FACE BASED BIOMETRIC AUTHENTICATION WITH CHANGEABLE AND PRIVACY PRESERVABLE TEMPLATES

Yongjin Wang, K.N. Plataniotis

The Edward S. Rogers Sr. Department of Electrical and Computer Engineering,
University of Toronto,
10 King's College Road, Toronto, ON, Canada, M5S 3G4

ABSTRACT

Changeability, privacy protection, and verification accuracy are important factors for widespread deployment of biometrics based authentication systems. In this paper, we introduce a method for effective combination of biometrics data with user specific secret key for human verification. The proposed approach is based on discretized random orthonormal transformation of biometrics features. It provides attractive properties of zero error rate, and generates revocable and non-invertible biometrics templates. In addition, we also present another scheme where no discretization procedure is involved. The proposed methods are well supported by mathematical analysis. The feasibility of the introduced solutions on a face verification problem is demonstrated using the well known ORL and GT database. Experimentation shows the effectiveness of the proposed methods comparing with existing works.

1. INTRODUCTION

Traditional methods of identity verification are based on knowledge (e.g., passwords), or possession factors (e.g., ID cards) [1]. Such methods afford low level of security since passwords can be forgotten, acquired by covert observation, while ID cards can be lost, stolen, and forged. Biometrics based authentication systems confirm an individual's identity based on the physiological and/or behavioral characteristics of the individual. Biometrics based method provides direct link between the service and actual user. With biometrics, there is nothing to lose or forget, and it is relatively difficult to circumvent [2].

A biometrics verification system is a one-to-one match that determines whether the claim of an individual is true. A feature vector $\mathbf{x}_P$ is extracted from the biometrics signal of the authentication individual $\mathcal{U}^{'}$, and compared with the stored template $\mathbf{x}_I$ of the claimed identity $\mathcal{U}$ through a similarity function S. The evaluation of a verification system can be performed in terms of hypothesis testing [3]: $\mathbf{H}_0$: $\mathcal{U}^{'} = \mathcal{U}$, the claimed identity is correct, $\mathbf{H}_1$: $\mathcal{U}^{'} \neq \mathcal{U}$, the claimed identity is not correct. The decision is made based on the system threshold t: $\mathbf{H}_0$ is decided if $S(\mathbf{x}_P, \mathbf{x}_I) \leq t$ and $\mathbf{H}_1$ is decided if $S(\mathbf{x}_P, \mathbf{x}_I) > t$. A verification system makes two types of errors: false accept (deciding $\mathbf{H}_0$ when $\mathbf{H}_1$ is true), and false reject (deciding $\mathbf{H}_1$ when $\mathbf{H}_0$ is true). The performance of a biometrics verification system is usually evaluated in terms of false accept rate (FAR, $P(\mathbf{H}_0|\mathbf{H}_1)$), false reject rate (FRR, $P(\mathbf{H}_1|\mathbf{H}_0)$), and equal error rate(EER, operating point where FAR and FRR are equal). The FAR and FRR are closely related functions of the system decision threshold t.

While biometrics technology provides various advantages, there exist some major problems. 1. *Changeability*: Biometrics can not be easily changed and reissued if compromised due to the limited number of biometrics traits that human has. Ideally, just like password, the users should use different biometrics representation for different applications. When the biometrics template in one application is compromised, the biometrics signal itself is not lost forever and a new biometrics template can be issued [2]. 2. *Privacy*: Biometrics data reflects the user's physiological/behavior characteristics, if the storage device of biometrics templates is compromised, the user's privacy may be revealed. The biometrics templates should be stored in a format such that the user's privacy is preserved even the storage device is compromised. 3. *Accuracy*: Unlike knowledge or token based systems where exact match can be obtained, biometrics systems are based on fuzzy match due to the noisy nature of biometrics data. This fuzzyness deteriorates the performance of biometrics systems, and in general zero error rate can not be achieved by using biometrics alone. This characteristic of biometrics limits the widespread deployment in large scale and high security.

Existing solutions for changeable and privacy preservable biometrics are intentional transformation [3] or binding of biometrics with random cryptographic keys [2]. The major challenge in the former lies in the difficulty of preserving the verification performance in the transformed domain, while the latter in the error tolerant capability to retrieve the key from noisy biometrics data. A common problem with existing works is the lack of strong verification accuracy. In this paper, we propose an approach for strong combination of biometrics with user specific secret key to generate changeable and privacy preservable biometrics, while producing zero error rate. To elaborate our approach, we also discuss another scheme where no discretization is applied on the transformed features. In this scheme, the template has the same level of security as that of the secret key, but it provides good property that exactly the same performance can be preserved as the original features in the stolen key scenario.

