A Key Generation Scheme Based on Face for Online Authentication

Xingsheng LIU Peng ZHOU Lifang WU
Beijing University of Technology, Beijing
E-mail: Lfwu@bjut.edu.cn Tel: +86-10-67396151

Abstract—**In this paper, we propose a key generation scheme based on face and apply it in online authentication. In order to improve security and the tolerance to intra-class variation, randomization and statistically optimal algorithm are chosen to generate the key. In enrollment stage, a 128-dimensional principal component analysis (PCA) feature vector is firstly extracted from the face image. And a randomized feature process is utilized to improve the security and control the intra-class variations of biometric data to the minimal level. From the binary vector of 128 bits we select the statistically distinguishable bits to form bio-key. Furthermore, an error-correct-code (ECC) is generated using Reed-Solomon algorithm. In authentication stage, the same procedure is implemented to extract PCA features and randomize the feature vectors. Then a bio-key is generated using the look-up table and auxiliary code. The on-line authentication mainly relies on checking the validity of the bio-key. The experimental results using ORL face database shows that our algorithm is more effective.**

I. INTRODUCTION

The rapid development of E-commerce requires reliable identity management system. Password based authentication is becoming insufficient for reliable authentication because password can be easily shared or stolen and a long and complex password is a challenge for users to memorize. Biometrics-based authentication is comparatively reliable, convenient and universal [3] .Unlike passwords, the biometric template is not changeable when it is compromised. Furthermore, the same biometric template allows cross-matching across databases. These problems cause new problem of privacy and security. Therefore, Biometrics template protection has attracted active research recently.

An ideal biometric template protection scheme should have the following four properties [3].

1) Diversity: The secure template must not allow cross-matching across databases, thereby ensuring users' privacy.

2) Revocability: It should be straightforward to revoke a compromised template and reissue a new one based on the same biometric data.

3) Security: It must be computationally hard to obtain the original biometric template from the secure template. This property can prevent an adversary from creating a physical spoof of the biometric trait from a stolen template.

4) Performance: The biometric template protection scheme should not degrade the recognition performance (FAR (false accept rate) and FRR (false reject rate)) of the biometric system.

Existing biometric template protection approaches can be broadly categorized into transformation-based approaches and biometric cryptosystem approaches [3]. The main idea of the transformation-based method is to transform the original biometric template into a new transformed template (or secure template). The transformed templates, instead of the original templates, are stored. The same transformation is applied to the query biometric data for authentication. Biometric cryptosystems [1, 2] were originally developed for the purpose of either securing a cryptographic key using biometric features or directly generating a cryptographic key from biometric features. It can be further divided into key binding and key generation approaches. Biometric-key generation is defined as a process to convert live biometric data into bit-string representation (key) using auxiliary information.

Cryptographic key generation from biometrics is an attractive but difficult problem because of the large intra-user variation of biometrics [3]. Monrose et al. [4, 5] combined a user's typing patterns with his password to generate a hardened password. The password is convincingly more secure than conventional passwords against both online and offline attackers. Chang et al. [6] and Veilhaueret et al. [7] proposed a user-specific quantization base key generation schemes. Information on quantization boundaries is stored as helper data which is used during authentication to account for intrauser variations. Dodis et al. [8, 9] introduced the concepts of secure sketch and fuzzy extractor in the context of key generation from biometrics and proposed secure sketches for three different distance metrics, namely, Hamming distance, set difference, and edit distance. Li et al. [10] introduced a two-level quantization-based approach for obtaining secure sketches. Sutcu et al. [11] discussed the practical issues in secure sketch construction and proposed a secure sketch based on quantization for face biometric. The problem of generating fuzzy extractors from continuous distributions was addressed by Buhan et al. [12]. Secure sketch constructions for other modalities such as fingerprints [13, 14] have also been proposed based on the fuzzy extractor scheme. Zhang et al. [15] studied biometric key generation using face images. They proposed generalized optimal shareholding method to improve the reliability and security of generated key. Chen et al. [16] used an entropy-based feature extraction method coupled with Reed-Solomon error correcting codes to generate deterministic bit-sequences from the output of an iterative one-way transform. They evaluated the method using 3D face data. The experimental results showed that can produce reliably keys of suitable length for 128-bit Advanced Encryption Standard (AES). Andrew et al. [17] introduced a novel biometric key generation scheme based on a

