Hybrid Biometric System Using Iris and Speaker Recognition

Gökçen ÇETİNEL *[1], Llukman ÇERKEZİ [1], Doğukan UZUNER[1], Barış YAZAR[1]

Abstract: In this study, a hybrid security system is proposed. The proposed system is composed of two subsystems namely iris recognition system (IRS) and speaker recognition system (SRS). Pre-processing, feature extraction and feature matching are the main steps of these systems. In IRS subsystem, Gaussian filter, Canny edge detector, Hough transform, and histogram equalization is performed for pre-processing, respectively. After that, by applying 4-level Discrete Wavelet Transform (DWT) to pure iris image, the iris image is decomposed into four sub-bands (LL4, LH4, HL4 and HH4). In order to extract the feature vector from iris pattern, the LH4, HL4 and HH4 sub-bands (matrices) are merged into one matrix. Finally the matrix is transformed in vector to obtain the feature vector of iris image. For SRS subsystem, the pre-processing step includes spectral arrangement, silence part removing and band limitation operations. After pre-processing, frame blocking and windowing are applied to the long-term speech samples and then Fast Fourier Transform (FFT) is performed for the each short-term speech segments (frames). Finally, the Mel Frequency Cepstral Coefficients (MFCC) technique is performed in order to obtain feature vector of the speech. The feature matching step of both IRS and SRS is implemented with Dynamic Time Warping (DTW) which is an efficient algorithm to measure the distance between two vectors. According to the DTW results, the false acceptance rate (FAR) is zero and false rejecting rate (FRR) is about 4 % for the proposed hybrid system.

Keywords: Biometric systems, iris recognition, speaker recognition, DTW, security.

1. Introduction

Nowadays, with the emerging of information technology the security of digital data has gained a great importance. One way to protect the digital data from unauthorized persons is to use identification and access control mechanisms. In order to provide these mechanisms one emerging technology that becomes more widespread is biometrics.

Biometric systems are designed based on people's physiological or behavioural characteristics. These systems can be defined as a pattern-recognition system which tries to recognize a person based on feature vector derived from her/his specific physiological or behavioural characteristics.

The physiological characteristics include fingerprint, face, iris, DNA, hand geometry whether the behavioural characteristics include gait, voice. Depending on application, a biometric system can operate for two main purposes: verification and identification. Verification refers the case where the user desires to be recognized by using an identity such as personal identification number (PIN), login name, smart card; and the system makes one-to-one comparison to determine whether this claim is true or not. The main goal is to prevent multiple people using the same identity. Identification refers the case where the system makes one-to-many comparison of user to determine who is exactly the user. The main goal is to prevent a single person from using multiple identities. One of the key advantages of biometrics over the traditional methods such as knowledge based systems is that biometrics cannot be lost or forgotten; they are difficult for attackers to forge and for user to repudiate [6].

Among biometric systems, iris recognition system is very popular due to significant reasons such as, uniqueness, non-invasiveness and high stability. Iris is part of human eye which lies between pupil and sclera. Iris cannot be copied so iris recognition system is the most reliable system in comparison with other biometric systems [2].

In recent years, speaker recognition also gained increased significance. Searching or controlling based on speaker identity is a growing interest in today's technologies. Speaker recognition is commonly used in the applications of personal authentication, national security and general forensics. In this study, speech analysis and decision process is performed by computer analysis automatically. The main goal of the proposed speaker recognition system is to increase the security of entire hybrid system [5].

Speaker recognition consists of two main tasks: identification and verification. In speaker identification we intend to identify a speaker from a closed or open set of speakers. For speaker verification, the objective is to verify the claim of an unknown speaker is true or not. This verification is performed by comparing the speech samples of claimer and other speaker's speech samples available in database. The characteristic of each speaker voice is typical and secret in her/his voice box. So, speaker recognition systems are the reliable biometric systems, too.

