CLASS-DISTRIBUTION PRESERVING TRANSFORM FOR FACE BIOMETRIC DATA SECURITY

Y C Feng & Pong C Yuen

Department of Computer Science
Hong Kong Baptist University
Email: ycfeng@comp.hkbu.edu.hk, pcyuen@comp.hkbu.edu.hk

ABSTRACT

This paper addresses the face data variations problem in biometric cryptosystems in which the cryptographic technique is applied to biometric system. To overcome the limitation, this paper introduce a new class-distribution preserving transform to biometric cryptosystems. The basic idea is to transform a real value face feature vector to a binary feature vector using a random points set. The proposed transform is integrated into a BCH coding technique. Fisherface algorithm is used for feature extraction and ORL face database is selected for experiments. It is shown that only around 0.8% accuracy is degraded in comparing with the original Fisherface algorithm while the system security can be enhanced by 126 bits.

Keywords: Biometric data security, biometric cryptosystems

1. INTRODUCTION

While biometric recognition systems have been in use for almost forty years, research on biometric data security [1, 7] is relatively new. A comprehensive analysis of different types of attacks in a biometric system has been reported. Recent studies show that simple attacks on a biometric system, such as hill climbing, are able to recover the biometric templates (biometric template refers to the extracted biometric features stored in the database or smartcard). In turn, cryptographic techniques are employed to protect the template and biometric cryptosystems [12, 13] have been proposed. Figure 1 shows the block diagram of a typical biometric cryptosystem, which consists of enrollment and authentication stages and includes sensor (acquisition), feature extraction, encryption and matching. The data flow is indicated by dotted lines.

In order to protect the biometric data, most (if not all) biometric cryptosystems perform matching process in encrypted space. Since most of the cryptographic techniques assume that the query data is exactly the same as the one at enrollment stage, the algorithms were designed that a small change of query data will result in a large change in encrypted domain. However, it is well-known that biometric data do have a certain degree of variations and the query biometric data will be different from the one obtained during the enrollment stage. Therefore, matching of biometric templates in encrypted domain is not reliable. To solve this problem, error correcting coding approach has been used in order to compensate the small variations of biometric data. Algorithms have been developed to protect fingerprint and iris biometrics [4, 3, 5, 6]. As far as we know, there are not much research articles reported on protecting face biometric. This may be due to the fact that face biometric data has a relative large within-class variations. One related article is to protect face photo on ID cards [8]. The second method computes cryptographic key from face images [10]. Recently, we have also developed a new scheme using Reed-Solomon codes [2].

The fuzzy vault scheme and reed-solomon codes in error correcting approach are good to handle small biometric data variations. The performance of the biometric cryptosystems will be degraded dramatically if the variations increase. In order to solve this problem, this paper proposes to transform the face feature vector representation. The objective of this process is to maintain the feature vector discriminability after transformation while the new representation can be input to existing error correcting coding methods. Based on this idea, this paper proposes a Class-Distribution Preserving (CDP) Transform which changes face feature vectors from real value coefficients to tri-state (0, 1 and ϕ) representation. BCH codes in error correction coding approach is then used to encrypt the tri-state data. The solid line in Figure 1 shows data flow with the added transformation process. This paper mainly focuses on the proposed CDP transform.

The rest of this paper is organized as follows. Section 2 gives the description of our proposed scheme. Experimental results and the Security analysis will be given in Section 3 and Section 4. Finally conclusion is given in Section 5.

2. CLASS-DISTRIBUTION PRESERVING TRANSFORMATION

2.1. Basic idea

The objective of the transformation is to change the representation of a real value feature vector to a binary feature

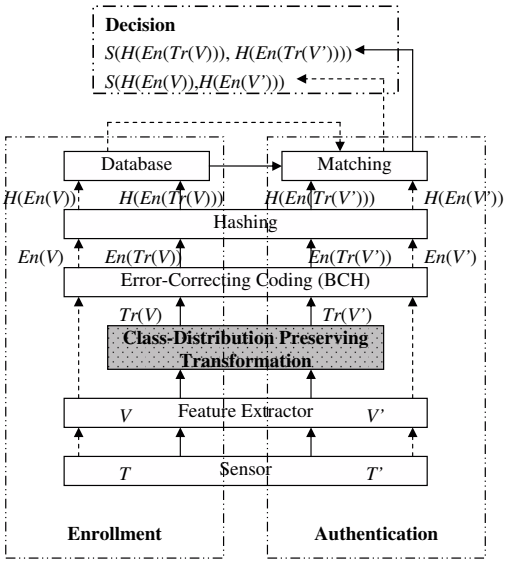

Fig. 1. Bolck diagram of an error-correcting coding (ECC) based biometric cryptosystem with CDP transform.

vector, which can be protected (encoded) using BCH coding method. Moreover, the class distribution after transformation should remain unchanged as as to maintain the face feature vector discriminability. The basic idea of our proposed class-distribution preserving transformation is shown in Figure 2 using a two class problem.