In this paper, we demonstrate the analysis in a face verification scenario due to high user acceptability, easy to capture, and low cost properties of face biometrics. The proposed framework can find wide applications in physical access control, ATM, and computer/network login. However, the methods are general enough and can also be used in conjunction with other biometrics signals. The remainder of this paper is organized as follows. In section 2, we review the related works. Section 3 introduces proposed methods and

1-4244-1549-7/07/$25.00

provides probabilistic analysis. Experimental results along with detailed discussion are presented in Section 4. Finally, conclusion and future works are provided in Section 5.

2. RELATED WORKS

A number of research works have been proposed in recent years to address the changeability and privacy problems of biometrics systems. Among the earliest efforts, Soutar et al [4] presented a correlation based method for fingerprint verification, Davida et al [5] proposed to store a set of user specific error correction parameters as template for an iris based system. However, Soutar et al, and Davida et al's words are lack of practical implementation and can not provide rigorous security guarantees [2].

In [6], Juels and Wattenberg introduced a error correction based method, fuzzy commitment scheme, which generalized and improved Davida's methods. Feng et al [7] and Kevenaar et al [8] subsequently implemented similar schemes on iris and face biometrics respectively. Later, a polynomial reconstruction based scheme, fuzzy vault, is proposed by Juels and Sudan [9], and a few implementation works have been reported in [10][11] based on fingerprints. In general, these methods [6-11] provide enhanced security by combining biometrics features with randomly generated keys. However, except Feng et al's method for iris, the rest of the works all produce unacceptable high FRR.

Recently, Teoh et al [12] introduced a BioHashing method which produces changeable, non-invertible biometrics template, and also claimed good performance, near zero EER. The BioHashing method is a two factor authenticator based on iterated inner product between tokenised pseudo-random number and user specific biometrics features [12]. The technique has been applied on various biometrics traits [13][14] and demonstrates zero or near zero equal error rate. Kong et al [15] points out that the good performance of BioHashing are based on impractical assumption that the secret key can not be stolen. They also showed that the performance will be degraded if the key is stolen through experimental results. Lumini et al [16] introduce some ideas to improve the performance of BioHashing in case of stolen token by utilizing different threshold values and fuse the scores.

The BioHashing method provides significant improvement in terms of verification accuracy. However, as shown in the analysis and experiments in later sections, the performance of BioHashing depends on the characteristics and dimensionality of biometrics features, and generally can not produce zero EER. In this paper, we introduce methods that produce zero EER and are independent of characteristics and dimensionality of the extracted features. Experimental results show that the proposed methods outperforms the existing works.

3. METHODOLOGY

This section presents the proposed methods for face based human verification. Fig. 1 depicts the diagrammatic representation of the proposed solution. A set of biometrics features is first extracted from the user's face images. The feature extraction module provides discriminant and low dimension biometrics representation. The extracted features are then combined with user specific inputs, which is associated with a secret key, and the generated templates are stored for authentication. To produce changeable and privacy preservable

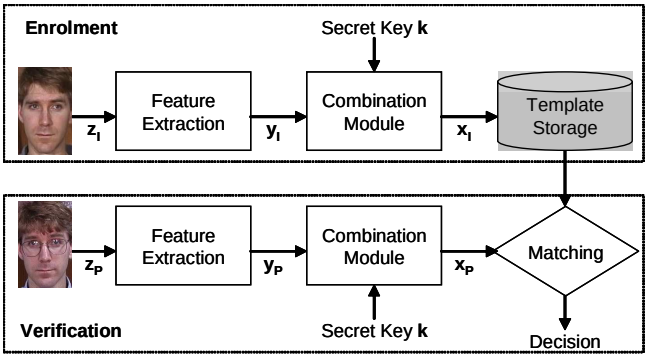

Fig. 1. General framework of proposed verification system

biometrics template, the combination should be performed such that the original face features will not be revealed if the templates are compromised. The verification will be successful if and only if both the correct biometrics and secret key are presented.

In this section, we first give a brief description of the applied feature extraction methods. We then detail the proposed schemes for combination of biometrics and user specific secret key to produce changeable and privacy preservable biometrics template. Specifically, we present two schemes that are both based on random orthonormal transformation, while differ in a discretization procedure and corresponding performance enhancing methods.

3.1. Feature Extraction

To study the effects of different feature extractors on the performance of proposed methods, we compare Principal Component Analysis (PCA) and Kernel Direct Discriminant Analysis (KDDA) .PCA is an unsupervised learning technique which provides an optimal, in the least mean square error sense, representation of the input in a lower dimensional space. In the Eigenfaces method [17], given a training set $\mathcal{Z} = \{\mathcal{Z}_i\}_{i=1}^{C}$, containing C classes with each class $\mathcal{Z}_i = \{\mathbf{z}_{ij}\}_{j=1}^{C_i}$ consisting of a number of face images $\mathbf{z}_{ij}$, a total of $M = \sum_{i=1}^{C} C_i$ images, the PCA is applied to the training set $\mathcal{Z}$ to find the M eigenvectors of the covariance matrix,

