Extracting Discriminative Binary Template for Face Template Protection

Yicheng Feng

Abstract

This paper addresses the security issues of the face biometric templates stored in a database. In order to improve the security level of the stored face templates, cryptographic techniques are commonly employed. Since most of such techniques require a binary input, thresholding is usually employed to binarize the real valued face features. While binary templates are obtained, the discriminability of the original templates may be affected and so is the recognition performance. In order to overcome this limitation, this paper proposes a new approach to discriminative extract binary templates from original face templates. A projection is applied before thresholding such that the thresholding can fit the projected data distribution better, which makes the extracted binary templates more discriminative. A discriminability index is constructed to measure the discriminability of the extracted binary templates and the projection matrix is optimized. The proposed method is evaluated on three public domain databases, namely FERET, CMU-PIE and FRGC. Experimental results show that, in comparison to the existing thresholding-based methods, the proposed method improves the GAR by around $11\% - 13\%$ *at a FAR of* 1%.

1 Introduction

Biometrics is a reliable, robust and convenient way for person authentication [9, 10, 6]. With the success of the biometrics research in the last two decades, several large scale recognition systems have been successfully deployed. With the growing use of biometrics, there is a rising concern about the security and privacy of the stored biometric templates (which refer to a set of features extracted from raw biometric data) stored in a database or a smartcard. Recent studies [11] show that simple attacks on a biometric system, such as hill climbing, are able to recover the raw biometric data from a stolen biometric template. Moreover, the attacker may be able to make use of the stolen template to access the system or cross-match across databases. A comprehensive analysis of eight types of attacks [6] on a biometric system has been reported. Therefore, biometric template security [9, 10, 6, 18] has been an important issue in deploying a biometric system.

In order to overcome the security and privacy problems [6, 7, 9], a number of biometric template protection algorithms have been reported in the last few years. 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 biometric template, the transformed/encrypted template 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. Generally speaking, the transform-based approach suffer from a trade-off between dicriminability and security of the transformed templates while biometric cryptosystems may provide both enhanced security and acceptable discriminability. Cryptographic technique is employed in the last step in the cryptosystems approach to enhance the security of the templates. Error-correcting schemes are applied to deal with the intra-class variance thus the discriminability of the encrypted templates will not degrade too much. However, most of the protection algorithms in biometric cryptosystem approach require a binary template or integer data (most for the fuzzy vault scheme [5]) for encryption. That means, the input template has to be converted into a binary or integer template before encryption. The fuzzy vault scheme is mostly applied to fingerprint. The minutiae of fingerprint are unordered set of points, which the fuzzy vault scheme is just able to encrypt. For face recognition algorithms like PCA or LDA schemes, the situation is much different because the extracted templates are ordered real value vectors with large range, which causes a problem to apply the fuzzy vault scheme. On the other hand, there are schemes to transform the original face templates into binary strings for protection. In order to satisfy the input requirements, thresholding is a typically employed in existing algorithms [12, 13, 14, 15, 16, 17]. While the binary template can be obtained, some useful and discriminative information in the original (real valued) template may be lost after thresholding leading to degraded matching accuracy (discriminability) [12, 13]. And existing approaches lack of discriminability evaluation of the thresholding process.

In view of the limitations on existing thresholding-based algorithms, this paper provide a discriminability optimized thresholding scheme. Directly optimize the thresholds may have some problems. For security issues, the information content of the quantized binary templates should be maximized such that a brute-force attack will be most expensive to break the system. To maximize the information content of the bits, the thresholds should be set to make half of the corresponding bits in the binary templates to be “0” and half to be “1”. Which implies the thresholds should be set as the mean value of the corresponding elements of the original templates (called max-entropy rule). It means that the thresholds are already determined. And binary templates with maximum information content imply that they already get certain discriminability, which are shown clearly in experiments. So thresholds optimization makes little sense. Here we choose another approach. Before thresholding, an orthonormal projection process is applied to rotate the data position such that the thresholding fits the data distribution better. And the projection matrix is optimized such that the output binary templates after thresholding will have maximum discriminability. And we turn the whole thresholding process including the projection into an approximation process. The output binary templates is approximated by a linear function of original templates such that we can avoid the non-linear thresholding function while constructing the objective function.

The rest of this paper is organized as follows. Section 2 gives a brief review of the existing binarization schemes. Our proposed algorithm is then reported in Section 3. Experimental results and analysis are given in Sections 4. Finally, Section 5 gives the conclusion.

