An Efficient Biocryptosystem Based on the Iris Biometrics

Ali Shojaee Bakhtiari, Ali Asghar Beheshti Shirazi, and Babak Zamanlooy

Department of Electrical Engineering
Iran University of Science and Technology
Narmak, 16846, Tehran, Iran
{Ali_Shojaeebakhtiari, Babak_Zamanlooe}@ee.iust.ac.ir,
Abeheshti@iust.ac.ir

Abstract. A new and efficient method for combining iris biometrics with custom cryptographic schemes to obtain an efficient biocryptosystem is proposed in this paper. Though the method structure is basically derived from a previously described biocryptosystem scheme, the introduction of new image processing methods alongside with efficient utilization of traditional methods show promising developments compared with the previous biocryposystem especially in the field of generating longer cryptographic key strings while keeping the system quality**.**

Keywords: Authentication system, biocryptosystem, error correction coding, iris segmentation.

1 Introduction

Traditionally the research fields of biometric and cryptography authentication have been considered as two distinct areas with different operational backgrounds. Whereas Biometric authentication systems are generally based on fuzzy comparison of the claimed data with a previously stored reference data and adequate similarity between the claiming and the original data results in positive authentication, the logic behind the cryptography based authentication on the other hand is absolutely exact. Therefore only the exact matching between the claiming and reference data, authentication key in this aspect, results in positive authentication. The mentioned deep difference between the natures of the two eras one relying on fuzzy and the other on exact comparisons, has made the effective combination of the two eras facing various odds. The various obstacles facing the combined method and the proposed counteractions to overcome each of the obstacles were first analyzed in [1]. Overall three different challenges facing the combined structure have been classified and different approaches have been taken to overcome each of them.

The first work to take into considers proposing an effective and concrete implementation of a biocryptosystem on this basis was proposed by Hao, Anderson and Daugman [1]. The proposed system acts considerably well with key lengths of up to 140 bit long however as shown in [1] the efficiency of the system starts to deteriorate once the chosen key length becomes longer.

D. Mery and L. Rueda (Eds.): PSIVT 2007, LNCS 4872, pp. 334 – 345, 2007.

Therefore in order to achieve longer key length, which is the requirement of the ever growing need for more security, in this paper we propose a new biocryptosystem based on the iris biometrics. Considering the successful implementation results of the system proposed in [1], the basis of our system is derived from the system, which from now shall be called the Hao's system, However in order to reach the desired objectives new code extraction and generation methods are designed and utilized in our designed system.

The paper is organized as follows in the 2nd section the Hao's system [1] is described and the strength and weaknesses of the method from the view of the authors are analyzed, afterwards in the 3rd section the proposed iris image processing and the code generation methods are introduced and analyzed, in the 4th section the proposed biocryptosystem is introduced and its different modules are described and finally in the 5th section the results of the designed system are analyzed and compared with other systems.

2 Hao's System

Fig.1 shows this basic two-factor scheme of the Hao's system. In this system the key depends on a combination of a biometric and a token, in which the information required for error correction of the received code, is stored. In order to make the system resilient against the three mentioned drawbacks in the introduction of this paper the designers of the Hao's system have considered a series of provisions in the system. In order to overcome the contradicting natures of the fuzziness of biometrics and exactitude of cryptography a set of error correction coding, which effectively detects and corrects the errors caused by the presence of noise, is added to the design.

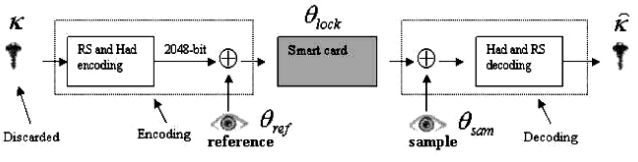

Fig. 1. Two factor scheme for biometric key generation [1]

The Hao's system works as follows: In the first step a random key k is generated in the random key generation module. Afterwards the generated key is entered to the coding unit which encodes the random key string in to a 2048 bit string, adapted to the Daugman's iris code length [7]. The encoder block consists of two blocks of random error correction coding and burst error correction coding units.

