GMM-SVM Fingerprint Verification Based on Minutiae Only

Berkay Topcu\textsuperscript{1,2}, Yusuf Ziya Isik\textsuperscript{1,2}, and Hakan Erdogan\textsuperscript{2}
\textsuperscript{1}\textsc{TUBITAK} BILGEM, Kocaeli, Turkey
\textsuperscript{2}Sabancı University, Istanbul, Turkey
\{berkay.topcu,yusuf.ziya\}@tubitak.gov.tr, haerdogan@sabanciuniv.edu

Abstract

Most fingerprint recognition systems use minutiae information, which is an unordered collection of minutiae locations and orientations. Template protection algorithms such as fuzzy commitment and other modern cryptographic alternatives based on homomorphic encryption require a fixed size binary template. However, such a template is not directly applicable to fingerprint minutiae representation which by its nature is of variable size. In this study, we introduce a novel method to represent a minutiae set with a rotation invariant fixed-length vector. We represent each minutia according to its geometric relation with neighbors and use Gaussian mixture model (GMM) to model its feature distribution. A two-class linear SVM is used to create a model template for the enrollment fingerprint sample, which discriminates impressions of the same finger from other fingers. We evaluated the verification performance of our method on the FVC2002DB1 database.

1. Introduction

Biometric recognition systems enable fast, reliable, and secure electronic authentication, however, their large scale deployment in real world applications causes privacy and security concerns \cite{10,23,24}. In the literature, several biometric template protection methods have been proposed \cite{11} (e.g., fuzzy commitment scheme \cite{15} and biohashing \cite{13}) to overcome these concerns by securing biometric templates (e.g., face and fingerprint).

Recent template protection schemes require either a fixed length feature vector representation or a binarized string as input. Thus, a variable length minutiae representation of a fingerprint cannot be directly used in combination with these schemes. In addition, some template protection schemes designed specifically to work with unordered sets of varying number of minutiae (e.g., fuzzy vault \cite{14}) experience degradation in matching accuracy due to alignment issues and nonlinear distortion \cite{12}.

One of the earliest works on fingerprint template protection has secured minutiae information \(x, y, \theta\) separately \cite{3}. In a later study, FingerCode feature (a texture based fingerprint representation without minutiae information) has been protected via biohashing \cite{13}. Another branch of research has focused on securing each minutia separately. Yang et al. \cite{28,29} have proposed methods to extract a binary secure hash bit string from each minutia and its vicinity using minutiae information only. A more recent study similarily has used neighboring minutiae information along with texture information around each minutia and secured each minutia feature vector by biohashing \cite{2}. Protected Minutiae Cylinder-Code (P-MCC) \cite{9}, one of the most accurate algorithms proposed recently, has secured each MCC structure that corresponds to a single minutia. All these studies have represented each minutia with a fixed length binary string therefore matching between variable length final templates has been addressed as a minutiae pairing problem.

In the spectral minutiae representation \cite{27}, each minutia location is coded by an isotropic two-dimensional Gaussian function in the spatial domain. Here, minutiae are represented as a magnitude spectrum and their orientations are incorporated by assigning each Gaussian a complex magnitude. Bringer et al. \cite{4} have characterized a fingerprint in terms of its similarity to each representative local minutiae vicinities in a set of fixed size. This fixed size set has been extracted from a representative database of all existing vicinities in the world of fingerprints. For a fingerprint, a feature vector that contains similarities of its vicinities to those of the representative set has been produced.

In this work, our ultimate goal is to describe an underlying framework that enables the generation of a fixed length feature vector representation for fingerprint minutiae. The framework draws upon the work of Campbell et al. on support vector machines using GMM supervectors for speaker verification \cite{6}. Each minutia and its neighbors within a specified radius are represented as a 2D image by placing two-dimensional Gaussians at the locations of neighbor minutiae. DCT coefficients of this patch image are rearranged based on zig-zag scanning and the first \(D\) DCT.
coefficients of this patch image are used to represent each minutia as a \( D \)-dimensional feature vector. A single user-independent GMM universal background model (UBM) is trained from a collection of fingerprints to represent the distribution of DCT features. A fingerprint is then represented with its probabilistic alignment into the UBM mixture components and a GMM supervector is created from the stacked first order statistics of the mixture components.