2 Review on Existing Binarization Schemes

A comprehensive survey on biometric template protection has been reported in [9, 8]. This section mainly reviews the existing binarization schemes in biometric template protection.

Monrose $et\ al.$ [12, 13] first proposed a binarization scheme, who described a cryptographic key generation scheme from biometrics. The biometric data is transformed into a binary string called ”feature descriptor” which has relatively small intra-class variance and large between-class variance. This binary string is generated by a thresholding technique based on mean and standard deviation of the biometric data.

Goh and Ngo [15] proposed a biohashing scheme. In their scheme, the original templates are first transformed using random mapping. Each element of the transformed template is thresholded to either 0 or 1, thus converting the template into binary form. Different versions of Biohashing algorithms [14, 17] have been proposed in the last few years but their binarization schemes are quite similar.

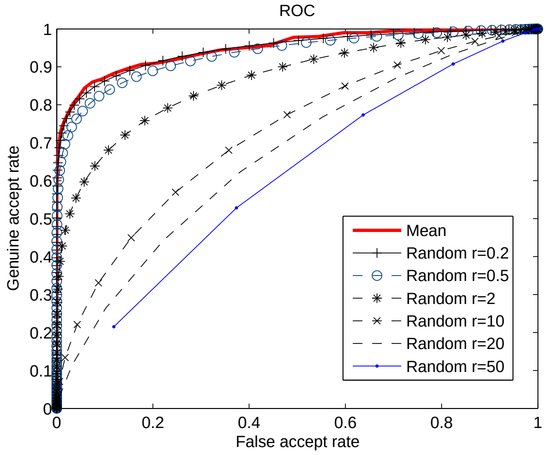

Figure 1. The effect of adjusting thresholds: The curves show performances in FERET database of binary templates extracted by a straightforward thresholding process. symbol “Mean" means the thresholds are set as the mean values of the original templates. “Random" means thresholds are randomly chosen with a Gaussian distribution, in which the mean of the distribution is mean of the original templates and variance of the distribution is r times of the variance of the original templates. The random thresholds are generated 100 times for experiment and the curves represent the average performance. It shows clearly that when r is closer to 0 the performance is higher, implying that thresholds with mean value of original templates are already close to optimal.

Chang and Roy [22] proposed a fingerprint binarization scheme. It draws lines to the original fingerprint. The difference of numbers of minutiae lying in each side of the lines are used to construct an integer number vector. PCA is applied decorrelate the vector and the deccorelated vector is then quantized to bits with threshold 0.

Nagar $et\ al.$ [24, 23] proposed a scheme to generate binary helper data from fingerprint. It first extract a minutiae descriptor from the minutiae of fingerprint with the orientation and ridge frequency information, then quantizes each element of the minutiae descriptor into 2^5 or 2^4 integer values. These values are converted to bits with Gray codes (because the converted Gray codes from neighbor value differ only 1 bit).

Kevenaar $et\ al.$ [25] proposed a face template binarization scheme. It statistically estimates the means and variances of the enrolled original templates. The means are used as thresholds for quantization which converts the original templates into bits. The scheme also constructs an index with the means and variances to measure the reliability of the converted bits. Only bits with high reliability are selected for recognition.

Linnartz and Tuyls [26] proposed a quantization scheme with quantization step size q. It did not directly extract the bit string from biometric data but link a binary secret to the biometric data. A binary string secret S is selected. For the original reference template X, a helper data W is constructed such that elements of $S+W$ are multiples of q and the quantized values of $X+W$ have the same parities of corresponding bits in S. For a testing template Y, $W+Y$ is quantized to an integer number vector and then the parity of the vector is checked to extract a bit string S'. S and S' is compared for decision.

Feng $et\ al.$ [19, 20, 21] proposed a class-distribution-preserving transform (CDP transform) for binarization. Distinguishing points are determined. The distance between each distinguishing point and the face template is calculated and thresholded. With optimal positioning of the distinguishing points, the transform optimizes the discriminability of the binary strings. Thus the discriminability-preserving ability of this scheme is justified. But note that it is still a thresholding-based approach.

3 Proposed Algorithm: Optimized Thresholding with Projection

3.1 Basic Idea

To binarize a biometric feature vector (template), the most straightforward and common way is to apply a quantization/thresholding algorithm, which is applied in most of the existing schemes. However, lack of discriminability analysis is still a problem for the binarization approach. Do the transformed binary templates have enough discriminability or not? In this paper we do a primary research to propose a new thresholding approach which provides a discriminability index, and optimize such index to extract binary templates having optimal discriminability. As we mentioned before, directly optimizing the thresholds is not effective. Here we choose a different approach (illustrated in Figure 2) to solve the optimization problem. Before thresholding, an orthonormal projection is applied to transform the input original templates into a new domain. The projection matrix is optimized such that the output binary templates after thresholding have optimal discriminability. The thresholds in the thresholding is fixed with the max-entropy rule.

