

Iris recognition model based on Principal Component analysis and 2 level Haar wavelet transform: Case study CUHK and UTIRIS iris databases

Mohammed Hamzah Abed

College of Computer Science and Information Technology
University of Al-Qadisiyah

A biometric recognition system provide automatic identification of human being based on some special and unique physical or behavioral features of the individual. One of the most reliable identification system is iris recognition system. This work aim to recognize and identify iris among many of images that have been save in databases. Each one of database that used manipulating in many steps starting with enhance the details of iris and segment the iris and pupil then extract the raw features based on 2D Haar wavelet transform to capture both global and local features of iris image. After that by Applying reduction step to select only the useful and unique features that belong to each person. In this work PCA used as a reduction method. Finally the minimum distance are used to check the similarity between the database's features training set and input image, also three similarity techniques are used between input iris image and the template that save in database. Weighted Euclidean distance, Hamming distance and Cosine distance. Cosine method achieve good result than other method by using reduction and without reduction.

Keywords: iris recognition, Haar wavelet, PCA, CUHK, UTIRIS, minimum distance

1- introduction

With a huge progress of information technology in last years and menace on the personal data and how to keep their data secure, so the requirement to design powerful and high accurate system for human Authentication and identification recognition become more important and most challenges topic. Authentication of a person can be done through pass code, smart card, inserting a physical key into the PC's USB port or by using a biometrics. Some of services provider for example the services that Google company provide: they used two verification steps to increase the security of personal email or data on G-drive. It is optional choice used when the user login for first time from any physical device like laptop or mobile to save

that device as an interested device belong to the user, then the user can enter user name and password for login. Common methods include entering a code sent to a user's mobile phone through a text message contain the passcode, entering a code using the Google Authenticator smartphone app. Remembering password or keeping smart card or physical security key from unauthorized people are difficult, compare with biometrics based authentication is highly preferred system. Some companies start adding biometric authentication iris, fingerprint and face to their product devices like apple and Samsung companies [1]. Pattern recognition based on iris of human eye for authentication is unique and more suitable method, but face various challenges in the recognition phase's viz. acquisition of high quality image, segmentation of the concentric boundaries, select the proper features that belong to each person and matching [2]. In this paper, a new scheme has been proposed with some addition features for iris recognition consist of multi-layer haar wavelet transform, used the coefficient after reduction by using PCA.

2-iris databases

In this paper work two databases are used to test the recognition method first is: CUHK Iris image dataset are authored by CHUN, Chun Nam Ben, Automation and Computer-Aided Engr Department, Computer Vision Laboratory. the Chinese University of Hong Kong, Hong Kong. The database consist of 254 images collected from 36 person for left and right eye. For each donator there are seven images. These database contain only the iris images [3]. The second database that have been used is UTIRIS dataset is authored by "Mahdi S. Hosseini, Babak N. Araabi, Hamid Soltanian-Zadeh".the database contain 1540 images collected from 79 person from both left and right eyes. The Database is published by "Control and Intelligent Processing Centre of Excellence (CIPCE) at School of Electrical and Computer Engineering (ECE) in the University of Tehran" [4] Figure no 1 show example of the CUHK iris image and UTIRIS iris image.