For a given enrollment fingerprint sample, a two-class linear SVM is trained in order to create a model template that discriminates positive samples from negative samples. The matching between a query fingerprint and the model template is performed by computing a single inner product of the mixture components.

The performance of our framework was evaluated between the target SVM model and the query GMM template is performed by computing a single inner product. This also enables two fingerprints to be matched by locally comparing patches pairwise and calculating their similarity score using the local scores of the best pairs. Although global coherency in the minutiae set is not utilized, the local approach has the advantage of limiting the crucial elastic distortion problem in fingerprint matching. In the local area of a patch, distortion due to the elasticity of the skin is negligible. The radius used in the local approach is of great importance. The neighborhood of the central minutia should contain several minutiae in order to be sufficiently discriminative but at the same time it should be small enough to be considered as a local area.

2.2. Gaussian minutia patch image

Within a specific radius \( R \), the number of neighbor points of a central minutia varies and this leads to a minutia representation of unknown length. In order to obtain a fixed length representation, one can use a rectangular grid of size \((2R + 1) \times (2R + 1)\) where the central minutia is at the center. Each neighbor minutia is then inserted into this grid with respect to its relative location to \( m_c \) on the fingerprint.

Representing a minutia with a single point in the spatial domain increases the sensitivity of minutiae positions to small variations and does not maintain direction information. Instead, each neighbor minutia is represented by a two-dimensional multivariate Gaussian function:

\[
 f(x, \mu, \Sigma) = \frac{1}{2\pi \sigma_1 \sigma_2} e^{-\frac{1}{2}(x - \mu)^T \Sigma^{-1} (x - \mu)} \tag{1}
\]

where \( \Sigma = \text{diag}(\sigma_1, \sigma_2) \) is the covariance matrix. A Gaussian is centered at the minutia location and its covariance matrix is selected such that the major axis coincides with the minutia orientation as illustrated in Figure 1(b). For a neighbor minutia \((x_i, y_i, \theta_i)\), a template Gaussian is translated to \((x_i, y_i)\) and rotated with \(\theta_i\). This makes the mean of Gaussian as \([x_i, y_i]^T\) and covariance matrix as \(\Sigma' = A \Sigma A^T\), where \(A = A(\theta_i)\) is a rotation matrix. The patch image is

\[\text{Figure 1. DCT representation of a minutia patch image}\]
then generated as a sum over these shifted Gaussians:

\[ I(x) = \sum_{i=1}^{N_p} f(x, \mu_i, \Sigma'_i) \]  

(2)

where \( N_p \) is the number of neighbor minutiae and \( \mu_i = [x_i, y_i]^T \) is the position of the neighbor minutia with respect to \( m_c \). Sample minutia patch images selected from a fingerprint are illustrated in Figures 2(b) and 2(c). Please note that, the central minutia is not directly included in this representation, but it defines the new coordinate system and the neighbors of the patch.

### 2.3. DCT representation for minutia patches

Although minutia patch images capture the required information for fingerprint matching, \((2R+1)^2\)-dimensional representation for each minutia brings heavy computation and storage requirements. Discrete cosine transform (DCT) is often used in image processing, especially for lossy compression (e.g., JPEG), due to its strong energy compaction property. It expresses a finite sequence of data points in terms of a sum of cosine functions oscillating at different frequencies. Since most of the signal information tend to be concentrated in a few low-frequency components of the DCT, discarding small high-frequency components results in compact representation of the signal. By keeping only the first \( D \) 2D-DCT coefficients after performing zig-zag scanning, each minutia patch image is represented as a \( D \)-dimensional feature vector.

We conducted an evaluation to assess the discriminative power of our DCT minutia patch representation. To compare two fingerprints, \( fp_1 \) and \( fp_2 \), a pairing matrix that contains similarity scores between patches of \( fp_1 \) and patches of \( fp_2 \) was constructed. The scores were computed using a decreasing function that converted the Euclidean distances between DCT coefficients to a score (i.e., \( g(x) = 1 / (1 + e^{x/\tau}) \)). A closest neighbor search algorithm was applied to the pairing matrix in order to select the best association of minutiae. At each turn, a minutia pair from \( fp_1 \) and \( fp_2 \) with the highest matching score was identified to the orientation of the central minutia \( \theta_c \).

and removed from the matrix. The final score between two fingerprints was computed by accumulating the matching scores of identified pairs during the search.

