A Hybrid Approach for Face Template Protection

Y C Feng[1], Pong C Yuen[1] and Anil K Jain[2]

[1]Department of Computer Science
Hong Kong Baptist University
Email: {ycfeng, pcyuen}@comp.hkbu.edu.hk

[2]Department of Computer Science and Engineering
Michigan State University
Email: jain@cse.msu.edu

Abstract

Biometric template protection is one of the important issues in deploying a practical biometric system. To tackle this problem, many algorithms have been reported in recent years, most of them being applicable to fingerprint biometric. Since the content and representation of fingerprint template is different from templates of other modalities such as face, the fingerprint template protection algorithms cannot be directly applied to face template. Moreover, we believe that no single template protection method is capable of satisfying the diversity, revocability, security and performance requirements. We propose a three-step cancelable framework which is a hybrid approach for face template protection. This hybrid algorithm is based on the random projection, class distribution preserving transform and hash function. Two publicly available face databases, namely FERET and CMU-PIE, are used for evaluating the template protection scheme. Experimental results show that the proposed method maintains good template discriminability, resulting in good recognition performance. A comparison with the recently developed random multispace quantization (RMQ) biohashing algorithm shows that our method outperforms the RMQ algorithm.

Keywords: Face template protection, Biometric data security, privacy, Face recognition

1. Introduction

Biometric recognition is a reliable, robust and convenient way for person authentication [1] [27]. With the growing concerns about security and terrorism, several large biometric systems such as US-VISIT program have been successfully deployed. Additionally, biometric systems are being developed for other applications [28] such as banking (for ATM machine), credit card industry and physical access control. With the growing use of biometrics, there is a rising concern about the security and privacy of the biometric data itself. Since each person is believed to have a unique biometric (e.g. fingerprint, face and iris), if this biometric data is compromised, it is not possible to replace it. Therefore, biometric data (template) security [1-4] is one of the most important issues in deploying a biometric system (biometric template refers to the extracted biometric features stored in a central database or a smartcard). Recent studies have shown that “hill climbing attacks” [2] on

biometric systems are able to recover the original raw biometric data from the biometric template. As a result, protection of biometric template is necessary.

In order to overcome the security and privacy problems [1-5], a number of biometric template protection algorithms have been reported. These methods can be broadly categorized into two approaches, namely biometric cryptosystem approach and transformation-based approach. The basic idea of both the approaches is that instead of storing the original template, it is the transformed/encrypted template that is stored. In case the transformed/encrypted biometric template is stolen or lost, it is computationally hard to reconstruct the biometric template and the original raw biometric data from the transformed/encrypted template. The advantage of biometric cryptosystem approach is that, since the output is an encrypted template, its ***security level*** is high. However, the error correction codes used in biometric cryptosystem may not be able to model large intra-user variations. Also, this approach is not designed to be revocable. In transformation-based approach, a transformed template is generated using a "one-way" transform and the matching is performed in the transformed domain. The transformation approach has a good ***cancelable ability*** (revocability), but the drawback of this approach is that a trade-off between performance and security is normally required.

In view of the limitations of existing approaches, we propose a hybrid framework for template protection. The proposed framework retains the advantages of both the transform based approach and biometric cryptosystem approach. Experimental results show that the hybrid algorithm is able to generate a secure and discriminative cancelable face template. The rest of this paper is organized as follows. In Section 2, a brief review of existing methods is presented. Our proposed framework together with the hybrid algorithm is reported in Section 3. Experimental results and conclusions are presented in Sections 4 and 5.