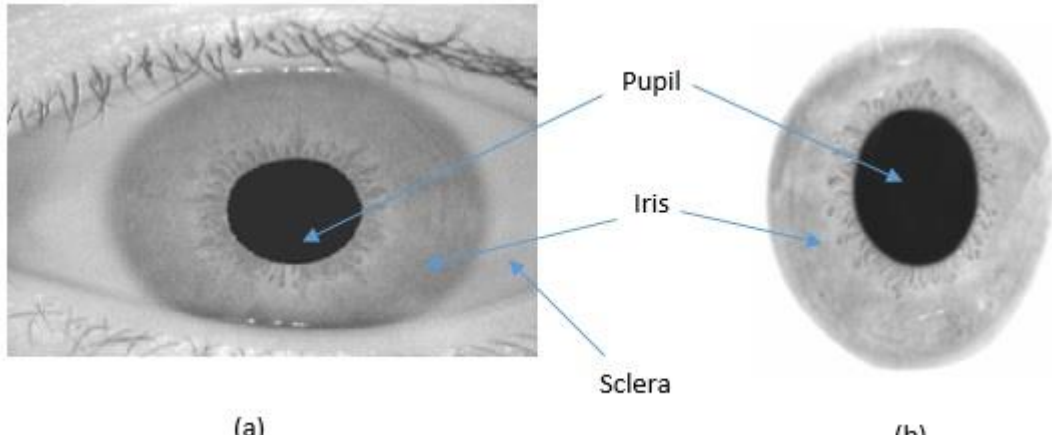


Figure .1 (a) sample of UTIRIS iris image (b) sample of CUHK

3- Proposed frame work

The proposed schema of iris recognition system consist of enrollment and verify phases. Each one of them contain four steps starting with resize and enhance the iris image, then iris localization and normalization, select set of features and check the similarity with the features dataset.

3-1 Enrollment phase

The first part of iris recognition system involves five steps: preprocessing, iris localization and normalization, features extraction and selection then features transformation by create features vector dictionary [5][6]. The figure number 2 show the steps of enrollment part.

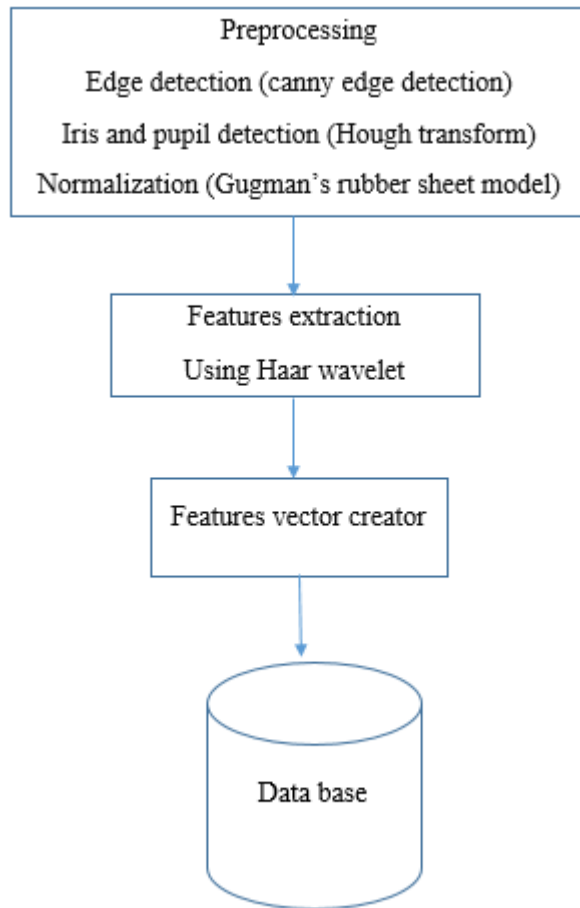


Figure .2 Enrollment phase

3-1-1 Iris Preprocessing , localization and normalization

In the first step the preprocessing are applies on all images in both databases by uniformed the image's size and dimensions then convert the images into gray scale after that adaptive histogram equalization are applied to enhance the images and to be ready to detect the edge by applying canny edge detection algorithm to detect the boundary of all eye. Then iris localization is applied to detect the edge of iris and pupil from the input image as shown in figure number three. The iris region can be approximated by two circles first one for pupil and second for iris as shown in figure three. The eyelids and eyelashes normally occlude the upper and lower

parts of the iris region [7]. There are many techniques that used for iris segmentation the most knowing techniques is Hough transform is a standard technique in computer vision algorithm that have been used to detect the simple geometric objects such as circle as in iris case or line in image. An automatic isolation algorithm based on the circular Hough transform is employed by many researcher. The segmentation method starting firstly, an edge map of the iris image is generated by calculating the first derivatives of intensity values in an input image then thresholding the result [8]. The parameters that used in this method to determine the circle geometric are the radius r and the center coordinates x_c and y_c . Which are able to define any circle according to the equation number 1.

