Soft Comput (2016) 20:4907–4918
DOI 10.1007/s00500-015-1778-2

METHODOLOGIES AND APPLICATION

High-dimension space projection-based biometric encryption for fingerprint with fuzzy minutia

Zhendong Wu[1,2] · **Bin Liang**[1] · **Lin You**[1] · **Zhihua Jian**[1] · **Jin Li**[3]

Published online: 14 July 2015

Abstract Biometric identification has caused a multitude of problems in the networking environment, such as the storage of user's biometric template and the leakage of user's privacy, therefore, biometric encryption has been the focus of the recent studies which are based on Fuzzy Vault, Fuzzy Commitment, and dynamic key generation. However, due to fuzzy information inherent biological characteristics, essentially deterministic analysis techniques, Fuzzy Vault, and Fuzzy Commitment have been accused of stored biometric templates and short keys. Fuzzy Information Processing needs suitable technology, such as fuzzy logic, to obtain better results. In this paper, we propose a new fingerprint encryption scheme which utilizes the high-dimension space projection. Unlike the reliance on biometric templates in Fuzzy Vault-based scheme or "encrypted" templates in Fuzzy Commitment-based scheme, this new scheme, similar to the dynamic biometric key generation scheme, protects biometric key in a polynomial, and hence saves "nothing" on the biometric characteristics. Thus, it integrates the advantages of Fuzzy Vault, Fuzzy Commitment, and dynamic key generation into one scheme.

Keywords Threshold (t, n) · Biometric encryption · Fuzzy information · High-dimension space projection quantization

Communicated by V. Loia.

✉ Zhendong Wu
wzd@hdu.edu.cn

Bin Liang
wz08lb@126.com

Lin You
youlin@hdu.edu.cn

Zhihua Jian
jianzh@hdu.edu.cn

Jin Li
jinli71@gmail.com

[1] School of Communication Engineering, Hangzhou Dianzi University, Hangzhou 310018, China

[2] Center for the Study of Language and Cognition, Zhejiang University, Hangzhou 310012, China

[3] School of Computer Science, Guangzhou University, Guangzhou 510006, China

1 Introduction

The rapid development of cloud computation and big data technology facilitates people's life, work, and study. The downside of this development is that the problems of information security and individual's privacy are becoming extremely serious. Biometric authentication can be used to ensure information security (Kikuchi et al. 2010; Choi et al. 2012; Tistarelli and Schouten 2011). However, it has a lot of problems, such as the finite number of biometric traits and the information leakage caused by stolen biometric templates. Taking these imperfections into consideration, biometric encryption (Bodo 1994) was proposed as an alternative. Compared with biometric authentication, apart from requiring sufficient similarity of biometric characteristics, biometric encryption can not only store a key securely and reliably, but also extract an identical and stable key dynamically. Therefore, it has better performance in terms of security, since it combines biometric identification with traditional cryptography (Cavoukian and Stoianov 2007). Recently, many researchers studied biometric encryption based on Fuzzy Vault (Uludag et al. 2005; Gaddam and Lal 2010; Li et al. 2010b), Fuzzy Commitment (Sutcu et al. 2008), and dynamic biometric key generation scheme (Li et al. 2011).

Biometric encryption based on Fuzzy Vault is a classic scheme. However, this scheme cannot work well in the network environment since the servers need to store biometric templates or the converted templates. Uludag et al. (2005) and Clancy et al. (2003) implement of fingerprint vaults, respectively, with the assumption that fingerprint features are pre-aligned. However, it is impractical. Then, helper data which contains maximum curvature points and maximum curvatures on fingerprint ridges were used to solve calibration problem (Nandakumar et al. 2007). Although alignment issues can be solved, the biometric cryptosystem still have security problems. Zhang et al. (2011) found that as long as two register vault templates and helper data are collected, the success rate of attacking a biometric cryptosystem can be up to 60 % in the case of nine order polynomial, such as when we analyze cross-matching loopholes of Fuzzy Vault scheme. This may lead to information leakage. Besides, a common problem is the storage of biometric templates. As noted in Scheirer and Boult (2007), if variable vaults of the same biometric characteristics are collected, user's biometric templates will be easily inferred by comparing these vaults. In order to better protect biological templates, cancelable biometrics technique is proposed (Rathgeb and Uhl 2011; Khan et al. 2015). Cancelable biometric template technology works through biological template deformation or salt, so the template cannot be restored, and information contained in biometric templates is protected from leakage. Although the deformable biological template can increase the security, it may also reduce the accuracy of biometric identification.

Another recent research topic is Fuzzy Commitment. It can deal with hamming errors between different biometric samples. It also demands a fixed-length binary biometric feature of high distinction. However, it is difficult to design an effective and stable key generation algorithm. The key generated by the algorithm is usually short and unstable. Sutcu et al. (2008) employed a user-specific cuboid to partition the minutia set, and they used principle component analysis (PCA) on the computed feature vectors to generate a binary output vector, which is combined with low-density parity check (LDPC) codes to obtain a secure fingerprint biometric. Nagar et al. (2010) extracted fingerprint features from minutias and ridges. Bringer et al. (2008) focuses on the selection of error correction code (ECC). Li et al. (2012) employed minutia triplets as the basic input features to extract feature strings and then used linear discriminant analysis (LDA) to reduce the dimension of these strings and to eliminate the correlation among fingers. Rathgeb et al. (2013) tried to use a feature fusion technology to achieve a higher efficiency of error correction code. Iris cryptosystems were also implemented (Zhou et al. 2012). Despite various efforts on this scheme, it still suffers from a short and unstable key. In addition, it has to save an "encrypted" template which is obtained by an XOR operation with the fix-length biometric feature and corresponding code-word. Usually, the code-word contains the secret and it is vulnerable to decodability attack (Kelkboom et al. 2011). Thus, this scheme cannot work very well in the network environment either.

Biometric encryption based on dynamic key generation is a promising scheme. It extracts a biometric key directly from the biometric template. The advantage of this scheme is that it does not need to store templates or biometric keys. Moreover, the dynamic keys binding to user's identity can work together with the current mainstream cloud storage security technologies, such as attribute-based encryption (Li et al. 2011), attribute-based signatures (Li and Kim 2010), attribute-based outsourcing (Li et al. 2014), and the derivative technology with attribute-based encryption (Castiglione et al. 2015; Wang et al. 2015; Esposito et al. 2013), providing more natural and flexible encryption methods. Li et al. (2014) proposed an outsourcing encryption scheme based on attribute-based encryption. If the attributes could be provided by dynamic biometric keys, the outsourcing encryption would use more flexible keys in a more natural way. Castiglione et al. (2015) proposed a cloud-based adaptive compression and secure management services for 3D healthcare data. If the services could bind user's identity in a more subtle way, the healthcare data would possess more security privacy. To generate a stable key, a biometric key requires highly consistent biometrics, even reproducible ones. However, biometric samples usually do not meet this requirement due to environmental and physiological factors. Hoque et al. (2005) partitioned feature space into subspaces and into cells with the purpose of deriving relatively long keys. Atah and Howells (2009) used a combination of stable features from the human voice to generate biometric keys directly with a novel method of feature concatenation. Sheng et al. (2008) modeled the intra and inter-user variation of statistical features extracted from metric samples by clustering the data into natural clusters using a fuzzy genetic clustering algorithm. Then, a reliable key was generated by selecting the most consistent features for each user individually. Li et al. (2010a) proposed a fuzzy keyword search technology over encrypted data. Esposito et al. (2015) proposed a service selection technology based on fuzzy logic. Lim et al. (2012) used a dynamic reliability-dependent bit allocation algorithm for biometric discretization to allocate bits dynamically to every feature element based on a binary reflected gray code. Although these attempts can extract a biometric key, they suffer from a short key or a relatively long key with high equal error rate (EER).

