Security of Iris Recognition and Voice Recognition Techniques

Ayad R. Raoof
Computer Center/College of Administration and Economics/University of Baghdad

Received in : 16 February 2012 , Accepted in:12 September 2012

Abstract

Recently, biometric technologies are used widely due to their improved security that decreases cases of deception and theft. The biometric technologies use physical features and characters in the identification of individuals. The most common biometric technologies are: Iris, voice, fingerprint, handwriting and hand print. In this paper, two biometric recognition technologies are analyzed and compared, which are the iris and sound recognition techniques. The iris recognition technique recognizes persons by analyzing the main patterns in the iris structure, while the sound recognition technique identifies individuals depending on their unique voice characteristics or as called voice print. The comparison results show that the resultant average accuracies of the iris technique and voice technique are 99.83% and 98%, respectively. Thus, the iris recognition technique provides higher accuracy and security than the voice recognition technique.

Keywords: Biometric technology; iris recognition; voice recognition, oversampling, feature vector, Gaussian Pyramid, Haar wavelet, DFT, MFCC, GMM, HMM, UBM, CASIA, MMU

Biometric Recognition Techniques

1. Iris Recognition Technique

Iris recognition technique is a biomedical technology that is utilized for the identification as well as verification of individuals by using the iris patterns. These patterns have high stability and it is uniqueness. Each person has unique iris. In this section, the main processes of the iris recognition technique are illustrated and explained. In addition, the implementation of the iris recognition security system by using MATLAB is explored. Image pre-processing techniques as well as normalization are an important parts of the iris recognition technique. Any modification in the lighting conditions causes a reduction in the recognition technique performance. [1, 2]

The main stages of implementation as shown in figure 1 are: Image acquisition, image manipulation, iris localization, mapping, feature extraction, Haar wavelets, binary coding scheme and test of statistical independence. [3]

A. Image acquisition

In this stage iris images are captured by using a CCD camera with 640x480 resolutions. The captured images are white and black in order to obtain more details. [3]

B. *Image manipulation*

Image manipulation is done by converting images form RGB into gray ones as well as from eight bits into double precision. This is done in order to simplify the image manipulation for the following steps. [3]

C. Iris localization

The iris boundaries which are must be detected which broadens from inside the limbus which is the edge among both the sclera and the iris to the outer of the pupil. The outer edge is determined by down-sampling images by factor value equals 4 in order to allow a rapid delay or processing by utilizing a Gaussian Pyramid. The gradient image is achieved by using the canny operator in the MATLAB. [4]

Circular summation which contains the summation of all the circle intensities by utilizing three overlapped loops is then used to move over all the probable radii and coordinates of center. Rescaling the achieved results is used to determine both the radius and center of the original image iris. After finding the outer edges, the iris pixels intensities are tested. By using the canny threshold, if the iris has a dark color, then a small threshold value is utilized to allow canny operator to discover the circle that separates between iris and pupil. In the other hand, if the iris has a color, then a high threshold is used. [3]

The center of the pupil is shifted by nearly 15% from the iris center as well as its radius is in the range between 0.1 and 0.8 of the iris radius. So the research time of the pupil center is small. Enhanced accuracy can be achieved by searching the original iris instead of the down-sample iris version. Figure 2 shows the resultant iris region. [5]

D. Mapping

The resultant iris edges must be secluded and stored in another image. The pupil size varies according to the light intensity. So, the coordinate system is modified by unpacking the iris lower area and then mapping all points in the iris edges into their polar equivalent as shown in figures 3 and 4. The mapped image size is constant. When the pupil size increase, then the same point will be mapped again. [3]

Unpacking the image is done by using the bilinear transformation in order to achieve the point's intensities in the new image. These intensities are resulted from the old image grayscales. [5]

E. Feature extraction

The iris patterns are extracted in this stage by considering the correlation among the neighboring pixels. In this stage, wavelets transform and Haar transform are used. Figure 5 shows the Haar wavelet. [3]

F. Haar wavelet

The computation time must be low. So, designing a neural network will take long time. To solve this problem, another wavelet will be chosen. A wavelet tree with 5-level is used with illustrating all the estimation and detail coefficients of a mapped image.