2. Review of Existing Methods

2.1 Biometric Cryptosystems

The basic idea of this approach is to integrate the cryptographic technique(s) into a biometric system to secure the biometric template. Davida et al. [6] applied the error correction code to generate a check data $\boldsymbol{K}$ and then hashed the original template $\boldsymbol{T}$ to $\boldsymbol{Hash(T)}$. Although this method has some limitations in terms of error tolerance and security level, it provides a good foundation and direction for future work. Along this line, Juels and Watternberg [7] proposed a fuzzy commitment scheme which considers the biometric template as a corrupted codeword. In this scheme, the security is linked to the codewords, but not the check vector $\boldsymbol{K}$, thereby increasing the security level. Juels and Sudan [8] proposed a fuzzy vault scheme which embeds a secret $\boldsymbol{S}$ in a fuzzy vault with a dataset $\boldsymbol{A}$. In order to extract the secret, one needs to present another set $\boldsymbol{B}$ to decrypt the vault $\boldsymbol{V}$. Clancy et al. [9] applied the fuzzy vault scheme to fingerprint biometric. Dodis et al. [10] presented some theoretical analysis of the fuzzy schemes and introduced fuzzy extractors.

Different from fuzzy schemes, Soutar et al. [11] proposed a biocrypto scheme which makes use of a filter function to transform the original template to a new representation. A secret key is then

inserted in the new representation. The randomized secret key makes it difficult to extract the information. Monrose et al. [12][13] proposed a cryptographic key generation scheme from biometrics. A two-stage structure was proposed. First, the biometric template is converted into a binary string, called feature descriptor. In the second stage, a cryptographic key is generated from the feature descriptor. Tuyls et al. [14] proposed a δ-contracting function algorithm to handle the intra-user variations. Helper data is also introduced to guide the matching process. Hao et al. [15] evaluated different types of errors introduced in iris recognition and proposed a two-layer error correction code to model the errors. Draper et al. [16] applied the distributed source coding to protect fingerprint and Sutcu et al. [17] proposed the use of sketch, which is an error tolerant cryptographic technique, for face biometric.

2.2 Transform-based Approach

The central problem of this approach is to find a "one-way" transformation such that the discriminative power of the transformed template is preserved. Ratha et al. [4] first proposed the concept of cancelable transform where the transformed template can be cancelled and re-issued by changing the transform parameters, if it is stolen or lost. Three different transformations [18] [19] were proposed for fingerprint biometric. Along the same line, Tulyakov et al. [20] employed a symmetric hash function as a cancelable transform. Ang et al. [21] proposed a key-dependent transformation algorithm for fingerprints. The original 2D space is divided into two parts which are key-dependent. The minutiae are scrambled between these two parts. Sutcu et al. [22] proposed a functional distortion of the original template. Teoh and his collaborators [23-24] developed different versions of biohashing algorithm for protecting fingerprint and face biometrics. The basic idea is to transform a template into another domain and then perform thresholding in the transformed domain. The resultant binary string is then used for authentication. In [24], the authors claim that a zero false acceptance rate can be obtained with the use of a token, but the template will not be secure if the token is known.

3. Proposed Framework

We propose to cascade the transformation approach with the biometric cryptosystem approach to form a new hybrid approach for face biometric as shown in Figure 1. The input is a face template extracted using face representation algorithms such as linear discriminant analysis or principle component analysis. In the first step *cancelable transform* is used to generate a *cancelable template*. A cancelable transform, normally, decreases the discriminative power of the original template. Therefore, a discriminability enhancement transform is then applied to compensate for the discriminative power lost in the first step. Another objective of the discriminability enhancement transform is to generate a binary template such that biometric cryptosystem method, e.g., hash function, can be employed in the final step. This way, the proposed three-step hybrid framework is able to satisfy the template protection requirements.

- ***Diversity and Revocability***: Different applications require different sets of parameters in cancelable transform. Therefore, the cancelable face templates and the secure face templates of an individual in different applications will be different. In turn, the cross-matching across

databases will not be feasible. Moreover, the secure face template can be cancelled and reissued by changing the cancelable transform parameters.