$$\mathbf{S}_{cov} = \frac{1}{M} \sum_{i=1}^{C} \sum_{j=1}^{C_i} (\mathbf{z}_{ij} - \bar{\mathbf{z}})(\mathbf{z}_{ij} - \bar{\mathbf{z}})^T \quad (1)$$

where $\bar{\mathbf{z}} = \frac{1}{M} \sum_{i=1}^{C} \sum_{j=1}^{C_i} \mathbf{z}_{ij}$ is the average of the ensemble. The Eigenfaces are the first $N (\leq M)$ eigenvectors corresponding to the largest eigenvalues, denoted as Ψ. The original image is transformed to the N-dimension face space by a linear mapping:

$$\mathbf{y}_{ij} = \Psi^T (\mathbf{z}_{ij} - \bar{\mathbf{z}}) \quad (2)$$

PCA produces the most expressive subspace for face representation, but is not necessarily the most discriminating one. This is due to the fact that the underlying class structure of the data is not considered in the PCA technique. Linear Discriminant Analysis (LDA), is a supervised learning technique that provides a class specific solution. It produces the optimal feature subspace in such a way that the ratio of between-class scatter and within-class scatter is maximized. PCA and LDA are linear solutions, and provides good performance

1-4244-1549-7/07/$25.00

in many cases. However, as the complexity of the face pattern increases, linear methods may can not provide satisfying performance. In such a case, nonlinear models are introduced to capture the complex distribution. In [18], different linear and nonlinear methods were compared in a complex generic database. It was shown that KDDA outperforms other techniques in most of the cases. Therefore we also adopt KDDA in this paper.

KDDA was proposed by Lu et al [19] to address the nonlinearities in complex face patterns. Kernel based solution find a nonlinear transform from the original image space $\mathcal{R}^J$ to a high-dimensional feature space $\mathcal{F}$ using a nonlinear function $\phi(\cdot)$. In the transformed high-dimensional feature space $\mathcal{F}$, the convexity of the distribution is expected to be retained so that traditional linear methodologies such as PCA and LDA can be applied. The optimal nonlinear discriminant feature representation of $\mathbf{z}$ can be obtained by:

$$\mathbf{y} = \Theta \cdot \nu(\phi(\mathbf{z})) \tag{3}$$

where Θ is a matrix representing the found kernel discriminant subspace, and $\nu(\phi(\mathbf{z}))$ is the kernel vector of the input $\mathbf{z}$. The detailed implementation algorithm of KDDA can be found in [19].

3.2. Random Orthonormal Transformation (ROT)

To produce changeable biometrics representation, the extracted face features $\mathbf{y}$ is converted to a new feature vector $\mathbf{x}$ by a repeatable transformation. The first scheme is based on random orthornormal transformation (ROT) of shifted biometrics features. The procedure of producing the shifted ROT feature vector is as follows:

1. Extract feature vector $\mathbf{y} \in \Re^N$ from the biometrics data
2. Generate a new feature vector $\mathbf{y}_s = \mathbf{y} + \mathbf{d}$, $\mathbf{d} \in \Re^N$ and the elements $\mathbf{d}_i >> t$, where t is the system threshold.
3. Use a user specific key $\mathbf{k}$ to generate a pseudo-random matrix, and apply the Gram-Schmidt method to transform it into an orthogonal matrix Q of size $N \times N$.
4. Compute shifted ROT feature vector $\mathbf{x} = Q^T\mathbf{y}_s$.

In this scheme, Euclidean distance is used as the similarity measure function S. Throughout this paper, we use the subscripts P and I to represent the authenticate individual and the template of the claimed identity respectively. In a true user authentication scenario, the correct key is presented, then $Q_P = Q_I$. Since $Q_P Q_I^T = I$, where I is the identity matrix, we have:

$$\begin{aligned} S(\mathbf{x}_P, \mathbf{x}_I) &= \|Q_P^T(\mathbf{y}_P + \mathbf{d}) - Q_I^T(\mathbf{y}_I + \mathbf{d})\|^2 \\ &= \|Q_P^T\mathbf{y}_P - Q_I^T\mathbf{y}_I\|^2 \\ &= \|Q_P^T\mathbf{y}_P\|^2 + \|Q_I^T\mathbf{y}_I\|^2 - 2(Q_P^T\mathbf{x}_P)^T(Q_I^T\mathbf{y}_I) \\ &= \|\mathbf{y}_P\|^2 + \|\mathbf{y}_I\|^2 - 2\mathbf{y}_P^T Q_P Q_I^T \mathbf{y}_I \\ &= \|\mathbf{y}_P\|^2 + \|\mathbf{y}_I\|^2 - 2\mathbf{y}_P^T\mathbf{y}_I \\ &= \|\mathbf{y}_P - \mathbf{y}_I\|^2 \end{aligned} \tag{4}$$

As shown in Equation 4, the ROT exactly preserves the similarity of original face feature. This also accounts for the stolen key scenario, where an imposter steals the secret key of the claimed identity, and use his own biometrics for verification. In this case, the verification performance will be the same as the original face features.