By contrasting the results of the combination between Haar transform and wavelet tree with the result of other wavelets, it is shown that Haar transform offers the best results. The used mapped image size is 100x402 pixels where it can be decayed by utilizing Haar wavelet into five levels. These levels are the following: [3]

- Horizontal coefficients: cD^h_1, cD^h_5
- Vertical coefficients: cD^v_1, cD^v_5
- Diagonal coefficients: cD^d_1, cD^d_5

The coefficients that stand for the iris pattern core must be chosen and the ones that expose disused data must be deleted. Figure 6 shows the chosen coefficients, which are: cD^h_1, cD^h_2, cD^h_3, cD^h_4. But only one must be used. [3]

cD^h_4 is chosen as an envoy of all the carried data by the four levels since it replicates the same patterns like the proposed horizontal coefficients and it has the smallest size. The fifth level has not the same textures, so it will be chosen like a whole. In the vertical and diagonal coefficients, the fourth and fifth ones will be selected. The resultant coefficients are joined to form a vector that describes the patterns of the iris where it is called feature vector with size equals 702 elements. [6]

G. Binary coding scheme

The resultant vector will be converted into binary since finding the difference among two binary words is simpler than finding the difference among two vectors. Coding the feature vector depends on the detection of its main features. The resultant vectors maximum value is bigger than zero, while their minimum value is smaller than zero. Their average is in the range from -0.08 to -0.007 and the standard variation is in the range from 0.35 to 0.5. Let "Coef" is an image feature vector, then the way to convert it into binary code is as the following: [8]

- When Coef(i) >= 0, then Coef(i) = 1
- When Coef(i) < 0, then Coef(i) = 0

After the binarization process, the resultant two code words are compared in order to find if these codes stand for the same individual or not. [3]

H. Test of statistical independence

In this stage, two iris patterns are contrasted with each other. The Hamming distance among two feature vectors and the divergence among them are directly related. In other words, two irises are the same when the Hamming distance is small. The hamming distance is the number of equivalent bit positions that differ.

The hamming distance is similar to the correlation coefficient in which the original image and the corresponding image in the database are converted into vectors and then each two related vectors are compared and the distance between them is calculated. The smallest resultant distance is called the Hamming distance. Daugman found that the minimum value of Hamming distance among two irises of the same individual is 0.32. So, the related binary feature vectors of the two contrasted images of iris are conceded to a function that corresponds to the computing of the Hamming distance among them. The following shows when these two images are related to the same individual or not: [3, 5]

- If HD <= 0.32, then the two images are for the same person.
- If HD > 0.32, then the two images are for different persons.

Voice Recognition Technique

Voice recognition technique consists of two main stages, training and recognition. In the training stage, an identified voice is inserted into a database. In the recognition stage, an undefined voice signal will be identified. Voice recognition method is used for both verification as well as identification. Figure 7 shows the main system structure. [7]

A. Feature extraction

The first step in the voice recognition system is the feature extraction. It converts the input voice signal into an acoustic feature vectors series by using the signal wave division technique. Most of the vectors depend on the voice signal cepstral representation. The purpose of this stage is to find a novel illustration which is more compressed, more appropriate for the statistical modeling and less disused. [6, 7]

The main processes that are applied on the recorded voice signal are: The voice signal is initially separated into a number of time windows. After that, the Discrete Fourier Transform (DFT) is used in order to convert each one of the resultant time windows into the spectral domain. Then, each one of the magnitude spectrums are smoothed out by using band pass-filters. Each one of these filters measures the sun-band mean. This process is called the extraction of Mel Frequency Cepstral Coefficients (MFCC) parameters as shown in figure 8. [7]

Energy and derivatives parameters are also added to the cepstral parameters. Energy is the summation of samples power over time in a frame. Speech signals are fixed from one frame to another one. So, features that are corresponding to the modifying in cepstral features are added in which a velocity feature is added into each one of the vector features. [8]

B. Speaker molding

Speaker model is constructed by the use of acoustic vectors that are found out from each signal part in the training stage. The main two methods of voice recognition are: Deterministic method which is the dynamic comparison in addition to vector quantization and statistical method which is the Gaussian Mixture Model (GMM) and Hidden Markov Model (HMM). The statistical method is the most utilized one. [7, 9]

GMM is an unverified leaning method or a clustering method. In addition, it is a parametric PDF which is signified as the weighted sum of all the densities of the Gaussian component. This method can be used in order to build flexible boundaries of clustering, such as the points in space join a class based on a known probability. As well, it is utilized as a probability distribution parametric model for either the biometric system features, such as the speaker recognition features or the continuous measurements. The parameters of this method can be evaluated by using a training dataset with the utilization of the iterative Expectation Maximization (EM) method. The following equation explores the general form of the GMM: [10, 11]

$$p(x) = \frac{1}{(2\pi)^{\frac{d}{2}}|\Sigma|^{2}} e^{\left[\frac{-1}{2}(x-\mu)^{T}\Sigma^{-1}(x-\mu)\right]}$$