$$x_c^2 + y_c^2 + r^2 = 0 \quad ..(1)$$

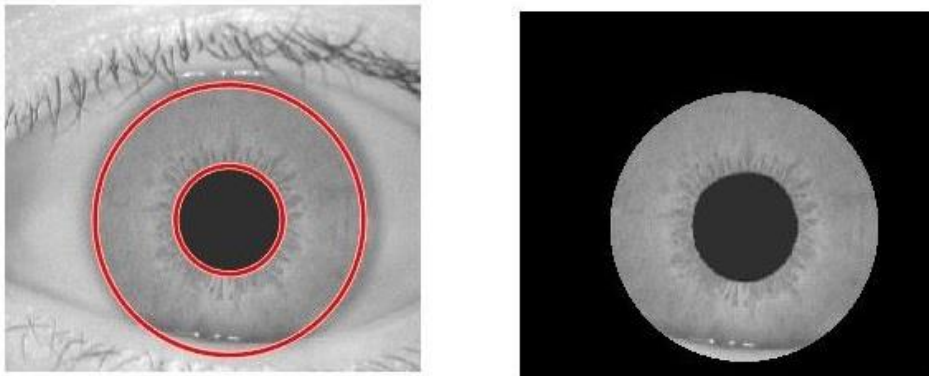


Figure .3 iris localization and segmentation

To extract iris region from the images and normalized into rectangular shape will applied Daugman's rubber sheet model as shown in figure 4. The rectangular form proper for features vector generator. This method implemented by remaps the homogenous of images by remapping each pixel within the iris region to a pair of polar coordinates (r, θ) where r is on the interval $[0,1]$ and θ is angle $[0,2\pi]$.

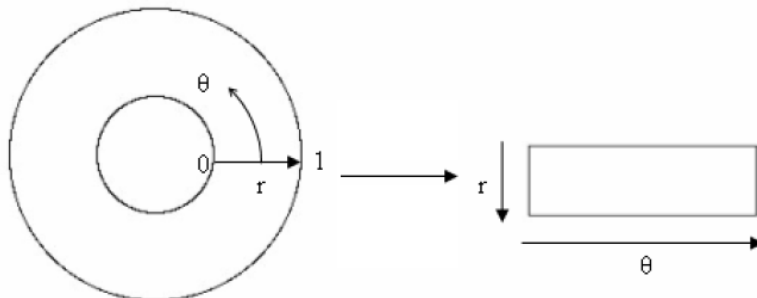


Figure .4 Daugman's rubber sheet model

The represented of the iris region from (x,y) Cartesian coordinates to the normalized non-concentric polar representation is modelled as $I(x(r, \theta), y(r, \theta)) \rightarrow I(r, \theta)$ (2)

Where

$$x(r, \theta) = (1 - r)x_p(\theta) + rx_1(\theta)$$

$$y(r, \theta) = (1 - r)y_p(\theta) + ry_1(\theta)$$

Where (x,y) are the original Cartesian coordinates, $I(x,y)$ is the iris region image, (r, θ) are the corresponding normalized polar coordinates, and x_p, y_p and x_1, y_1 are the coordinates of the pupil and iris boundaries along the θ direction. This method takes into consideration the pupil size to produce a normalize representation with constant dimensions.