Let's consider a scenario where an imposter tries to authenticate as the true user. Since different users are associated with distinct keys, therefore $Q_P \neq Q_I$. To quantify the probability of error and illustrate the importance of shifting the face features (step 2), we first consider a case where ROT is applied on the extracted face features directly, i.e., $\mathbf{x} = Q^T\mathbf{y}$. The FAR corresponds to the probability of deciding $\mathbf{H}_0$ when $\mathbf{H}_1$ is true, $P(\mathbf{H}_0|\mathbf{H}_1)$, and the FRR corresponds to $P(\mathbf{H}_1|\mathbf{H}_0)$. Let's select the system threshold t such that $P(\mathbf{H}_1|\mathbf{H}_0)$=0. Since the transformation is orthonormal and random, the ROT of a point in N-D space corresponds the rotation of point in the hyper-sphere whose radius is specified by the length of the point. We have:

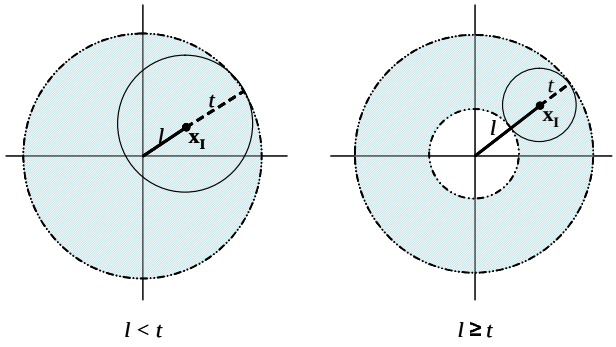

Fig. 2. Demonstration of computing probability of error in 2-D space

$$P(\mathbf{H}_0|\mathbf{H}_1) = P(l_{\mathbf{x}_I} - t \leq l_{\mathbf{x}_P} \leq l_{\mathbf{x}_I} + t, S(\mathbf{x}_I, \mathbf{x}_P) \leq t) \tag{5}$$

where $l_{\mathbf{x}_I}$ and $l_{\mathbf{x}_P}$ represent the length of the template and authenticate vector respectively. As shown in Fig. 2, the computation of Equation 5needs to be split into two cases: $l_{\mathbf{x}_I} \leq t$ and $l_{\mathbf{x}_I} > t$. In 2-D space, $P(S(\mathbf{x}_P, \mathbf{x}_I) \leq t | l_{\mathbf{x}_I} - t \leq l_{\mathbf{x}_P} \leq l_{\mathbf{x}_I} + t) = \frac{\pi t^2}{\pi(l_{\mathbf{x}_I}+t)^2}$ when $l_{\mathbf{x}_I} \leq t$, and $P(S(\mathbf{x}_P, \mathbf{x}_I) \leq t | l_{\mathbf{x}_I} - t \leq l_{\mathbf{x}_P} \leq l_{\mathbf{x}_I} + t) = \frac{\pi t^2}{\pi(l_{\mathbf{x}_I}+t)^2 - \pi(l_{\mathbf{x}_I}-t)^2}$ when $l_{\mathbf{x}_I} > t$. This can be easily extended to N-D space, where the volume of a N-D hypersphere with radius r is defined as [20]: $V_N = \frac{S_N r^N}{N}$, where S_N is the hyper-surface area of an N-sphere of unit radius. In N-D space, we have:

$$\begin{aligned} P_1 &= P(S(\mathbf{x}_P, \mathbf{x}_I) \leq t | l_{\mathbf{x}_I} - t \leq l_{\mathbf{x}_P} \leq l_{\mathbf{x}_I} + t, l_{\mathbf{x}_I} \leq t) \\ &= \frac{\frac{S_N t^N}{N}}{\frac{S_N (l_{\mathbf{x}_I}+t)^N}{N}} = \frac{t^N}{(l_{\mathbf{x}_I}+t)^N} \end{aligned}$$

$$\begin{aligned} P_2 &= P(S(\mathbf{x}_P, \mathbf{x}_I) \leq t | l_{\mathbf{x}_I} - t \leq l_{\mathbf{x}_P} \leq l_{\mathbf{x}_I} + t, l_{\mathbf{x}_I} > t) \\ &= \frac{\frac{S_N t^N}{N}}{\frac{S_N (l_{\mathbf{x}_I}+t)^N}{N} - \frac{S_N (l_{\mathbf{x}_I}-t)^N}{N}} = \frac{t^N}{(l_{\mathbf{x}_I}+t)^N - (l_{\mathbf{x}_I}-t)^N} \end{aligned}$$

$$\begin{aligned} P(\mathbf{H}_0|\mathbf{H}_1) &= P(l_{\mathbf{x}_I} \leq t) P(l_{\mathbf{x}_P} \leq l_{\mathbf{x}_I} + t | l_{\mathbf{x}_I} \leq t) P_1 \\ &\quad + P(l_{\mathbf{x}_I} > t) P(l_{\mathbf{x}_I} - t \leq l_{\mathbf{x}_P} \leq l_{\mathbf{x}_I} + t | l_{\mathbf{x}_I} > t) P_2 \end{aligned} \tag{6}$$