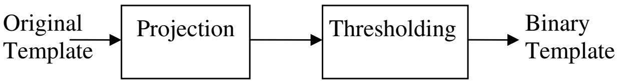

Figure 2. The proposed thresholding scheme with projection.

An orthonormal projection will not change the relative positions of the original templates in the feature space, thus after the projection the discriminability of the original templates will be completely preserved. This is quite important because if the projection cause discriminability lost, it will contradict our purpose to maximize the discriminability. And an orthonormal projection can rotate the original templates in feature space to new positions, such that the thresholding can fit the data distribution better. The two-dimensional case is illustrated in Figure 3. The thresholding in each dimension (abscissa and ordinate) can be treated as a line that divides the feature space into two halves. Given two classes of points as Figure 3 illustrated, directly adjust the abscissa and ordinate thresholds would not completely separate the two classes (illustrated in sub-figure (a)). However, after projection, thresholding in abscissa can completely separate the two classes (illustrated in sub-figure (b)). So, when the thresholds are fixed, we can optimize the projection matrix such that the thresholding well separates different classes of rotated templates, therefore provide high discriminability to the quantized binary templates. Since thresholding will cause rather complicated formulation to construct the discriminability index, the whole thresholding (including projection) process is converted to an approximation process in latter steps.

3.2 Optimize Thresholding through Approximation

To measure the discriminability of the binary templates, we adopt the concept of within-class variance and between-class variance. The within-class variance and the between-class variance are defined for each class respectively, therefore different optimized projection matrices can be chosen for different classes such that they can fit the distribution of each class better than a matrix for all classes. The within-class variance and between-class variance are defined as Equation (1) and (2):

$$D_W(\Omega) = \frac{\sum_{p\in\Omega} ||w(p) - w_\Omega||^2}{\sum_{p\in\Omega} 1} \tag{1}$$

$$D_B(\Omega) = \frac{\sum_{p\notin\Omega} ||w(p) - w_\Omega||^2}{\sum_{p\notin\Omega} 1} \tag{2}$$

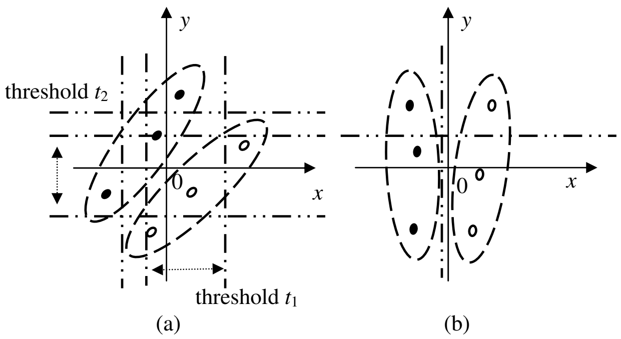

Figure 3. The effort of projection to thresholding (a) scenario of the original thresholding, thresholds are directly optimized; (b) thresholding with projection. The solid an hollow points represent templates in two different classes (illustrated with ellipses). It shows clearly that the thresholding in (b) separates two classes.

where $D_W(\Omega)$ and $D_B(\Omega)$ represents the within-class variance and between-class variance of class Ω. w_Ω denotes the reference binary template that represents Ω and $w(p)$ represents the binary template transformed from original template p. With the defined within-class variance and between-class variance, the discriminability is defined as $D_B(\Omega) - D_W(\Omega)$. Here w_Ω needs to be determined and we do not directly generate w_Ω from the training data, but find it through optimization together with the projection matrix such that the discriminability is maximized. Assume the projection matrix is M. $D_B(\Omega)$ and $D_W(\Omega)$ are subject to M and w_Ω. Then the objective function is constructed as Equation (3):

$$(M_{opt}, w_{opt}) = \underset{M, w_\Omega}{\operatorname{argmin}}(D_B(\Omega) - D_W(\Omega)) \qquad (3)$$

As we mentioned before, directly optimize such an objective function is unconvenient because of the thresholding function. So before optimization, we first turn the thresholding function to an approximation process.