- ***Security:*** This framework produces a three-stage template security. First, since the main objective of the cancelable transform is to provide cancelable ability, a "near one-way" non-invertible transform can be used. Second, the discriminability enhancement transform normally is a non-linear transform and the output is a binary face template. It provides additional protection. Finally, biometric cryptosystem algorithms store the template in hashed form. The security strength will be equal to 2^n, where n is the dimension (width x height) of binary face template. If the dimension of the binary face template is longer than 100, the template security strength will be sufficient. Combining the three-stage protection makes it computationally hard to reconstruct the original face template from the secure face template.

- ***Performance:*** Our framework is not simply a combination of the transformation and biometric cryptosystem algorithms, but proposes a new *discriminability enhancement transform* in between. This step is very important to compensate the loss in the discriminative power of the original face template in the cancelable transform step. It can even increase its discriminative power if the face template representation is not optimized. Moreover, most of the existing biometric cryptosystem algorithms are able to model intra-user variations so as to increase the system performance. The face protection algorithm developed based on our framework should not adversely affect the recognition performance of the original biometric system. On the other hand, there is a possibility that the system performance can be increased if original face template representation is not optimized.

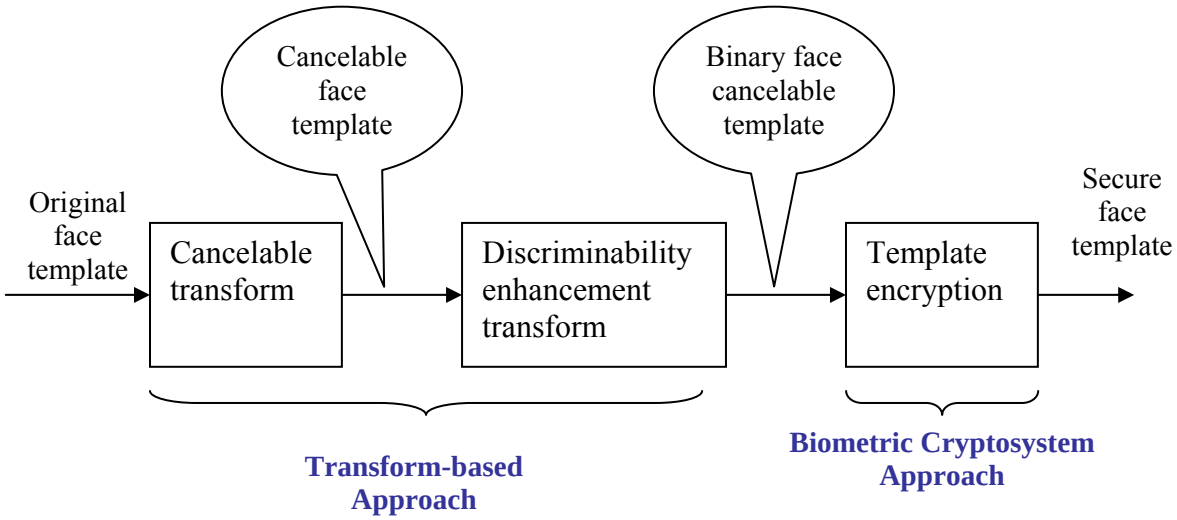

Figure 1: Proposed hybrid framework for protecting face biometric template

3.1 An example: Three-step hybrid algorithm

Based on the three-step framework, a hybrid algorithm is developed. Figure 3.2 shows the block diagram of the hybrid method. Like traditional biometric system, it consists of two phases, namely enrollment and query phases. In enrollment phase, each user's face image is captured and the face biometric template is extracted. In this paper, Fisherface [31] algorithm is employed in the feature extraction step and the LDA feature vector is considered as the original biometric template. In the first step, random projection is employed as a cancelable transform to project the original template into a subspace to generate a cancelable template. By using different transforms of random projection or changing the parameters of the same transform, the cancelability property can be achieved. In the second step, we apply the class distribution preserving (CDP) transform [26] to enhance the cancelable template discriminability and convert the real value cancelable template into a binary template. Finally, a hash function [33] is employed in the third step to encrypt the binary face template.