During the decoding process the sample iris code is delivered to the system to unlock the locked iris code. The obtained pseudo iris code (not necessarily equal to the original one) is next entered to the decoding module. The decoding module applies the two random error correction and burst error correction decoding to regenerate the random key string $\hat{k}$. Should the claiming iris and the reference iris match each other to the extend that the minute differences between the two codes are

correctable by the decoding module, the obtained key $\hat{k}$ would be equal to the original key string k. A simple comparison between the hash value of $\hat{k}$ and the stored hash value of k reveals if the authentication is positive or not. The entire decoding process is formulized as below:

$$< \theta_{sam}, \tau > \Rightarrow \hat{k} \tag{1}$$

Table 1 shows the implementation result of the Hao's method for different keylenghts.

Table 1. Performance of the Hao's system for $k_{hadamard} = 6$ [1]

RS Corrected blocks	Length of biocrypto key	FRR% (False reject rate)	FAR% (False accept rate)
0	224	12.22	0
1	210	6.5	0
2	196	3.65	0
3	182	2.06	0
4	168	1.26	0
5	154	0.79	0
6	140	0.47	0

The Hao's system shows great improvement compared with the previously proposed biocryptosystems notably in the field of FRR%. Also the obtainable key length in the system has improved greatly compared with the previous works.

However from the authors' point of view with some reconsiderations which are the basic of this paper, more proper results are obtainable. The following section introduces the reconsiderations proposed by the authors.

3 The Proposed Image Processing and Code Generation

In order to fulfill the needs of the proposed system as mentioned in section (2) a group of image processing methods are utilized for our own purposes. A new iris segmentation method designed for the purpose of this paper is introduced in this section. The aim of the segmentation method is to choose the areas most suitable for biocryptosystem purposes.

3.1 Image Processing and Feature Enhancement

After the iris image has been captured and the initial preprocessing is done on the image, classic approach is to apply an efficient segmentation method such as the Hough transform or the Canny operator [11] on the image to separate the iris data from the rest of the image. The segmentation method developed by the authors for

selecting the desired regions of the iris automatically omits the need for this step. However the segmentation method which shall be explained later in this section and which shall be called local entropy method from now initially requires special features of the image to be enhanced in order to operate properly.

For two distinct reasons related to the objectives of this work the Phase Congruency based iris feature enhancement as described in [4], based in the work of Kovesi [5],[12] and Morrone [8] is chosen as the basis for the feature enhancement of this work. The first reason for selecting the method is the robustness of the method against the unwanted brightness and contrast changes in the received image, which acts as an important source in generating burst errors in the biocryptosystem. The other reason is the ability of the method to omit the unwanted random noise while simultaneously enhancing the required edge features of the image which is a unique feature of the phase based methods compared with the majority of spatial methods.

In order to support the claim of the efficiency of the phase congruency based feature enhancement a group of experiments on the ability of the method to recover the desired features of the image in the presence of different interferences was performed so that initially a group of the features in the unenhanced images were chosen and afterwards the unenhanced images were distorted by different kinds of distortions and again the selected features were sought in the images. In the next step a group of selected features

Table 2. Different distortions and related percent of recovered feature points

Distortion	Recovered feature points	
	no enhancement %	Phase congruency %
Addition of 2% Noise	50%	93.35%
Addition of 10% Noise	28.57%	38.71%
Rotation by 2 degrees	35.71%	61.62%
Rotation by 10 degrees	33.24%	46.86%
Contrast reduction by 10%	21.42%	95.2%
Contrast reduction by 50%	20.87%	92.98%
Contrast Increase by 10%	42.57%	93.35%
Contrast Increase by 50%	22.58%	52.39%

in the enhanced images were chosen and afterwards the enhanced images were distorted by different kinds of distortions and again the selected features were sought in the images. The results of which are presented in Table 2.

As can be seen from Table 2 the phase congruency method works actually well in the presence of deep contrast changes and is able to recover to majority of the assigned feature points correctly. Also the method shows some improvement in the feature extraction in the presence of image noise which considering the fact that the method is enhancing the edges of the image is quite noteworthy.