In our proposed thresholding process, the original template p is first projected with an orthonormal matrix M to $u = M^T p$, and then thresholded with Equation (4):

$$b_j = \begin{cases} 1 & : \ h_j \geq t_j \\ -1 & : \ h_j < t_j \end{cases} \qquad (j = 1, 2 \ldots k) \qquad (4)$$

where h_j are components in u. Because of the max-entropy rule, t_j should be the mean value of h_j. A normalization process with Equation (5) can turn the equation to be simpler with all thresholds 0:

$$q = \frac{p - \overline{p}}{||p - \overline{p}||} \qquad (5)$$

where $\overline{p}$ denotes the mean vector of all the original templates. With such a normalization, Equation (4) turns to Equation (6):

$$b_j = \begin{cases} 1 & : \ a_j \geq 0 \\ -1 & : \ a_j < 0 \end{cases} \qquad (j = 1, 2 \ldots k) \qquad (6)$$

where a_j are the elements in $v = M^T q$.

Equation (6) can be further turned into an approximation process (Equation (7)):

Given an unit vector $v = (a_1, a_2 \ldots a_k)$, it is transformed to a binary string $w'(v) = (b_1, b_2 \ldots b_k) \in \{1, -1\}^k$ with the following equation:

$$w'(v) = \underset{\omega}{\operatorname{argmin}} ||v - \frac{\omega}{\sqrt{k}}|| \qquad (7)$$

where $w'(v)$ represents the binary template transformed from v.

It can be shown that

$$||v - \frac{\omega}{\sqrt{k}}||^2 = \sum_{j=1}^{k} (a_j - \frac{b_j}{\sqrt{k}})^2$$

is minimum if and only if b_j satisfies Equation (6). So Equation (7) is equivalent to Equation (6), thus $w'(v) = w(p)$.

Substitute $v = M^T q$ into Equation (7), we have

$$w''(q) = w'(v) = \underset{\omega}{\operatorname{argmin}} ||M^T q - \frac{\omega}{\sqrt{k}}|| \qquad (8)$$

So our binarization scheme $w(p)$ can be turned into an approximation process $w''(q)$ as Equation (8) describes. Denote $e = \frac{M w_\Omega}{\sqrt{k}}$. Therefore,

$$\begin{aligned} & ||w_\Omega - w''(q)|| \\ = \ & \sqrt{k} || \frac{w_\Omega}{\sqrt{k}} - M^T q + M^T q - \frac{w''(q)}{\sqrt{k}} || \\ \leq \ & \sqrt{k} (|| \frac{w_\Omega}{\sqrt{k}} - M^T q || + || M^T q - \frac{w''(q)}{\sqrt{k}} ||) \\ \leq \ & 2\sqrt{k} || \frac{w_\Omega}{\sqrt{k}} - M^T q || \\ = \ & 2\sqrt{k} || \frac{M w_\Omega}{\sqrt{k}} - M M^T q || \\ = \ & 2\sqrt{k} || e - q || \end{aligned} \qquad (9)$$

Notice here $||\frac{w_\Omega}{\sqrt{k}} - M^T q||$ is no less than $||\frac{w''(q)}{\sqrt{k}} - M^T q||$ because $\frac{w''(q)}{\sqrt{k}}$ is closet to $M^T q$ (Equation (8)).

On the other side,

$$\begin{aligned} & ||w_\Omega - w''(q)|| \\ = \;& \sqrt{k}||\frac{w_\Omega}{\sqrt{k}} - M^T q + M^T q - \frac{w''(q)}{\sqrt{k}}|| \\ \leq \;& \sqrt{k}||\frac{w_\Omega}{\sqrt{k}} - M^T q|| + \sqrt{k}||M^T q - \frac{w''(q)}{\sqrt{k}}|| \\ = \;& \sqrt{k}||q - e|| + \sqrt{k}||M^T q - \frac{w''(q)}{\sqrt{k}}|| \end{aligned}$$

implies

$$\begin{aligned} & |(||w_\Omega - w''(q)|| - \sqrt{k}||q - e||)| \\ \leq \;& \sqrt{k}||M^T q - \frac{w''(q)}{\sqrt{k}}|| \end{aligned} \tag{10}$$