In this paper, we propose a fingerprint encryption scheme based on high-dimension space projection (HDSP). Unlike the reliance on templates in the Fuzzy Vault-based scheme or the "encrypted" templates in the Fuzzy Commitment-based scheme, it protects biometric key in a polynomial, and thus saves "nothing" about biometric characteristics, being simi-

lar to dynamic biometric key generation scheme. Moreover, HDSP can extract a relatively long key, achieve a higher secure performance, and therefore can be deployed on-line conveniently.

2 Models

2.1 Encryption fingerprint with minutia

Cryptographic keys extracted from biometric templates are named "Biometric keys". It lays out security considerations that not only can protect the information confidentiality, but also protect the privacy of biometric information by itself. In the current biometric key extraction, the fingerprint is the most versatile and accurate biometric feature. Normally, we use fingerprint minutia to extract "Biometric keys". But the minutias are not so stable, but always fuzzy. Extracting stable "Biometric keys" from fuzzy minutias is the goal of this article.

2.2 Threshold (t, n)

After preprocessing, the fingerprint thinning image could be got, and the minutias are extracted from the thinning image. Normally, the minutias of the same fingerprint with different sampling images have a large difference. That is, the minutias in a sample may not exist in the other sample, although they come from the same fingerprint. A fault-tolerant mechanism is required for extracting stable "Biometric keys" from fuzzy minutias to tolerate possible minutias that may not occur. Threshold (t, n) is a possible choice.

A threshold (t, n) scheme can be explained as follows: dividing a key (or something else) into n pieces, then it will be easily computed if any t or more pieces are given, and it will be undetermined if t−1 or fewer pieces are given. This is a useful scheme to share a key.

An order n polynomial $P(x)$ in a finite field GF(p) can be written as:

$$P(x) = K_0 + K_1 \cdot x + \cdots + K_n \cdot x^n \tag{1}$$

For a given set $T = \{x_i | i = 1, 2, \ldots, n+1\}$, $P(T) = \{P(x_i) | i = 1, 2, \ldots, n+1\}$ and a set

$$\mathrm{PV} = \{(x_i, P(x_i)) | \, i = 1, 2, \ldots, n+1\}$$

will be evaluated. As we can see, for another given set

$$\mathrm{PQ} = \left\{ \left(x_i', y_i\right) \middle| \, i = 1, 2, \ldots, n+1, \ldots, \gamma \right\},$$

the polynomial $P(x)$ is able to be reconstructed according to the set PQ despite some elements not lying on the polynomial,

Fig. 1 the schematic plot → the aligned fingerprint image

if and only if PQ $\supseteq$ PV . That is to say, if we traverse the whole set PQ, the set PV will be selected ultimately, and $P(x)$ will be obtained correctly. This can be regarded as a fault-tolerant mechanism of threshold (t, n).

2.3 Fingerprint image alignment

Strictly speaking, the biometric key needs to be extracted from blind fingerprints, so there is no reference template for feature alignment. And in actual situation, there are a lot of non-alignment phenomena such as translation and rotation in the fingerprint sample. This will make it very difficult to extract the stable key. The sample alignment is essential for the biometric key extraction task at this stage. It can bring us more consistency, and more advantageous for extracting "Biometric keys". So fingerprint image alignment is a critical technique in the now field of biometric encryption. One disadvantage of current alignment algorithms is that they need to store some Help Data, which could be the potential source of user's fingerprint information leakage. To avoid this, a novel alignment algorithm which is suitable for our biometric encryption scheme is presented below.

Our algorithm transforms all fingerprint images into a valid status. Figure 1 shows the schematic plot of this algorithm. The detailed steps are as follows:

(1) Calculate the horizontal and vertical translations between the core coordinate and the center of the fingerprint image. Then translate the minutia coordinates based on them.
(2) Scan each coordinate in the annulus whose center is the center of the image and radius is 15–60 pixs. Calculate the angle between the horizontal and the line, which connects the core coordinate to the scanned coordinate. Record the coordinate and the difference between the angle and the orientation of the scanned coordinate.
(3) Find the smallest difference and calculate the mean of corresponding coordinates. Then the rotation angle θ between the vertical and the line is calculated. The line connects the core coordinate to the mean coordinate. If

the smallest difference is close to 0, the rotation angle θ can be regarded as an accurate rotation parameter.

(4) Rotate the minutia coordinates (and orients, directions, if exist) according to the rotation angle θ. Then alignment fingerprint features are obtained.

This algorithm highly relies on the core coordinate. Thus, if the core coordinate is inaccurate, the alignment algorithm may be invalid.

2.4 Feature high-dimension space projection

As we know, a feature vector $\mathbf{t}$ from one fingerprint can be expressed in different forms (n feature vectors, for example, $\overrightarrow{t_i}, i = 1, 2, \ldots, n$) when noise is present within a certain range. Then, the idea is that if some transformation can convert these noisy feature vectors to a stable one, the stable biometric key will be extracted from the stable vectors. This can be described as Eq. (2), where $\mathrm{PRO_{MAT}}$ is a projection matrix:

$$\begin{bmatrix} \overrightarrow{t_1} \\ \overrightarrow{t_2} \\ \vdots \\ \overrightarrow{t_n} \end{bmatrix} \cdot \mathrm{PRO_{MAT}} = \begin{bmatrix} \overrightarrow{t} \\ \overrightarrow{t} \\ \vdots \\ \overrightarrow{t} \end{bmatrix} \tag{2}$$

In fact, it is impossible to calculate the exact noiseless feature vector $\mathbf{t}$. Moreover, extracting a number of noisy feature vectors from $\mathbf{t}$ is impractical. Alternatively, n random feature vectors based on the extracted noisy feature vector within a certain error band, a mean vector of these vectors, are generated. Let $\overrightarrow{t_i}$ and $\mathbf{t}$ be the random vectors and mean vector, respectively. Then, $\mathbf{t} = \sum_1^n \overrightarrow{\mathrm{t}_i} / n$ will be obtained. Figure 2 depicts the distribution of feature vectors before and after they are projected. The vectors tend to be projected to the mean vector by the matrix $\mathrm{PRO_{MAT}}$. Therefore, for a feature vector that is close to $\mathbf{t}$, it will tend to be projected to $\mathbf{t}$ within a slight tolerance boundary by corresponding $\mathrm{PRO_{MAT}}$. Normally, Eq. (2) may be contradiction, and the matrix $\mathrm{PRO_{MAT}}$ is the least square solution of linear equations or contradiction equations. It means that in the linear space of the known training samples, Eq. (2) can get the optimal solution based on Euclidean distance. Then, if the testing sample falls into the linear space of the training samples, we can get the optimal solution of the linear projection of the testing sample.

2.5 Fingerprint feature quantitation

Due to various environmental and physiological factors, the same feature would be expressed in different coordinates. This variance makes it difficult to extract the exactly identical minutia coordinates. Hence, either the polynomial $P(x)$ or the key is undetermined according to threshold (t, n). To obtain the identical minutia coordinates as much as possible, error correction code (ECC) and quantitation method are used to compensate deformation of minutia features and to stabilize them. In this paper, we use quantitation method to stabilize minutia coordinates. The quantitation formula can be written as Eq. (3).

The quantitation threshold D is the range of an interval that the number lies in. For example, if $D = 10$ and a minutia feature coordinate is (156, 39), 156 will be quantified to 159 as it is in the interval (154 164], and 39 will be changed to 38 for the interval (33 43]. At last, (156, 39) is quantified to (159, 38).