In this study, by connecting iris and speaker recognition systems in sequence, a hybrid high security biometric system is obtained. Both IRS and SRS are designed by applying pre-processing, feature extracting and feature matching (decision) steps. Used techniques in order to achieve these steps are explained in the corresponding sections. By performing computer simulations, it is shown that the proposed hybrid biometric system (HBS) is more robust against to the intruders than separate biometric system. To evaluate the performance of the system FAR and FRR values are also calculated in the simulations. According to the results, we can say that the proposed systems ensure security requirements.

The paper is organized as follows. In Section 2, IRS and SRS are explained with all steps, respectively. A brief information about used techniques in these steps are also given in Section 2. HBS which consist of IRS and SRS is introduced in Section 3. To evaluate the performance of the system, simulation results are

---

[1] *Electrical and Electronics Engineering Department, Engineering Faculty, Sakarya University, Sakarya/TURKEY*
* Corresponding Author: Email: gcetinel@sakarya.edu.tr

given in Section 4. Finally, discussion and conclusion part is stated in Section 5.

2. Subsystems

In this section, two subsystems of the HBS are explained in details. Pre-processing, feature extraction and feature matching are the main phases of both subsystems.

2.1. Iris Recognition System (IRS)

Pre-processing: In pre-processing, eye image is transformed into an appropriate model to extract the distinctive features for recognition. Iris localization and normalization are two basic processes in pre-processing. In our study, before localization and normalization 2-D Gaussian filter is applied for noise reduction. Then, the edges are detected with Canny edge detector.

Iris localization is crucial for iris recognition systems. As we see, there are two contours in an eye. The first one is inner contour separating the iris and pupil; second one is the outer contour separating the iris and sclera. By using Circular Hough Transform (CHT), we obtain the inner and outer boundaries of iris. In other words, we estimate the centre and radius of contours.

Iris normalization is an operation converting iris region from Cartesian coordinates to polar coordinates. Thus, circular image turns into a rectangular form and now is suitable for feature extraction. As a last step, histogram equalization is used for enhancing the contrast of iris pattern. In Figure 1, pre-processing is illustrated.

Feature Extraction: In order to use iris signature for recognition systems, the representing of iris in an appropriate way is very important. To extract the specific features from iris, Discrete Wavelet Transform (DWT) is used. DWT is an efficient tool, which transforms the signal from time domain into frequency domain without losing time information.

In general, by performing DWT to 1-D signal, the original signal is divided in low and high frequency parts. This process is called as decomposition and it is provided through special analysis filters. The outputs of filter are referred as DWT coefficients and the original signal can be reconstructed from these coefficients by using appropriate synthesis filters (inverse DWT, IDWT). For 2-D signal case, after performing DWT, image will be divided into four sub-bands corresponding to low frequency (LL_1), middle-frequency (LH_1, HL_1) and high frequency (HH_1) components, respectively. DWT can be applied to the LL_1 sub-band which concentrates the maximum energy of the signal, one more times until we get the desired level of decomposition. The middle and high frequency (LH_1, HL_1, and HH_1) sub-bands represent the details of image such as edges, outline and texture.

In our study, we used 4-level DWT in order to extract the correct features from iris texture. By performing 4-level DWT we get LL_4, LH_4, HL_4 and HH_4 sub-bands. Since the distinctive features of iris are hidden in the high- frequency components, we constitute the feature matrix by merging LH_4, HL_4, and HH_4 sub-bands in one matrix. Finally the matrix is transformed into a vector to obtain the feature vector of iris image.

2.2. Speaker Recognition System

In SRS, at the beginning we record the speech of the users under same conditions. Since the SRS is text dependent recognition system, the users say the same words for database. The block diagram of the proposed SRS system is given in Figure 2. The steps of SRS can be explained as follows.

Pre-processing: In order to process a digital signal, first we must sample the analog signal by Nyquist frequency rate. Then, pre-processing step including spectral arrangement, silence part removing and band limitation operations is applied to the digitized signal.