Suppose O_1 and O_2 are the center of two class A_1 and A_2 and we have a random point set $S = \{B_1, B_2, B_3...B_p\}$. For any point Q belongs to A_1 and A_2, the distance between Q and B_i, $d(Q, B_i)$, $i = 1, 2, ...p$, is calculated and thresholded as follows.

$$m_i = \begin{cases} 0 & \text{if } d(Q, B_i) \leq t \\ 1 & \text{if } d(Q, B_i) > t \end{cases}$$

After this transformation, all points belonging to the same class will also cluster together, with minor distortion. Moreover, each point will be in binary representation. In the simplest two-class problem with only two-dimensional feature vector, if the bit is 0, the point belongs to class A_1, otherwise belongs to class A_2. It is obvious that the performance of this classification depends on the position of B_i and the threshold t. If B_i lies near extension line O_2O_1 and $t = B_iC$ (C is the center of line Q_1Q_2), the performance will be optimal. For the natural condition there are many classes. If there are more points, it is more possible that any pair of classes will find a point in good position, thus well classified.

2.2. Variation considering

The transformation process basically solve the biometric data variation problem. However, there is a situation that the proposed transformation may not work well. Considering two feature vectors v_1 and v_2 belonging to the same class. Suppose the distance between these two vectors and a random point B in set S, $d(v_1, B)$ is larger than threshold t but $d(v_2, B)$ is smaller than t, v_1 and v_2 will be treated from different classes. To overcome this drawback, a variation range r is defined and the thresholding criteria are modified as follows.

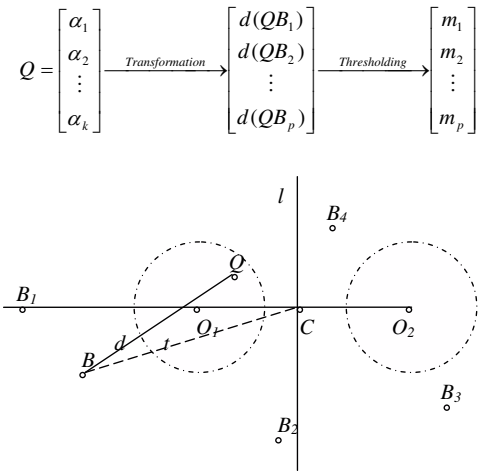

Fig. 2. The simple two classes situation for the class-distribution preserving transformation.

$$m_i = \begin{cases} 0 & \text{if } d(v, B_i) < t - r/2 \\ 1 & \text{if } d(v, B_i) > t + r/2 \\ \phi & \text{if } t - r/2 \leq d(v, B_i) \leq t + r/2 \end{cases}$$

The transformed vector will then be a tri-state feature vector. In matching of two tri-state vectors, if ϕ is found at the i_{th} entry of either one feature vector, the i_{th} entry is ignored and not counted in the hamming distance measurement.

2.3. Threshold Selection

One of the key factors in the proposed transformation is to determine the thresholds t_i. Basically, we can use the same t for all random points in set B or compute the average distance from the random point set to the feature vectors as the the threshold. However, these two methods are not directly linked with the system performance. They are determined subjectively without theoretical justification. In turn, we propose another method as follows.