Here $||M^T q - \frac{w''(q)}{\sqrt{k}}||$ is minimized with optimal $w''(q)$. To let the approximation error as small as possible, Equation (5) makes $M^T q$ an unit vector and scalar $\sqrt{k}$ is used such that $\frac{w''(q)}{\sqrt{k}}$ is a normalized vector too. Without $\sqrt{k}$,

$$||M^T q - w''(q)|| \geq ||w''(q)|| - ||M^T q|| = \sqrt{k} - 1$$

causing large approximation error. With $\sqrt{k}$, $\frac{w''(q)}{\sqrt{k}}$ is a unit vector thus the norms of $M^T q$ and $\frac{w''(q)}{\sqrt{k}}$ provide no contribution to the approximation error, resulting in small approximation error. Then Equation (10) implies that

$$||w_\Omega - w''(q)|| \approx \sqrt{k}||\frac{w_\Omega}{\sqrt{k}} - M^T q|| \tag{11}$$

Denote $m = \sum_{p \notin \Omega} 1$ and $n = \sum_{p \in \Omega} 1$. Substitute Equation (9) into Equation (1) and (11) into (2). Since $w''(q) = w'(v) = w(p)$, we have

$$D_W(\Omega) = \frac{\sum\limits_{p\in\Omega} ||w(p) - w_\Omega||^2}{n} \leq 4k \frac{\sum\limits_{q\in\Omega} ||q - e||^2}{n} \tag{12}$$

and

$$D_B(\Omega) = \frac{\sum\limits_{p\notin\Omega} ||w(p) - w_\Omega||^2}{m} \approx k \frac{\sum\limits_{q\notin\Omega} ||q - e||^2}{m} \tag{13}$$

Equation (13) implies that a large $D'_B(\Omega)$ leads to large $D_B(\Omega)$. Equation (12) indicates that a small $D'_W(\Omega)$ leads to a small $D_W(\Omega)$. So the objective function is changed to

$$(M_{opt}, w_{opt}) = \underset{M, w_\Omega}{\text{argmin}}(D'_B(\Omega) - D'_W(\Omega)) \tag{14}$$

where

$$D'_B(\Omega) = \frac{\sum\limits_{q\notin\Omega} ||q - e||^2}{m}$$

and

$$D'_W(\Omega) = \frac{\sum\limits_{q\in\Omega} ||q - e||^2}{n}$$

.

From Equation (14) we can see that with the approximation process, the terms including binary templates in the objective functions are replaced by terms with original templates. Therefore, there is no need to consider the correlation between bits and the thresholding process, thus solves the problems mentioned before. In Equation (14) M and w_Ω are only related to e. So we can first optimize e and then choose M and w_Ω to fit the optimized e:

$$e_{opt} = \underset{e}{\text{argmax}}(\frac{\sum\limits_{q\notin\Omega} ||q - e||^2}{m} - \frac{\sum\limits_{q\in\Omega} ||q - e||^2}{n}) \tag{15}$$

with constraint

$$||e|| = 1 \tag{16}$$

.

The constraint is because $e = M\frac{w_\Omega}{\sqrt{k}}$ is an unit vector. After e_{opt} is found, w_{opt} is randomly generated and M_{opt} is constructed such that $\frac{M_{opt} w_{opt}}{\sqrt{k}} = e_{opt}$.

However, we should notice that the objective function can only ensure $D'_B(\Omega)$ is large, but can not ensure that $D'_W(\Omega)$ is small. So the following constraint

$$e \cdot q_\Omega \geq \theta \tag{17}$$

is required. Here q_Ω is the normalized mean vector of Ω and θ is a threshold smaller than 1. This constraint ensure that $D'_W(\Omega)$ is relatively small.

3.3 Solve the Objective Function

The objective function is first expanded as follows:

$$\begin{aligned} & \sum_{q\notin\Omega}(||q - e||^2)/m - \sum_{q\in\Omega}(||q - e||^2)/n \\ = \;& (\sum_{q\notin\Omega}||q||^2 - 2\sum_{q\notin\Omega} q \cdot e)/m + ||e||^2 \\ & -(\sum_{q\in\Omega}||q||^2 - 2\sum_{q\in\Omega} q \cdot e)/n - ||e||^2 \\ = \;& (\sum_{q\notin\Omega}||q||^2/m - \sum_{q\in\Omega}||q||^2/n) \\ & +2(\sum_{q\in\Omega} q/n - \sum_{q\notin\Omega} q/m) \cdot e \end{aligned}$$

where $\cdot$ denotes inner product. Assume r is the normalized vector of $\sum_{q\in\Omega} q/n - \sum_{q\notin\Omega} q/m$. According to the above equation, to maximize the objective function is equivalent to maximize $r \cdot e$.