3-1-2 Features extraction

The second major part in enrollment phase is features extraction and transformation. After segment iris image and detect boundary of iris and pupil based on Hough transform and normalization the iris region. Now the iris images are ready to extract the features. By applying Haar Wavelet packets as well as the energy of the packets sub images to extract the features of texture. In this work the 2 levels wavelet packet decomposition of Haar wavelet transform are employed to extract the texture of unwrapped iris region image [9]. In the Haar wavelet transformation method, low-pass filtering is calculated by averaging two adjacent pixel values, whereas the difference between two adjacent pixel values is signed for high-pass filtering. The Haar wavelet applies a pair of high pass and low-pass filters to image decomposition first in image columns and then in image rows independently [9]. As a first result the image divided into four sub bands as the

first level's output of Haar wavelet. The four sub-bands are Low Low 1 (LL1), High Low1 (HL1), Low High1 (LH1), and High High1 (HH1). Upto two levels of decomposition are done to get the detail image. As shown in Figure number 5.

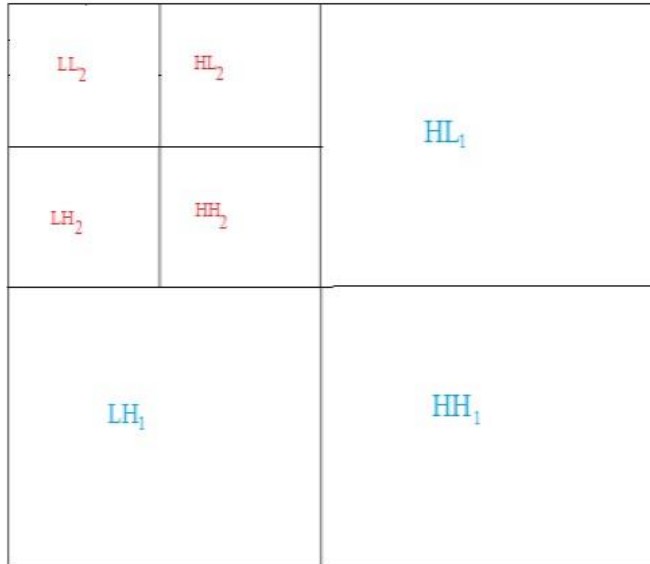


Figure .5 structure of 2 level Wavelet decomposition

3-1-3 features reduction and transformation

When the vector of raw features are extracted and select only the low low part of the haar wavelet transform in this step the reduction are ready to applied to minimize the number of redundancy features to increase the efficiency of recognition by choice the useful characterize of each iris image in database. In this phase PCA are applied as a reduction method [10]. The selected features will be used to generate new features of each iris image to each person in the database. The new features that will generated are mean, median, standard deviation, variance, skewness , kurtosis ,min and max. To generate this features will used the following equations:

$$\text{Mean } \mu_X = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad (3)$$

$$\text{Standard deviation } \sigma_X = \sqrt{\frac{1}{n} \left\{ \sum_{i=1}^n X_i^2 - \frac{1}{n} \left(\sum_{i=1}^n X_i \right)^2 \right\}} \quad (4)$$

$$\text{Kurtosis } k = n \frac{\sum_{i=1}^n (x_i - x_{avg})^4}{(\sum_{i=1}^n (x_i - x_{avg})^2)^2} \quad (5)$$

$$\text{Skewness } s = \sqrt{n} \frac{\sum_{i=1}^n (x_i - x_{avg})^3}{(\sum_{i=1}^n (x_i - x_{avg})^2)^{3/2}} \quad (6)$$

3-2 Verify phase

The second part of the iris recognition system is verifying phase, in this phase the unknown entering iris image will be process to check the similar pattern from the dataset that contain all most known iris image. The verify stages can be shown in figure number 6.

The first step in this phase is preprocessing to enhance the query image to be ready to extract the features by using 2 levels of Haar wavelet transform and using the coefficients as a set of features and by using PCA to reduce the unwanted features. In this steps the features are transform into another pattern by select only the useful features that can be help in recognition part. Then when all the features of the unknown iris are generated, now the matching phase is ready to found the similar or closer result with the dataset. In this work will used two methods of similarity measurement: Weighted Euclidean distance and Hamming distance, and compare the result of them in the experiment part.