Assume there are m classes and their corresponding cluster centers are $O_1, O_2, O_3 \ldots O_m$. The proposed scheme generate mp random points $B_1, B_2,$
$B_3 \ldots B_{mp}$ as the distinguish points. The corresponding mp thresholds are determined by

$$t_i = |O_q B_i|, (q = int((i-1)/m) + 1.)$$

Thus, for cluster center O_q, its distances to p points $|O_qB_{(q-1)p+1}|$, $|O_qB_{(q-1)p+2}|$, $|O_qB_{(q-1)p+3}|\ldots|O_qB_{qp}|$ are set as the corresponding thresholds. The situation for $p = m = 4$ is shown in figure 3: After this setting, consider a

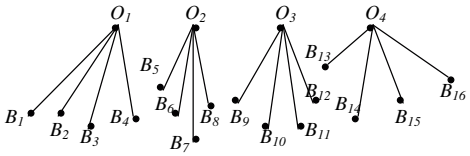

Fig. 3. Thresholds specifying with average feature vectors.

feature vector (point) P in class 1. P belongs to class 1 means that P is close to O_1, thus,

$$|PB_i - O_1B_i| < \epsilon, i = 1, 2, p.$$

in which ϵ is a small scalar. If $r/2 > \epsilon$, then $|PB_i - t_i| = |PB_i - O_1B_i| < r/2$, $i = 1, 2, \ldots p$, thus, the first p bits of the transformed vector should be ϕ. Then all the feature vectors in class 1 will be transformed to binary strings with first p bits ϕ, thus, the same. As the same way, the $p + 1$ to $2p$ bits of the binary strings transformed from class 2 will be the same as ϕ, the $2p + 1$ to $3p$ bits of the binary strings transformed from class 3 will be the sameand so on. In other words, this thresholds setting method can make sure p bits in the transformed binary vectors from the same class to be the same and thus, decreases the FRR.

3. EXPERIMENTAL RESULTS

In our experiments, fisherface [9] algorithm is employed for feature extraction and the ORL database is used with 40 individuals and 10 images per person is used for evaluation.In the experiment, length of the feature vector is 39, the parameter p is equal to 10.

Figure 4 shows the ROC curve with original Fisherface algorithm and algorithm with our CDP transform, with parameter r = 100, 120, 140, 160 and 180. The dashed diagonal shows the position where ERRs should lie on. From the figure we see that the ERR of original algorithm and algorithm with CDP transform are 5.7% and 6.5% respectively. The curve using the original Fisherface algorithm is very close to the curves with CDP transforms. This implies that performance with and without CDP transform are very similar. It shows that the CDP transform do affect the performance of the original system much.

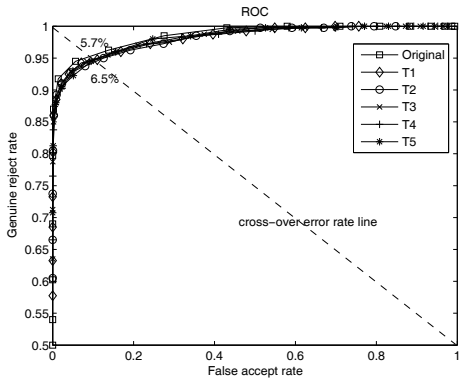

Fig. 4. Results without/with transformation. Lines "T1" to "T5" separately represents transform with $r = 100, 120, 140, 160, 180$ and line "original" represents original algorithm. The diagonal means the position where cross-over error rates should lie on.

4. SECURITY ANALYSIS

4.1. System design

After transformation, the tri-state feature vectors should be protected. The BCH code is chosen to do the encoding/decoding process from the fuzzy scheme. [3] In enrollment, assume the extracted binary string (ignore the ϕ) is s. A BCH codeword c is randomly generated. Store the hashing $Hash(c)$ and $s - c$ in database. In authentication, a new binary string s' is extracted from the query biometric data. Compute $s' - (s - c)$ and do BCH decoding [11] to this $s' - (s - c)$ and we get c'. If $Hm(s', s) \leq th$ (Hm denotes Hamming distance, th denotes threshold), then $Hm(s' - (s - c), c) \leq th$. And the decoding can correct $s' - (s - c)$ to c. That is c' equals c. Comparing $Hash(c')$ and $Hash(c)$, the decision is obtained as shown in figure 1.

4.2. Security level analysis

It is obvious that the security of our scheme depends on how many bits used in the transformed bit strings (except the bits of value ϕ). If an attacker wants to access our system and he claims that he belongs to a certain class (suppose class 1), he should try to present a biometric data v belonging to the class and v will be transformed into a binary string b with some bits ϕ. If the binary string b has a Hamming distance no more than threshold t to the stored string s that represents class 1, b will be treated as class 1. Assume the length of the binary string is n, there are q bits with value "0" or "1" in the string b and pp bits of ϕ, the possibility that $Hm(b, s) \leq t$ is

$$Pr(Hm(b, s) \leq t) = (\sum_{x=0}^{t} C_q^x)/2^q.$$

The reciprocal is the security level of our system. From the

equation, we know that the security level depends on q and t. Because different transformed binary strings may have different number of ϕ, thus different q because q equals to $n - pp$, we should analysis what's the distribution of pp's value after transformation. It depends on parameter r. With different r, the mean, maximum and minimum values of pp according to the binary strings are computed. Also, we can choose suitable threshold t from experiment result to different r to get an error rate near the crossover error rate. Thus, with every r we get a t and pp, result in a security level, which is shown in figure 5.