3.2 Image Segmentation and the Local Entropy Method

After the features of the image have been enhanced it is time for the segmentation of the iris image to derive the segments suitable for the biocryptosystem purposes. In order to choose the best regions for the paper purpose, the authors have proposed a method named the local entropy method. The method which is explained in this subsection is based on locating the high entropy regions of the surface of the iris for code generation purposes.

According to Shannon's 2nd theorem [2] if the event *i* occurs from a set of valid events, with the probability p_i the amount of uncertainty related to the event is equal to:

$$H_i = -\log_2(p_i)(bits/Symbol) \ . \tag{2}$$

And also the amount of the uncertainty that the source of the events generates is equal to:

$$H = -\sum(p_i \log_2(p_i))(bits) \ . \tag{3}$$

From equation 2 it can be seen that the highest amount of uncertainty from an information source is realized when the output symbols of the source are equally probable.

The idea behind local entropy method is to divide the processed image into separate regions and then to analyze each region separately as information source. The amount of entropy calculated for each region gives an overview about the level of correlation between individual blocks (bits) in the selected region.

A research by J.Daugman [3] has shown that in average the iris image has discrimination entropy of 3.2 bits per square millimeters which is indeed a suitably high value for identity recognition purposes and comparison based structures.

The process of deriving the local entropy of the image begins with the acquisition of the image, after this step the initial preprocessing is performed on the image to prepare the image for the main processing, after preprocessing, the extracted image is further processed for revealing its hidden features, After the features are revealed and enhanced, the local entropy segmentation process divides the obtained result into separate regions each containing a portion of the enhanced features, In this step, assuming each section as an information source, the entropy of each of the segments is separately calculated and finally the obtained entropy values are sorted to deliver the entropy function of the image. In the rest of the paper, unless stated, fig.2 is considered as the reference image.

Fig. 2. Reference image

Fig. 3 shows the result of calculating the entropy curve of the reference image for 100 segments.

After the entropy of the surface is calculated it is the necessity of the application that dictates how much of the iris surface is to be segmented for the best performance. Another question that arises here is the size of the segmentation blocks. Apparently the finer the blocks are chosen the better the entire surface is segmented, however making the segmentation too small also results in the selection of the regions where image noise is still present as the presence of noise acts as a source of entropy in the region.

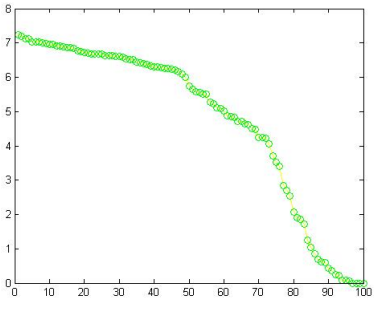

Fig. 3. Sorted Local entropy in bits/segment, for 100 local segments of fig.2

In biocryptosystem applications it is necessary not to let the phony regions to enter the process of key generation therefore experiment results show that the best practice in segmenting the iris for such purposes is to choose the highest entropy regions and to use large segments in order to prevent the noisy regions from masquerading themselves as the desirable regions. For comparison based purposes in which not every single bit, but the iris surface as a whole is important it is therefore a good practice to choose fine regions and also a larger portion of the entropy curve to segment the most of the iris surface. Table III shows the experimentally obtained results for best segmentation practices and percent of the local entropy chosen for different applications.

One point that is necessary to be mentioned here is the potential weakness of the method at the presence of eyelashes in the captured image. As the eyelashes generally occur in the image with a random pattern, this pattern generally leads to unwanted uncertainty inside the image and therefore an undesired source of entropy is generated

by the eyelashes. Moreover in designing biocryptosystems, the presence of eyelashes are modeled as channel burst errors [1] and to overcome their negative effects it is necessary to sacrifice a great deal of system capacity for covering the effect of burst noise. Therefore it is necessary to detect those areas of the image affected by the presence of eyelashes and put them in the blacklist. Various eyelash detection methods have been proposed in iris processing literature from which the method proposed by Kong and Zhang [6] is chosen by the authors for its efficiency and ease of implementation.