Random Projection

Random projection is a popular dimension reduction technique and has been successfully applied to many computer vision and pattern recognition applications. Recently, it has also been employed as a cancelable transform [24] for face biometric. While random projection provides a good cancelable ability, there is a trade-off between the system performance and the template security. Under our proposed framework, the performance and security problems are not the main issues to be handled in this step and can be addressed in the subsequent steps.

Class Distribution Preserving Transform

Class distribution preserving (CDP) transform [25],[26] enhances the template discriminability and converts a real value template into a binary template. The basic idea is to make use of a set of distinguishing points, a distance function and thresholding. For each template, the distances between the template and each distinguishing point are calculated. If the distance is below the threshold, a bit "0" is generated; otherwise, a bit "1" is generated. In this way, if the set contains k distinguishing points, each template will be converted into a k-bit binary template. It has been shown [26] that the CDP transform can enhance the template discriminability and improve the system performance.

Hash Function

The hash function [33] is applied for biometric data protection. It encrypts the original template s to a hashed codeword *Hash(s)* which is stored in the database; the matching process is done in hash space. While hash function is not considered to be good for biometrics because of its sensitivity to facial variations due to illumination, pose and facial expression, the use of CDP transform increases the template discriminative power enabling the use of hash function for face biometric. In our CDP transform, the templates are well classified. In most of the cases, the binary strings transformed via CDP transform are identical if they belong to the same class. In other words, the intra-class variation

after transform is essentially eliminated. Thus, the sensitivity to facial variations is no longer a problem in applying hash functions.