For spectral arrangement, to enhance the high-frequency energy of the signal, we use a high-pass finite impulse response pre-emphasize filter. The transfer function of this filter is given as follows,

$$H(z) = 1 - \mu z^{-1}, 0.95 \leq \mu \leq 0.97$$
(1)

In silence part removing differentiation operation is performed on speech samples from left-to-right and right-to-left to eliminate the non-variant (silence) parts. Then accordance with the human audio system, frequency components outside of 1KHz-5KHz band is stopped. This can be referred as band limitation.
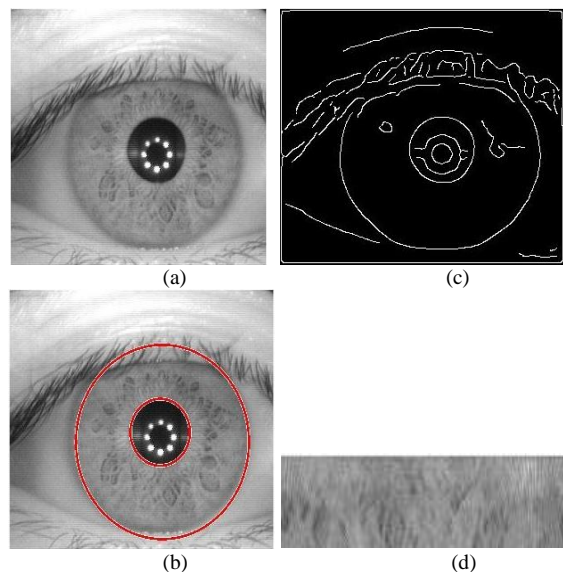
Figure 1. Pre-processing operations: (a) Original iris image, (b) inner and outer boundaries of iris (c) Canny edge detector applied to iris image, (d) iris template in polar form.
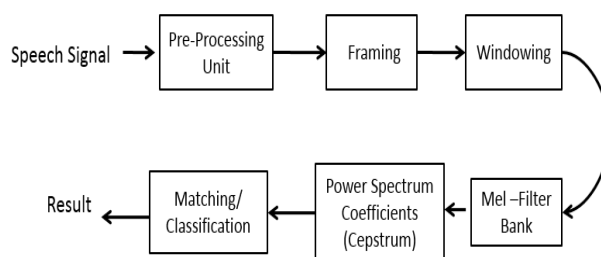
Figure 2. Block diagram of SRS.

Feature extraction: In general speaker recognition system, the critical step is feature extraction step. To achieve high recognition performance, it is known that the features should be extracted form short-time (20-25ms) speech segments. One of the most commonly used short-time features are Mel Frequency Cepstral Coefficients (MFCC). To calculate MFCC values, speech signals are divided into short overlapping frames. Then each frame is multiplied by a window function. In our study, we use Hamming window given as,

$$w(n) = 0.56 - 0.46 \cos\left(\frac{2\pi n}{N-1}\right)$$
(2)

where N is the length of the window. Then with Fast Fourier Transform (FFT), Fourier power spectrum is obtained. In the next step, the logarithm of the spectrum is calculated a nonlinearly

spaced triangular band-pass Mel-space filter-bank (illustrated in Figure 3) analysis is performed. With filter-bank analysis, spectrum energy for each channel i.e. filter-bank energy coefficients are obtained. By performing Discrete Cosine Transform (DCT) on filter-bank energy coefficients we attain to MFCC.
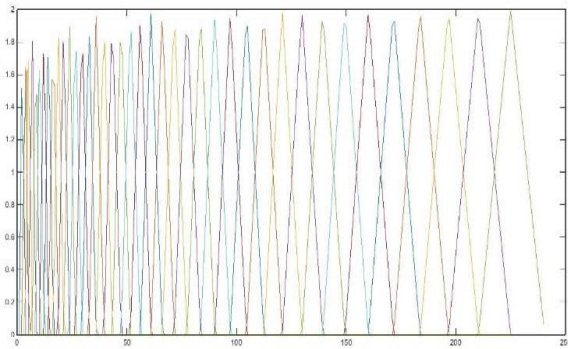

Figure 3. Mel Filter Bank

3. Proposed System

In this section, we explain how the hybrid system is designed. In addition, the last step of both IRS and SRS referred as feature matching step is discussed.