randomized biometric helper. This scheme consisted of a randomized feature process and a code redundancy construction. In our previous work, we proposed an optimal statistics based key generation scheme. We chose more effective face features by statistic optimality. And a cryptosystem is formed using the key for message encryption and decryption. These schemes had such a problem: the intra-class variations of biometric data are relatively great, while the inter-class variations of biometric data are relatively little, therefore, the FAR and FRR are bigger.

In this paper, we propose a novel approach to biometric key generation. Firstly a Randomization and Binarization is used to generate binary bits. Then the distinguishable binary bits are generated by statistical optimality. Then Reed-Solomon Code is used to generate robust bio-key. Our approach alleviates intra-class variations of biometric data to the minimal level and increases the inter-class disparity.

The remaining parts of this paper are organized as follows: In Section 2, we describe the framework of proposed approach, where we give an application model. We describe the details of proposed approach section 3 through 4. In Section 5, we provide and analyze the experimental results. In section 6, we conclude the paper.

II. FRAMEWORK OF THE PROPOSED APPROACH

In an authentication system based key generation. The key generated usually suffers from low discriminability which can be assessed in terms of key stability and key entropy [3]. Key stability refers to the extent to which the key generated from the biometrics is repeatable. Key entropy is related to the number of generated keys. If a scheme generates the same key irrespective of the input template, it has high key stability but zero entropy leading to high false accept rate. On the other hand, if the scheme generates different keys for different templates of the same user, the scheme has high entropy but no stability. It results in high false reject rate.

In a key generation authentication system, the authentication key should be exactly identical with the enrollment key. However, the features extracted from different face images of a subject are not definitely the same. It usually is caused by noise of camera, pose or illumination variation and so on. To solve these problems, we propose a novel key generation scheme. The framework of our proposed scheme is shown in Fig. 1.

In enrollment stage, PCA features of a face image are extracted. A randomized feature process is utilized to control the intra-class variations of biometric data to the minimal level. Using these randomized features, we generate a binary string by binarization. And then the stable bio-key is generated by statistical optimization and the order number of these optimal bits is saved in a look-up table. The bio-key is encoded using Reed – Solomon code [19] to further reduce the errors. The optimal order number, ECC code and bio-key are stored in the server.