From Equation 6, it is clear that the probability of false accept depends on the characteristics and dimensionality of the features. In general, zero error rate can not be achieved by directly apply ROT

1-4244-1549-7/07/$25.00

on the extracted face features. However, since $P(l_{\mathbf{x}_P} \leq l_{\mathbf{x}_I}+t|l_{\mathbf{x}_I} \leq t)P_1 \leq 1$, and $P(l_{\mathbf{x}_I} > t)P(l_{\mathbf{x}_I} - t \leq l_{\mathbf{x}_P} \leq l_{\mathbf{x}_I} + t|l_{\mathbf{x}_I} > t) \leq 1$, Equation 6 can be simplified as:

$$P(\mathbf{H}_0|\mathbf{H}_1) \leq P(l_{\mathbf{x}_I} \leq t) + \frac{t^N}{(l_{\mathbf{x}_I}+t)^N - (l_{\mathbf{x}_I}-t)^N} \quad (7)$$

This probability can be minimized by adding an extra vector $\mathbf{d} \in \Re^N$, $\mathbf{d}_i >> t$, to the extracted face features, $\mathbf{y}_s = \mathbf{y} + \mathbf{d}$, such that after ROT, $P(l_{\mathbf{x}_I} < t) = 0$. We have:

$$P(\mathbf{H}_0|\mathbf{H}_1) \quad \leq \quad \frac{t^N}{(l_{\mathbf{x}_I}+t)^N - (l_{\mathbf{x}_I}-t)^N} \quad (8)$$

and

$$\lim_{\frac{t}{l_{\mathbf{x}_I}} \to 0, \forall N} P(\mathbf{H}_0|\mathbf{H}_1) = 0 \quad (9)$$

By using the proposed method, both zero FAR and FRR can be achieved. It should be noted that the stolen biometrics scenario also complies with the above analysis, since the $\mathbf{x}_P$ in Equation 5 can also be generated from the true user's biometric features. Therefore it has the same performance as both-non-stolen scenario. This also explains the changeability of our methods. After generating a new biometric templates, the old templates can not be used for successful authentication.

3.3. Discretized Random Orthonormal Transformation (DROT)

The proposed shifted ROT method provide changeable biometrics template, and produce zero error rate. However, it only offers limited security since the ROT is invertible. If the storage and the secret key are both compromised, the original face features of the user will be revealed. To overcome this problem, we propose another scheme which discretizes the random orthonormal transformation of the original features. The discretization is non-invertible, therefore this method provides more rigorous security.

The procedure of producing the discretized ROT feature vectors are as follows:

1. Extract feature vector $\mathbf{y} \in \Re^N$ from the biometrics data
2. Use a user specific key **k** to generate a pseudo-random matrix, and apply the Gram-Schmidt method to transform it into an orthogonal matrix *Q* of size $N \times N$.
3. Compute feature vector $\mathbf{u} = Q^T\mathbf{y}$.
4. Compute the *N* bits code b_i,$i = 1, ..., N$, according to:

$$b_i = \begin{cases} 0, & if \mathbf{u}_i < \tau \\ 1, & if \mathbf{u}_i \geq \tau \end{cases}$$

   where τ is a preset threshold (usually 0).
5. Use key **k** to generate a set of *M* random bits **d**, $M \gg N$, .
6. Generating $N+M$ dimension code by concatenating **b** and **d**, $\mathbf{x} = [\mathbf{b}\ \mathbf{d}]$.

The first four steps in the above procedure correspond to the best performance scenario in the BioHashing method [12]. Unlike the shifted ROT method, the discretized ROT method utilizes Hamming distance as the metric to measure the distance between two bit strings. To quantify the probability of error, let's first consider the case where **b** is used for verification (i.e., BioHashing). Since the orthonormal transformation is random, we can assume each bit in **b** is random. Let *t* be the system threshold in terms of Hamming distance, and the *t* is selected such that $P(\mathbf{H}_1|\mathbf{H}_0)$=0, then the probability of false accept $P(\mathbf{H}_0|\mathbf{H}_1) = \frac{\sum_{i=0}^{t}\binom{N}{i}}{2^N}$. This probability (therefore the performance of BioHashing) depends on two factors, the system threshold *t* and dimension *N*. The system threshold *t* depends on the separability of biometrics features in terms of Hamming distance. It is also not suitable to increase the dimension of extracted face features as will since increase of feature dimension may also increase system threshold. However, $P(\mathbf{H}_0|\mathbf{H}_1)$ can be minimized by appending $M, M \gg N$, extra random bits associated with each secret key to vector **b**, such that $P(\mathbf{H}_0|\mathbf{H}_1) = \frac{\sum_{i=0}^{t}\binom{N}{i}}{2^{N+M}}$, and we have $\lim_{m\to\infty,\forall N,t} P(\mathbf{H}_0|\mathbf{H}_1) = 0$.