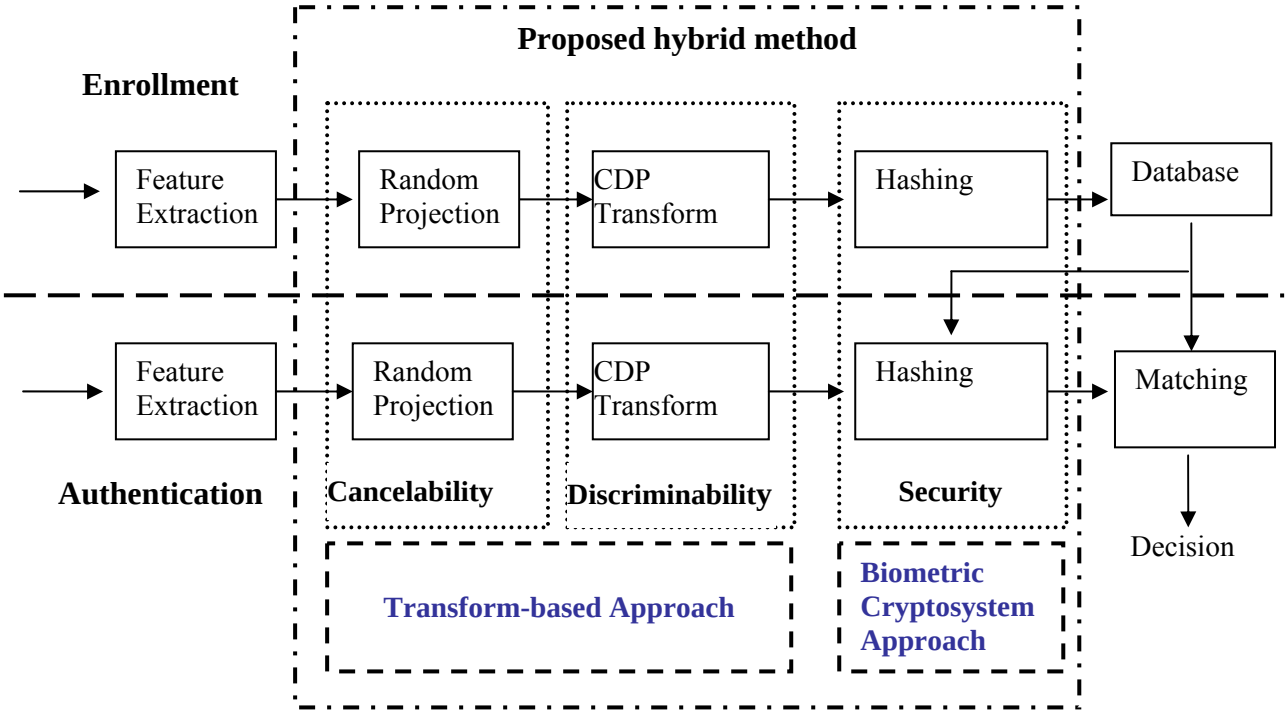

Figure 2: Block diagram of the proposed method

3.2 Security Analysis

The security of our algorithm is strengthened by the hash function while the random projection and CDP transform are able to increase the template security strength. For the random projection matrix, although the inverse may not exist, an attacker might guess its pseudo-inverse for recovering the original template from the projected templates. In the CDP transform process, the transformed template is converted into a binary string which supposes to be secure. However, the recovery of the transformed template from the binary string may still be feasible because the binary strings contain discrimination information about the original templates and the parameters (distinguishing points and thresholds) of the CDP transform are not protected. Therefore, in order to generate a secure template, hashing step is mandatory.

The security strength of the hashed template is analyzed as follows. If we assume that each bit in the binary bit string is random, using the MD5 hashing algorithm [33] as an example, the security of string is 2^n, where n is the bit length. In the CDP transform, the positions of the distinguishing points are generated with a random variable. Therefore, in the proposed algorithm, it is reasonable to assume that the bit string is randomly generated. Moreover, the bit length in our proposed algorithm is equal to 210. In turn, the security strength of the proposed method is equal to 2^{210} which is a very high security standard given the state of the art.

4. Experimental Results

Two public domain face databases, namely FERET and CMU-PIE, are used to evaluate the proposed method. In the FERET database, 250 individuals are selected and each individual has 4 different facila images. There are small pose, illumination, age and occlusion (glasses) variations in the FERET database. Images of one individual are shown in Figure 3(a). In the experiments, we randomly select 2 images per user for training and the remaining two images for testing.

There are 68 individuals in the CMU-PIE database which contains images with pose, illumination and expression variations. In this paper, 105 images for each individual with multiple poses (± 15 degrees) and large illumination variations are selected for experiments. Images of an individual are shown in Figure 3(b). We randomly select 10 images per individual for training while the remaining images are used for testing.

The experiment settings are as follows. For FERET database, the data dimension after the random projection is 40 while it is 150 for CMU-PIE database. Experiments with different number of dimensions after the class distribution preserving transform (*kc*) as shown in Table 1 are performed.

Two classification algorithms are selected for comparison with our proposed method. The first one is the original Fisherface with 1NN as a classifier. This is used as a benchmark. The second is one of the recent face template protection algorithms, namely the random multispace quantization (RMQ) biohashing algorithm [24]. The experimental settings for the RMQ biohasing algorithm are as follows. The transformed vector length *(kr)* is chosen as 40 and 150 for CMU PIE database and FERET database, respectively. Other settings are the same as described above.

Database	kc_1	kc_2	kc_3	kc_4
FERET	120	150	180	210
CMU	120	150	180	210

Table 1: Different values for parameter *kc* in the proposed hybrid algorithm

The experimental results on the FERET and CMU PIE databases are shown in Figures 4 and 5, respectively. The ROC curves using the proposed hybrid algorithm are labeled as “SRC” while it is labeled as “RMQ-S” for the RMQ biohashing algorithm where the same random projection subspace is used for training and testing. The result of using Fisherface with 1NN classifier is labeled as “original”. It can be seen from Figures 4 and 5 that the hybrid algorithm outperforms the RMQ biohashing algorithm. For most values of FAR, there is ~5% improvement in the recognition accuracy. It is important to note that the hybrid algorithm outperforms the classification method based on (Fisherface + 1NN). This result suggests that the original face representation may not be the most discriminative; the proposed hybrid approach is able to further improve the discriminative power of the original representation.

