1-Experimental results show that the proposed methods are most effective in CASIA version 4.0 as compared to other existing iris databases.

2-Result shows that the system accuracy is 99.4% with less error rate when using first order statistic method in iris feature extraction stage.

3-The use of blocking iris region with block size (15×15) pixels is very convenient to compute the statistical measurements by using the measures (mean, standard deviation, coefficient of variation, entropy) to extract feature vectors. This selected block size results in feature vectors which have an appropriate size which makes storing and retrieving processes very efficient.

4-The results show that the iris recognition system is less effective when using second order statistic method in iris feature extraction stage, it requires huge storage and many calculations to gray level co-occurrence matrix, and it is slower than first order statistic method, so it is more useful in small enterprises where there are few individuals.

5-The results show that the proposed methods gives faster performance when using of SAD measure than using SSD measure in all tested databases so it is an efficient measure for iris recognition system.

6-To calculate FAR and FRR measurements of the proposed system, the following equations are used:

$$\text{FAR} = \frac{No.of\ false\ acceptance}{No.of\ imposter\ verification\ attempt}$$

$$\text{FRR\%} = \frac{No.of\ false\ rejections}{No.of\ enrolee\ verification\ attempt}$$

The obtained recognition result is compared with the recognition results of other existing methods. The results of this comparison are given in Table (2).

Table (2) The accuracy of the proposed system based on first order statistic measurement and other existing methods.

Methods	Database	Recognition rate
Daugman	-	100%
Wildes [8]	-	99.82%
K. Gulmire [12]	CASIA v1	89.5%.
The proposed method	CASIA v1	98.5 %
The proposed method	CASIA v4	99.4 %
The proposed method	CASIA v4	80 %

Table (1) The experimental results of the proposed iris recognition system.

Eye image DB	Feature extraction methods	No. of images	No. of persons	No. of image correctly recognized	No. of image faulty recognized	No. of image faulty accepted (imposter)	Average Time of matching eye (SSD)	Average Time of matching eye (SAD)	FAR %	FRR %	Average accuracy %
CASIA v4.0	FOS	360	30	358	2	3	~1 sec	~0.94 ms	0.0083	0.0055	99.4
CASIA v1.0	FOS	400	100	394	6	0	~1 sec	~0.98 ms	0	0.015	98.5
CASIA v4.0	SOS	30	10	26	4	2	~1.25 sec	~1.08 ms	0.067	0.13	86.67

4. CONCLUSION

A new iris feature extraction and pattern matching methods has been introduced,the conclusions are drawn:

1-Most of the exiting iris recognition systems are based on binary representation in the iris feature extraction, while in the proposed methods the grayscale representation has been adopted, and gives, very good results.

2-Two types of Statistical measurements (first order statistics and the second order statistics)are used to

3-extracting iris feature,each type used alone to extract feature vectors which are used in pattern matching measures.

4-The proposed weights (45%, 7%, 45%, 3%) that are assigned to the four iris regions in the matching measures increase matching rate compared with the matching rate without using those weights.

5-The proposed system is influenced by the type of digital camera used and the lighting conditions when capture an eye images, it is also influenced by the number of samples stored in database.

6-The proposed methods has short execution time where theaverage time consumed to perform the test phase is less than two seconds.

References

[1] Franjie El Khoury, “Iris Biometric Model for Secured Network Access”, CRC Press/Taylor & Francis Group, LLC, 2013.

[2] Charles A. Shoniregun, Stephen Crosier, “Securing Biometrics Applications”, Springer Science+Business Media, LLC, 2008.

[3] Surjeet Singh, Kulbir Singh "Segmentation Techniques for Iris Recognition System", International Journal of Scientific & Engineering Research, Vol. 2, Issue 4,pp.1-8, April 2011.
[4] J. Daugman, "How iris recognition works?", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 14, No. 1, pp. 21–30, January 2004.
[5] J. Daugman, "High Confidence Visual Recognition of Persons by a Test of Statistical Independence," IEEE Transactions on Pattern Analysis and Machine Intelligence", vol. 5, No. 11, pp. 1148-1161, November 1993.
[6] J. Daugman, "The importance of being random: statistical principles of iris recognition", Pattern Recognition, No. 36, pp. 279– 291, 2003.
[7] J. Daugman, "New Methods in Iris Recognition",IEEE Transactions on Systems,Vol. 37, No. 5, pp. 1167-1175, October 2007.
[8] R. P. Wildes, "Automated Noninvasive Iris Recognition System and Method", United States Patent, no. 5572596, 1996.
[9] B. Son, H. Won, G. Kee, Y. Lee, "Discriminant Iris Feature and Support Vector Machines for Iris Recognition", in Proceedings of International Conference on Image Processing, vol. 2, pp. 865–868, 2004.
[10] Aparna G. Gale, S. S. Salankar, " A Review On Advance Methods Of Feature Extraction In Iris Recognition System", IOSR Journal of Electrical and Electronics Engineering,PP. 65-70, 2014 (www.iosrjournals.org).
[11] Sheeba Jeya Sophia S. , Veluchamy S. ,"Security System based on Iris Recognition", Research Journal of Engineering Science ,Vol .2 ,PP:16-21 ,March 2013.
[12] K. Gulmire, S. Ganorkar, " Iris Recognition Using Independent Component Analysis", International journal of emerging technology and advance engineering, Vol. 2, Issue 7, PP.433-437, July 2012.

AUTHOR

Assistant Prof Dr. Abdulamir Abdullah Kerim received the B.S. in Statistics from Baghdad University, M.S. in artificial Intelligent from Baghdad University and Ph.D in Digital Image Processing degrees from University of technology, Iraq, in 1985, 1992 and 2006, respectively. In 2007 He was the head of artificial intelligent branch, in 2009 he was the head of the computer security branch, and in 2012 he was the head of the software branch. In 2013 he was the assistance of the head of the computer science Department\ University of Technology.

Sarah j. Mohammed received the B.S. degrees in computer science from University of Technology, Baghdad, Iraq in 2007. She is working as Lecturer, in Mechanical Engineering Department, University of Technology, Baghdad, Iraq. She is currently pursuing the M.S. degree in computer science at Department of Computer Science, University of Technology, Baghdad, Iraq.