The attachment of random bits does not increase the system threshold since **d** is unique for every user. For different users, the added bits are different, which is equivalent to increase the Hamming distance between different users. Therefore, by adding sufficiently large number of random bits, we can produce zero error rate. For example, even in relatively low dimension and high threshold scenario, let $N = 20$ and $t = 10$, then $P(\mathbf{H}_0|\mathbf{H}_1) = 0.5881$. If we add 100 random bits to the bit string, then $P(\mathbf{H}_0|\mathbf{H}_1) = 1.53 \times 10^{-17} \approx 0$.

4. EXPERIMENTS AND DISCUSSION

To evaluate the performance of proposed methods, we conducted our experiments on two sets of face databases: ORL [21] and GT [22]. The ORL database contains 400 face images from 40 subjects with 10 images each. The GT database contains 750 images of 50 people with 15 images each. The face images in GT database have larger pose and illumination variation than the ORL database. The original images in GT database were taken on cluttered background. In this work, we use the cropped data set generated by manually determined label filters. In both database, the first five images of each subject are used as training samples as well as gallery sets. The rest images of each subject are used as probe samples. The classification is based on nearest neighbor.

Our evaluation is based on equal error rate (EER), which is defined as the operating point at which false accept rate (FAR) and false reject rate (FRR) are equal, i.e., $EER = (FAR + FRR)/2$ [12]. As illustrated in Section 3, the stolen biometrics scenario is the same as the both-non-stolen case. Therefore only analyzing both-non-stolen and stolen key scenarios will be sufficient. A description of the abbreviations of the terminologies used in the paper is given in Table 1. In the shifted ROT method, all extracted features are shifted by $\mathbf{d}_i = 10^{10}$, while $M = 200$ random bits are added in the discretized ROT method. To minimize the effect of randomness, all the experiments were performed 5 times, and the average of the results are reported.

Fig. 3 depicts the EER as a function of feature dimensions when PCA and KDDA are used as feature extractors respectively. The EER obtained at the highest dimensionality of our experimental setting are reported in Table 2. In general, The ROT-O and DROT-BH methods can not produce zero EER, while ROT-S and DROT-RB achieve zero EER in all dimensions. This complies with our analysis in Section 3. In the stolen key scenario, The ROT based methods exactly preserve the performance of original face features, but

1-4244-1549-7/07/$25.00

Name	Description
ROT-O	ROT on original face features
ROT-S	ROT on shifted face features
ROT SK	ROT stolen key scenario
DROT-BH	BioHashing method
DROT-RB	DROT with added random bits
DROT SK	DROT stolen key scenario

Table 1. Description of abbreviations of terminologies

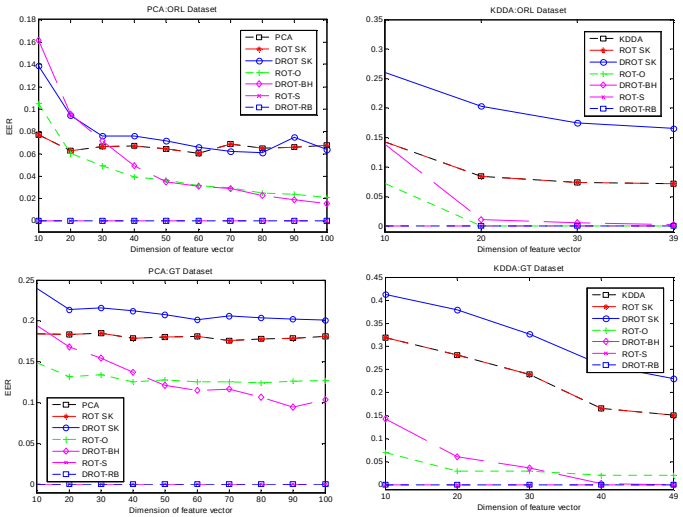

Fig. 3. EER obtained as a function of feature dimension by using PCA and KDDA as feature extractors

the performance of DROT methods degrades since the discretization procedure corrupts the representation of biometrics features. The performance of DROT based methods is improved as the dimensionality increases, but leveled off after a certain dimension. This is due to the inherent discriminant capability of the face features.

