International Journal of Pattern Recognition
and Artificial Intelligence
Vol. 25, No. 8 (2011) 1243–1260

DOI: 10.1142/S0218001411009135

MODIFIED PRINCIPAL CURVES BASED FINGERPRINT MINUTIAE EXTRACTION AND PSEUDO MINUTIAE DETECTION

HONGYUN ZHANG* and DUOQIAN MIAO†

Department of Computer Science and Technology
Tongji University
Shanghai 201804, P. R. China

Key Laboratory of Embedded System & Service Computing
Ministry of Education of China
Shanghai 201804, P. R. China
**zhanghongyun583@sina.com*
†*miaoduoqian@163.com*

CAIMING ZHONG

College of Science and Technology, Ningbo University
Ningbo 315211, P. R. China
zhongcaiming@nbu.edu.cn

It is difficult but crucial for minutiae extraction and pseudo minutiae deletion of low quality fingerprint images in auto fingerprint identification systems. Traditional methods based on thinning images or gray-level images are, however, susceptible to noise. Reference 14 indicated that principal curves based fingerprint minutiae extraction was feasible to overcome the drawback, but the extended polygonal line (EPL) principal curves algorithm used in the paper extracted the principal curves ineffectively. As the fingerprint data sets are usually large, the original EPL principal curves algorithm is time-consuming. Meanwhile, scattered fingerprint data lead to the deviation of fingerprint skeleton. In this paper, the algorithm is modified, and a fingerprint minutiae extraction and pseudo minutiae detection method based on principal curves is proposed. Experimental results show that the modified EPL principal curves algorithm outperforms the original EPL algorithm both in efficiency and quality, and the proposed minutiae extraction method outperforms the methods proposed by Miao under noise conditions.

Keywords: Principal curves; skeletonization; fingerprint minutiae extraction; pseudo minutiae detection.

1. Introduction

The uniqueness of fingerprints is determined by the characteristics and the relationships of local ridges, which are also called minutiae.[6] Minutiae are local

*Author for correspondence

discontinuities in terms of ridge endings and bifurcations of ridge flow patterns that constitute a fingerprint. The ridge ending is defined as the ridge point at which a ridge terminates, while the ridge bifurcation is defined as a point at which a single ridge splits into two. These two types of minutiae have been considered by the Federal Bureau of Investigation for the purpose of identification.[22] The main task of the fingerprint minutiae extraction algorithm is to identify the quantity, type, position and direction of the minutiae. In a traditional automatic fingerprint identification system, fingerprint minutiae are mainly extracted based on thinning images or gray-level images.[5,10,13,15,17,19,21,23] However, the thinning images and gray-level images are a set of pixels. While looking at the image of a fingerprint, it is often regarded as a collection of curves instead of a set of pixels. Furthermore, traditional methods based on thinning images or gray-level images are sensitive to noise pixels. Therefore, we try to figure out a method to depict fingerprint images with a collection of curves rather than a set of pixels. In the paper, we employ principal curves to describe the structural information of fingerprint, since they can reflect the genuine structure of a data set.[20,21,26] Currently, considerable work has been reported concerning the applications of principal curves.[7,20,28] The definition of principal curves is provided by Hastie and Stuetzle as a set of self-consistent smooth curves which pass through the "middle" of a multidimensional probability distribution or data cloud.[4,18,25] Considering a fingerprint as a data cloud, a set of principal curves can be used to extract the skeleton of the fingerprint with the following advantages:

(1) Since principal curves are self-consistent smooth curves, they can be employed to form a fingerprint skeleton from a set of pixels, and the skeleton can describe the fingerprint information more accurately. For example, the endings and directions of the skeleton can be used to identify the pseudo minutiae.
(2) Some useful information, such as the topology, etc., can be perceived from principal curves. With the information, the fingerprint minutiae and veins can be extracted easily.
(3) Principal curves contain the global fingerprint information, such as the information of adjacent principal curves, involving direction, length, location, etc., which can benefit for fingerprint feature extracting and matching.