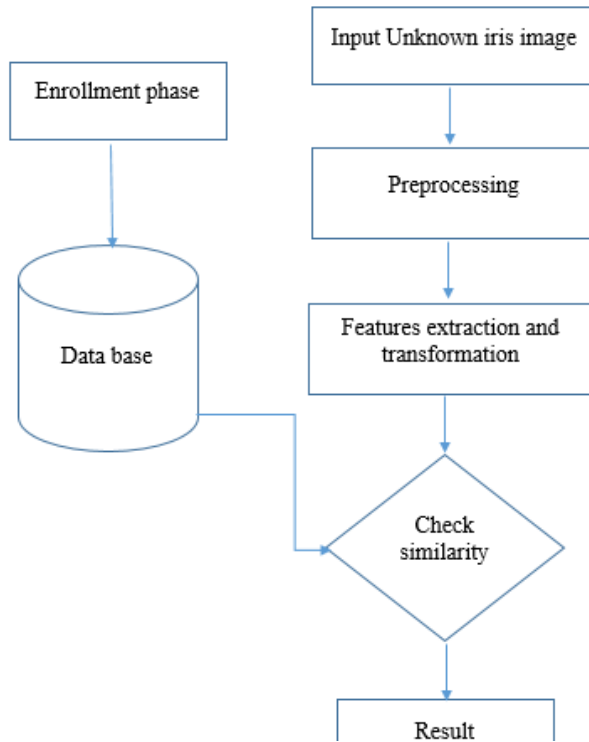


Figure. 6 verifying phase of iris recognition system

3-2-1 Matching measurement

To check the similarity between the query unknown iris image and the pattern in the dataset, will use two methods to found the best matching result.

1- Weighted Euclidean distance

It is a technique can be used to compare between two vector of features with same size represented as a vector, the WED gives a measure of how set of values are

closer or similar, between two vectors of features [11][10]. The Weighted Euclidean distance can be calculated as an equation number (7)

$$WED = \sum_{i=1}^N \frac{(f_i - f_i^{(k)})^2}{(\delta^{(k)})^2} \quad (7)$$

Where

f_i is the i^{th} feature of the query input iris.

$f_i^{(k)}$ is the i^{th} feature of the iris template that save in dataset.

$(\delta^{(k)})$ is standard deviation of the i^{th} feature of the iris template k .

The unknown iris found the matching template k , when WED is a minimum at k .

2- The Hamming distance

It is one of the similarity measurement methods gives a measure of how many bits are similar between two vectors of bits. Clearly the Hamming distance cannot be negative, and if it is zero, then the vectors are identical. The distance does not depend on which of two vectors we consider first, the Hamming distance, HD , is defined as the sum of the exclusive-OR between X and Y bits over N , the total number of bits in the bit vector [11].

$$HD = \frac{1}{N} \sum_{j=1}^N X_j (XOR) Y_j \quad (8)$$

In some cases the Hamming distance measure value between the two patterns equal to half. This case occurs because independence implies the two bit patterns will be totally random, so there is half chance of setting any bit to 1, and vice versa[12].

3- Cosine distance

The Cosine distance measure the similarity between two vector with non-zero vector integer or Boolean component, in the space a point may be through of as a direction. The cosine measurement distance between two vectors is an angle between point and vector that make. This angle's degree between 0 to 180, depend on how many dimensions the space has. Where the Cosine of 0 is 1 and it is less than one for any other angle in positive space. If we have two vectors with the same values a cosine similarity of 1, and two vectors at 90° have a similarity of 0, and two vectors diametrically opposed have a similarity of -1 [11]. The equation number (9) show the formula of Cosine distance.

$$\cos(A, B) = \frac{A \cdot B}{\|A\| \|B\|} \quad (9)$$

4- Experiment result

The entire proposed frame work of iris recognition system was implanted by using Matlab R2015a. The experiment of this proposed work was investigated with two different databases CUHK and UTIRIS. The first dataset consist of 254 iris images for left and right eye from 36 person. And the other contain 1540 eye image collected from 79 person for left and right eye. Different experiment were performed for those two databases. For CUHK database seven iris's image belong to each person select four images from them for training set and the remain used for testing.