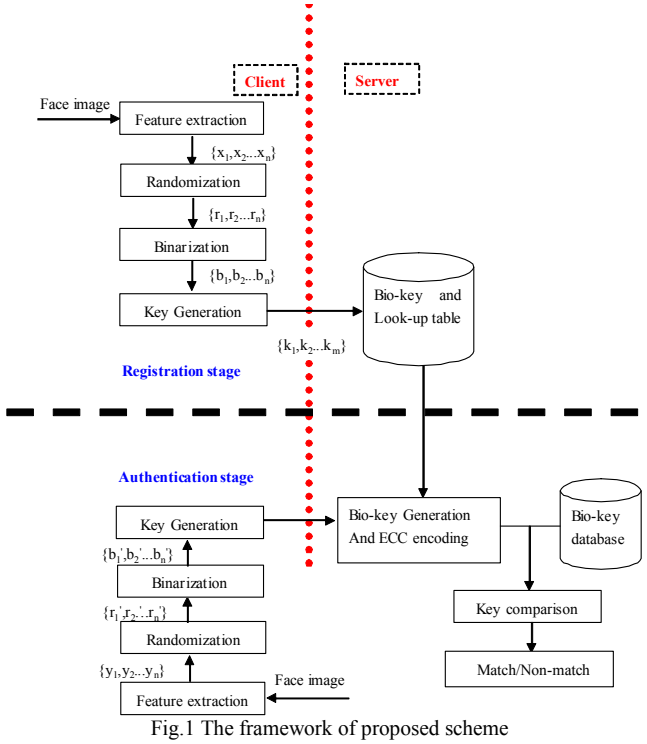

Fig.1 The framework of proposed scheme

In the authentication stage, the PCA features of the query face image are extracted. And the corresponding bio-key is generated by randomization and binarization. At the server side, the key is generated by the optimal order number. And it is further corrected from the generated key and auxiliary information. Finally the authentication is implemented by key comparison.

III. FEATURE EXTRACTION

In our scheme, we extract face features using PCA (Principal Component Analysis)[21]. Given a total of M images with $(N_x * N_y)$ pixels, we convert them into training set $\mathbf{\Gamma} = [\mathbf{\Gamma}_1, \mathbf{\Gamma}_2 \ldots, \mathbf{\Gamma}_M]$ with lexicographic ordering the pixel elements. Ψ is the mean vector of all the data. Φ is the result of images data subtract the mean vector, as given by the following

$$\mathbf{\Psi} = \frac{1}{M_t} \sum_{i=1}^{Mt} \mathbf{\Gamma} i \quad . \tag{1}$$

$$\mathbf{\Phi}_i = \mathbf{\Gamma}_i - \mathbf{\Psi} \quad . \tag{2}$$

Furthermore, the covariance matrix C of size $(N_x * N_y) * (N_x * N_y)$ is computed:

$$\mathbf{C} = \frac{1}{M_t} \sum_{i=1}^{Mt} \mathbf{\Phi}_i \mathbf{\Phi}_i^T \quad . \tag{3}$$

However, dimension of covariance matrix C is very high. Unwieldy of this excessive data has been solved by [21] who uses a small dimension covariance, $\mathbf{L}$ to replace $\mathbf{C}$. Eigenvectors, v computed from covariance $\mathbf{L}$ are multiplied

with $\mathbf{A}$ to obtain principal component υ which is able to represent the actual eigenvectors of covariance $\mathbf{C}$, as given by

$$\mathbf{A} = [\Phi_1, \Phi_2 \ldots, \Phi_m] . \quad (4)$$

$$\mathbf{L} = \mathbf{A}^T\mathbf{A} = \frac{1}{M}\sum_{i=1}^{M} \Gamma_i{}^T\Gamma_i \quad . \quad (5)$$

$$\Upsilon_i = \mathbf{A} \bullet \upsilon_i \quad . \quad (6)$$

υ becomes PCA projection basis in which both training and testing samples are project on it, yielding corresponding weight of each subject, ω_k which are arranged in sets, as in

$$\omega_k = \upsilon_k^T \bullet \Phi = \upsilon_k^T \bullet (\Gamma - \Psi) . \quad (7)$$

$$\Omega = [\omega_1, \omega_2, \ldots \omega_M] . \quad (8)$$

Equation (8) shows M sets of sample weights with a dimension of M. This dimension can be reduced to M' if we select only M' principal components in (6). Through this method, computational cost can be greatly reduced to $m * n$. A dimension of 128 PCA features is used in this paper.

IV. Bio-Key Generation

In this section, a bio-key is generated from a set of randomized PCA feature vectors. It includes distinguishable binary feature generation and key generation, as shown in Fig.2.

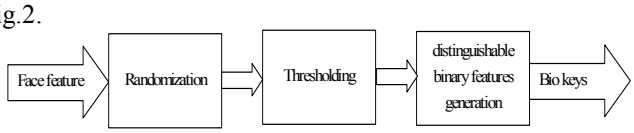

Fig.2 The diagram of the Statistic optimal Key generation

A. Thresholding

The corresponding binary vectors can be obtained from PCA vector by thresolding. The threshold is the average of the corresponding features of all the training samples.