Although our previous work indicated that principal curves based fingerprint minutiae was feasible,[14] the original principal curves algorithm proposed by Kegl [8,9] cannot extract the fingerprint skeleton effectively. Since a fingerprint dataset is usually large, the original principal curves algorithm is time-consuming. At the same time, scattered fingerprint data cause the deviation of fingerprint skeleton. In order to address the two problems, the penalty function, fitting-smoothing step and projection step of the original algorithm are improved. Furthermore, a feature extraction algorithm based on modified principal curves is proposed. Experimental results demonstrate that the modified EPL principal curve algorithm is much more efficient. The obtained fingerprint skeleton is more accurate than the original one. The feature extraction algorithm based on modified principal curves shows better performances than the method proposed by Miao[14] and the traditional algorithms based on thinning and gray-level images.

The rest of the paper is organized as follows. In Sec. 2, the modified EPL principal curve algorithm is introduced, and a comparison between the modified algorithm and the original one is also given. In Sec. 3, the approach of minutiae extraction and pseudo minutiae elimination is provided in detail. In Sec. 4, the experimental results and analysis are presented. Lastly, in Sec. 5, the main conclusions are covered.

2. Definition of Principal Curves and Modified EPL Principal Curve Algorithm

2.1. *Definition of principal curves*

Hastie and Stuetzle defined a principal curve, and emphasized its self-consistency property (property 3) as follows.[4,25]

Definition 1. The smooth curve $f(s)$ is a principal curve if and only if:

(1) $f(s)$ does not intersect itself,
(2) $f(s)$ has finite length inside any bounded subset of R^d,
(3) $f(s)$ is self-consistent, i.e. $f(s) = E[X|s_f(X) = s]$, and

$$s_f(X) = \sup\left\{s : \|X - f(s)\| = \inf_{\tau} \|Y - f(\tau)\|\right\} \quad (1)$$

The definition indicates that any point of a principal curve is the conditional expectation of those points that project to this point, and a principal curve satisfies the property of self-consistency. The theoretical foundation of a principal curve is a low-dimensional nonlinear manifold embedded in a high-dimensional space.[25] Principal curves are a nonlinear generalization of principal component analysis. Figure 1 shows a first principal component line and a principal curve.[8] Compared

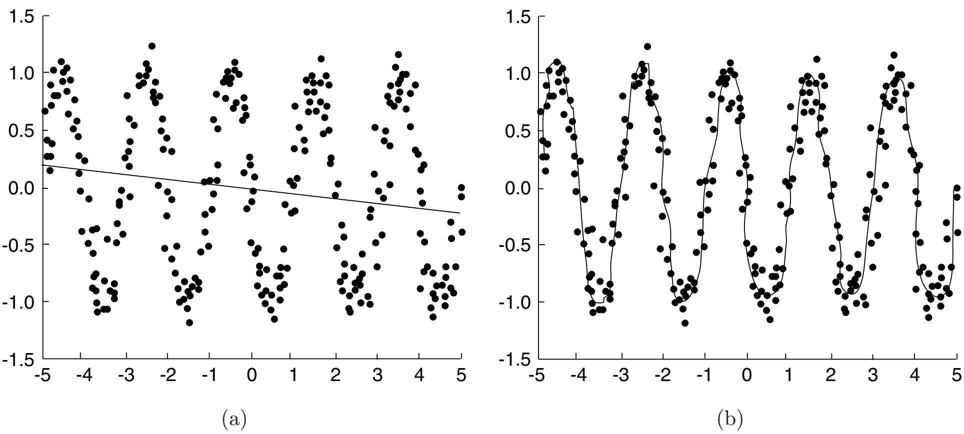


Fig. 1. (a) First principal components, (b) principal curves.

with corresponding first principal component, two obvious advantages of a principal curve can be observed. Firstly, a principal curve can keep more information of data; and secondly, it can describe the outline of primitive information better. However, Hastie and Stuetzle's principal curves cannot depict fingerprint skeletons directly. From the work of Zhang,[26,27] we learn that the extended polygonal line (EPL) algorithm can deal with the dataset which are concentrated along a highly curved or self-intersecting curve. In this paper, therefore, EPL principal curves algorithm is employed to describe fingerprint skeletons.