(a) (b)

Figure 3: Sample images of an individual in (a) the FERET database (b) the CMU-PIE database.

The equal error rate (EER) of each method is also recorded and shown in Table 2. The EER of the proposed method on FERET and CMU-PIE databases are 8.55% and 6.81%, respectively. This is significantly better than the EER of the biohashing algorithm (12.83% and 11.93%) and the original Fisherface method (12.58% and 18.18%).

	Fisherface + 1NN	**Proposed Hybrid Method**				**RMQ**
		kc_1	kc_2	kc_3	kc_4	
FERET	***12.58%***	9.52%	8.86%	8.61%	***8.55%***	***12.83%***
CMU-PIE	***18.18%***	7.61%	7.30%	6.95%	***6.81%***	***11.93%***

Table 2: EER on FERET and CMU-PIE databases

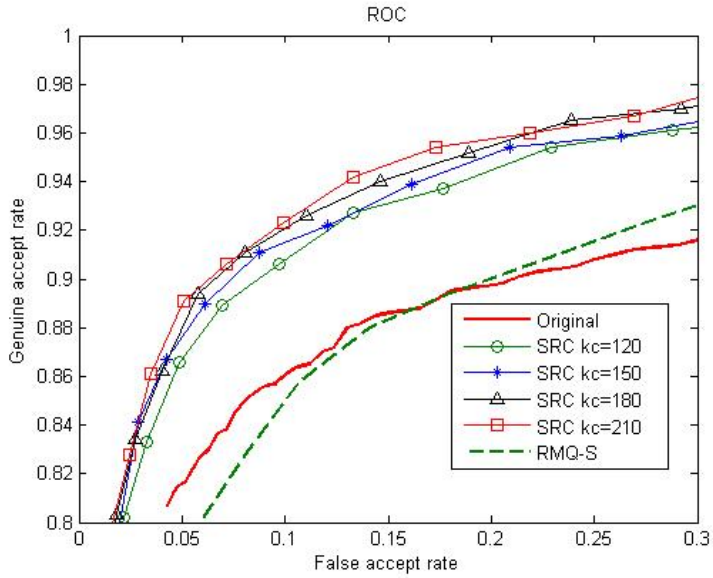

Figure 4: Experimental results on the FERET database

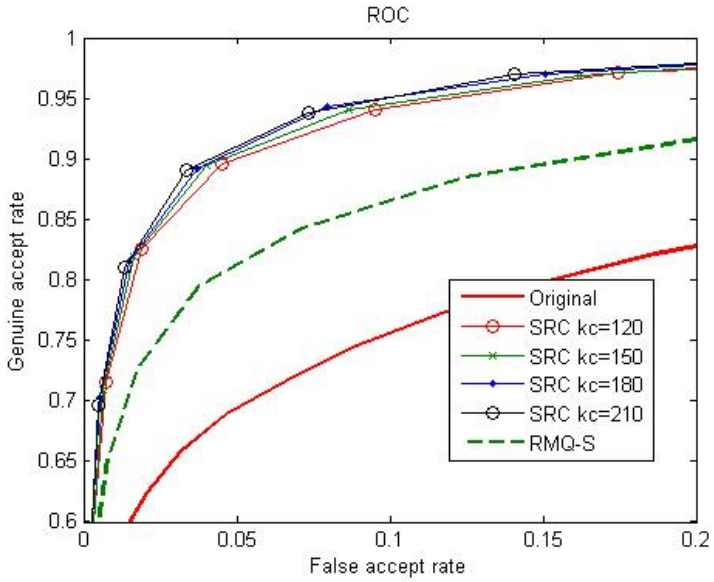

Figure 5: Experimental results on the CMU PIE database

5. Conclusions

A hybrid face protection framework for face biometric protection has been designed and evaluated. The new framework has the advantages of both the transform-based approach and biometric cryptosystem approach. It consists of three parts, namely cancelable transform, discriminability enhancement transform and template protection. Each part provides the cancelable ability, discriminability and security, respectively. Based on the proposed framework, a hybrid method is also developed. Two public domain face databases have been used to evaluate the proposed method. Experimental results show that the proposed method not only protects the template but in fact is able to increase the template discriminability. A comparison with the random multispace quantization (RMQ) biohashing algorithm shows the superiority of the proposed method for face template protection.

The preliminary results of the proposed three-step algorithm are good. However, the CDP transform is designed for authentication. In future, we will extend the CDP transform for the identification and further study the other alternative methods for the discriminability enhancement transform in the proposed framework.

Acknowledgement

This project was partially supported by the Science Faculty of the Hong Kong Baptist University. The authors would like to thank NIST and CMU for the face databases used in this paper.

References

1. S. Prabhakar, S. Pankanti, and A. K. Jain, "Biometric Recognition: Security and Privacy Concerns," *IEEE Security and Privacy Magazine*, Vol. 1, No. 2, pp. 33-42, March-April 2003.
2. A. Adler, "Images can be regenerated from quantized biometric match score data," *Proceedings of Canadian conference of Electrical and Computer Engineering*, pp. 469-472, 2004
3. A. Adler, "Vulnerabilities in Biometric Encryption Systems," *Audio- and Video-Based Biometric Person Authentication*, vol. 3546, pp. 1100-1109, 2005.
