# Principal Component Analysis for Minutiae Verification on Fingerprint Image

CHING-TANG HSIEH, SHYS-RONG SHYU
Department of Electrical Engineering
Tamkang University
151 Ying-chuan Road, Tamsui, Taipei County, Taiwan 25137
TAIWAN
Tel: +886-2-2620-5031, Fax: +886-2-2620-9814

*Abstract:* Minutiae are the two most prominent and well-accepted classes of fingerprint features arising from local ridge discontinuities: ridge endings and ridge bifurcations. However, the preprocessing stage doesn't eliminate all possible defects in the original gray-level image and the orientation estimation in a poor image is extremely unreliable. In order to further eliminate the false minutiae caused by low quality, a minutiae verification mechanism is proposed to improve the identification per- formance. The minutiae verification mechanism uses the concept of eigen-codebook to find the optimal projection bases for true minutiae regions and false minutiae regions. Experimental results show that the performance is improved efficiently even less training data.

*Keywords*: Fingerprint image postprocessing, minutiae-based, principle component analysis, local ridge orientations, automatic fingerprint identification system

## 1 Introduction

A fingerprint is the pattern of ridges and valleys on the surface of a fingertip. A total of eight different types of local ridge/valley descriptions have been identified [1], and most of them are not used in automatic fingerprint identification system (AFIS). Instead, in accordance with the representation of fingerprints in the U.S. Federal Bureau of Investigation (FBI) [2], ridge endings and bifurcations, called minutiae, are taken as the distinctive features of the fingerprints. The corresponding coordinate, angle, and type of the features are used to represent the fingerprint in the matching process. The most prominent classes of fingerprint matching methods include minutiae-based approaches and image-based approaches [3]-[9]. Minutiae-based approaches are the most popular ones to identify and verify the fingerprint patterns. Due to imperfections of the acquired image, the false minutiae appear as the disconnection or misconnection of the ridge in a fingerprint image.

Typical approaches of minutiae-pruning in the fingerprint image postprocessing stage can be categorized as: without training structure-based and supervised training classifier-based. O'Gorman and Nickerson [4] calculated the statistical characteristics in an n-by-n matrix along the rows and columns of the image to remove the false minutiae. Xiao and Raafat [3] described methods to identify spurious minutiae and eliminate them using the structural definition of minutiae. For each minutia, statistics of ridge length, ridge direction and minutiae direction were used to decide the spurious minutiae. Ratha et al. [6] used heuristics based on the structural and spatial relationship of the minutiae to validate spurious feature points. Deriche [8] used the inter-ridge pixel distance as a guide to eliminate these false minutiae. The process of elimination took the neatest neighboring minutiae pairs into account to accurately discard false minutiae. These researchers have proposed minutiae- pruning based on without training structure or is so called rather ad-hoc technology in the postprocessing stage to eliminate spurious minutiae.

Prabhakar et al. [9] proposed a minutiae verification based on reexamining the gray-scale profile in a detected minutiae's spatial neighborhood in the sensed image. The minutiae verification used the minutiae classifier, which based on supervised training with Learning Vector Quantization (LVQ) [10] to learn the characteristics of minutiae in gray level image. Prabhakar's algorithm is easily implemented, but it has the following drawbacks: (1) it needs large training data, (2) reference template depends on the initialization and input sequence of training data, (3) owing to the multi-variances of minutiae, it's difficult to find the representative and optimal reference template for the system. In this paper, we use the concept of eigen-codebook for the minutiae verification mechanism of fingerprint image. We utilize the principal component analysis (PCA) technology to find the optimal projection bases for true minutiae regions and false minutiae regions. This paper is organized as follows. In Section II, we briefly introduce a sequential algorithm of minutiae extraction. Minutiae verification mechanism using PCA is described in Section III. Section IV presents experimental results. Concluding remarks are given in Section V.

## 2 Minutiae Extraction

The flow chart of minutiae extraction algorithm is shown in Fig.1. First, the fingerprint image enhancement algorithm was proposed by authors [11] that is based on global textural filtering and local oriented averaging is used to improve the quality of ridge-valley structure of the input fingerprint image. Then, a simple regional average thresholding (RAT) scheme [12] is used to segment the ridge region from the input image. Once the ridges are located, directional smoothing is applied to smooth the ridges. Ratha et al. [6] use a $3\times 7$ mask to smooth the ridge according to the orientation field of each block. After this smoothing process, some of the lost ridges can be reconstructed and the spurious ridges can also be deleted. Before the features can be extracted, the fingerprints have to be thinned or skeletonized so that all the ridges are one pixel thick. In our evaluation, the thinning algorithm proposed by Zhou et al. [13] is used because of the satisfactory performance and the reasonable computational time. Finally, The minutiae can be extracted from the thinned image by using the Crossing Number (CN) at a point P, which is expressed as:

$$CN = \sum_{i=1}^{8} |P_i - P_{i-1}| \qquad P_8 = P_0 \qquad (1)$$

Where $P_i$ is the pixel value (0 or 1) in a $3\times 3$ neighborhood of P, as shown in Fig.2. According to the characteristics of CN, the minutiae can be easily obtained and the minutiae points are recorded using three-dimensional feature vector, namely, the x-coordinate, the y-coordinate and the local ridge direction $\theta$.

## 3 Minutiae Verification mechanism

### 3.1 Principal Component Analysis

The PCA method uses the statistical distribution of input samples to find the best projection bases. It is widely used in the computer vision application [14]-[16]. The advantages of PCA method are that the principal eigenvectors are orthogonal and represent the directions where the signals have maximum variation. This property will speed up the convergence of model training and improve the system performance. The PCA method tries to find the projection of the feature vector on a set of base vectors. Let X={ $x_t$, $t$=1, 2… $M$ } be a set of $M$ $n$-dimensional feature vectors. Hence, the covariance matrix C of X can be found as follows:

$$C = \frac{1}{M}\sum_{t=1}^{M}(x_t - \bar{x})(x_t - \bar{x})^T \qquad (2)$$

Where

$$\bar{x} = \frac{1}{M}\sum_{t=1}^{M} x_t \qquad (3)$$

And $M$ is the number of n-dimensional feature vectors. The projection vector, $\Omega = \{\omega_1, \omega_2, ......, \omega_m\}$, describes the distribution of eigenvectors on m-dimensional linear subspace (eigenspace). This weight is used to represent the input minutiae image and can be obtained as follows:

$$\omega_k = \phi_k^T\ (x_k - \bar{x}) \quad ,k = 1,\ldots\ldots M \quad (4)$$

Where $\phi_k$ is the eigenvector. The transformation matrix $\Phi$ is formed by the principle eigenvectors $\phi_1$ ,$\phi_2$ … $\phi_m$ (m < M) of the covariance matrix C, m is the number of projection bases, and the eigenvectors $\phi_k$ are organized in $\Phi$ in such a way that all eigenvectors indices are in a descending order corresponding to their respective eigenvalues ($\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_m$).

### 3.2 Eigen-Codebook Vector Quantization

The well-known technology of VQ is called the Linde-Buzo-Gray (LBG) algorithm [17], which is based on two necessary conditions for optimality: the centroid and the nearest neighbor conditions. In this paper we use the modified VQ algorithm [18], which was proposed for text-independent speaker identification experiments. While sample vectors including false minutiae region, ridge endings region and ridge bifurcations region belong to a code word have been processed by PCA method, the evaluated mean vector $\bar{x}$ and eigenvectors $\vec{\phi}_1, \vec{\phi}_2, ..., \vec{\phi}_m$ may effectively describe the distribution of the samples. We propose a modified vector quantization methodology called eigen-codebook vector quantization (ECVQ) as the identifier for minutiae verification. This model uses the principal component analysis (PCA) method to evaluate the codebooks for capturing more details of true minutiae and false minutiae region's characters.

If there is a new sample $x_{new}$ for classification, it will be adjusted by the mean vector and project to the m orthonormal eigenvectors as follows:

$$p_i = (x_{new} - \bar{x})^T \cdot \vec{\phi}_i \qquad i = 1,2,...,m \qquad (5)$$

Where $p_i$ is the projection of $x_{new}$ on the $i$th projection base $\vec{\phi}_i$, and $m$ is the number of the projection bases. Finally, the error vector $\varepsilon$ between $x_{new}$ and the projection of $x_{new}$ on the eigenspace of the code word is evaluated as follows:

$$\varepsilon = x_{new} - \bar{x} - \sum_{i=1}^{m} p_i \vec{\phi}_i \qquad (6)$$

If the norm $\|\varepsilon\|$ is smaller, the new sample $x_{new}$ is more closing to the codeword.

The training procedure of the proposed ECVQ is described as follows:

Step 1: set the number of the projection bases m, number of code words $C_{no}$ and maximum of training times N.

Step 2: use modified LBG method to evaluate the centroids of all code words and classifies all training samples.

Step 3: use PCA method to evaluate the mean vector $\bar{x}$ and eigenvectors $\vec{\phi}_1, \vec{\phi}_2, ..., \vec{\phi}_m$ for every code word.