The first experiment was implemented to determine the pupil and iris circle for both databases by using Hough transform with different radius and center. And the accuracy of the detect pupil and iris shown in table number 1.

Table 1. Iris segmentation by using Hough transform

| database | Number of image | Right iris segmentation | Wrong iris segmentation |
|----------|-----------------|-------------------------|-------------------------|
| CUHK | 254 | 98.425% | 1.575% |
| UTIRIS | 790 | 98.734% | 1.276% |

The second experiment in this work tested the features of iris image that extracted by using Haar Wavelet Transform in both cases: first one used all features of iris image and second one using PCA as reduction methods to select only the good features of the iris image. And by using three methods to check the similarity of the input image with iris images in databases. Table number 2 show the accuracy of the reduction to get a best result of similarity and matching with the training part of the databases using WED, Hamming and Cosine distance of the CUHK database , and Table number 3 show the accuracy of UTIRIS database.

Table 2. CUHK database with different reduction and similarity methods

| Database | Reduction techniques | Similarity techniques | Right matching | Wrong matching |
|----------|----------------------|-----------------------|----------------|----------------|
| CUHK | Without reduction | WED | 80.7% | 19.3% |
| CUHK | PCA | WED | 85.433% | 14.567% |
| CUHK | Without reduction | Hamming | 81.496% | 18.504% |
| CUHK | PCA | Hamming | 88.976% | 11.024% |
| CUHK | Without reduction | Cosine | 82.677% | 17.323% |
| CUHK | PCA | Cosine | 92.519% | 7.481% |

Table 3. UTIRIS database with different reduction and similarity methods

| Database | Reduction techniques | Similarity techniques | Right matching | Wrong matching |
|----------|----------------------|-----------------------|----------------|----------------|
| UTIRIS | Without reduction | WED | 84.81% | 15.19% |
| UTIRIS | PCA | WED | 87.341% | 12.659% |
| UTIRIS | Without reduction | Hamming | 86.7% | 13.3% |
| UTIRIS | PCA | Hamming | 88.607% | 11.393% |
| UTIRIS | Without reduction | Cosine | 89.24% | 10.76% |
| UTIRIS | PCA | Cosine | 91.139% | 8.861% |

The figure 7 and 8 show the Receiver Operating Characteristic (ROC) curve Of the CUHK and UTIRIS databases with different threshold.