4. N. Ratha, J. Connell and R. Bolle, "Enhancing security and privacy in biometric-based authentication systems," *IBM Systems Journal,* Vol. 40. No. 3, pp. 614 - 634, 2001.
5. I. R. Buhan, and P. H. Hartel, "*The state of the art in abuse of biometrics",* Technical Report TR-CTIT-05-41 Centre for Telematics and Information Technology, University of Twente, Enschede. ISSN 1381-3625
6. G. Davida, Y. Frankel, and B. Matt, "On enabling secure applications through off-line biometric identification," *IEEE Symposium on Privacy and Security*, pp. 148-157, 1998.
7. A. Juels and M. Wattenberg, "A fuzzy commitment scheme", *Sixth ACM Conf. on Comp. and Comm. Security,* pp. 28-36, 1999.
8. A. Juels and M. Sudan. "A Fuzzy Vault Scheme", *IEEE International Symposium on Information Theory,* 2002.
9. T. C. Clancy, N. Kiyavash, and D. J. Lin, "Secure smartcard-based fingerprint authentication", *Proc. ACMSIGMM 2003 Multimedia, Biometrics Methods and Applications Workshop,* pp. 45-52, 2003.
10. Y. Dodis, L. Reyzin, and A. Smith, "Fuzzy extractors: How to generate strong keys from biometrics and other noisy data," *In Proc. Advances in Cryptology—Eurocrypt '04*, 2004.
11. C. Soutar, D. Roberge, A. Stoinav, G. Gilroy and V. Kumar, "Biometric Encryption Using Image Processing," *Proc. SPIE, vol. 3314*, pp. 174-188, 1998.
12. F. Monrose, M. K. Reiter and S. Wetzel, "Password Hardening Based on Key Stroke Dynamics," *Proc. ACM Conf. Computer and Comm. Security*, pp. 73-82, 1999.
13. F. Monrose, M. Reiter, Q. Li and S. Wetzel, "Cryptographic Key Generation from Voice," *Proc. IEEE Symp. Security and Privacy*, pp.202-213, May 2001.
14. P. Tuyls and J. Goseling, "Capacity and Examples of Template-Protecting Biometric Authentication Systems," *ECCV Workshop BioAW 2004*, pp. 158-170, 2004.
15. F. Hao, R. Anderson and J. Daugman, "Combining cryptography with biometric effectively", Technical Report, University of Cambridge, UCAM-CL-TR-640, ISSN 1476-2986, 2005.
16. S C Draper, A Khisti and E Martinian, A Vetro and J S Yedidia, "Using distributed source coding to secure fingerprint biometric", *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, 2007.
17. Y Sutcu, Q Li and N Memon, "Protecting biometric template with sketch: theory and practice", *IEEE Transactions on Information Forensics and Security*, Vol. 2, No. 3 Part 2, pp. 503-512 , 2007.
18. N Ratha, J Connell, R Bolle, S Chikkerur, "Cancelable biometrics: A case study in Fingerprints", *Proceedings of International Conference on Pattern Recognition,* 2006.
19. N Ratha, S Chikkerur, J Connell, R Bolle, "Generating Cancelable Fingerprint Templates", *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol. 29, pp. 561-752, 2007