Apply the Gram-Schmidt algorithm to randomly generate an orthonormal basis $\{s_1, s_2 \dots s_k\}$ where $s_1 = q_\Omega$. Denote Q to be the orthonormal matrix using $s_1, s_2 \dots s_k$ as columns. Then From Equation (17), we have

$$e^T Q Q^T q_\Omega \geq \theta \tag{18}$$

Denote $z = Q^T e = (x_1, x_2 \dots x_k)$. According to Equation (16), z is an unit vector. Because of the construction of Q, $Q^T q_\Omega = (1, 0, \dots 0)$. Substitute these two terms into Equation (18), we have

$$x_1 \geq \theta \tag{19}$$

Denote $l = Q^T r = (y_1, y_2 \dots y_k)$. Since r is an unit vector, l is an unit vector. Therefore,

$$r \cdot e = r^T Q Q^T e = l^T z = \sum_{j=1}^{k} x_j y_j$$

And

$$\begin{aligned}
\sum_{j=1}^{k} x_j y_j &= x_1 y_1 + \sum_{j=2}^{k} x_j y_j \\
&\leq x_1 y_1 + (\sum_{j=2}^{k} x_j^2 \cdot \sum_{j=2}^{k} y_j^2)^{\frac{1}{2}} \\
&= x_1 y_1 + \sqrt{1 - x_1^2}\sqrt{1 - y_1^2}
\end{aligned}$$

$\sum_{j=1}^{k} x_j y_j$ will achieve its maximum if and only if

$$x_j = \frac{\sqrt{1 - x_1^2}}{\sqrt{1 - y_1^2}} y_j \qquad j = 2 \dots k$$

Denote $\phi = \arccos x_1$, $\varphi = \arccos y_1$. Then,

$$\begin{aligned}
\sum_{j=1}^{k} x_j y_j &= x_1 y_1 + \sqrt{1 - x_1^2}\sqrt{1 - y_1^2} \\
&= \cos\phi \cos\varphi + \sin\phi \sin\varphi \\
&= \cos(\varphi - \phi)
\end{aligned}$$

From Equation (19), we have the restriction for ϕ:

$$0 \leq \phi \leq \arccos\theta$$

If $\varphi \leq \arccos\theta$ (that is, $y_1 \geq \theta$), set $\phi = \varphi$, $r \cdot e$ can achieve the maximum value 1. If $\varphi > \arccos\theta$, Then $0 < \varphi - \phi < \pi$. $\cos(\varphi - \phi)$ will be maximum when $\varphi - \phi$ is minimum. So ϕ should be $\arccos\theta$. That is, $r \cdot e$ will achieve its maximum when

$$x_1 = \begin{cases} y_1 & : \quad y_1 \geq \theta \\ \theta & : \quad y_1 < \theta \end{cases}$$

and

$$x_j = \frac{\sqrt{1 - x_1^2}}{\sqrt{1 - y_1^2}} y_j \qquad j = 2 \dots k.$$

Then, $e_{opt} = Qz = Q[x_1, x_2 \dots x_k]^T$.

After e_{opt} is found, w_Ω is randomly generated and M_{opt} is constructed such that $M_{opt} w_\Omega / \sqrt{k} = e_{opt}$.

3.4 Algorithm Implementation

In enrollment,

- Normalize the original feature vectors with Equation (5).
- Construct M and w_Ω for class Ω.
- w_Ω is used as the reference template and encrypted with the fuzzy commitment scheme for protection.
- The encrypted w_Ω and M are stored in database.

In authentication,

- when input a query p, the corresponding M is released.
- Normalize p to q with Equation (5). And then q is projected to $v = M^T q$.
- Threshold q to $w''(q) = (b_1, b_2 \dots b_k)$ with Equation (6).
- Compare $w(p)$ with w_Ω to make a decision.

4 Experimental Results

In the experiments, three public databases namely CMU PIE, FERET and FRGC are applied to evaluate the performance of our proposed thresholding with orthonormnal projection (TOP) algorithm. The Fisherface [1] scheme is used for feature extraction. The detailed parameters settings are shown in Table 1, where n_c is the number of individuals in the database, n_p denotes the total number of images from each individual, n_t is the number of images used for training from each individual. In our TOP algorithm, the paramter θ is chosen 0.8.