In the evaluation, the FVC2002DB1 database \(^{20}\) which has 8 impressions of 100 different fingerprints captured with an optical sensor was used. Following the performance evaluation protocol of FVC2002 \(^{21}\), 2800 genuine and 4950 imposter comparisons were performed. An Equal Error Rate (EER) of 4.46\% was achieved for \( D = 50 \). Although, the achieved EER was below the state of the art \(^{20,21}\), it arguably confirmed the discriminative capability of minutia patches.

### 3. GMM supervector training

Gaussian mixture models (GMMs) have been dominantly used for modeling in text-independent speaker verification. The distribution of features extracted from speech segments (i.e., frames of an utterance) is modeled by performing background model adaptation of GMMs. First, a universal background model (UBM) is trained from set of frames and then the speaker model for the \( i^{th} \) speaker is derived by adapting the universal background model to match the observations of the speaker. Recently, the use of GMM for modeling feature distribution has also become an effective approach for face verification \(^{26}\).

Similar to the frames of a speech utterance or the blocks of a face image, minutiae points of a fingerprint are separate observations of the same underlying signal. DCT patch representation of minutiae is used to train a universal background minutiae model. The UBM is a large GMM trained to represent the distribution of features. From a huge database of fingerprints, a large number of minutiae patches are extracted as training data and they are pooled to train the UBM via EM (expectation maximization) algorithm:\(^2\)

\[ g(x) = \sum_{i=1}^{N} w_i p_i(x) \]  

(3)

where \( w_i \) are the mixture weights and \( p_i(x) \) is the unimodal Gaussian density with mean \( \mu_i \) and covariance \( \Sigma_i \) (diagonal covariance is assumed as this requires fewer observations).

Given a fingerprint with \( T \) minutiae, \( x_t, t = 1, ..., T \) are the DCT minutia patches for each minutia. The estimates of first order statistics for the fingerprint data are computed for mixture \( i \) in the UBM as:

\[ E_i(x) = \sum_{t=1}^{T} Pr(i|x_t)(x_t - \mu_i) \]  

(4)

\[ Pr(i|x_t) = \frac{w_i p_i(x_t)}{\sum_{j=1}^{M} w_j p_j(x_t)}, \]  

(5)

\(^{2}\)For further details, please refer to \(^{25}\).
Using only the first order statistics \((E_i(x))\), a GMM supervector is formed by concatenating the first order statistics of each mixture. The GMM supervector maps a fingerprint to a high-dimensional vector of length \(D \times N\), where \(D\) is the number of DCT coefficients and \(N\) is the number of Gaussians in the mixture. Please note that, we do not perform MAP adaptation as done in [6, 13, 23, 29] for adapting a speaker model. Our experiments showed that using first order statistics without MAP adaptation performed better, so we employed first order statistics only.

4. Linear SVM training for template generation

An SVM is a two-class linear classifier constructed from sums of a kernel function \(K\)

\[
f(x) = \sum_{i=1}^{L} \alpha_i t_i K(x, x_i) + d
\]

(6)

where \(t_i\) are the ideal outputs (either 1 or -1), \(d\) is a learned constant, \(\sum_{i=1}^{L} \alpha_i t_i = 0\), and \(\alpha_i > 0\). The vectors \(x_i\) are support vectors and obtained from the training set by an optimization process. The kernel \(K\) is constrained to have certain properties so that \(K\) can be expressed as an inner product, \(K(x, x_i) = \Phi(x_i) \cdot \Phi(x)\), where \(\Phi(\cdot)\) is a mapping to a higher dimension.

SVM provides a suitable solution to fingerprint verification problem, since it is fundamentally a two-class problem. We aim to decide whether the fingerprint comes from the user or the fingerprint belongs to someone else. As the number of features is large in our problem \((D \times N)\), we do not need to map data to a higher dimensional space and use linear kernel (i.e., \(K(x, x_i) = x_i^T x\)). In practice, the linear kernel tends to perform very well when the number of features is large. In addition, GMM supervector has already been employed as a linear kernel with a simple diagonal scaling [5, 6]. The SVM in [6] can be expressed as:

\[
f(x) = \sum_{i=1}^{L} \alpha_i t_i x_i^T x + d = w^T x + d
\]

(7)

which reduces two-class classification to an inner product between the classifier model \(w\) and GMM supervector \(x\). The model \(w\) is solved by minimizing:

\[
\min_{w,d} \left( \frac{1}{2} \|w\|_2^2 + C \sum_i H_1(t_i(w^T x_i + d)) \right)
\]

(8)

where \(H_1(z) = \max(0, 1 - z)\) is the "Hinge Loss" and \(C\) is a regularization parameter that controls a tradeoff between a low error on the training data and the ability to generalize well.