Where: x represents the dimensional component vector of the features, μ represents the dimensional component vector of the features means and Σ represents the covariance matrix. [11]

GMM depend on the Universal Background Model (UBM) which constructed by utilizing all the database recordings. This model is used due to the flexibility of the modeling by using GMM and the utilization of GMM offers a well cooperation among performances as well as the system complication. [7]

C. Pattern matching and decision

The voice recognition method recognizes the voice signal if it belongs to the assumed speaker. Let Y is the voice segment and S is the assumed speaker, then two hypotheses are defined as follows: [7]

- H0: Y is from the assumed speaker S
- H1: Y is not from the assumed speaker S

To make a decision between these two hypotheses, the ratio between the probability density function of both hypotheses is calculated, where if it is bigger than a certain value, then H0 will be accepted, otherwise, it will be rejected. [12]

Results

1. Iris recognition technique

Capturing iris images is done by using a device which offers a standard complex video signal that is in combination with a frame of a grabber board.

The captured image by the proposed hardware is inserted into the MATLAB as shown in figure 9 via the image acquisition toolbox which offers an interface for several devices of image acquisition. This toolbox also allows the collection of images by using the frame grabber. This is done by using two methods, get-snapshot function and triggering. [13]

Capturing a video is important in order to make sure that images are achieved. This is done by using capture video button which starts the image acquisition toolbox which in turn starts commands and saves data. When the capturing process is completed, then program control is reinstated and the collected data are transmitted to the MATLAB workspace.

The transmitted data into the workspace is a four dimensional array. The first two dimensions are related to each one of the person frame rows and columns. The third one is related to the components of red, green, and blue colors. The fourth dimension is related to the number of the frame. Imaqmontage command is used in order to show the images to perform quality corroboration.

Persons move their eyes during capturing. The resultant data are utilized in the analysis of liveness testing algorithms as well as location tracking algorithms. After the proposed processes of acquisition, data are stored. When the images have high quality, images are saved.

Several experiments are done in order to evaluate the iris recognition technique. So, this technique is implemented by using MATLAB software program. The utilized databases in this experiment are from Chinese Academy of Sciences Institute of Automation (CASIA) as well as Multimedia University (MMU).

Databases of CASIA contain eye images which were captured for nearly 108 individuals with seven images for each one of them from two assemblies, where each assembly was taken with one month time period. Databases of MMU contain eye images that were captured for both eyes of individuals. [14, 15]

The left image that is shown in figure 10 shows the iris image after the iris localization process, while the middle one shows the iris segmentation second phase or as called noise reduction. The right image displays the failed segmentation of the iris which results from several reasons, such as the upper eyelid as well as eyelashes impeded with the iris localization process, persons did not support their irises with the used scanner of iris in a proper way, or there was not adequate characteristic difference among both the iris and the sclera. Figure 11 shows the resultant image after normalization. [16]

Table 1 shows the iris coding process results. As shown below, for the CASIA databases, all irises are coded in a successful way, while for the MMU databases, two iris images from

450 images are not coded successfully. So the total rate of coding success is 99.83% and the needed average time for the process is 708 ms for each iris. [2]

Table 2 shows the experimental results of iris recognition technique. The resultant accuracy is 84.45% for the used CASIA databases. The data of the second assembly are utilized as reservation and the first one is utilized as the database of the system. The resultant accuracy for both the left and right eyes of the used MMU databases are 77.78% and 86.67%, respectively. For the databases of the MMU, the first two image acquisitions are utilized as reservations, while the other three ones are utilized as the system database. The total accuracy is 82.9%. [2]

2. Voice recognition technique

Speech processing as well as language consists of several concepts and terminologies. Figure 12 shows speech signals of vowels in both time and frequency domains. As shown, the pitch as well as rate for each one of the vowels is the same which creates major worry for the voice recognition systems. They have high rejection rate due to the background interference.

The simple voice recognition system is implemented by using MATLAB/SIMULINK. Voice reference template is implemented and can be contrasted adjacent to subsequent voice recognitions. Reference template is built by let a person speak his name and this will be recorded in form of .wav.