2.2. *Modified EPL principal curves algorithm*

The EPL principal curves algorithm is proposed by Kegl to extract the skeleton of digit images. Since a fingerprint dataset is different from a digit dataset, in the experiment of extracting the fingerprint skeleton, we find that the EPL principal curves algorithm is time-consuming and cause the deviation of fingerprint skeleton. In this paper, the EPL principal curves algorithm is modified to solve the problems. In this section, the EPL principal curves algorithm is introduced and analyzed based on which the modified EPL principal curves algorithm is presented.

2.2.1. *EPL principal curves algorithm*

In this paper, we adopt EPL principal curves algorithm proposed by Kegl[8,9] to extract the skeleton of fingerprints. The algorithm contains the following mains steps:

Step 1 (Initialization). A thinning algorithm is adopted to obtain the approximate initial skeleton of a fingerprint image. The initial skeleton captures the approximate topology of the fingerprint, and it roughly follows the medial axis of the fingerprint. However, it is not smooth and usually contains a number of spurious branches and inadequate structural elements. The skeleton is denoted by G_{VS} which consists of V and S, where $V = \{v_1, \ldots, v_m\}$ is a set of *vertices*, and $S = \{(v_{i1}, v_{j1}), \ldots, (v_{ik}, v_{jk})\} = \{s_{i1,j1}, \ldots, s_{ik,jk}\}$, $1 \leq i1, j1, \ldots, ik, jk \leq m$ is a set of *edges*, such that s_{ij} is a line segment that connects v_i and v_j.

Step 2 (Fitting-smoothing). Iteratively fit and smooth the skeleton by repeatedly projecting data point and optimizing vertex until convergence is achieved, while keeping the skeleton approximately equidistant from the contours of the fingerprint.

Step 2.1 (Projection). Given a data set $X_n = \{x_1, \ldots, x_n\}$, scan the whole skeleton for every data point x_i, the data point x_i is partitioned into "the nearest neighbor regions" according to which segment or vertex projects. This step is time-consuming, since thousands of scans are required.

Step 2.2 (Optimization). Every vertex v_i in the skeleton is optimized by using a gradient method to adjust the positions of vertexes and segments for finding a local minimum of $E(G)$.[8] The penalized distance function $E(G)$ is:

$$E(G) = \Delta(G) + \lambda P(G) \tag{2}$$

$$\Delta(G) = \frac{1}{n}\sum_{i=1}^{n} \Delta(x_i, G) \tag{3}$$

$$P(G) = \frac{1}{m}\sum_{i=1}^{m} P_v(v_i) \tag{4}$$

Step 2.3. If the adjusted skeleton meets the convergent condition, go to Step 3, and otherwise go to Step 2.1.

Step 3 (Restructuring). Rectify the structural imperfections of the skeleton graph by deleting short paths and small loops to get more accurate skeleton.

Step 4 (Fitting-smoothing). This step is the same as Step 2.

Note that in Step 2.2, $\Delta(G)$ is the average squared distance of all points X_n from G_{VS}. $P(G)$ is a penalty on the total curvature of the skeleton. The smaller the value of $\Delta(G)$, the better G_{VS} fits the data. On the contrary, the smaller the value of $P(G)$, the smoother G_{VS} is. λ is a penalty coefficient that determines the trade-off between the accuracy of the approximation and smoothness of the curves. $\Delta(x_i, G)$ is the Euclidean squared distance between a point x_i and the nearest point of the skeleton to x_i. n is the number of the data points. $P_v(v_i)$[8] is the curvature penalty at vertex v_i. m is the number of vertices. In general, $P_v(v_i)$ is small if line segments incident to v_i join smoothly at v_i.

2.2.2. *EPL principal curves algorithm analysis*

As a special information carrier, the fingerprint dataset has a lot of specific characteristics. When the original EPL principal curves algorithm is used to extract the fingerprint skeleton, those specific characteristics may cause several problems as follows:

(1) A fingerprint image contains large volume of information, and consists of thousands of data points. When EPL algorithm is applied, all the data points can be considered as candidates of skeleton vertices. Consequently, the original EPL algorithm is time-consuming.

(2) The data distribution in a fingerprint image is very scattered. In other words, the ratio of skeleton vertex number to the data point number, which is denoted as α, is small. In general, the value of α is between 4 and 6, sometimes even only between 2 and 3. This indicates that only a small number of data points on average are involved in detecting skeleton vertices, and result in the deviation of the adjusted skeleton.