Table 3. Experimentally obtained results for best segmentation and surface selection

Application	No. of Segmentation blocks	Percent of sorted curve used. (From higher to lower entropy)
Biocryptosystem	100 (10 by 10)	10%
Comparison based purposes (in the presence of eyelashes in the image database)	900(30 by 30)	30%
Comparison based purposes (in the absence of eyelashes in the image database)	900(30 by 30)	50%

3.3 Iris Code Generation

After the local entropy segmented iris data is chosen it is time for the generation of the iris code from the sorted local entropy data. The process begins with the normalization of the obtained sorted local entropy data for obtaining a normalized data value between 0 and 1. Afterwards the data quantities are quantized to a group of assigned thresholds to limit the number of output bits. Depending on the thresholds chosen for the code generation a group of nonoverlapping bit strings are related to each of the thresholds, for example a 10 level thresholding requires the assignment of 4 bit strings to each quantized value. For the database consisting of images with the size 320×280 this results in 89600 bits of data. As it shall be mentioned in the next section for the system to work with the standard 2048 bit strings a set of provision should be considered for adoption. The detailed procedure for obtaining the iris-code by this method is presented in [9] and is briefly described in the next section.

4 The Biocryptosystem

In this section the proposed biocryptosystem is introduced and analyzed. As it has already been mentioned in the previous sections the basis of the design is from the biocryptosystem proposed by Hao, Anderson and Daugman in [1]. However in order to fulfill the objectives of this paper different modules of the system have been altered to fit our desires. Fig.4 shows the block diagram of the system proposed by the authors.

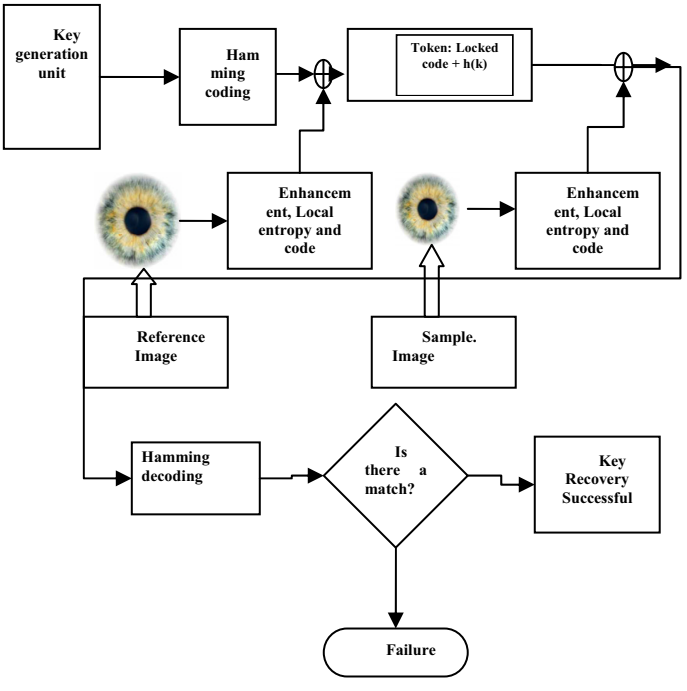

Fig. 4. Block diagram of the proposed biocryptosystem

As can be seen by comparison between the Hao's system and the presented system it can be seen that some modules of the Hao's system have been either omitted or altered in our proposed system. The first and the most important alteration towards the Hao's system is the replacing of the iris-code module in the Hao's with the local entropy code based code generation unit. The 2nd change is the omission of the burst error correction coding and decoding modules from the Hao's system. As is already mentioned and will be shown in the results section the omission of the burst error correction coding results in a free hand in choosing the arbitrary though limited longer key string lengths. Also in order to analyze the efficiency of the image processing part of the system the strong Hadamard code described in the Hao's system has been replaced by the simple Hamming code.