B. Feature Randomization

The process of randomization comprises of two steps: face feature data normalization and subspace projection. In the normalization stage, the face image PCA feature is first normalized as $\hat{x} = x/||x|| \in R^n$, then $\hat{x}$ is projected on an orthonormal subspace, φ is given by

$$\mathbf{v} = \varphi\hat{\mathbf{x}} \quad (9).$$

Where φ is a $m * n$ ($m \le n$) orthonormal matrix with entries φ_{ij} such that $\varphi\varphi^T = I$ is an identity matrix. The matrix with orthonormal entries can be obtained from kernel matrices of classical transforms [23] such as Discrete Cosine Transform (DCT), $\varphi_{ij} = \sqrt{\frac{2}{n}}\cos[\frac{\Pi(2i+1)j}{2n}]$, and Discrete Sine Transform (DST), $\varphi_{ij} = \sqrt{\frac{2}{n+1}}\sin[\frac{\Pi(i+1)(j+1)}{n+1}]$, and orthogonal wavelet transform kernels [24]. In this paper, we select DCT as orthonormal matrix entries.

C. Distinguishable Bits Generation

The object of distinguishable bit generation is to find these bits which are distinguishable to separate the authentic user from potential imposter users. Generally, according to the degree of distinguishability, each effective bit may contribute one or bit of information to the cryptographic key generation [6]. Furthermore, the optimal bits are selected to form the bio-key. At the same time, a look-up table is generated by the order number of optimal bits.

D. ECC encoding

We adopt the Reed-Solomon code due to its powerful error correction capacity [19]. The Reed-Solomon code is denoted as $RS(n_b, k_b, t_b)$ where $n_b \le 2m_b-1$ and $2t_b \le n_b-k_b$. n_b is the number of blocks after encoding, k_b represents the number of blocks before encoding, t_b is the number of error blocks that can be corrected and m_b is the bits number per block. Compute the Reed-Solomon codes $RS(n_b, k_b, t_b)$ of the feature vector. Reed-Solomon codes usually contain two parts: vector data and correction codes. We only keep the correction codes as our error-correct-code (ECC).

V. Experimental Results

Our proposed method is tested using the ORL face database, which comprises 400 face images from 40 subjects with 10 face images each subject. Images of some subjects were taken at different times; some vary with illumination, expression and pose variation.

We compare our approach with our previous proposed approach, the first 5 images of each subject are used as training set, and the rest are testing set. The first 128 PCA features are used. The performance is evaluated using FAR (False accept Ratio) and FRR (False Rejected Ratio). Table I shows the compared results with the number T. T is the number of fault-tolerant.

TABLE I

FAR and FRR with Different

	Scheme in Ref[25]		Our scheme		T
key	FRR	FAR	FRR	FAR	
64	97.0%	0.0%	95.0%	0.0%	0
	72.0%	0.0%	65.0%	0.0%	4
	39.5%	0.0%	37.5%	0.0%	8
	20.5%	0.0%	18.5%	0.0%	12
	13.0%	0.0%	11.5%	0.0%	16
	8.5%	0.0%	6.5%	0.0%	20
	4.0%	5.0%	2.5%	2.5%	24
	1.0%	12.5%	1.0%	7.5%	28
	0. 5%	15.0%	0.25%	12.5%	29
	0.0%	15.0%	0.0%	15.0%	30
	0.0%	15.0%	0.0%	15.0%	31
	0.0%	17.5%	0.0%	17.5%	32
	0.0%	27.5%	0.0%	27.5%	36
	0.0%	32.5%	0.0%	32.5%	40
	0.0%	47.5%	0.0%	37.5%	44

	0.0%	52.5%	0.0%	40.5%	48
	0.0%	67.5%	0.0%	47.5%	52

Table I shows that the FAR increase and FRR decrease as T increase. From table I, we can see the following: the EER (equation error rate) of our approach is 2.5%, which is smaller than 4.5% of our previous approach. And when FAR is 0, the corresponding FRR of our approach is 6.5%, which is also smaller than 8.5% of our previous work. When FRR is 0, our approach has the same FAR 15.0% as our previous work.