Table 1. The experiment settings

Database	n_c	n_p	n_t
CMU PIE	68	105	10
FERET	250	4	2
FRGC	350	40	5

Our proposed TOP algorithm is also compared with the recent developed Random Multi-space Quantization

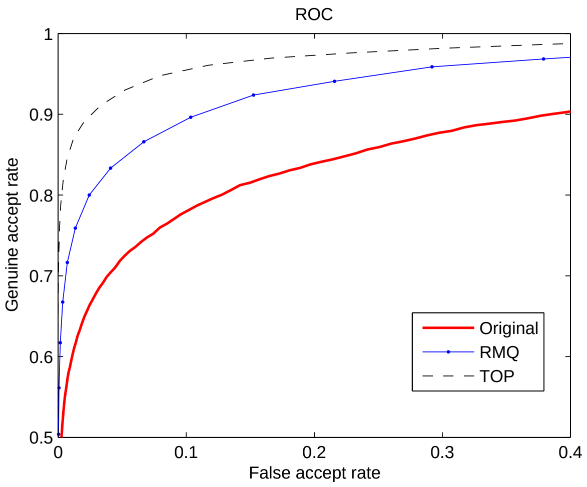

Figure 4. Experimental results on the CMU PIE database.

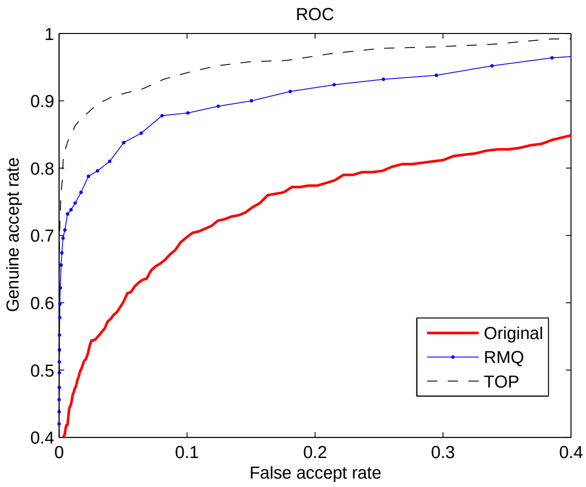

Figure 5. Experimental results on the FERET database.

Table 2. The GARs(%) of the experiments with fixed FAR=0.01

Database	Original	RMQ	TOP
CMU PIE	59.26	73.53	85.99
FERET	45.47	74.01	85.09
FRGC	26.28	67.40	78.35

Table 3. The EERs(%) of the experiments

Database	Original	RMQ	TOP
CMU PIE	17.32	10.37	6.30
FERET	21.66	11.29	7.44
FRGC	31.75	13.38	10.05

(RMQ) algorithm. The RMQ algorithm, as a binarization scheme, is also used for face recognition and has similar procedure with our scheme. It also does a projection before thresholding, though the purpose of the projection is much different from ours. For fair comparison, we use a randomly generated orthonormal matrix in the projection step of the RMQ algorithm, therefore the projection step does not reduce the dimension of the face template and will not cause information lost. So it will get higher performance than a normal RMQ algorithm.

The experimental results are shown in Figure 4-6. Symbol "Original" represents the original Fisherface algorithm without protection. The figures show clearly that our algorithm outperforms the RMQ algorithm. Following the popular setting, we fix the FAR at 0.01 and compare the GAR. The results are shown in Table 2. We also measure the equal error rate (EER) and the results are shown in Table 3. In all three databases, our proposed method gives the highest GAR and lowest EER.

The security of the proposed scheme depends on the security of the extracted binary templates, which are protected by cryptosystems scheme like the fuzzy commitment scheme. Since this part is not our main concern, we assume the protecting scheme provides enough security thus it is very computationally expensive for attackers to extract the binary templates from the protected data. Then attackers would have to guess the binary template with what they got. It should be mentioned that the projection matrix M_{opt} is unprotected and may be exposed to attackers. However, this would not help the attackers because the only relationship between M_{opt} and w_Ω is equation $M_{opt}w_\Omega/\sqrt{k} = e_{opt}$, where e_{opt} is kept secret to attackers. And as we know, w_Ω is randomly generated. So the entropy of w_Ω is the length of it, that is, k bits. And this is also the security level of our proposed algorithm.