The voice recognition system consists of several steps, which are: Energy levels measurement of the signal small period, removing noises as well as unrequited signals via using the digital filter design block that acts as discrete FIR band-pass filter and generating the signal frequency components. After that, the feature extraction is done which determining the pitch contour, format frequencies and density of average energy spectral by utilizing both FFT and autocorrelation. Sizes of both input patterns and reference are calculated in order to perform the comparison. Final results are verified by measuring several parameters, such as standard deviations.

The speech voice is a waveform that is called a voice pattern. Figure 13 and 14 illustrates two voice patterns of a person. One of them is utilized as voice reference pattern. The proposed voice samples were recorded in several times. The voice recognition system will calculate the differences between them. Reference Voice as well as Voice ID is considered as nonparametric approximations of the voice energy bands. The resultant Reference Voice value is 5.28 and the resultant Voice ID value is 5.959. The difference among them is 0.6792. The obtained standard deviations that contrasted with Reference Voice and Voice ID are 12.79% and 11.33%, respectively. These two percentages are less than 15%; these two voices are for the same person.

Figures 15, 16 and 17 show speech signals for dissimilar males as well as females. The percentage of both standard deviations is bigger than 15%. The average accuracy of the voice recognition system is 98%.

Comparison between the Securities of the Two Techniques

In this project, two biometric identification methods are explained, which are: Iris recognition technique and voice recognition technique. The resultant average accuracies of the iris technique and voice technique are 99.83% and 98%, respectively.

So, the iris recognition technique offers higher accuracy than the voice recognition technique. In other words, iris recognition technique provides more security and safety for the data of individuals.

Conclusion

Biometric techniques are widely used due to their security as well as high accuracy. These techniques use several physical features in the recognition of individuals. In this paper, two recognition methods with their main stages are explained extensively, which are the iris recognition technique and voice recognition technique. These two techniques are implemented and simulated by using MATLAB software program. The resultant average accuracies of the iris recognition technique and voice recognition technique are 99.83% and 98%, respectively.

Iris recognition technique as shown offers higher accuracy as well as security of person's data than the voice recognition technique.

References

1. Chaskar, U. M. and Sutaone, M. S. (2010) A Novel Approach for Iris Recognition, 2nd International Conference on Computer Technology and Development (ICCTD 2010), 495-500.
2. Agus Harjoko, Sri Hartati, and Henry Dwiyasa, (2009), A Method for Iris Recognition Based on 1D Coiflet Wavelet, World Academy of Science, Engineering and Technology 56:123-129.
3. Chen, Y.; Dass, S. C. and Jain, A. K., (2006),Localized iris image quality using 2-D wavelets, Springer Lecture Notes in Computer Science 3832: International Conference on Biometrics, 373-381.
4. Daugman, J. How Iris Recognition Works".
5. Gonzalez, R.C., Woods, R.E, (2002), Digital Image Processing, 2rd ed., Prentice Hall.
6. Lim, S.; Lee, K.; Byeon, O.and Kim, T, (2001),Efficient Iris Recognition through Improvement of Feature Vector and Classifier", ETRI Journal, 23 (2): 61-70.
7. Mohamed Chenafa, Dan Istrate, Valeriu Vrabie, and Michel Herban, '' Biometric system based on voice recognition using multi-classifiers''
8. Kinnunen, T.; Hautamki, V. and Franti, P.(2004), Fusion of spectral feature sets for accurate speaker identi_cation. In 9th International Conference Speech and Computer (SPECOM), 361-365.
9. Jackson Zhang, and Bruce Wang, (2010), A novel voice recognition model based on HMM and fuzzy PPM, 637-640.
10. César Souza, (2010), Gaussian Mixture Models and Expectation-Maximization''.
11. Bertrand Scherrer,(2007), Gaussian Mixture Model Classifiers, 1-4 .
12. Doddington, G.(1985),Speaker recognition identifying people by their voices", In Proc. Of the IEEE, 73, 1651-1664.
13. Robert, C. Schultz, Robert W. Ives, (2005), Biometric Data Acquisition using MATLAB GUIs'', 35th ASEE/IEEE Frontiers in Education Conference, Indianapolis, IN, October 19 – 22.
14. Masek, L.(2003),Recognition of human iris patterns for biometric identification", Bachelor's Thesis, University of Western Australia.
15. Masek, L. and Kovesi, P.(2003), MATLAB Source Code for a Biometric Identification System Based on Iris Patterns, The University of Western Australia.
16. Kaleb Waite and Amanda Eure, (2009), Iris Recognition Optimized for Information Assurance, Proceedings of The National Conference On Undergraduate Research (NCUR) (2009), University of Wisconsin La-Crosse, La-Crosse, Wisconsin