2.2.3. *The improved EPL principal curve algorithm*

In order to address the problems, firstly, Vertices-Merge step is added between the Initialization Step and the Projection Step. The reason is as follows: (a) A fingerprint image includes thousands of vertices, and the Vertices-Merge step can effectively reduce the number of vertices which need to be adjusted. As a result, the efficiency of the algorithm is improved. (b) Vertices-Merge step increases the proportion of data points to skeleton vertices. More data points on average are involved in adjusting a skeleton vertex so that the deviation of skeleton can be controlled to a certain degree.

Secondly, considering the scattered distribution of fingerprint data points, the Projection Step of original algorithm is modified. The step only scans certain areas of the vertices around the data points instead of the whole skeleton. It costs much less time than the original Projection Step. Furthermore, according to the distribution of fingerprint data points, the penalized distance function $E(G)$ is redefined in Vertex-Optimization Step. $D(G)$ is introduced to reduce the skeleton deviation. Since $D(G)$ only involves simple additions, it is more efficient than $P(G)$ which uses a triangle function.

Finally, in terms of the topology of a fingerprint, Step 4 of the original approach is deleted. Two reasons are given as below: (a) The fitting-smoothing step is the most time-consuming part in a principal curve algorithm. Since an ideal skeleton can be generated after the first fitting-smoothing step, it does not need to do this step twice. (b) Short branches and small enclosures are common in the topology of a fingerprint. In the restructuring step, they are deleted. If we do another fitting-smoothing step based on the restructured skeleton, it will sabotage the skeleton. One of the experimental results is shown in Fig. 2. The flow chart of the improved principal graph algorithm is shown in Fig. 3. The modified algorithm is composed of the following steps:

Step 1. The Initialization Step: The step is the same as Step 1 of the original EPL principal curve algorithm.

Step 2. Vertices-Merge Step: The step merges the adjacent vertices in terms of distance and curvature. The distance criterion makes sure that a number of vertices are retained within a certain area coverage, while the curvature criterion is set to reduce the number of vertices merged in the area of a big curvature.

Step 2.1. Merging based on distance criterion: v_l and v_r are the left adjacent vertex and the right adjacent vertex of the vertex v respectively. We denote the distance between v and v_l by $d(v_l, v) = \|v_l - v\|$. Let d_t denote threshold. For every vertex v in G, if $d(v_l, v) + d(v, v_r) < 3d_t$ or $d(v_l, v)/d(v, v_r) > 4$ or $d(v_l, v)/d(v, v_r) < 0.25$, then v_l and v_r are connected, and v is removed. Threshold d_t is computed by

$$d_t = \frac{1}{n}\sum_{i=1}^{n} d(v_i, G) \tag{5}$$

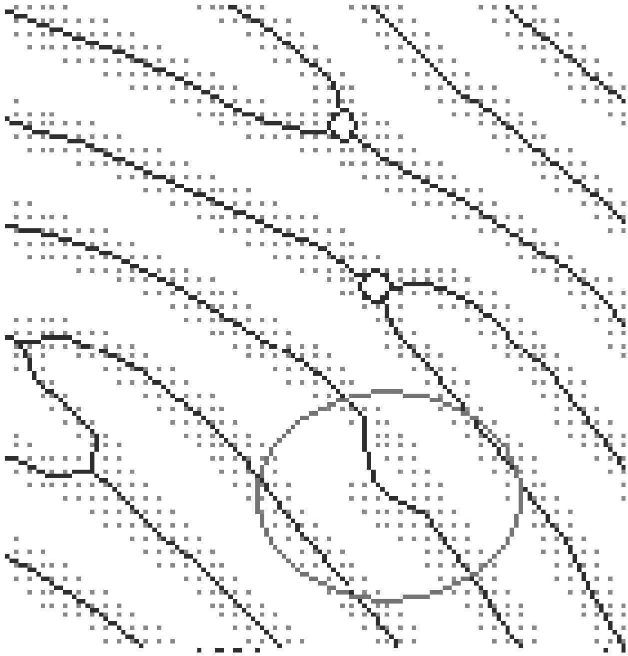


Fig. 2. Bad skeletons caused by the second fitting-and-smoothing after deleting the short branch.