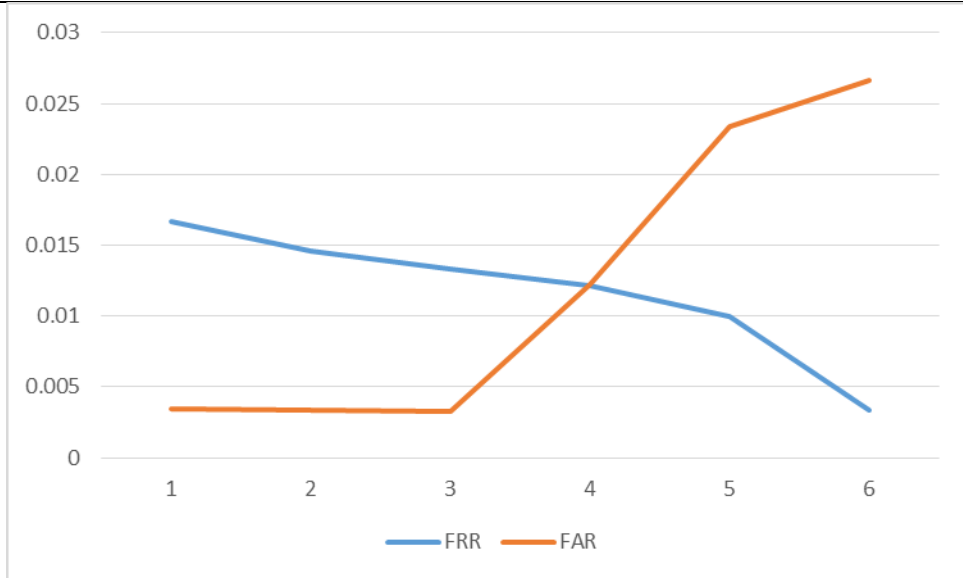


Figure. 7 ROC of CUHK database

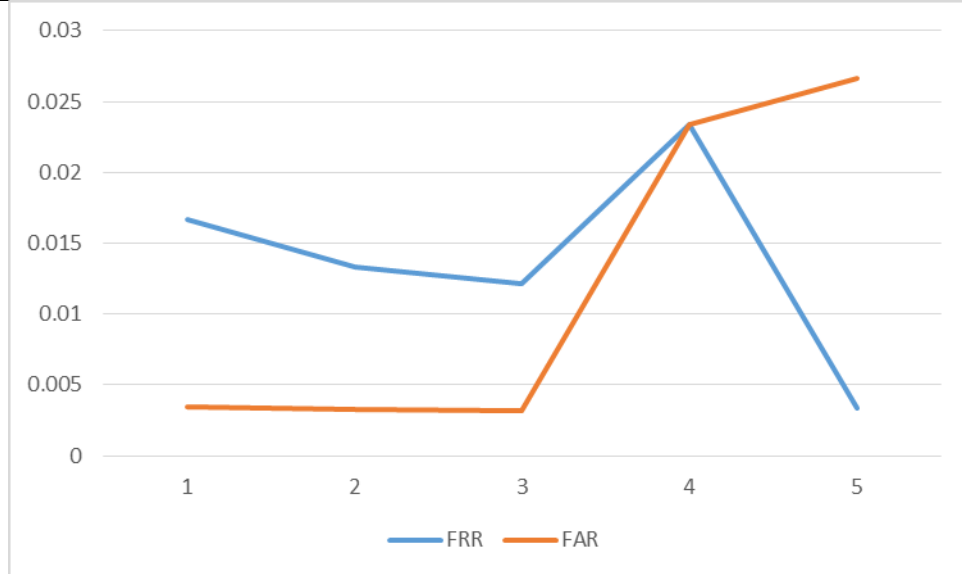


Figure. 8 ROC of UTIRIS database

5- Conclusion

This work are presented iris segmentation and recognition system, which used two iris databases CUHK and UTIRIS gray scale eye image .In this paper work the automatic segmentation are achieved by using Circular Hough transform method for iris segmentation and pupil detection and to enhance the iris images in both databases adaptive histogram equalization are applied on all images. In the second part the features are extracted based on Haar wavelet transform to extract the raw features of iris gray scale image for many reasons it's simple and fast to compute and need less memory compare with another methods of wavelet. By using PCA as a reduction method to select only useful features the recognition rate are increasing. Depend on the minimum similarity measurement are used the experiment result show the Cosine distance score good result more weighted Euclidean distance and Hamming distance .in conclusion Haar wavelet transform is efficient method for recognition and sensitive to noise under different cases and conditions with Cosine distance.