Step 4: use (6) to evaluate the error vector $\varepsilon$ and classify the training samples by $\|\varepsilon\|$.

Step 5: If the training times are less than N and samples classification is not convergent, go to step 3.

Step 6: store the mean vector $\bar{x}$ and eigenvectors $\vec{\phi}_1, \vec{\phi}_2, ..., \vec{\phi}_m$ for every code word.

## 4 Experimental Result

The proposed algorithm is evaluated using NIST4 database. Samples include 500 fingerprint images. The interested fingerprint region of 512 x 480 pixels within the original image is segmented and the divided block size n is set as 25 x 25. After capturing the required fingerprint region from the input image, the minutiae extraction algorithm described in sec. 2 is applied to extract the minutiae points. We randomly choose 10 fingerprint image minutiae for training to label the ridge ending, ridge bifurcation and false fingerprint minutiae by visual inspection and then find the optimal projection bases of each labeling with PCA. Through the matrix projection mechanism, the large dimensional $n \times n$ matrix is transformed to a small dimensional $T \times T$ matrix. Hence, the computation time is largely decreased. Fig.3 shows some examples of the training samples and its corresponding projection bases of fingerprint image minutiae in left and right side, respectively. Moreover, these figures represent ridge endings region, ridge bifurcations region, and false minutiae region from top to bottom, respectively.

Fig.4(a) describes that the initial filter mechanism of minutiae can effectively remove the false minutiae whose are near the edge region of fingerprint. But it's not satisfactory result only using the filtering mechanism for the noise region of fingerprint. After minutiae verification, the proposed algorithm can identify the spurious minutiae in the noise region, but also it can maintain the valid features in the clear region, as shown in Fig.4(b).

## 5 Conclusion

In this paper, we use the concept of eigen-codebook for the minutiae verification mechanism of fingerprint image. We utilize the principal component analysis (PCA) technology to find the optimal projection bases for true minutiae regions and false minutiae regions. Experimental result show that the performance is improved efficiently even less training data.

*References:*


[1] Henry C. Lee and R. E. Gaensslen, editors, "*Advances in Fingerprint Technology*", Elsevier, New York, 1991.

[2] K. Hrechak and J. A. McHugh, "Automated fingerprint recognition using structural matching", *Pattern Recognition*,Vol.23, pp.893-904. , 1990

[3] Q. Xiao and H.Raafat, "Fingerprint image processing: A combined statistical and structural approach", *Pattern Recognition*, Vol.24, No.10, 985-992, 1991.

[4] L. O'Gorman and J. Nickerson," An approach to fingerprint filter design", *Pattern Recognition*, Vol.22, No.1, pp.29-38, 1989.

[5] D. C.D. Hung, "Enhancement and feature purification of fingerprint images", *Pattern Recognition*, Vol.26, No.11, pp.1661-1671, 1993.

[6] N. Ratha, S. Chen, and A.K. Jain, "Adaptive flow orientation-based feature extraction in fingerprint

images", *Pattern Recognition*, Vol.28, No.11, pp.1657-1672, 1995.

[7] A.K. Jain, L. Hong, S. Pankanti and R. Bolle, "An identity authentication system using fingerprints", *Processing of the IEEE*, Vol.85, No. 9, pp.1365-1388, 1997.

[8] M. Deriche, "An algorithm for reducing the effect of compression/decompression technique on fingerprint minutiae", *Seventh Australian and Newzealand Intelligence information systems Conference*, 18-21, pp.243-246, Nov. 2001.

[9] S. Prabhakar, A.K. Jain, J. Wang, S. Pankanti, and R. Bolle, "Minutia verification and classification for fingerprint matching", *IEEE. Conf.* pp.25-29, 2000.

[10] T. Kohonen, J. Kangas, J. Laaksonen, and K. Torkkola, "LVQ_PAK: A program package for the correct application of learning vector quantization algorithm", *in Proc. Int. Joint Conf. On Neural Networks*, (Baltimore), pp.1725-1730, June 1992.

[11] C.T. Hsieh, E. Lai, and Y.C. Wang, "An effective algorithm for fingerprint image enhancement based on wavelet transform", *Pattern Recognition*, Vol.36, No.2, 2003.

[12] Emiroglu, M. B. A khan, "Pre-processing of Fingerprint Images", *ECOS 97, European Conference on Security and Detection.*, 28-30, pp. 147-151, 1997.

[13] R. W. Zhou, C. Quek and G. S. NG, "Novel single-pass thinning algorithm", *Pattern Recognition. Lett.*, Vol.16, No.12, pp.1267-1275, 1995.