3 High-dimension space projection-based fingerprint cryptosystem implementation

As it has been described before, feature projection tends to fix fingerprint features. This is available to quantitate the noisy features, which originates from the identical features, to a stable one. Along with a fault-tolerant scheme threshold (t, n), a cryptosystem is implemented in this section. The fingerprint feature vector used in this system is defined as

$$\{x, y, \mathrm{dir}, O, d_1, d_2, d_3, d_4\},$$

where $\{x, y\}$ is the minutia coordinate of the feature, dir is the direction of the coordinate; O is the average orientation of the rectangle decided by the present coordinate and the fingerprint's core coordinate; d_1, d_2, d_3, d_4 are the distance from the current coordinate to the center, the top left, the top right, and the bottom left of a fingerprint image, respectively. The system is composed of two parts: enrollment phase and recall phase. The overall flowchart of the cryptosystem is illustrated in Fig. 3.

$$\Lambda(x) = \begin{cases} \frac{D}{2} + (D+1) \cdot i & (D+1) \cdot i < x \le (D+1) \cdot i + D \, (i = 0, 1, \cdots) \ , \ mod\,(D, 2) = 0 \\ \frac{D-1}{2} + D \cdot i & (i-1) \cdot D + 1 < x \le D \cdot i \, (i = 1, 2, \cdots) \ , \ mod\,(D, 2) = 1 \end{cases} \tag{3}$$

3.1 Enrollment phase

In this phase, the true fingerprint feature vectors, denoted as

$$T = \left\{ \overrightarrow{t_i} = (x_i, y_i, \mathrm{dir}_i, o_i, d_{1i}, d_{2i}, \ d_{3i}, d_{4i}, \cdots) \mid i = 1, 2, \cdots, s \right\}$$

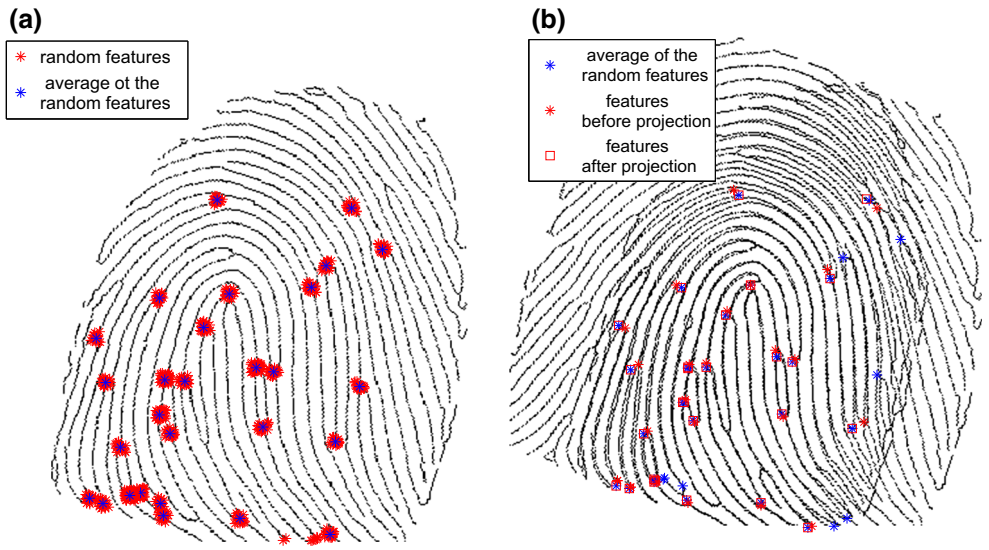

Fig. 2 Feature projection (**a**) random feature vectors and mean-feature vector (**b**) feature vectors after projection

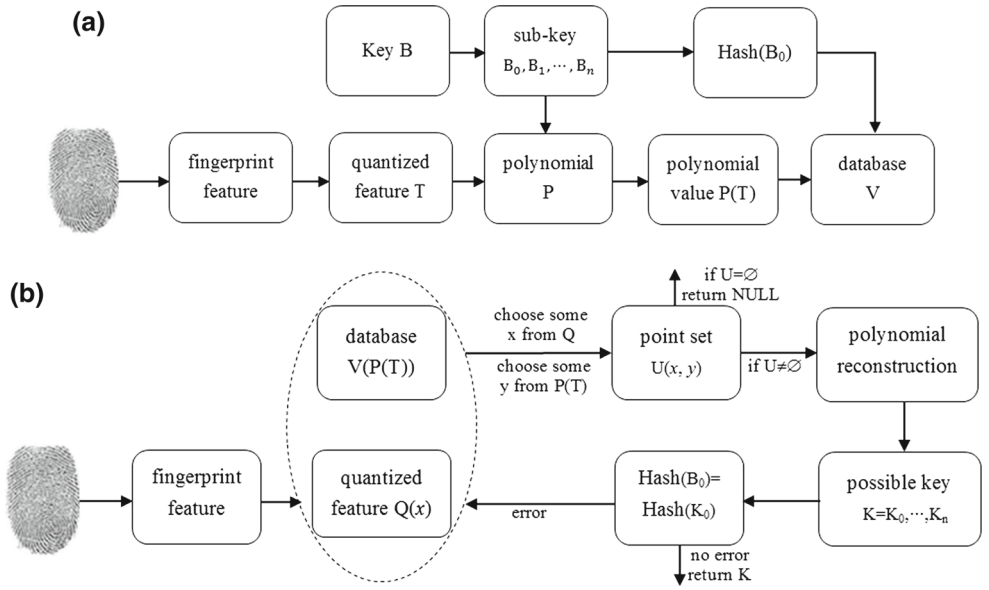

Fig. 3 Framework of threshold (t, n) based fingerprint cryptosystem. **a** Enrollment phase, **b** recall phase

with s coordinates, are extracted from two finger images of the same finger. The overall flowchart of the enrollment phase is illustrated in Fig. 3a. The detailed steps of enrollment algorithm are as follows.

(1) For each $\overrightarrow{t_i}$, it generates dim -1 feature vectors randomly based on $\overrightarrow{t_i}$ with a certain error boundary ROM. Get a matrix of dim $\times 8$ dimension.Then these dim vectors are extended to dim dimension using nonlinear method. Get a matrix of dim $\times$ dim dimension. Calculate the average vector-matrix of these extended vectors.

$$\overrightarrow{t_{ai}} = \{(x_{ai}, y_{ai}, \mathrm{dir}_{ai}, O_{ai}, d_{1ai}, d_{2ai}, d_{3ai}, d_{4ai}, \cdots) \,|\, i = 1, 2, \ldots, \mathrm{dim}\}$$

Calculate the projection square matrix $\mathrm{PRO}_{\mathrm{MAT}i}$ according to Eq. (2). Define the whole projection square matrixes as

$$PM = \{\mathrm{PRO}_{\mathrm{MAT}i} \,|\, i = 1, 2, \ldots, \mathrm{dim}\}$$

and the projected fingerprint feature vectors as

$$T = \left\{\overrightarrow{t_{ai}} = (x_{ai}, y_{ai}, \mathrm{dir}_{ai}, O_{ai}, d_{1ai}, d_{2ai}, d_{3ai}, d_{4ai}, \ldots) \,|\, i = 1, 2, \ldots, \mathrm{dim}\right\}$$

(2) Quantify the minutia coordinates of T. If the quantified feature is repeated, only one will be recorded. Due to the

relatively small variance of the same minutia feature of two finger images coming from the same fingerprint, it is possible to quantify the minutia coordinates to reduce difference and stabilize fingerprint features. Choose an appropriate quantization threshold D, and then quantify T according to Eq. (3). For convenience, the quantified features will be denoted as $T_q\{(x_i, y_i)\,|i = 1, 2, \ldots, s\}$ with s features.