	PCA		KDDA	
	ORL(100)	GT(100)	ORL(39)	GT(49)
ROT SK	6.78	18.09	7.19	15.08
DROT SK	6.35	20.13	16.53	23.03
ROT-O	2.09	12.75	0	2.01
DROT-BH	1.52	10.39	0.21	0.06
ROT-S	0	0	0	0
DROT-RB	0	0	0	0

Table 2. EER (%) obtained by using PCA and KDDA as feature extractors (with feature dimension in (-))

KDDA has similar performance as PCA in the ORL dataset, but offers improvement in the GT dataset. This is in line with the experiments shown in [18], that KDDA is a more advanced technique, and particularly as the complexity of dataset increase, the nonlinearity becomes more severe. Furthermore, it can be observed that the BioHashing method produces near zero EER at appropriate high dimensions. However, the tradeoff of the improvement in BioHashing is the significant degradation in the stolen key scenario, which in some cases is even worse than PCA. This is due to the fact that thresholding method adopted in BioHashing is equivalent to use only the angle information between vectors for classification purpose. More precisely, the classification is based on the closeness of the orthants that the feature points fall into. For better discretization consequence, the feature points should be well spread over the whole plane with respect to each dimension.

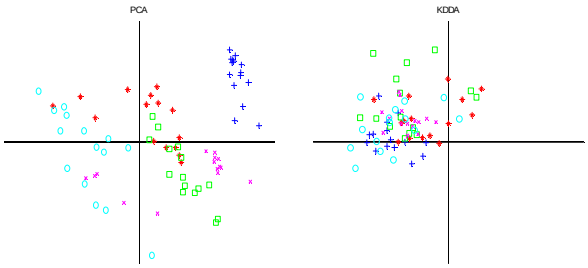

Fig. 4. Distribution of PCA and KDDA coefficients

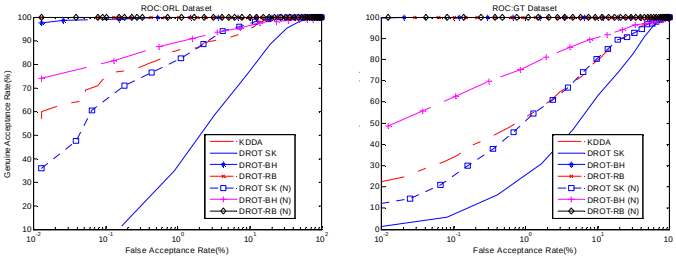

Fig. 5. ROC curve of DROT v.s. normalized DROT

The distributions of the first two PCA and KDDA coefficients of five subjects in GT dataset are plotted Fig. 4. It is clear that PCA coefficients are well spread in the plane since PCA has a normalization procedure to produce zero mean along each dimension (see Equation 2). KDDA has a more compact representation since no such normalization is performed (see Equation 3). In BioHashing, The compact representation of KDDA produces smaller system threshold *t*, and therefore better performance. But this compact representation also corrupt the separability of discretized code. To produce good discretization in case of stolen key, it is important to normalize the KDDA features. In this paper, we normalize the KDDA features by subtracting the mean vector of the training data. We perform experiments on the maximum dimension of each dataset, i.e., 39 for ORL and 49 for GT. Fig. 5 shows the ROC curve of different methods when KDDA and normalized KDDA features are used, while Table 3 details the results in terms of FAR, FRR, and EER.

It can be seen that the performance of DROT in stolen key scenario approaches and sometimes even outperforms that of KDDA after the normalization procedure. The BioHashing results degrade using the normalized features. This is due to the normalization procedure increase the system threshold *t* and therefore the error. However, by utilizing the proposed methods of adding random bits, zero EER can be achieved.

1-4244-1549-7/07/$25.00

	ORL(39)			GT(49)		
	FAR	FRR	EER	FAR	FRR	EER
KDDA	6.88	7.5	7.19	14.76	15.4	15.08
DROT SK	16.8	16.25	16.53	22.13	23.92	23.03
DROT-BH	0.17	0.25	0.21	0.08	0.04	0.06
DROT-RB	0	0	0	0	0	0
DROT SK(N)	7.26	6.1	6.68	14.23	16.2	15.21
DROT-BH(N)	6.21	5.35	5.78	8.83	10.68	9.75
DROT-RB(N)	0	0	0	0	0	0

Table 3. Experimental results (in %) of different methods on KDDA and normalized KDDA features ("N" denotes normalized)

5. CONCLUSION

This paper introduced a systematic framework for addressing the challenging problem of template changeability and privacy protection in biometrics-enabled authentication systems. The proposed method is based on discretized random orthonormal transformation, which is associated with a user specific secret key. By using different keys, distinct biometric templates can be generated. The discretization procedure is non-invertible, therefore the privacy of users can be protected. Our method provides functional advantage in that zero error rate can be achieved. In the stolen key scenario, we show that the proposed method maintains the performance of original features at appropriate high dimension. In addition, we also introduce another method where random orthonormal transformation is applied on shifted biometric features. This method is less secure since the transformation is invertible, but it provides exactly the same performance as the original features in the stolen key scenario regardless of the characteristics and dimensionality of the biometrics features.