Apart from the structural changes the operation procedure of the proposed system is identical to the Hao's system. The same as the Hao's system the presented system initially generates a random key string which is passed through the bank coders to generate a pseudo iris-code string. This string is afterwards Xored by the iris-code to generate the locked iris-code. During the decoding phase the sample iris-code is again Xored with the locked iris-code stored in the token to reveal the supposed pseudo-iris code. Afterwards the supposed pseudo-iris code is delivered to the decoding module. If the difference between the supposed pseudo-iris code and the original pseudo-iris code is in the acceptable range of the decoding ability of the decoding bank the correct key string is extracted and a comparison between the hash value of the derived key and the stored hash value of the original key confirms the whole process.

It should be mentioned here that as the standard iris-code generation method has not been used in our method and therefore, we are not obliged to follow the traditional 2048 bit scheme. However in order to make comparison with the Hao's work the iris code generation module is designed to generate the standard 2048 bit iris-code length for fair comparison.

The procedure for generating the 2048 bit string is as follows. The CASIA database consists of images with the size of 320×280 Pixels each, or equivalently 89600 total pixels. The Local entropy method selects 10% of the total pixels which results in 8960 total pixels. In this work the threshold step is chosen to be 0.1 in 0 to 1 normalized space which contributes to 10 threshold levels and therefore 4 bits of data for each pixel which results in the total of 35840 bits of data. However considering the (15, 11) Hamming code used in the system, the system requires only $\frac{2048\times11}{15}$ or around 1502 bits of data. In order to reduce the available data to the needed number of bits a 6×4 averaging mask is applied on the sorted local entropy data. This operation results in a 1493 bit string. The difference between the 1502 and the 1493 bit strings is filled with a random generated bit string or a simple null string for simulation purposes. The obtained string is entered to the encoding block and the 2048 bit string is obtained.

Excluding the burst error correction module from the system structure immediately arises the question of how to deal with the present burst errors. As will be shown in the next section the image processing methods applied in addition to the local entropy method results in a great reduction in the negative effects of typical causes of the burst noise.

5 Experimental Results

In order to analyze the efficiency of the proposed system a series of experiments has been done on the system. The CASIA database [10] has been used as the image database. The database consists of 108 identical iris images each taken in 7 different states, which in total has 756 images.

The result of applying the system on the image database for different chosen key lengths is brought in Table 4.

In designing the system 3 new alterations are proposed for improving the operation of the Hao's system. As has already been mentioned in the paper, the first alteration is

the replacing of the Hao's image processing method with the phase based feature enhancement method, the 2nd is the introduction of the local entropy method combined with eyelash detection methods for choosing the suitable regions for biocrypto purposes and the 3rd alteration is the replacement of the heavy loading error correction blocks of the Hao's system with simple linear error correction codes to reduce the overload imposed to the system by the coders.

Table 4[1]**.** False reject rate of the system for different chosen key lengths

Length of biocrypto key	FRR% before applying error correction coding(Hamming)	FRR% after applying error correction coding(Hamming)
224	16.6	1.6
210	23.6	1.3
196	29.3	2
182	31.3	1.3
168	24	2.6
154	22	1.3
140	23	1.6

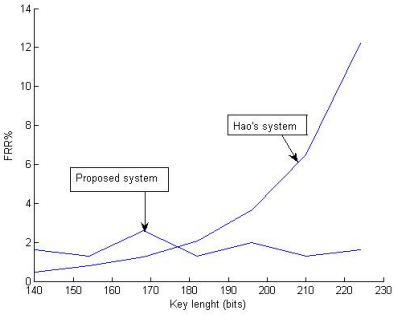

Fig. 5. Comparison of the FRR% of the Hao's system and the proposed system

Analyzing Table 4 and Fig. 5 and comparison of the results of the Table 4 with the results of Table 1 results in the following conclusions:

To analyze the efficiency of the image processing methods and the burst prevention methods proposed in the system a fair point in the Hao's system must be chosen. It's because in the system the burst error correction module is practically omitted. The only key length of the Hao's system in which the effect of burst correction coding is not present is the 224 bit key length. Comparison between the

[1] It must be noted and emphasized here that the results brought in Table. 4 are strictly for qualitative comparison, and as the authors of [1] have refused to share the original database to the public, no quantitative comparison between the results in this paper and reference [1] is logically valid.

results in Table V and Table II shows that the system error in the key length of 224 is in the scale of 2% compared with the error level of 12% for the Hao's system. As the main cause of the error in the Hao's system is because of the presence of the burst error this result shows that the system has been able to overcome the negative effect of unwanted burst noise.