(3) Concatenate the ordinate of minutia features, then the features are denoted as

$$T_c = \{XY_i = x_i|y_i; i = 1, 2, \ldots, s\}$$

For example, if $x_i = 100$ and $y_i = 200$, then turn x_i and y_i to binary, we can obtain

$$XY_i = x_i\,|y_i = (01100100)|\,(11001000) = (25800)_{10}$$

(4) $n + 1$ random numbers are generated as the sub keys, denoted as $B = \{B_i|i = 0, 1, \ldots, n\}$. And the sub keys determine the polynomial $P(x)$, which is expressed in Eq. (4)

$$P(x) = B_0 + B_1 \cdot x + \cdots + B_n \cdot x^n \tag{4}$$

(5) Calculate $h(B_0)$ with the one-way hash function h and the projection value $P(T_c) = \{P(t_i)\,|i = 1 \ldots, s\}$ according to Eq. (4).

(6) Delete all the intermediate data T, T_c, T_q, $\{B_i|i = 1, 2, \ldots, n\}$, $\{C_i|i = 1, 2, \ldots, n\}$, and only save the common parameters n, D, GF(p), P (T_c), $h(B_0)$ and PM in a database V.

3.2 Recall phase

In this phase, query fingerprint features are defined as

$$Q = \left\{\overrightarrow{q_i} = \left(x_i, y_i, \mathrm{dir}_i, O_i, d_{1i}, d_{2i}, d_{3i}, d_{4i}|i = 1, 2 \ldots, s'\right)\right\}$$

with s' features. The overall flowchart of the enrollment phase is illustrated in Fig. 3b. The detailed steps of recall algorithm are as follows:

(1) For each $\overrightarrow{q_i}$, the same nonlinear extension method in enrollment phase is applied, and the extended feature vector

$$\overrightarrow{q_{ei}} = \left\{(x_i, y_i, \mathrm{dir_i}, \mathrm{o_i}, \mathrm{d_{1i}}, \mathrm{d_{2i}}, d_{3i}, d_{4i}, \cdots)\,|i = 1, 2, \cdots s'\right\}$$

with dim elements is obtained. Then calculate the projected vector by the following equation:

$$\overrightarrow{q_{ei}} \cdot \mathrm{PRO}_{\mathrm{MAT_i}} = \overrightarrow{R_i}, i = 1, 2, 3 \ldots, s, \tag{5}$$

where

$$\overrightarrow{R_i} = \left(x_{aij}, y_{aij}, \mathrm{dir}_{aij}, O_{aij}, d_{1aij}, d_{2aij}, d_{3aij}, d_{4aij}, \ldots\right)$$

$i = 1, 2, \ldots, s'$, For each i, if $\overrightarrow{q_{\mathrm{ei}}} - \overrightarrow{R_i}$ is within the boundary, $\overrightarrow{q_{\mathrm{ei}}}$ will be considered as close to an average feature vector. And $\overrightarrow{R_i}$ can be used to recover the very key. For all the $\left\{\overrightarrow{R_i}|i = 1, 2, \ldots, s'\right\}$ within that boundary, record their minutia coordinates $\{(x_{ai}, y_{ai})\,|i = 1, 2, \ldots, s'\}$. For convenience, the recorded coordinates are overwritten as $Q = \{(x_i, y_i)\,|i = 1, 2, \ldots, \}$.

(2) Quantify the query minutia coordinates s Q by Eq. (3) with the parameter D got from the database V. If the quantified feature is repeated, only one will be recorded. For convenience, the quantified query features will be denoted as $Q_c = \{(x_i, y_i)\,|i = 1, 2, \ldots, m\}$ with m features.

(3) As we know, reconstruct an order n polynomial $P(x)$ needs at least n+1 points. Therefore, if $m < n+1$, $P(x)$ is unable to be reconstructed, then the system returns S =NULL and then aborts. Or else it will go on.

(4) Concatenate the ordinate of minutia features in Q_c. then the features are denoted as $Q_s = \{q_i = x_i|y_i; i = 1, 2, \ldots, m\}$.

(5) Select $n+1$ numbers from Q_s and $P(T)$ in the database V, respectively. And then a set

$$U = \left\{\left(q_i, P\left(t_j\right)\right)|i, j = 1, 2, \ldots, n+1\right\}$$

can be obtained. There are $(n+1)!$ combinations in total in U. Assure the polynomial reconstruction is

$$P^*(x) = k_0^* + k_1^* \cdot x + \cdots + k_n^* \cdot x^n \tag{6}$$

Then the possible key is $k' = k'_0||k'_1||\cdots||k'_n$ where $||$ denotes a connector. If $h(B_0) = h\left(k'_0\right)$, it means the key k is the key needed and it returns k. Or if $U = \emptyset$, return K = NULL. If $U = \emptyset$, repeat the step 5). Since only part of the information stores hash value, replacing the use of B with $B0$ can improve security and computing speed, but it will make the integrity testing of the hash function imprecise. However, if we allocate enough random keys, so that no fingerprints correspond to a user of the same $B0$, then the test will always return the correct key.

4 Experimental results and analysis

4.1 Databases and evaluation indicators

We use two fingerprint databases to evaluate the performance of the proposed fingerprint cryptosystem, including: (1) an