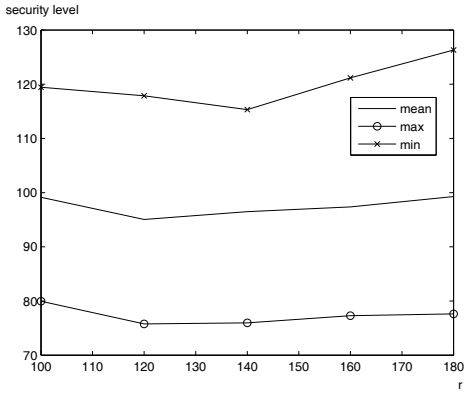

Fig. 5. security level tested with ORL database and LDA algorithm. Line "max", "mean", "min" separately means the security levels computed from maximum, minimum and mean value of *pp*.

From the experiment results we know that the security level of our scheme highly depends on the string length, that is, how many random points are used. This is depending on the feature vector length. Algorithm applied in ORL database with LDA algorithm can get a security level of about 78$\sim$126 bits.

5. CONCLUSION

A class-distribution preserving (SDP) transform has been desiged and reported in this paper. The proposed CDP transform also applied to a human face biometric cryptosystem using BCH code and evaluated. A system level security analysis is also reported. A popular ORL face dataset is selected to evaluate the performance of the proposed transform in the cryptosystem. It is found that, by introducing the CDP transform into the system, only 0.8% accuracy is degraded while the system security can be increased upto 126 bits. This result is very encouraging. Moreover, the computational time is less than 0.2 second in a typical personal computer.

6. REFERENCES

[1] N. K. Ratha, J. H. Connell, and R. M. Bolle, "Enhancing security and privacy in biometrics-based authentication systems," *IBM Systems Journal*, Vol. 40. No. 3, pp. 614 - 634, 2001.

[2] Y. C. Feng and P. C. Yuen, "Protecting Face Biometric Data on Smartcard with Reed-Solomon Code," *IEEE CVPR Workshop on Biometrics*, 2006.

[3] A. Juels and M. Wattenberg, "A Fuzzy Commitment Scheme," *Sixth ACM Conference on Computer and Communications Security*, pp. 28-36, ACM Press. 1999.

[4] G.I. Davida, Y. Frankel, and B.J. Matt, "On enabling secure applications through off-line biometric identification," *IEEE Symposium on Privacy and Security*, pp. 148-157, 1998.

[5] A. Juels and M. Sudan, "A Fuzzy Vault Scheme," *Proceedings of IEEE Internation Symposium on Information Theory*, p.408, 2002.

[6] T. C. Clancy, N. Kiyavash, and D. J. Lin, "Secure smartcard-based fingerprint authentication," *Proceedings of IEEE Internation Symposium on Information Theoryin Proc. ACMSIGMM 2003 Multimedia, Biometrics Methods and Applications Workshop*, pp.45-52, 2003.

[7] U. Uludag, S. Pankanti, S. Prabhakar and AK Jain, "Biometric Cryptosystems: Issues and. Challenges," *Proc. of the IEEE, Special Issue on Multimedia Security for Digital Rights*, vol. 92, no. 6, pp. 948-960, June 2004.

[8] D. Kirovski, N. Jojic, and G. Jancke, "Tamper-Resistant Biometric IDs," *in Information Security Solutions. Europe*, September 2004.

[9] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection.," *IEEE Trans. on PAMI, 19(7)*, pp. 711-720, 1997.

[10] A. Goh and D. Ngo, "Computation of Cryptographic Keys from Face Biometrics," *Communications and Multimedia Security*, 2003.

[11] J. L. Massey, "Shift-Register Synthesis and BCH Decoding," *IEEE Trans. Inform. Theory,*, vol. 15, no. 1, pp. 122-127, January 1969.

[12] 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.

[13] 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.