<table>
<thead>
<tr>
<th>DB Name</th>
<th>#Fingers</th>
<th>#Fingers × #Samples/Finger</th>
</tr>
</thead>
<tbody>
<tr>
<td>FVC2002DB2</td>
<td>800</td>
<td>100 × 8</td>
</tr>
<tr>
<td>FVC2002DB3</td>
<td>800</td>
<td>100 × 8</td>
</tr>
<tr>
<td>FVC2002DB4</td>
<td>800</td>
<td>100 × 8</td>
</tr>
<tr>
<td>FVC2004DB1</td>
<td>800</td>
<td>100 × 8</td>
</tr>
<tr>
<td>FVC2004DB2</td>
<td>800</td>
<td>100 × 8</td>
</tr>
<tr>
<td>FVC2004DB3</td>
<td>800</td>
<td>100 × 8</td>
</tr>
<tr>
<td>FVC2004DB4</td>
<td>800</td>
<td>100 × 8</td>
</tr>
<tr>
<td>FVC2006DB1</td>
<td>1648</td>
<td>140 × 12</td>
</tr>
<tr>
<td>FVC2006DB2</td>
<td>1680</td>
<td>140 × 12</td>
</tr>
<tr>
<td>FVC2006DB3</td>
<td>1680</td>
<td>140 × 12</td>
</tr>
<tr>
<td>FVC2006DB4</td>
<td>1680</td>
<td>140 × 12</td>
</tr>
<tr>
<td>IN-HOUSE</td>
<td>3520</td>
<td>440 × 8</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>15808</strong></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Number of fingerprints used in GMM training

We use SVM to create a model \(w\) (which we also refer to as a template) for an enrollment fingerprint sample \(f_{enroll}\). This is achieved by training an SVM using the GMM supervector of \(f_{enroll}\) as a positive sample (labeled as +1) and GMM supervectors of fingerprints from example imposters as negative samples (labeled as -1). Given a query fingerprint sample \(f_{query}\), its matching score for the subject \(i\) is the inner product between \(w_i\) and \(x_{query}\), where \(w_i\) is the SVM classifier model for the subject \(i\) and \(x_{query}\) is the GMM supervector of \(f_{query}\). The verification decision is based upon whether the score \(w_i^T x_{query}\) is above or below a threshold.

This approach provides 1-to-1 fingerprint matching since only one single training sample for each class is used to train the template model. It corresponds to comparing a gallery fingerprint to a query fingerprint as done in all other fingerprint verification systems.

5. Experiments and Results

We performed 1-to-1 fingerprint verification experiments on the FVC2002DB1A fingerprint database [29]. For minutiae extraction, a commercial fingerprint minutiae extractor (which participated in FVC-onGoing [19], Ongoing MINEX [17] and FpVTE 2012 [16]) was used to obtain minutiae information in ISO 19794-2 format \((x, y, \theta)\) [3].

In order to create patches for each minutia, all neighbor minutiae within a radius \(R = 60\) pixels were used. This resulted in minutia patch images of size \(121 \times 121\) pixels. For DCT representation of patches, the first 50 DCT coefficients after zig-zag scanning were kept (i.e., \(D = 50\)), which means that a minutia was represented along with its local information via a feature vector of length 50.

We used 15804[3] fingerprints from publicly available FVC databases and an in-house fingerprint database col-

---

32 minutiae in FVC2006DB1 have zero neighbors within \(R\), therefore they were not used in GMM training.
lected via an optical reader. The details of these databases (the number of fingers and samples per finger) can be found in Table 1. Our target database, FVC2002DB1, was not included in GMM training to prevent any bias that might favor supervector representation in the advantage of the FVC2002DB1 database. The GMMs were trained for different number of Gaussians (1024, 2048, and 4096) and their results were reported separately. Once the universal models were trained, we extracted first order statistics of fingerprint samples from FVC2002DB1 and produced supervectors for GMMs with different number of Gaussians, which resulted in supervectors of dimensions 51200 (1024 × 50), 102400 (2048 × 50), and 204800 (4096 × 50).