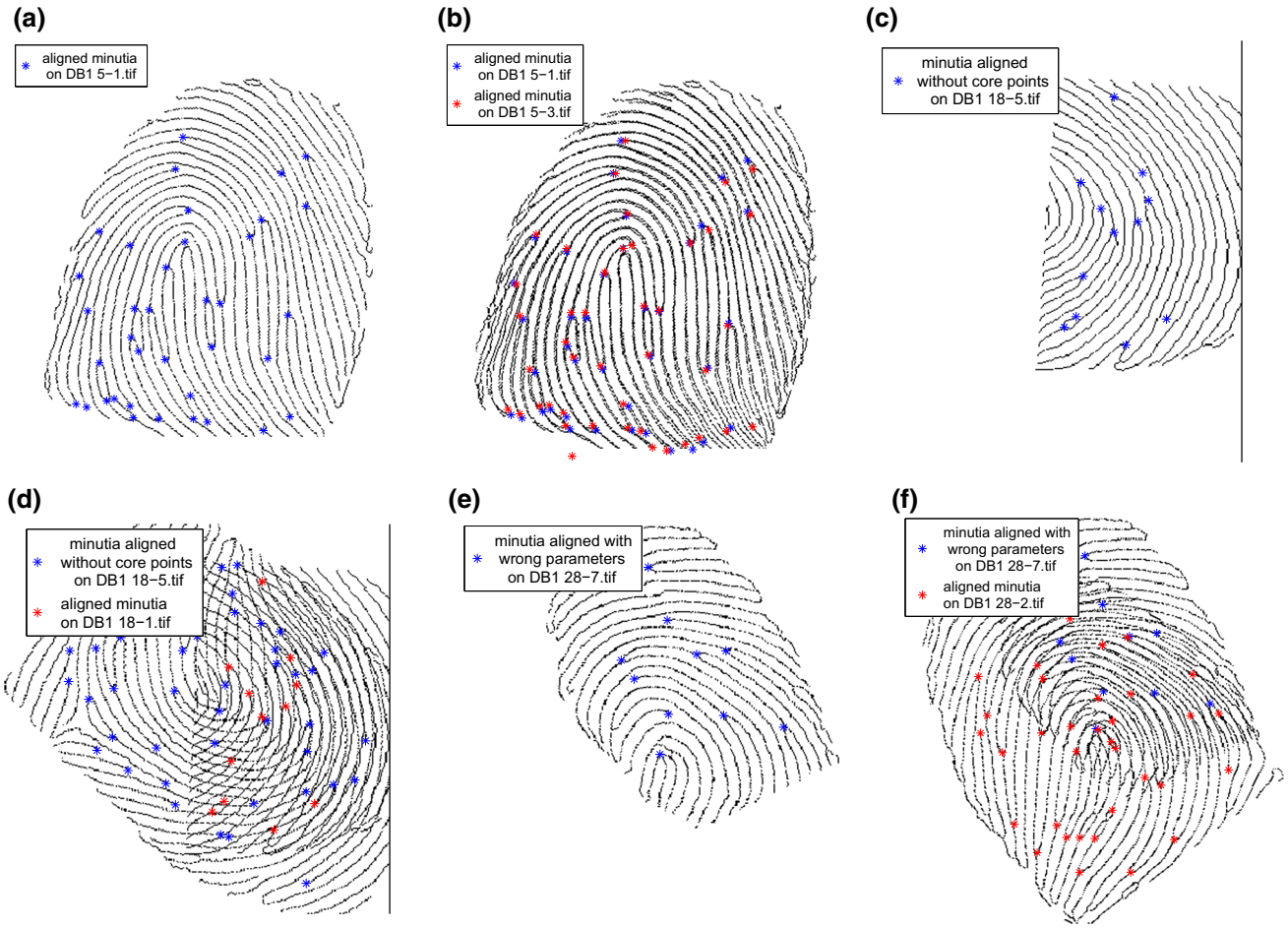

Fig. 4 Aligned image and aligned minutias on DB1 (**a**, **b**) accurate translation and rotation parameter (**c**, **d**) an inaccurate core coordinate (**e**, **f**) wrong rotation parameter

in-house database (SF for short); (2) FVC2002 DB1 database (DB1 for short). SF database is an in-house database with 42 fingers and 25 samples for each finger. DB1 is a public database with 100 fingers and eight samples for each finger. The image in DB1 database has a medium quality, and SF database has a relatively better quality. To assure the robustness of the fingerprint cryptosystem, common feature vector of one fingerprint is extracted from two samples whose common minutias are between 20–28 (if all are less than 20, search the maximum, if all are more than 28, search the minimum.) common feature vectors of each finger within the two databases are used in the enrollment phase, and the rest for the recall phase. To evaluate the performance, genuine accept rate (GAR) and false accept rate (FAR) are used as the main indicators.

4.2 Analysis of fingerprint image alignment

Figure 4a–f shows the alignment results. Figure 4a, c, e are single fingerprint images, (b, d, f) are the synthesis of two samples into one image to illustrate the different situations encountered during the process of fingerprint alignment. Figure 4a, b shows the aligned images and minutia (DB1 5_1.tif and 5_5.tif). We find that almost all the aligned minutias on 5_1.tif and 5_5.tif are in a small error boundary. However, the alignment algorithm in Sect. 2 highly relies on the core coordinate of fingerprint image. If the core coordinate is incorrect, the alignment algorithm will be invalid. There are $70\tilde{8}0$ inaccurate core coordinates on DB1 and 50 ones on SF. The inaccuracy of these core coordinates is the main reason why most unaligned feature vectors cannot be adjusted correctly, as the translation and rotation of these feature vectors are calculated according to the wrong core coordinates. Figure 4c, d show the aligned image and minutia with an inaccurate core coordinate (DB1 18_5.tif). Figure 4e, f show the aligned image and minutia with wrong rotation parameter (DB1 28_7.tif). There are two reasons that explain the incorrect rotation parameter: (1) the core coordinate is close to the edge of the corresponding fingerprint image, thus it cannot find a point in the fixed annulus to meet the requirement that the rotation angle needs to be close to 0; (2) As many points that make close to 0 may be found, the

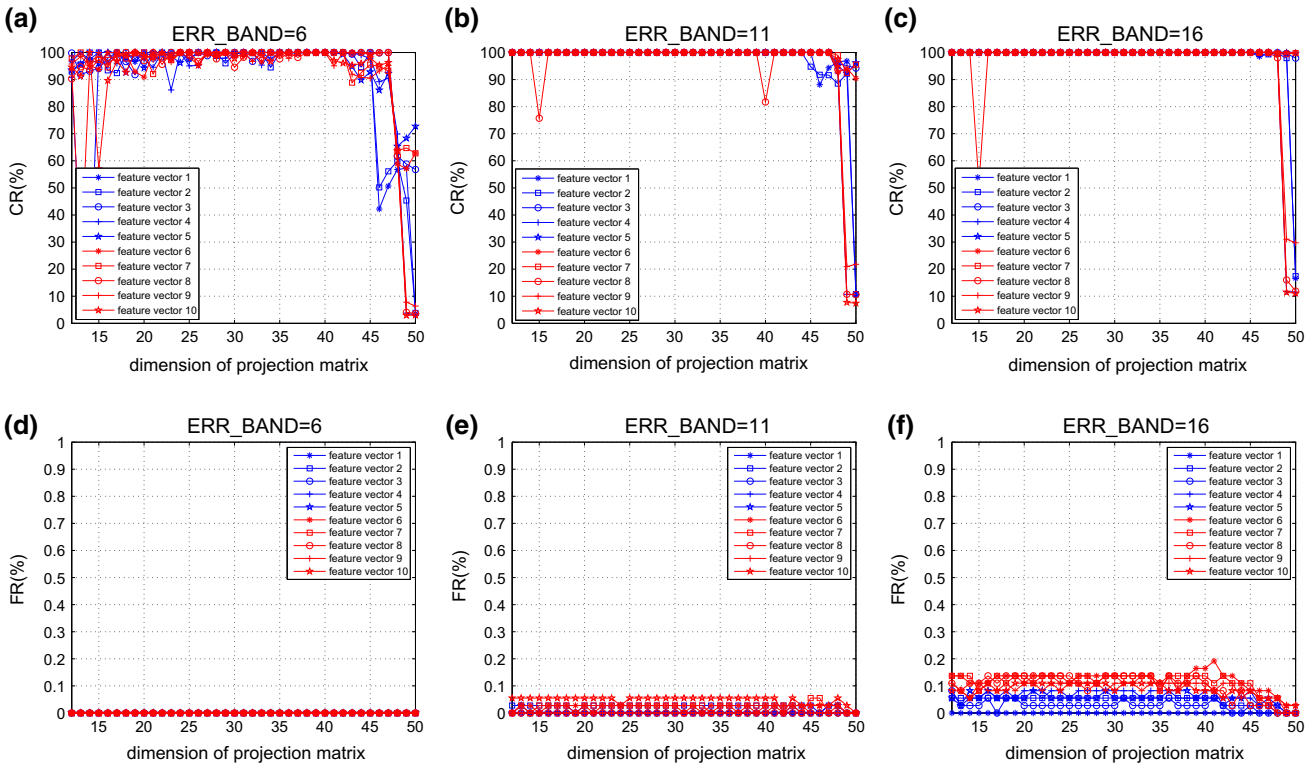

Fig. 5 Average GR(%) & FR with different ERR_BAND and dim

average coordinate that acts as a substitution may introduce some errors. Considering all the errors, a genuine alignment rate of DB1 and SF is about 80 and 95 %.

4.3 Analysis of HDSP performance

Equation (2) in Sect. 2.4 shows that if a feature vector is close to **t**, it tends to be projected to **t** by corresponding PRO_{MAT}. However, the PRO_{MAT} does not always work well, it may cause an incorrect projection. In this section, an experiment is conducted to test the correct rate (CR) and false rate (FR) of projected feature vectors that are close to **t**. In this experiment, two parameters, feature vector dimension dim and error boundary ERR_BAND are chosen. The former parameter determines the projection matrix of dimension, and the latter is the error boundary before and after projection to determine whether it is the same feature point.