In table I, when T is 30 the FRR is 0, but the corresponding FAR is 15%. When T increases from 0 to 20, the FRR decreases to minimum and the corresponding FAR still remains 0. In an online authentication system, the FAR is 0 to ensure that the system is safe. Therefore, we select RS (104, 64, 20) in the Reed-Solomon code.

We further compare our scheme with some previous schemes; the compared result is shown in Table II.

Table II shows that our scheme can get both the lowest FRR and FAR. And our approach can keep FAR is 0.

TABLE II
COMPARISON OF FAR AND FRR

Schemes	FAR	FRR
Zhang's [15]	0.06%	9.3%
Chen's [16]	1.22%	28%
Tuyls's[17]	-	35%
Ours	0%	6.5%

VI. CONCLUSIONS

In this paper, we have proposed a key generating biometric online authentication system based on face features. We extract PCA features from face images, and a randomized feature process is utilized to control the intra-class variations of biometric data to the minimal level. Then we get binary represents by thresholding, and then we generate statistically optimal bio-keys from the binary represents. Finally Reed-Solomon code is used for increase the error tolerance of generated keys. Our experimental results show that our scheme can get the smallest FRR when FAR is zero. In the future work, we will improve our method of key generation and generate more stable and secure bio-keys.

ACKNOWLEDGMENT

This paper is supported partially by the Key project of Beijing municipal Nature Science Foundation under Grant No 4091004 and program of Beijing Municipality excellent under Grant No 2009D005015000010.

REFERENCES

[1] Uludag U, Pankanti S, Prabhakar S, et al. "Biometric Cryptosystems: Issues and Challenges," Proc. of the IEEE VOL.92, NO.6, JUNE 2004, pp.948-960.

[2] A. Cavoukian and A. Stoianov, "Biometric encryption: a positive-sum technology that achieves strong authentication, security and privacy," Tech. Rep., Office of the Information and Privacy Commissioner of Ontario, Toronto, Ontario,Canada, March 2007.

[3] Anil K. Jain, Karthik Nandakumar, and Abhishek Nagar. "Biometric Template Security," Hindawi Publishing Corporation EURASIP Journal on Advances in Signal Processing Volume 2008.

[4] F. Monrose, M. K. Reiter, and S. Wetzel, "Password hardening based on keystroke dynamics," in Proceedings of the 6th ACM conference on Computer and Communications Security (ACM CCS '99), pp. 73–82, Singapore, November 1999.

[5] Monrose F, Reiter M K, Li Q, "Cryptographic Key Generation from Voice , " Proceedings of the 2001 IEEE Symposium on Security and Privacy, 2001,pp.202-213.

[6] Y.-J. Chang, W. Zhang, and T. Chen, "Biometrics-based cryptographic key generation," in Proceedings of the IEEE International Conference on Multimedia and Expo (ICME '04), vol. 3,pp. 2203–2206, Taipei, Taiwan, June 2004.

[7] Clause.Vielhauer, R.Steinmetz, and A.Mayerhofer, "Biometric hash based on statistical features of online signatures," in Proceedings of the International Conference on Pattern Recognition,vol. 1, pp. 123–126, Quebec, QC, Canada, August 2002.

[8] Y. Dodis, L. Reyzin, and A. Smith, "Fuzzy extractors: how togenerate strong keys from biometrics and other noisy data," in Proceedings of International Conference on the Theory and Anil K. Jain et al. 17 Applications of Cryptographic Techniques: Advances in Cryptology (EUROCRYPT '04), vol. 3027 of Lecture Notes in Computer Science, pp. 523–540, Interlaken, Switzerland, May 2004.