Another supporting evidence for showing the strength of the method against the negative effects of the burst noise can also be deduced from Table V. As it can be seen from the 2nd column of the Table V before applying the error correction coding block the FRR% of the system is in the scale of about 25%, however after the introduction of the error correction coding module which simply consists of a simple Hamming error correction coding, The error rate is reduced to the scale of about 2%. Considering the fact the Hamming code is not at all capable of detecting burst errors logically reveals that the burst error noise should not have been present in the first. Otherwise the negative effects were reflected in the system output.

As it can be seen from the Table V the FRR of the system for different key length is nearly constant and in the scale of about 2%. However the Hao's system shows a FRR % ranging from 0.47% for 140 bit length key to 12.22% for 224 bit length chosen key. Therefore it can be logically deduced that the Hao's system acts more superior to the proposed system in shorter key lengths but as the chosen key length begins to increase the proposed system shows superiority compared with the Hao's system

From the complexity point of view the system is superior to the Hao's system in the encoding process because of the replacement of the relatively computationally complicated Hadamard & Reed-Solomon Codes with the simple Hamming and permutational coding.

6 Conclusion

In this paper the possibility of the implementation of an effective biocryptosystem was analyzed and confirmed. Comparison between the system and the previously implemented systems with different structural backgrounds shows promising results. The system also shows a strong background for obtaining longer key lengths required for higher security applications.

Acknowledgment

This paper is supported by Iran telecommunication research center under the contract agreement number T/500/150/50

References

1. Hao, F., Anderson, R., Daugman, J.: Combining crypto with biometrics effectively. IEEE Transactions on Computers 55, 1081–1088 (2006)
2. Shannon, C.E.: A Mathematical Theory of Communication. Bell System Technical Journal 27, 379–423, 623–656 (1948)
3. Chen, W.: Linear Networks and Systems, Belmont, CA: Wadsworth, USA (1993)

4. Daugman, J.: How Iris Recognition Works. IEEE Transactions on Circuits and Systems for Video Technology 14, 21–30 (2004)
5. Shojaee Bakhtiari, A., Beheshti, A.-A, Zamanlooy, B.: Phase congruency based image enhancement method and its application in enhancing iris feature extraction. In: Proceedings of Iranian conference on electrical engineering (2007)
6. Kovesi, P.: Phase Congruency Detects Corners and Edges. School of Computer Science & Software Engineering, The University of Western Australia (2003)
7. Kong, W.-K, Zhang, D.: Accurate iris segmentation based on novel reflection and eyelash detection model. In: Proceedings of International Symposium on Intelligent Multimedia, Video and Speech processing (2001)
8. Daugman, J.: Biometric personal identification system based on iris analysis. US Patents 291560 (1994)
9. Morrone, M.C., Owens, R.A.: Feature detection from local energy. Pattern Recognition Letters 6, 303–313 (1987)
10. Shojaee Bakhtiari, A.: Design of a Biocryptosystem Based on the Key Extracted from the Iris Biometrics. School of Electrical Engineering, MSc Thesis, Iran University of Science and Technology, Tehran, Iran (2007)
11. National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, http://www.cbsr.ia.ac.cn
12. Canny, J.: A computational approach to edge detection. IEEE Transactions on Pattern Analysis and Machine Intelligence 8, 679–714 (1986)
13. Kovesi, P.D.: MATLAB Code for Calculating Phase Congruency and Phase Symmetry Asymmetry (1996), http://www.cs.uwa.edu.au/_pk/Research/MatlabFns/