3.1. Hybrid Biometric System (HBS)

The HBS is designed by connecting IRS and SRS subsystems as demonstrated in Figure 4. The HBS works as follows; at first the user tries to pass through IRS. If IRS denies his/her access, the system will automatically reject this attempt. On the other hand, if the user passes through IRS successfully, SRS is put into use. With the acceptance of SRS, the user will be able to enter.
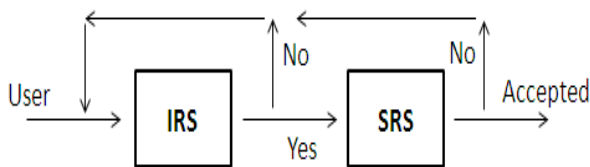

Figure 4. Hybrid Biometric System (HBS)

3.2. Feature Matching

In this study, we use Dynamic Time Warping (DTW) technique to implement feature matching process. DTW is applied to feature vector which is constituted by combining iris image feature vector and MFCC. This technique is very efficient to find optimal distance between two vectors that may differ both in phase and length. The first step in algorithm is to determine local cost measure $c(x, y)$ between each element of the X and Y vectors by calculating Manhattan distance. If $c(x, y)$ is small, the similarity between X and Y is high and vice versa. In the second step, the cost matrix is obtained by evaluating the local cost for each pair of elements in X and Y feature vectors [7].

Algorithm tries to find a warping path $p= (p_1,...,p_u)$ which satisfies the following three conditions:

i)   Boundary condition: $p_1=(1,1)$ and $p_u=(N,M)$,
ii)   Monotonicity condition: $n_1<n_2<....<n_u$ and
         $m_1<m_2<.....<m_u$
iii)   Step size condition: $p_{u+1} - p_u \in \{(1,0),(0,1),(1,1)\}$ for
         $u \in [1:=L-1]$

The total cost $c_p(X, Y)$ of a warping path between X and Y with respect to the local measure c is defined by

$$c_p(X,Y) := \sum_{u=1}^{U} c(x_{n_u}, y_{m_u})$$
(3)

An optimal warping path between X and Y is a warping path p* having minimal total cost among all possible warping paths. The DTW distance DTW(X, Y) between X and Y is then defined as total cost p*:

DTW(X, Y):= c_{p*}(X, Y) = min {c_p(X, Y) | p is an (N, M) - warping   path}
(4)

DTW algorithm is widely used in speech processing applications due to the nature of speech waves. As far as we know, DTW algorithm is not applied for iris recognition systems. In our study, we used DTW both IRS and SRS and we obtained very promising results.

4. Experimental Results

In this section, the simulation results are given for IRS, SRS and HBS systems respectively. Simulations are performed by using MATLAB computer program. The performance of system is evaluated by calculating two rates defined as follows:

i)   False Acceptance Rate (FAR): the probability of identifying an intruder as an enrolled user,
ii)   False Rejection Rate (FRR): the probability of rejecting an enrolled user as if he was an intruder.

Since the aim of HBS system is to ensure high security especially in a house, the database of our system is modelled for 10 people. For each user, 10 images of right iris and 10 sample voices are recorded. Two iris images and two sample voices are used for training among them. The iris images are provided by Casia- Iris V4- Interval-R database. The size of iris images is 280x320. After applying pre-processing and feature extraction processes to the iris images we get a feature vector of size 1x84 for each image.

As discussed before, since our SRS sub-system is text dependent, the sentence "open the door" in Turkish is used as a password. The recording time is two seconds and the sampling frequency is 16 kHz. After pre-processing and feature extraction we obtain a feature vector of size 1x40. Finally, we form feature vector for whole system by combining the feature vectors of IRS and SRS subsystems. The resulting vector's size is 1x40.

After applying DTW algorithm to obtained feature vector we calculate the FAR and FRR values. These values are given in Table 1 for two subsystems and hybrid system. As can be seen from Table 1, the FAR vales for these systems are zero. On the other hand FRR values are different in each system. By connecting IRS and SRS sub-systems we increase the security with HBS. A user must pass through both sub-systems in order to enter so it is almost impossible to cheat the system.