For the enrollment of target fingerprints, we trained an SVM for each fingerprint sample using the target GMM supervector and the set of imposter GMM supervectors labeled as -1, using the first impression of each subject as imposters. The weight vector of the SVM classifier model was the template for the enrolled fingerprint sample. During verification, GMM supervector of the query fingerprint was compared to the template of the claimed identity and their inner product was used to give accept or reject decision.

The verification protocol was as follows:

i) Each impression was matched against the remaining impressions of the same finger. The total number of genuine tests was 5600 (8 × 7 × 100).

ii) The first impression of each finger was matched against the first impression of the remaining fingers. The total number of impostor tests was 9900 (99 × 100).

For both cases, symmetric matching (i.e., \( f_{p1} vs f_{p2} \) and \( f_{p2} vs f_{p1} \)) was executed as \( w^T f_{p1} x_{f_{p2}} \) is different from \( w^T f_{p2} x_{f_{p1}} \).

Equal error rates (EERs) for GMMs with different number of Gaussians are shown in Table 2. The optimal \( C \) value for training SVMs corresponding to different number of Gaussians were found by grid search and best \( C \) values were 10, 0.001, and 1 for 1024, 2048, and 4096 Gaussians, respectively. As the number of Gaussians in the GMMs increases, our method performs better in representing feature distribution which eventually leads to lower error rates.

In order to provide comparison with our system, we also performed direct minutiae matching with the commercial algorithm which we also used for minutiae extraction. It obtained 0.50% EER on FVC2002DB1 and performed better than our proposed method. This difference stems from the facts that we can neither perform minutiae pair search, which is a crucial step for minutiae matching, nor include singular point information. However, the main purpose of this study is to present a fixed length fingerprint representation and this performance drop was expected. Two other fixed length approaches that can be compared with our system are the spectral minutiae representation \( [27] \) and binary feature vector representation \( [4] \). However, they do not report error rates for the FVC2002DB1 database. When we analyzed their reported results on the FVC2002DB2 database (2.48% \( [27] \) and 3.88% \( [4] \) EERs compared to minutiae matching 1.0% on FVC2002DB2), we also observed similar performance drop.

### Table 2. Equal error rates for GMMs with different number of Gaussians

<table>
<thead>
<tr>
<th># Gaussians</th>
<th>1024</th>
<th>2048</th>
<th>4096</th>
</tr>
</thead>
<tbody>
<tr>
<td>EER</td>
<td>2.23%</td>
<td>1.85%</td>
<td>1.63%</td>
</tr>
</tbody>
</table>

6. Conclusion

The GMM-SVM based feature representation is a novel method to create a fixed length feature vector for fingerprint minutiae. Although minutiae-based matching is the most widely used technique in fingerprint verification/identification, the increasing security and privacy concerns make minutiae template protection one of the most crucial tasks. The main motivation of this study is to obtain a fixed length feature vector for fingerprints so that minutiae based fingerprint verification can be combined with template protection schemes. In addition, our method avoids the difficulties of minutiae registration by representing minutiae patches on a normalized coordinate system defined by the orientation of the central minutia. Also, the major problem of elastic distortion in fingerprint matching is compensated with the local representation of the minutiae neighborhoods.

This paper introduces a fixed length feature representation for variable length minutiae of a fingerprint. In order to combine our method with the cryptographic primitives for template protection, such as \( [8] \), one should extract bits that are stable for genuine users and completely random for arbitrary users. Another possible direction for template protection might be random projection-based biometric hashing \( [13] \), which cannot be directly applied to minutiae templates. In the future, we will work towards protecting feature vectors that are created by our approach and include the binarization of the GMM-SVM feature vectors. We also conjecture that enriching the database that is used in training GMMs and using random resampling \( [22] \) for addressing data-imbalance problem in SVM will be possible improvements to our GMM-SVM minutiae representation.

---

4 Additional fingerprint features that are not defined in ISO minutiae template were not used in any of the experiments.

5 We plan to analyze the performance of our system and make comparison on the FVC2002DB2 database in the future.
References