Table(I): Results of iris coding test [2]

Database	Number of Data	Number of Failure	Time (second)	Average (second)
CASIA 1.0	756	0	553.15	0.7317
MMU 1.0	450	2	301.96	0.6725

Table(II): Iris identification result [2]

Database	Number of Query	Number of persons identified	Number of person correctly identified
CASIA 1.0	432	108	91
MMU 1 Left	89	45	35
MMU 1 Right	89	45	39

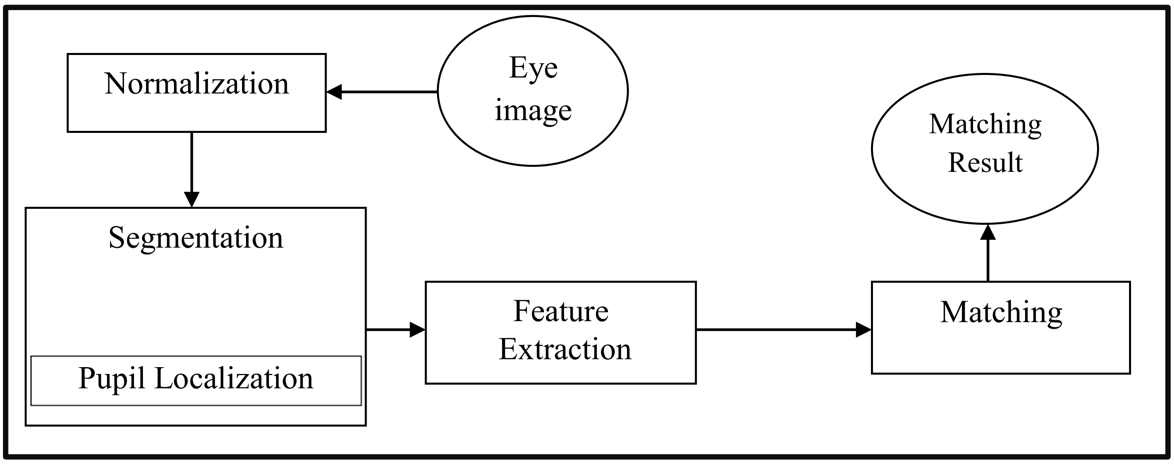

Fig. (1): Iiris recognition stages [2]

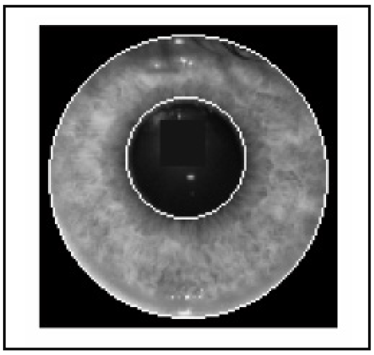

Fig. (4): Localized Iris [3]

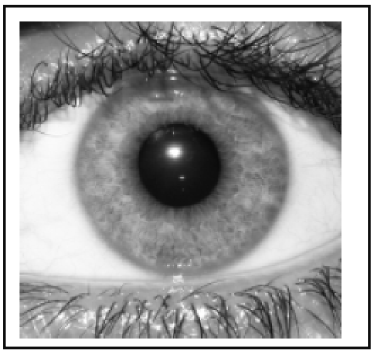

Fig. (4): Original image [3]

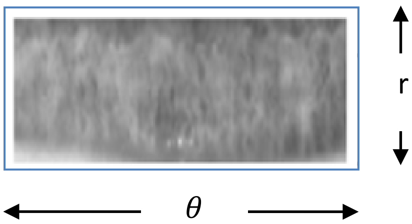

Fig. (4): Iris isolated image [3]

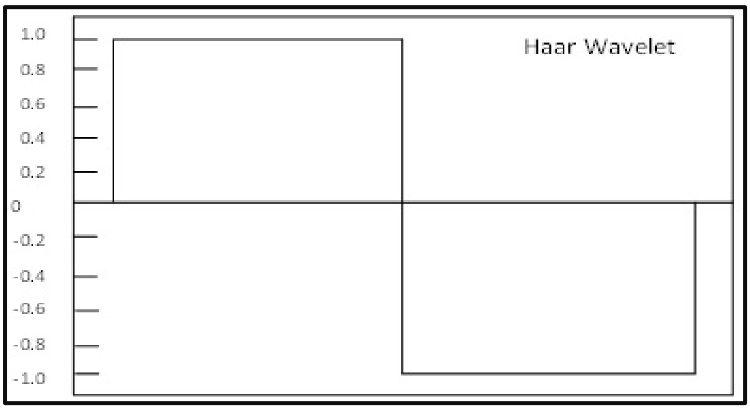

Fig. (5): The Haar wavelet [3]