A detailed mathematical analysis on the proposed framework was provided in this work. The experiments demonstrated the effectiveness of the proposed approaches comparing with existing works. Although we focus on face based verification, the proposed methods are general and can also be applied to other biometrics. In the future, we are going to work on more advanced feature extraction techniques to improve the performance in the stolen key scenario. Discretization methods that preserves the representation of features, while provide non-invertible properties will also be investigated.

6. REFERENCES

[1] A. K. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition", IEEE Trans. on Circuits and Systems for Video Technology. vol. 14, no. 1, pp. 4-20, January 2004

[2] U. Uludag, S. Pankanti, S. Prabhakar, and A. K. Jain, "Biometric Cryptosystems: Issues and Challenges", Proc. of the IEEE, vol. 92, no. 6, pp. 948-960, 2004

[3] R. M. Bolle, J. H. Connel, N. K. Ratha, "Biometric perils and patches", Pattern Recognition, vol. 35, pp. 2727-2738, 2002

[4] C. Soutar, D. roberge, A. Stoianov, R. Gilroy, and B. V. K. Vijaya Kumar, "Biometric Encryption", ICSA Guide to Cryptography, McGraw-Hill, 1999

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

[6] A. Juels, and M. Wattenberg, "A fuzzy commitment scheme", Proc. of sixth ACM Conf. on Computer and Communication Security, pp. 28-36, 1999

[7] F. Hao, R. Anderson, and J. Daugman, "Combining crypto with biometric effectively", IEEE Trans. on Computers, vol. 55, no. 9, pp. 1081-1088, 2006

[8] Kevenaar, T.A.M.; Schrijen, G.J.; van der Veen, M.; Akkermans, A.H.M.; Zuo, F.; "Face recognition with renewable and privacy preserving binary templates",Automatic Identification Advanced Technologies, 2005. Fourth IEEE Workshop on 17-18 Oct. 2005 Page(s):21 - 26

[9] A. Juels, and M. Sudan, "A fuzzy vault scheme", Proc. of IEEE International Symp. on Information Theory, pp. 408, 2002

[10] R. C. Clancy, N. Kiyavash, and D. J. Lin, "Secure smart card based fingerprint aithentication", Proc. of ACM SIGMM Workshop on Biometrics Methods and Applications, pp. 45-52, 2003

[11] U. Uludag, S. Pankanti, and A. Jain, "Fuzzy vault for fingerprints", Proc. of International Conf. on Audio and Video based Biometric Person Authentication, pp. 310-319, 2005

[12] A.B.J. Teoh, D.C.L Ngo and A. Goh, BioHashing: two factor authentication featuring fingerprint data and tokenised random number, Pattern Recognition, vol. 37, pp. 2245-2255, 2004.

[13] T. Connie, A. Teoh, M. Goh and D. Ngo, PalmHashing: A Novel Approach for Dual-Factor Authentication, Pattern Analysis and Application, vol 7, no. 3, pp. 255-268, 2004

[14] D. C. L. Ngo, A. B. J. Teoh, and A Goh, "Biometric hash: high-confidence face recognition", IEEE Trans. on Circuits and Systems for Video Technology, vol. 16, no. 6, June 2006

[15] k. H. Cheung, B. Kong, D. Zhang, M. Kanem. J. You, "Revealing the secret of FaceHashing", in ICB 2006, in Lecture Notes in Computer science, vol 3832, Springer, Berlin, 2006, pp. 106-112

[16] A. Lumini and L. Nanni, "An improved BioHashing for human authentication", Pattern Recognition vol. 40, pp. 1057-1065, 2007

[17] M. Turk, A. Pentland, "EigenFaces for recognition", Journal of Cognitive Neuroscience 13(1) (1991) 71-86

[18] Jie Wang, K. N. Plataniotis, Juwei Lu and A. N. Venetsanopoulos, "On Solving the Face Recognition Problem with One Training Sample per Subject", Pattern recognition 39(2006), pp. 1746-1762

[19] Juwei Lu, K.N. Plataniotis, and A.N. Venetsanopoulos, "Face Recognition Using Kernel Direct Discriminant Analysis Algorithms", IEEE Trans. on Neural Networks, Vol. 14, No. 1, Page: 117-126, January 2003.

[20] Wolfram MathWorld, "http://mathworld.wolfram.com/Hypersphere.html".

[21] ATT Laboratories Cambridge, ORL face databse, "www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html".

[22] Georgia Tech face database,"www.anefian.com/face-reco.htm".

1-4244-1549-7/07/$25.00