Ten feature points are selected randomly from one finger as the base vectors to generate other vectors and then they are extended to dim dimension. For each feature vector, its projection matrix is calculated. Assuming that the generated vectors of the 10 base vectors are $\overrightarrow{t_{ij}}, i = 1, 2, \ldots, 10, j = 1, 2, \ldots,$ dim. To evaluate the correct rate, 10,000 random feature vectors are generated for each base vector and also they are extended to dim dimension. Assuming that these vectors are $\overrightarrow{pt_{ij}}, i = 1, 2, \ldots, 10, j = 1, 2, \ldots, 10,000,$ and those after projection are $\overrightarrow{pt_{pij}}, i = 1, 2, \ldots, 10, j = 1, 2, \ldots, 10,000$, respectively. CR is conducted with

$$\left|\overrightarrow{pt_{pij}} - \overrightarrow{pt_{ij}}\right| < \text{ERR\_BAND}$$

for each i when j ranges from 1 to 10,000. FR is calculated as follows: select 50 fingers and eight images of each finger, extract ten feature points from each finger, and extend feature vector to dim dimension. Supposing that the extended feature vectors of image j of finger m of i-th point is $\overrightarrow{pt_{mji}}, m = 1, 2, \ldots, 50, j = 1, 2, \ldots, 8, i = 1, 2, \ldots, 10$. Then the projection is calculated by $\overrightarrow{pt_{mji}} \cdot \text{PRO}_{\text{MAT}ik} = \overrightarrow{pt_{mjik}}, i = 1, 2, \ldots, 10, k = 1, 2, \ldots, 10, i \neq k, m = 1, 2, \ldots 50, j = 1, 2, \ldots, 8$. For each i, m, j, k and each vector, if $\left|\overrightarrow{pt_{mji}} - \overrightarrow{pt_{mjik}}\right| <$ ERR_BAND, the values of FR are added by one.

Figure 5a–c are CRs with different ERR_BAND and dim. These graphs show that a high value of CR lies in the dim interval [20,40]. That is to say, it is possible to obtain a high CR under this situation. Moreover, CR would increase with the growth of EER_BAND when dim lies in the approximate interval [20,40]. Figure 5d–f are FRs with different ERR_BAND and dim. Although FR increases as ERR_BAND increases, it is still less than 0.2 %. Comparing CRs with FRs, a compromise can be reached between FR and CR with ERR_BAND from 11 to 16.

4.4 Analysis of HDSP fingerprint cryptosystem

Actually, the base feature vectors extracted in the enrollment phase cannot be obtained in the recall phase. Thus, it needs to determine whether a feature vector is close to some vector using dynamic range. In addition, every element in the feature vector has different dynamic ranges and using only one parameter may lack flexibility. In this section, parameters D, ROM, X_Y, DIR, O, and DIS are designed to get a better GAR and a lower FAR. D denotes the quantitation parameter, ROM denotes the error boundary when generating random feature vectors, X_Y denotes the error boundary of x and y of feature vector, which is similar to the ERR_BAND, DIR, and O denoting dir and o of feature vector, respectively, and DIS are the thresholds of four distances in a raw feature vector $\{x, y, \text{dir}, \text{o}, d_1, d_2, d_3, d_4\}$. D, ROM, XY, DIR, O, and DIS are all distances, and measure the different parameter space, respectively. With the increase of distance, the probability of mapping the two points to one value is increased. It means the GAR will increase. But, meanwhile, the FAR will also increase. The best balance between GAR and FAR need to be obtained by experiment.

Figure 6 shows the distribution of GAR and FAR of SF with different parameters of D, ROM, X_Y, DIR, O, and DIS. From Fig. 6a–c, we can find that the method of high dimension space projection works well. A better performance can be achieved by adjusting the dimension range in [10,40]. Also, the performances of two databases are different in GAR, referring to Figs. 6a–c and 7a–c, although the parameters are roughly the same. The main reason is that SF has a relatively better image quality than that in the DB1. Thus, the fingerprint images can be aligned better and there are more matched features. From the genius alignment rate of DB1 and SF in Sect. 4.2 and the distribution of GAR in Fig. 5, we can infer that the GAR will perform better if the core points of these fingerprint images can be extracted more accurately and stably. The FAR of Fig. 7d–f are obviously higher than the FAR of Fig. 6. This is mainly due to the difference of image quality of two databases. We preprocess the images using the general method of fingerprint image preprocessing, and there are some "bubbles" in the results of DB1, which will bring many excrescent minutias. Too many minutias would seriously affect the FAR recognition of fingerprint. By adjusting the parameters we can obtain different recognition rates, but the rate of change is different in the two databases. DB1 changes more dramatically than SF does, which shows that the parameters of DB1 reach its critical case.

In addition, the key size in our algorithm is $16 \cdot (n+1)$. And this is superior to the key size of the present Fuzzy Commitment and dynamic key generation scheme.

5 Security analysis

In this section, brute force attack, cross-match of different databases from the same finger and the security of projection matrix will be discussed.

5.1 The security of projection matrix

Equation (2) shows that **t** can work as a noisy feature vector. It follows:

$$\overrightarrow{v} \cdot \text{PRO}_{\text{MAT}_i} = \overrightarrow{v} \tag{7}$$

That is, if a feature vector **v** satisfies:

$$\overrightarrow{v} \cdot \left(\text{PRO}_{\text{MAT}_i} - E\right) = \overrightarrow{0}, \tag{8}$$

where E is an identity matrix with the same dimension of PRO_MAT_i, it will be a possible vector to recover the very key. However, if $\text{PRO}_{\text{MAT}_i} - E$ is nonsingular and Eq. (8) would have only one solution with $\mathbf{v} = \mathbf{0}$. Thus, searching the feature vectors according to Eq. (8) would fail. Therefore, if the projection matrix is perturbation sensitive and nonsingular, the projection matrix is strong enough to resist attack.

5.2 Brute force attack

We assume an attacker try to recover the biometric key by brute force attack and he can obtain the parameters n, D, GF(p), and $P(T_c)$ in the database V. Firstly, the attacker calculates the numbers of possible quantified results in the interval [0,255] and $P(T_c)$. They are denote as ξ (the numbers under each parameter D can be calculated by Eq. (2)) and s, respectively. The analysis is shown in Table. 1. As the finger features must be quantified, the attacker may regard all the possible quantified results in the interval [0,255] as valid user's features. That is, ξ^2 features are stored equivalently in the database V after the attacker combining the ordinates. Then, the attacker picks out $n+1$ features and $P(t_i)$ to recover the key. For an order n polynomial, $P(T)$ and all the possible quantified results, its combinations can be as much as $\binom{\xi^2}{n+1} \cdot \binom{s}{n+1} \cdot (n+1)!$. But only $\binom{s}{n+1}$ combinations can recover the very biometric key. The probability to recover it by trying one combination is

$$\frac{\binom{s}{n+1}}{\binom{\xi^2}{n+1} \cdot \binom{s}{n+1} \cdot (n+1)!} = \frac{1}{\binom{\xi^2}{n+1} \cdot (n+1)!}$$