[14] Hurase, H., and Nayar, S., "Visual learning and recognition of 3D objects from appearance," *Int'l J. Computer Vision*, vol.14, pp. 5-24, 1995.

[15] Belhumeur, P. N., Hespanha, J. P., and Kriegman, D. J., "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.19, no.7, pp. 711-720, July 1997.

[16] Martinez, A. M., and Kak, A. C., "PCA versus LDA," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.23, no.2, pp. 228-233, Feb. 2001.

[17] Linde, Y., Buzo, A., and Gray, R. M., "An algorithm for vector quantizer design," *IEEE Trans. Commun.*, vol.20, pp. 84-95, 1980.

[18] Lee, D., Baek, S., and Sung, K.: "Modified K-means algorithm for vector quantizer design," *IEEE. Signal Processing Letters*, Vol.4, No.1, pp. 1-5, 1997.

[19] M. Turk, and A. Pentland, "Eigenfaces for recognition", *Journal of Cognitive Neuroscience*, vol.3, no.1, pp. 71-86, 1991.

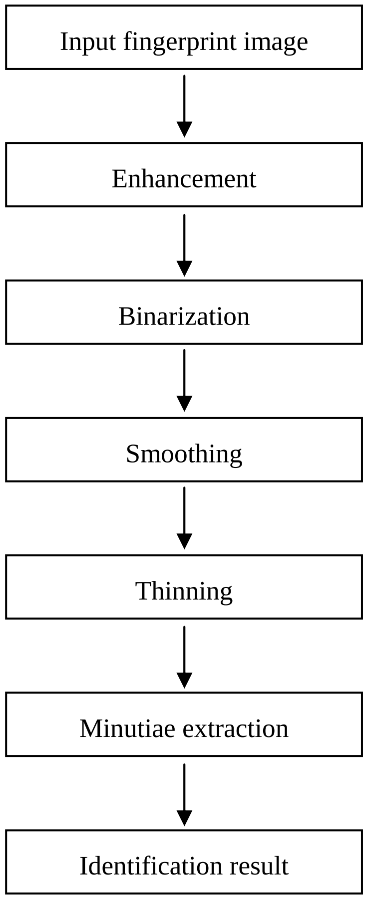
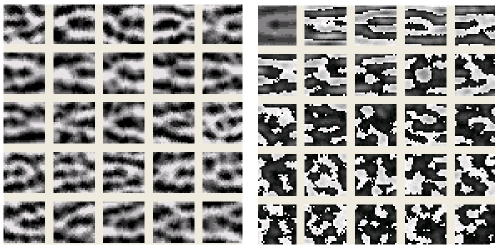


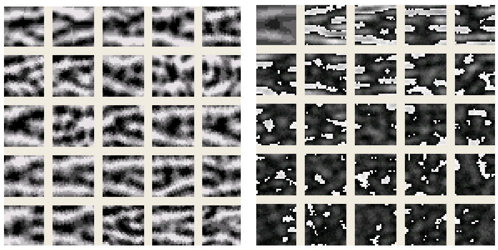

Fig.1 The flowchart of the minutiae extraction algorithm.

| | | |
|---|---|---|
| $P_0$ | $P_1$ | $P_2$ |
| $P_7$ | $P$ | $P_3$ |
| $P_6$ | $P_5$ | $P_4$ |

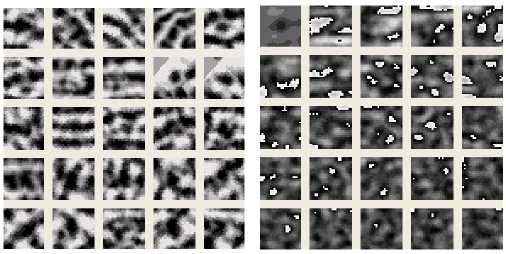
Fig.2 A $3 \times 3$ neighborhood mask of $P$ .



(a) ridge endings region

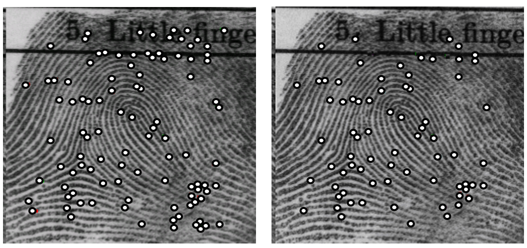


(b) ridge bifurcations region



(c) false minutiae region

Fig.3 Some examples of the training samples and its corresponding projection bases of the extracting minutiae.




(a) (b)

Fig.4 The results of (a) before and (b) after minutiae verification.