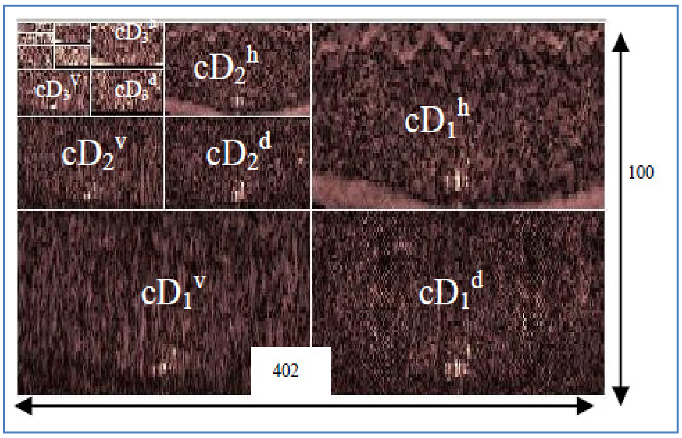

Fig. (6): Conceptual diagram for organizing a feature vector [3]

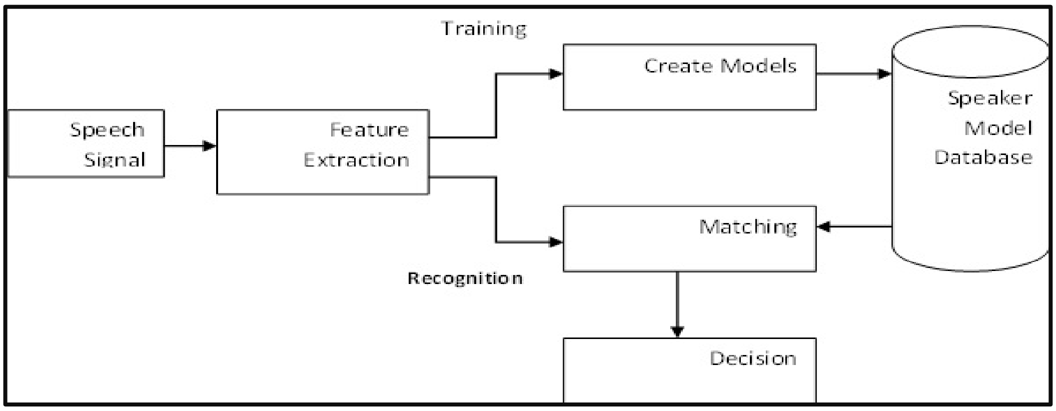

Fig. (7): Voice recognition system structure [7]

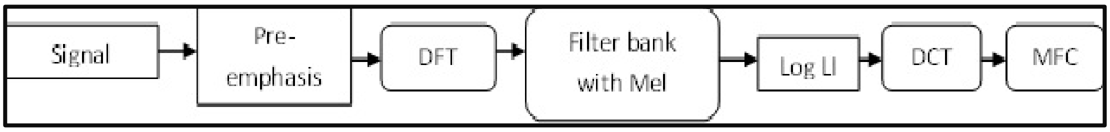

Fig. (8): The extraction of Mel Frequency Cepstral Coefficients (MFCC) parameters [7]

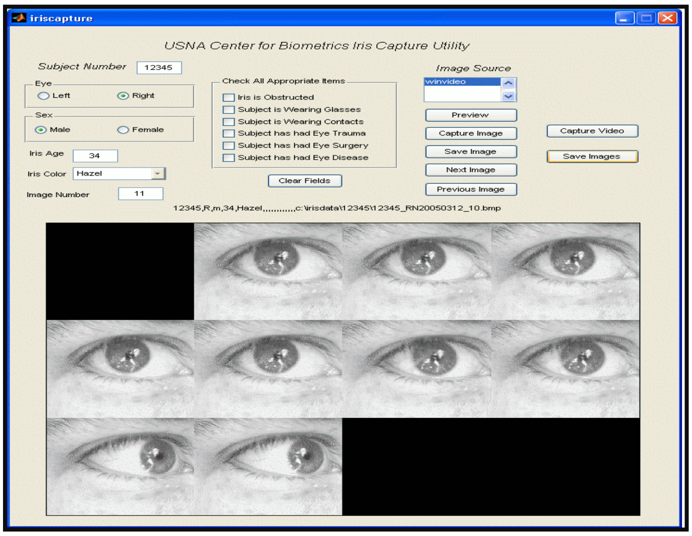

Fig. (9): Iris capture layout [13]