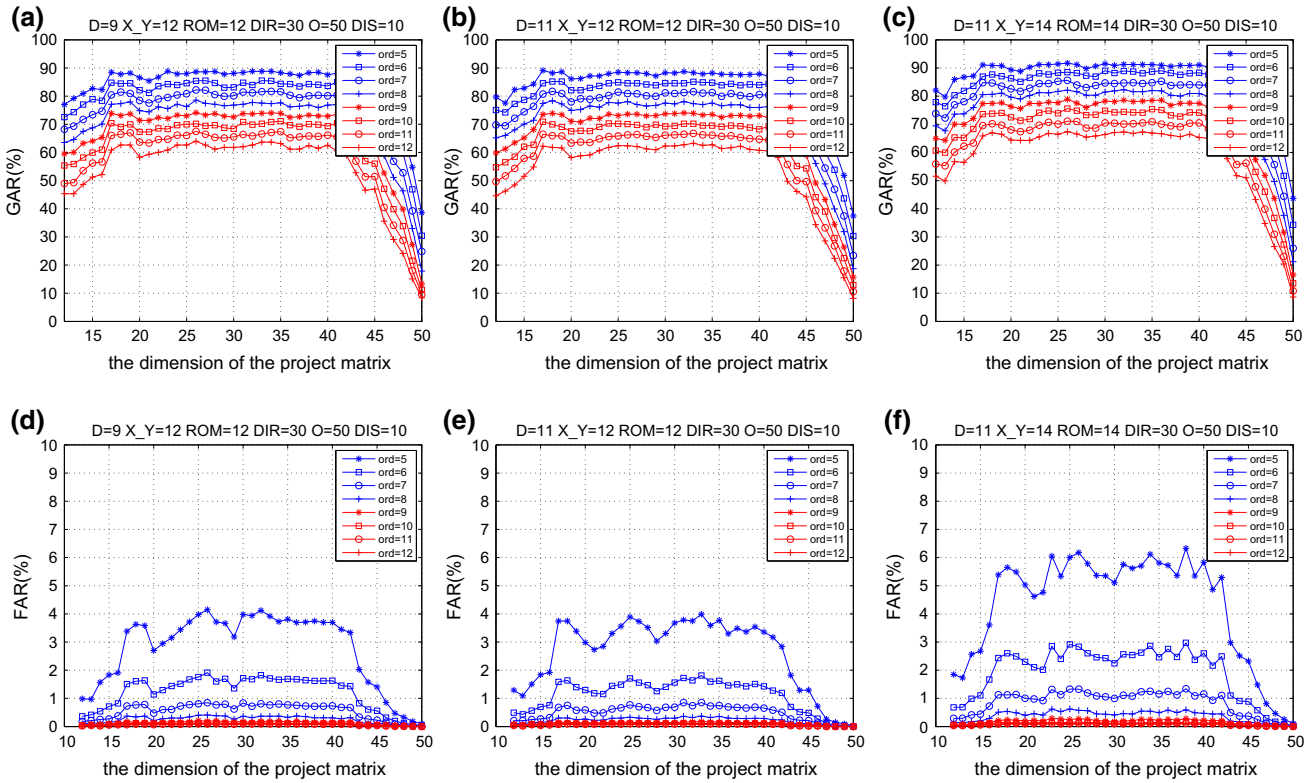

Fig. 6 Distribution of GAR and FAR of SF with the certain parameters *D*, *X-Y*, ROM, DIR, O, and DIS

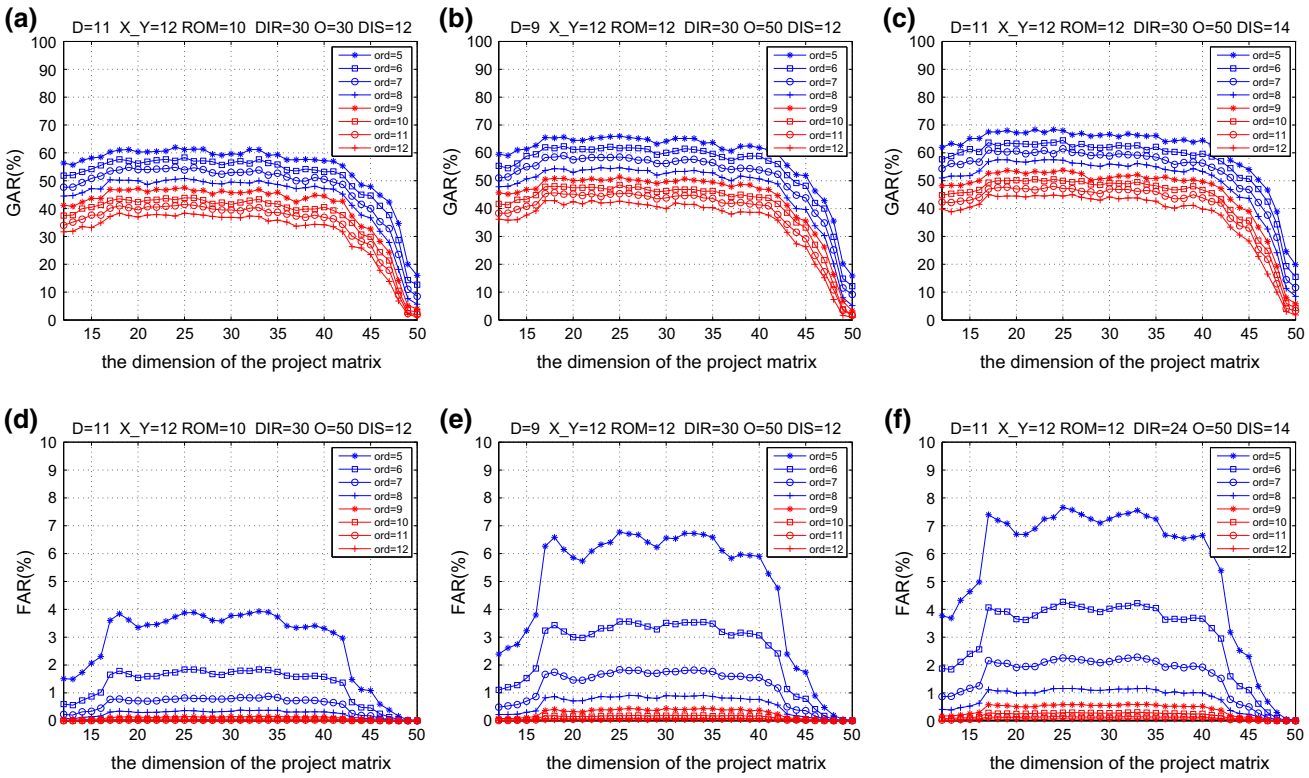

Fig. 7 Distribution of GAR and FAR of FVC2002.DB1 with the certain parameters *D*, *X-Y*, ROM, DIR, O, and DIS

Table 1 Average time (year) to recover the very key under each ξ when $s = 20$, according to Nandakumar et al. (2007)

	$n = 5$	$n = 6$	$n = 7$	$n = 8$	$n = 9$	$n = 10$	$n = 11$
$\xi = 23$	1.1×10^8	5.8×10^{10}	3.0×10^{13}	1.6×10^{16}	8.2×10^{18}	4.3×10^{21}	2.2×10^{24}
$\xi = 28$	1.2×10^9	9.2×10^{11}	7.2×10^{14}	5.6×10^{17}	4.3×10^{20}	3.3×10^{23}	2.6×10^{26}
$\xi = 36$	2.4×10^{10}	3.1×10^{13}	4.0×10^{16}	5.2×10^{19}	6.7×10^{22}	8.6×10^{25}	1.1×10^{29}
$\xi = 37$	3.4×10^{10}	4.6×10^{13}	6.3×10^{16}	8.6×10^{19}	1.2×10^{23}	1.6×10^{26}	2.1×10^{29}

And the average times to recover it is $\binom{\xi^2}{n+1} \cdot (n+1)!$.

Based on the fact that an attacker will take 13 years when trying 2.5×10^9 times using a 3.4 GHz processor (Nandakumar et al. 2007), the average time is shown in Table. 1. when $s = 20$. Apparently we can find that if ξ increases, its average attempts will increase, and it will be more difficult to recover the very key. Table. 1. shows that it is very difficult that the attacker recovers the biometric key by brute force attack.

5.3 Cross-match attack

With the random strings generated in enrollment phase and non-storage of any templates, cross-match of different databases from a same finger is impossible. That is, the sub keys (or say the coefficient) extracted from the same feature are different because of the random string. This leads to different polynomials. We assume two databases V_1 and V_2 are from the same finger, and their corresponding polynomials are $P_1(x)$ and $P_2(x)$, respectively. Because of the random strings, the order n polynomials related to V_1 and V_2 will be different with the probability of $1 - \frac{1}{(2^{16})^{n+1}}$, where n is the quantity of the coefficients. The probability shows that the two polynomials are hardly equivalent. Therefore, $P_1(T_c) \cap^{P_2} (T_c) = \emptyset$ or $P_1(T_c) \cap^{P_2} (T_c) \neq \emptyset$ will be possible when the same feature set T is given. Conversely, the same value in $P_1(T_c)$ and $P_2(T_c)$ may be mapped by the different features. Therefore, P (T_c) stored in the database V is a bunch of meaningless numbers for cross-match attackers. In other words, the proposed scheme can achieve a higher security performance.