References

- [1] Jain, A.K.; Bolleand, R.M.; Qankanti, S.; Biometrics: personal Identification in a networked society; Norwell, MA: kluwer; 1999.
- [2] Ashok K Bhateja , Shikhar Sharma, Santanu Chaudhury, Nitin Agrawal , Iris recognition based on sparse representation and k-nearest subspace with genetic algorithm, Pattern Recognition Letters, Volume 73 Issue C,2016 .
- [3] CHUN, Chun Nam Ben, CUHK Iris Image Dataset, Computer Vision Laboratory,2003.
- [4] Mahdi S. Hosseini, Babak N. Araabi and H. Soltanian-Zadeh, Pigment Melanin: Pattern for Iris Recognition, IEEE Transactions on Instrumentation and Measurement, vol.59, no.4, pp.792-804, April 2010.].
- [5] R. P. Ramkumar , Dr. S. Arumugam, Improved Iris Segmentation Algorithm without Normalization Phase, International Journal of Engineering and Technology,2014.
- [6] Surbhi Gaur, Vivek Agarwal, Performance of Iris databases for Authentication , International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 5, Issue 4, April 2016
- [7] Hui Zhang, Zhenan Sun, Tieniu Tan, and Jianyu Wang , Zhenan Sun Jianhuang Lai Xilin Chen Tieniu Tan (Eds.), Ethnic Classification Based on Iris Images, Biometric Recognition 6th Chinese Conference, CCBR 2011Beijing, China, December 3-4, 2011 Proceedings Springer
- [8] , Xiao Nan Liu and Wei Qi Yuan, Iris Plaque Detection Method Based on Level Set Zhenan Sun Jianhuang Lai Xilin Chen Tieniu Tan (Eds.), Biometric Recognition 6th Chinese Conference, CCBR 2011Beijing, China, December 3-4, 2011 Proceedings Springer
- [9] , Ali Abdulmunim Ibrahim, Iris Recognition using Haar Wavelet Transform
Journal of Al-Nahrain University, 2014.
- [10] Mohammed Hamzah Abed, Wrist and Palm Vein pattern recognition using Gabor filter, Journal of Qadisiyah Computer Science and Mathematics, 2017.
- [11] Shraddha Pandit , Suchita Gupta, A comparative study on distance measuring approaches for clustering, International Journal of Research in Computer Science ,2011.

[12] , Leila Baccour, Adel M. Alimi, and Robert I. John , Some Notes on Fuzzy Similarity Measures and Application to Classification of Shapes, Recognition of Arabic Sentences and Mosaic IAENG International Journal of Computer Science,2014.

[13] Md.Sajjad Hossain, Kandarpa Kumar Sarma, Iris Recognition Based Identification Using 2D-Discrete Cosine Transform and Self Organizing Map Neural Network, International Journal of Systems , Algorithms & Applications,2012.

نظام تمييز قزحية العين بالاعتماد على تحليل المكون

UTIRIS و CUHK :دراسة حالة قاعدتي بيانات Haarالرئيسي و المستوى الثاني من تحليل الموجي

محمد حمزة عبد الصالحي

قسم علوم الحاسوب ، كلية علوم الحاسوب وتكنولوجيا المعلومات

جامعة القادسية ، العراق

الخلاصة

نظام التعرف على الهوية الحيوية يوفر التعرف التلقائي للإنسان على أساس بعض الخصائص البدنية أو السلوكية الخاصة والفريدة للفرد. واحدة من أهم أنظمة التعرف الأكثر موثوقية هو نظام التعرف على الأشخاص أو الأفراد عن طريق قزحية العين. يهدف هذا العمل إلى التعرف على القزحية وتحديدتها بين العديد من الصور التي تم حفظها في قواعد البيانات التي تمت دراستها CUHK و UTIRIS . كل واحد من قواعد البيانات التي تستخدم تتبع العديد من الخطوات بدءاً من فصل تفاصيل القزحية وتقطيع القزحية ثم استخراج الميزات الخام على أساس تحويل الموجة هار للصورة الثنائية لانتقاط كل من الخصائص العالمية والمحلية لصورة قزحية العين. بعد ذلك عن طريق تقليل الخصائص لتحديد فقط ميزات مفيدة وفريدة من نوعها التي تنتمي إلى كل شخص. في هذا العمل PCA تستخدم كطريقة للتقليل. وأخيراً يتم استخدام المسافة الدنيا للتحقق من التشابه بين مجموعة التدريب ميزات قاعدة البيانات وصورة الإدخال، كما يتم استخدام ثلاثة تقنيات التشابه بين صورة قزحية الإدخال وال قالب الذي حفظ في قاعدة البيانات. المسافة الإقليدية المرجحة، مسافة هامنج ومسافة جيب التمام. طريقة التمام تحقيق نتيجة أفضل من الطرق الأخرى باستخدام الخصائص كاملة أو بدون تخفيض للخصائص.