The FRR value of HBS is evaluated by using total probability theorem as given in the following equation

FRR(HBS) = FRR(IRS) + FRR^C(IRS)* FRR(SRS| IRS= accepted)   =   0.0125   +   0.9875x0.025=   0.037
(5)

where FRR is false rejecting rate for HBS and FRR^C(IRS) is the probability of true decision in IRS. It can be seen from Equation (5) that the FRR (HBS) value is smaller than FRR(SRS) value. Thus, by using HBS we can provide small FRR value and increase the robustness of system against fake biometric attack.

Table 1. FAR and FRR values for IRS, SRS and HBS systems

	FAR	FRR

IRS	0	0.0125	
SRS	0	0.05	
SRS | IRS accepted	0	0.025	

5. Conclusion

In this study a novel Hybrid Biometric System (HBS) is proposed. First, the Iris Recognition System (IRS) and Speaker Recognition System (SRS) are designed and tested separately. The performances of sub-systems are evaluated by calculating FAR and FRR rates. Then, we connected two sub-systems to obtain the HBS system. The purpose of proposed HBS is to ensure high reliability. By means of the proposed HBS, to cheat system by spoof attacks is almost impossible.

Acknowledgements

References

[1] *L.V. Birgale, M. Kokare, "*Iris Recognition Using Discrete Wavelet Transform*",* International Conference on Digital Image Processing, 7-9 March 2009, Bangkok.

[2] J.G. Daugman," High Confidence Visual Recognition of Persons by a Test of analysis of Statistical Independence*",* IEEE Trans. On Pattern Analysis and Machine Learning, Vol. 15, Number 11, 1993, pp. 1148-1161.

[3] Zh. Lin, B. Lu, "Iris Recognition Method Based on the Imaginary Coefficients of Morlet Wavelet Transform ",2010 Seventh International Conference on Fuzzy Systems and Knowledge Discovery (FSDK), 10-12 August 2010, Yantai, Shandong.

[4] S.D. Thepade, P. Bidwai, "Iris Recognition Using Fractional Coefficients of Transforms, Wavelet Transforms and Hybrid Wavelet Transform", 2013 International Conference on Control Computing Communication & Materials (ICCCCM), 3-4 August 2013, Allahabad.

[5] J. Hansen, T. Hasan, "Speaker Recognition by Machines and Humans: A Tutorial Review", IEEE Signal Processing Magazine, Vol.32, Number 6, 2015, pp. 74-99.

[6] S. Prabhakar, S. Pankanti, A. Jain, "Biometric Recognition: Security and Privacy Concerns", IEEE Security & Privacy, Published by IEEE Computer Society, 8 April, 2003

[7] M. Müller, "Information Retrieval for Music and Motion", Springer, 2007.

[8] A. Sukhwal, M. Kumar, "Comparative Study of Different Classifiers Based Speaker Recognition System Using Modified MFCC for Noisy Environment" 2015 International Conference on Green Computing and Internet of Things (ICGCIoT), 8-10 October 2015, Noida.

[9] K. S. Ahmad, A. S. Thosar, J. H. Nirmal, V. S. Pende, "A Unique Approach in Text Independent Speaker Recognition Using MFCC feature Sets and Probabilistic Neural Network" 2015 Eighth International Conference on Advances in Pattern Recognition (ICAPR), 4-7 January 2015, Kolkata.

[10] H. B. Kekre, V. A. Bharadi, A. R. Sawant, O. Kadam, P. Lanke, R .Lodhija, *"*Speaker Recognition Using Vector Quantization by MFCC and KMCG Clustering Algorithm", 2012 International Conference on Communication, Information & Computing Technology (ICCICT), Oct. 19-20, 2012, Mumbai, India.

[11] J. Martinez, H. Perez, E. Escamilla, M. M. Suzuki, "Speaker Recognition Using Mel Frequency Cepstral Coefficients (MFCC) and Vector Quantization (VQ) Techniques", 22nd International Conference on Electrical Communications and Computers (CONIELECOMP), 27-29 February, 2012, Cholula, Puebla.