6 Conclusions

Fuzzy Vault-based biometric encryption has the risk of information leakage for the stored biometric templates and is not suitable for the on-line case. Fuzzy Commitment-based biometric encryption needs to store an "encrypted" template, and it has a short and unstable key and may be unavailable on-line either. Although some dynamic biometric key generation schemes need to store neither templates nor secrets, they suffer from a low number of effective bits. The proposed biometric cryptosystem does not need to store templates, and its biometric key is relatively long. If the projection matrix is strong enough to resist attack, it will be available to work on-line since its database does not leak information about biometric templates. But it suffers from a not so satisfied GAR, high time complexity and a coarse quantitation. For these deficiencies, our future work will focus on more accurate quantitation scheme and a stronger projection matrix design.

Acknowledgments This research was supported by Zhejiang Province Science and Technology Innovation Program under Grant Number (2013TD03) and the National Science Foundation of China (No. 61272045), (No. 61201301), (No. 61472091). And the Guangzhou Zhujiang Science and Technology Future Fellow Fund (Grant No. 2012J2200094), Distinguished Young Scholars Fund of Department of Education.(No. Yq2013126), Guangdong Province. We gratefully acknowledge funding support from the Major Program of National Social Science Foundation of China (No.11&ZD088).

References

Atah JA, Howells G (2009) Key generation in a voice based template free biometric security system. In: Biometric ID management and multimodal communication, Springer, pp 170–177

Bodo A (1994) Method for producing a digital signature with aid of a biometric feature. German Patent DE 42(43):908

Bringer J, Chabanne H, Cohen G, Kindarji B, Zémor G (2008) Theoretical and practical boundaries of binary secure sketches. IEEE Trans Inf Forensics Secur 3(4):673–683

Castiglione A, Pizzolante R, De Santis A, Carpentieri B, Castiglione A, Palmieri F (2015) Cloud-based adaptive compression and secure management services for 3d healthcare data. Futur Gener Comput Syst 43:120–134

Cavoukian A, Stoianov A (2007) Biometric encryption: a positive-sum technology that achieves strong authentication. Secur Priv, p 15

Choi K, Toh KA, Uh Y, Byun H (2012) Service-oriented architecture based on biometric using random features and incremental neural networks. Soft Comput 16(9):1539–1553

Clancy TC, Kiyavash N, Lin DJ (2003) Secure smartcardbased fingerprint authentication. In: Proceedings of the 2003 ACM SIGMM workshop on biometrics methods and applications, ACM, pp 45–52

Esposito C, Ficco M, Palmieri F, Castiglione A (2013) Interconnecting federated clouds by using publish-subscribe service. Clust Comput 16(4):887–903

Esposito C, Ficco M, Palmieri F, Castiglione A (2015) Smart cloud storage service selection based on fuzzy logic. IEEE Trans Comput. doi:10.1109/TC.2015.2389952

Gaddam SV, Lal M (2010) Efficient cancelable biometric key generation scheme for cryptography. IJ Netw Secur 11(2):61–69

Hoque S, Fairhurst M, Howells G, Deravi F (2005) Feasibility of generating biometric encryption keys. Electron Lett 41(6):309–311

Kelkboom EJ, Breebaart J, Kevenaar TA, Buhan I, Veldhuis RN (2011) Preventing the decodability attack based cross-matching in a fuzzy commitment scheme. IEEE Trans Inf Forensics Secur 6(1):107–121

Khan SH, Akbar MA, Shahzad F, Farooq M, Khan Z (2015) Secure biometric template generation for multi-factor authentication. Pattern Recognit 48(2):458–472

Kikuchi H, Nagai K, Ogata W, Nishigaki M (2010) Privacy-preserving similarity evaluation and application to remote biometrics authentication. Soft Comput 14(5):529–536

Li J, Kim K (2010) Hidden attribute-based signatures without anonymity revocation. Inf Sci 180(9):1681–1689

Li J, Wang Q, Wang C, Cao N, Ren K, Lou W (2010a) Fuzzy keyword search over encrypted data in cloud computing. In: INFOCOM, 2010 Proceedings IEEE, IEEE, pp 1–5

Li J, Wang Q, Wang C, Ren K (2011) Enhancing attribute-based encryption with attribute hierarchy. Mob Netw Appl 16(5):553–561

Li J, Huang X, Li J, Chen X, Xiang Y (2014) Securely outsourcing attribute-based encryption with checkability. IEEE Trans Parallel Distrib Syst 25(8):2201–2210

Li P, Yang X, Cao K, Tao X, Wang R, Tian J (2010b) An alignment-free fingerprint cryptosystem based on fuzzy vault scheme. J Netw Comput Appl 33(3):207–220

Li P, Yang X, Qiao H, Cao K, Liu E, Tian J (2012) An effective biometric cryptosystem combining fingerprints with error correction codes. Expert Syst Appl 39(7):6562–6574

Lim MH, Teoh ABJ, Toh KA (2012) An efficient dynamic reliability-dependent bit allocation for biometric discretization. Pattern Recognit 45(5):1960–1971

Nagar A, Rane S, Vetro A (2010) Alignment and bit extraction for secure fingerprint biometrics. In: IS&T/SPIE electronic imaging, international society for optics and photonics, pp 75,410N–75,410N

Nandakumar K, Jain AK, Pankanti S (2007) Fingerprint-based fuzzy vault: implementation and performance. IEEE Trans Inf Forensics Secur 2(4):744–757

Rathgeb C, Uhl A (2011) A survey on biometric cryptosystems and cancelable biometrics. EURASIP J Inf Secur 1:1–25

Rathgeb C, Uhl A, Wild P (2013) Iris-biometric fuzzy commitment schemes under image compression. In: Progress in pattern recognition, image analysis, computer vision, and applications, Springer, pp 374–381

Scheirer WJ, Boult TE (2007) Cracking fuzzy vaults and biometric encryption. In: Biometrics symposium, 2007, IEEE, pp 1–6

Sheng W, Howells G, Fairhurst M, Deravi F (2008) Template-free biometric-key generation by means of fuzzy genetic clustering. IEEE Trans Inf Forensics Secur 3(2):183–191

Sutcu Y, Rane S, Yedidia JS, Draper SC, Vetro A (2008) Feature transformation of biometric templates for secure biometric systems based on error correcting codes. In: Computer vision and pattern recognition workshops, 2008. CVPRW'08. IEEE on computer society conference, IEEE, pp 1–6

Tistarelli M, Schouten B (2011) Biometrics in ambient intelligence. J Ambient Intell Humaniz Comput 2(2):113–126

Uludag U, Pankanti S, Jain AK (2005) Fuzzy vault for fingerprints. In: Audio-and video-based biometric person authentication, Springer, pp 310–319

Wang XA, Ma J, Yang X (2015) A new proxy re-encryption scheme for protecting critical information systems. J Ambient Intell Humaniz Comput pp 1–13. doi:10.1007/s12652-015-0261-3

Zhang R, Liu E, Zhao H, Pang L (2011) Improved cancelable fingerprint fuzzy vault system. J Xidian Univ 38(4):173–180

Zhou X, Kuijper A, Busch C (2012) Retrieving secrets from iris fuzzy commitment. In: 2012 5th IAPR international conference on biometrics (ICB), IEEE, pp 238–244