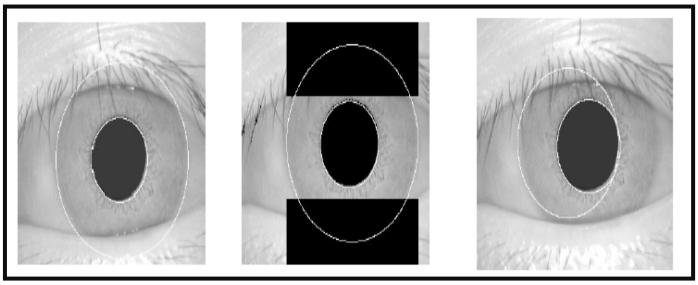

Fig. (10): Visual examples of segmentation process [16]

Fig. (11): Iris image after normalization [16]

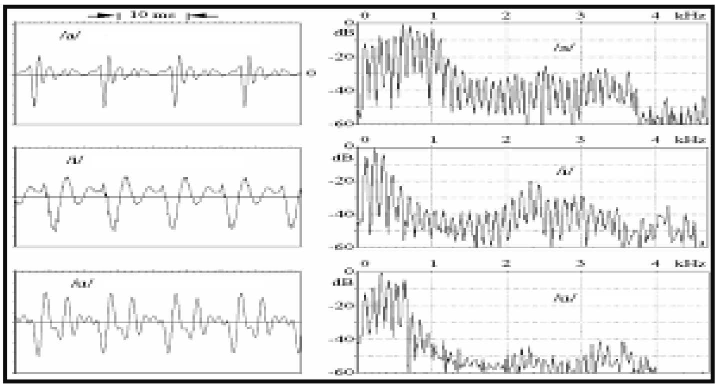

Fig. (12): The time-and frequency-domain presentation of vowels [16]

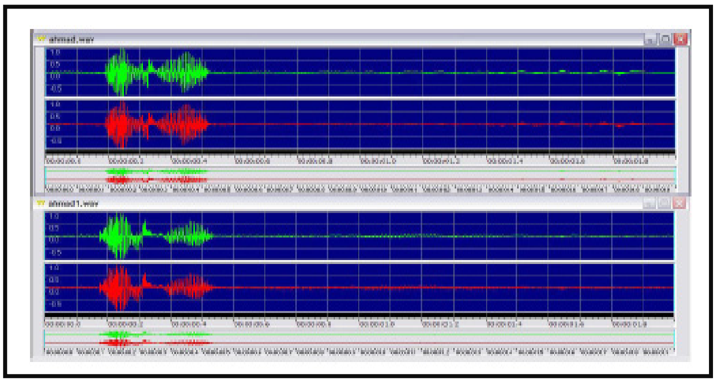

Fig. (13): Same Person Voice Patterns [9]

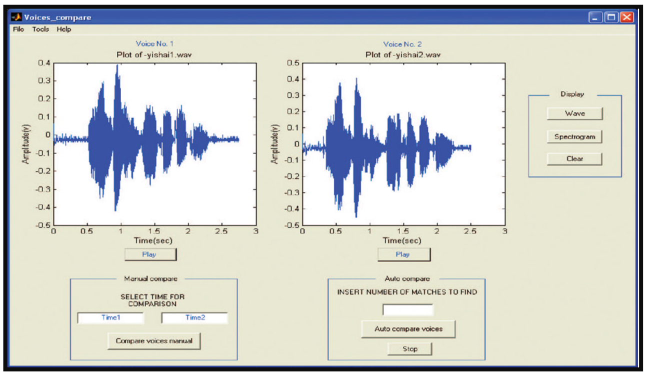

Fig. (14): Same Person Voice Patterns by using MATLAB [9]