20. S Tulyakov, V Chavan, and V Govindaraju, “Symmetric Hash Functions for Fingerprint Minutiae,” *Proc. Int’l Workshop Pattern Recognition for Crime Prevention, Security, and Surveillance,* pp. 30-38, 2005.
21. R Ang, R Safavi-Naini, and L McAven, “Cancelable Key-Based Fingerprint Templates,” *ACISP 2005*, pp. 242-252.
22. Y. Sutcu, H. Sencar, and N. Nemon, “A Secure Biometric Authentication Scheme Based on Roubst Hashing,” *Proc. Seventh Workshop Multimedia and Security*, pp.111-116, 2005.
23. D Ngo, A Teoh, and A Goh, “Biometric Hash: High-Confidence Face Recognition”, *IEEE Transactions onCcircuits andSsystems for Video Technology,* vol. 16, no. 6, 2006.
24. A. Teoh, A. Goh, and D. Ngo, “Random Multispace Quantization as an Analytic Mechanism for BioHashing of Biometric and Random Identity Inputs,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, pp. 1892-1901, Dec. 2006.
25. Y C Feng and P C Yuen, “Class-Distribution Preserving Transform for Face Biometric Data Security,” *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP),* pp. 141-144, 2007.
26. Y C Feng and P C Yuen, “Selection of Distinguish Points for Class Distribution Preserving Transform for Biometric Template Protection,” *Proceedings of IEEE International Conference on Biometrics (ICB),* 2007.
27. U Uludag, S Pankanti, S Prabhakar, and A K Jain, “Biometric cryptosystems: issues and challenges,” *Proceedings of the IEEE*, vol. 92, no. 6, pp. 948–960, 2004.
28. A. K. Jain and S. Pankanti, "A Touch of Money", *IEEE Spectrum*, pp. 22-27, July 2006.
29. A. K. Jain, K. Nandakumar and A. Nagar, "Biometric Template Security", EURASIP Journal on Advances in Signal Processing, January 2008.
30. D. Maltoni, D. Maio, A. K. Jain, and S. Prabhakar, *Handbook of Fingerprint Recognition*. Springer-Verlag, 2003.
31. P N Belhumeur, J P Hespanha, and D J Kriegman, “Eigenfaces vs. fisherfaces: Recognition using class specific linear projection”, *IEEE Trans. on Pattern Analysis and Machine Intelligence,* 19(7), pp. 711-720, 1997.
32. Karthik Nandakumar, Abhishek Nagar and Anil K. Jain, "Hardening Fingerprint Fuzzy Vault Using Password", in *Proceedings of International Conference on Biometrics*, LNCS 4642, pp. 927-937, 2007.
33. R. L. Rivest, “The MD5 Message-Digest Algorithm,” *RFC1321, Network Working Group, MIT Laboratory for Computer Science and RSA Data Security, Inc*., 1992.