[9] Y.Dodis, R. Ostrovsky, L. Reyzin, andA. Smith, "Fuzzy extractors: how to generate strong keys from biometrics and other noisy data," Tech. Rep. 235, Cryptology ePrint Archive, February 2006.

[10] Q. Li and E.-C. Chang, "Robust, short and sensitive authentication tags using secure sketch," in Proceedings of the 8th Multimedia and Security Workshop (MM and Sec '06), pp. 56–61, Geneva, Switzerland, September 2006.

[11] Li Q, Sutcu Y, Memon N. Secure sketch for biometric templates. In: Lai XJ, Chen KF. Proc. of Advances in Cryptology 12th Int'lConf. on the Theory and Application of Cryptology and Information Security (ASIACRYPT 2006). 2006. Pp.99−113.

[12] I. Buhan, J. Doumen, P. Hartel, and R. Veldhuis, "Fuzzy extractors for continuous distributions," in Proceedings of the 2nd ACM Symposium on Information, Computer and CommunicationsSecurity (ASIACCS '07), Singapore, March 2007, pp. 353–355.

[13] A. Arakala, J. Jeffers, and K. J. Horadam, "Fuzzy extractors for minutiae-based fingerprint authentication," in Proceedings of the 2nd International Conference on Biometrics, pp. 760–769, Seoul, South Korea, August 2007.

[14] E. C. Chang and S. Roy, "Robust extraction of secret bits frominutiae," in Proceedings of 2nd International Conference onBiometrics, pp. 750–759, Seoul, South Korea, August 2007.

[15] Wende Zhang, Tsuhan Chen, "Biometric Generalized optimal thresholding for biometrics key generation using face image," IEEE International Conference on Volume 3 .2005

[16] B. Chen, V. Chandran, "Biometric Based Cryptographic Key Generation from Faces ," Dec.2007 Proceedings of the 9th Biennial Conference of the Australian Pattern Recognition Society on Digital Image Computing Techniques and Applications.

[17] Beng, A., JinTeoh, Kar-AnnToh, Secure "Biometrics-key Generation with Biometrics Helper,"Industrial Electronics and Applications, ICIEA 2008. 3rd IEEE Conference on 3-5 June 2008 ,pp.2145 - 2150.

[18] Sashank Singhvi, R.; Venkatachalam, S.P.; Kenyan, P.M.; Palanisamy, "Cryptography Key Gneration Using Biometrics," V.;Control, Automation, Communication and Energy Conservation, 2009 International Conference on 4-6 June 2009.

[19] M. Purser, Introduction to Error-Correcting Codes, Artech House, Boston, 1995.

[20] W. Zhang, C. Zhang and T. Chen, "Security Analysis for Key Generation Systems using face images," IEEE Conference on Image Processing, 2004,pp.3455-3458.

[21] Turk, M. and A. Pentland. "Eigenfaces for recognition," Journal of Cognitive Neuroscience, 1991. 3(1): pp. 71-86.

[22] Andrew Beng Jin Teoh and Kar-Ann Toh. "Secure Biometric-Key Generation with Biometric Helper" 2008 , IEEE , Pg2145-2150

[23] S.A. Martucci, Symmetric convolution and the discrete sine and cosinetransforms, IEEE Trans. Sig. Process SP-42 (1994) 1038–1051.

[24] Daubechies, Orthonormal bases of compactly supported wavelets,Commun. Pure Apply. Math. 41 (1988) 909–996.

[25] Lifang Wu; Xingsheng Liu; Songlong Yuan; Peng Xiao; "A novel key generation cryptosystem based on face features" Signal Processing (ICSP), 2010 IEEE 10th International Conference on Digital Object Identifier: 10.1109/ICOSP.2010.5656719 Publication Year: 2010 , Page(s): 1675 - 1678