5 Conclusions

This paper has proposed a new method to generate a binary face template from a real valued face template. The original face templates are first projected with an orthonormal matrix, and then thresholded to binary templates. The orthonormal matrix is optimized such that the extracted binary templates can get highest discriminability. Three public domain available face databases have been used to evaluate the proposed method. The experimental results show that the proposed method has good performance and outperforms the RMQ algorithm for comparison. The security of

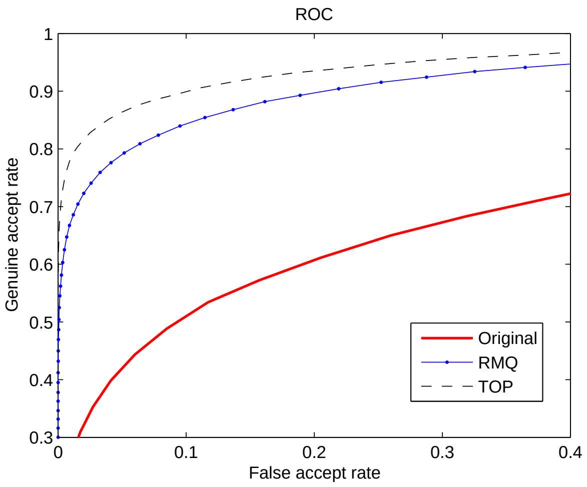

Figure 6. Experimental results on the FRGC database.

the proposed algorithm is just the length k of the extracted binary template, which is quite sufficient when k is large (say 40).

References

[1] 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.

[2] S Z Li, "Face Recognition Based on Nearest Linear Combinations," *Computer Vision and Pattern Recognition*, pp. 839-844,1998.

[3] R L Rivest, "The MD5 Message-Digest Algorithm," *RFC1321, Network Working Group, MIT Laboratory for Computer Science and RSA Data Security, Inc.*, 1992.

[4] J Daugman, "The Importance of Being Random: Statistical Principles of Iris Recognition," *Pattern Recognition*, Vol. 36, No. 2, pp. 279-291, 2003.

[5] A Juels and M Sudan. "A Fuzzy Vault Scheme", *IEEE International Symposium on Information Theory*, 2002.

[6] 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.

[7] 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.

[8] A K Jain, K Nandakumar and A Nagar, "Biometric template security," *EURASIP Journal on Advances in Signal Processing*, Vol. 8, 2008.

[9] 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.

[10] A K Jain, A Ross and S Pankanti, "Biometrics: A Tool for Information Security", *IEEE Transactions on Information Forensics and Security*, Vol. 1, No. 2, pp. 125-143, 2006.

[11] A Alder, "Images can be regenerated from quantized biometric match score data", *Proceedings of Canadian conference of Electrical and Computer Engineering*, pp. 469-472, 2004.

[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] A Teoh, D Ngo and A Goh, "Biohashing: Two Factor Authentication Featuring Fingerprint Data and Tokenised Random Number," *Pattern Recognition*, vol. 37, no. 11, pp. 2245-2255, Nov. 2004.

[15] A Goh and D C L Ngo, "Computation of cryptographic keys from face biometrics", *in Proc. 7th IFIP TC6/TC11 Conf. Commun. Multimedia Security*, vol. 22, pp. 1-13, 2003.

[16] D Ngo, A Teoh and A Goh, "Biometric Hash: High-Confidence Face Recognition", *IEEE transactions on circuits and systems for video technology*, vol. 16, no. 6, 2006.

[17] 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.

[18] N Ratha, S Chikkerur, J Connell and R Bolle, "Generating Cancelable Fingerprint Templates", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.29, no.4, pp. 561-572, 2007.

[19] Y C Feng, P C Yuen and A K Jain, "A Hybrid Approach for Face Template Protection," *Proceedings of SPIE Defense and Security Symposium*, 2008.

[20] 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.

[21] 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),* pp. 636-645, 2007.

[22] E C Chang and S Roy, "Robust extraction of secret bits from minutiae," *in Proceedings of 2nd International Conference on Biometrics*, pp. 750C759, 2007.

[23] A Nagar, K Nandakumar and A K Jain, "A Hybrid Biometric Cryptosystem for Securing Fingerprint Minutiae Templates," *Pattern Recognition Letters,* 2009.

[24] A Nagar, K Nandakumar and A K Jain, "Securing fingerprint template: Fuzzy vault with minutiae descriptors," *International Conference on Pattern Recognition,* pp. 1-4, 2008.

[25] T A M Kevenaar, G J Schrijen, M Veen, A H M Akkermans, "Face recognition with renewable and privacy preserving binary templates," *IEEE Workshop on Automatic Identification Advanced Technologies*, pp. 21-26, 2005.

[26] J P Linnartz, P Tuyls, "New Shielding Functions to Enhance Privacy and Prevent Misuse of Biometric Templates," *Audio and Video-Based Biometric Person Authentication,* 2003.