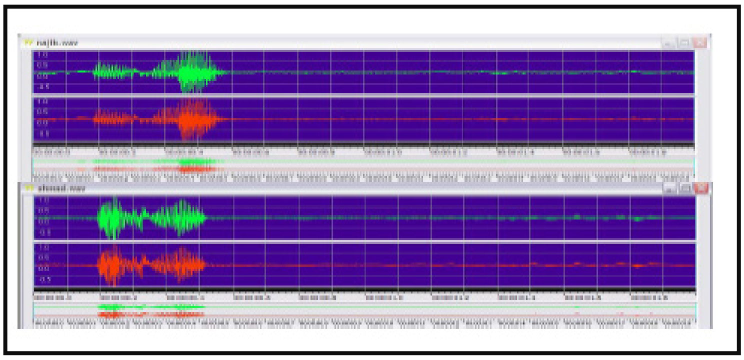

Fig. (15): Different Male Voice Patterns [9]

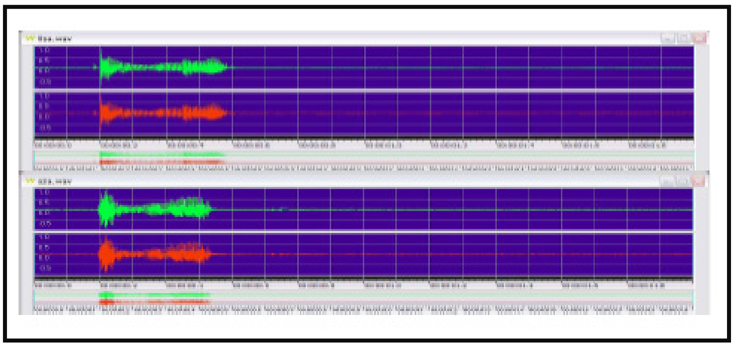

Fig. (16): Different Female Voice Patterns [9]

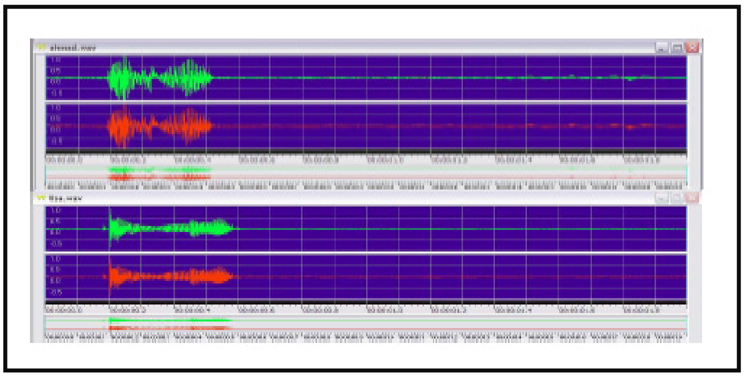

Fig. (17): Voice Patterns from a Male and Female [9]

التقنيات الامنية لتميز بين حدقية العين والصوت

أياد رفعت رؤوف
مركز الحاسبة الالكترونية / كلية الادارة والاقتصاد / جامعة بغداد

استلم البحث في 16 شباط 2012 ، قبل البحث في 12 ايلول 2012

الخلاصة

تقنيات التحقق من الهوية يتم استخدامها بشكل كبير نظرا لأمنها المحسن الذي يقلل من حالات الغش والسرقة. التقنيات البيومترية تستخدم مزايا مادية في تحديد هوية الأفراد. التقنيات البيومترية الأكثر شيوعا هي: القزحية ، والصوت، والبصمات، وكتابة اليد ،وبصمة اليد. في هذه الورقة، شرحت وقورنت تقنيتي تحقق من الهوية وهي تقنية تمييز القزحية، وتقنية تمييز الصوت، تفنية تمييز القزحية تميز الأفراد من خلال تحليل الأنماط الرئيسة في بنية قزحية بينما تقنية تمييز الصوت تميز الأفراد اعتمادا على خصائص الصوت الفريدة أو كما تسمى بصمة الصوت. نتائج المقارنة اشارت ان معدل دقة هذه التقنيتين هي 99.83% لتقنية تمميز القزحية، و 98% لتقنية تمييز الصوت.ومن ثم، تقنية تمييز القزحية توفر دقة وامنية اكثر من تقنية تمييز الصوت.

الكلمات المفتاحية : التكنولوجيا الحيوية، التعرف على القزحية، التعرف على الصوت، الافراط في العينات، ناقل الميزات، هرم غاوسي، موجات هار، تحويل فورييه المنفصل، معاملات التردد، نموذج غاوسي الخليط، نموذج ماركوف الخفي، النموذج الاساسي العالمي، قاعدة البيانات CASIA، قاعدة البيانات MMU.