Step 2.2. Merging based on curvature criterion: We denote the directional angle of v_l and v by $\theta_{v_l v}$; let β denote the included angle contained by vectors $v_l v$ and $v v_r$, where β is smaller than π. If $\pi - \beta < 20^\circ$, then v_l and v_r are connected, and v is removed.

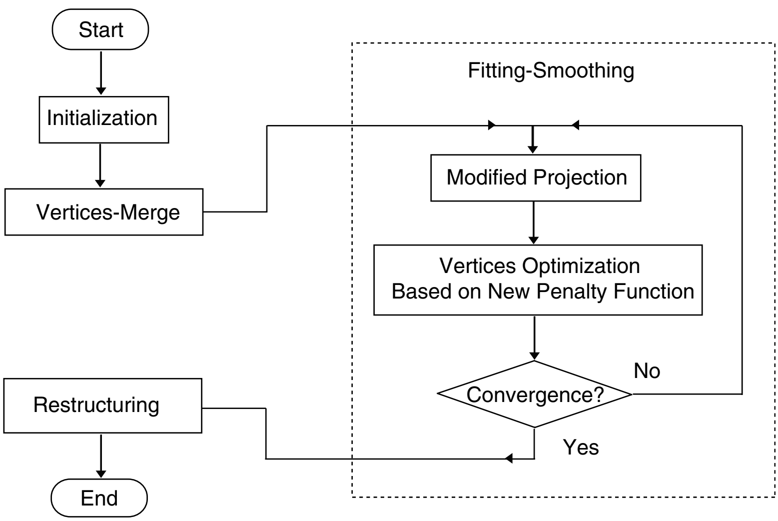


Fig. 3. The flow-chart of the modified algorithm.

Step 3. The Fitting-smoothing Step: Compared with Step 2 of the original algorithm, we improve the Projection Step and redefine the penalized distance function.
Step 3.1. The Improved Projection Step: Due to the scattered fingerprint data distribution, the step only scans certain areas of the vertices around the data point x_i and partitions it into the nearest neighborhood region instead of the whole skeleton. A data point belongs to a fingerprint line data whose width is normally 4 to 5 pixels. Since the skeleton passes through the middle of a fingerprint line data, we only need to search the area near the fingerprint line to find its "Voronoi partition." In our experiment, we set the width of 30 pixels which leads to good results.
Step 3.2. The Improved Vertex Optimization Step: We redefine the penalized distance function $E(G)$:

$$E(G) = \Delta(G) + \lambda D(G) \tag{6}$$

$$D(G) = \frac{1}{m} \sum_{i=1}^{m} \sum_{x \in V_i \cup S_i} \Delta(x, v_i) \tag{7}$$

Step 3.3. Judge whether the adjusted skeleton meets the convergent condition, that is whether a local minimum of $E(G)$ is found. If true, go to Step 4, else go to Step 3.1.

Step 4. The Restructuring Step: The step is the same as the restructuring step of the original algorithm.

Step 5. End.

Note that, in Step 3.2, we modify the second component of $E(G)$, $D(G)$ is a redefine penalty function on the total curvature of the skeleton, where, V_i denotes a set of points whose projection point is v_i; S_i denotes a set of points whose projection in edge s_i and point v_i is an endpoint of edge s_i, $\Delta(x_i, v_i)$ is the Euclidean squared distance of a point x_i and vertex v_i.

3. The Approach for Minutiae Extraction and Pseudo Minutiae Elimination

Before an effective minutiae extraction and pseudo minutiae elimination algorithm is constructed, we introduce some preprocesses on the fingerprint image, which include filtering and binarization. In the paper, we apply the enhancement algorithm proposed by Lin *et al.*[12] to enhance images followed by binaralization algorithm. As mentioned above, among various minutiae, ridge bifurcations and ridge endings can determine the uniqueness of fingerprint. As the degree of a ridge bifurcation is 3 and the degree of a ridge ending is 1, we add the vertices whose degree is 1 or 3 to a minutiae point set. Then the pseudo minutiae which are caused by a broken fingerprint ridge or two misconnected fingerprint ridges are deleted. Finally, a convex hull algorithm is adopted to delete the pseudo minutiae close to the border of fingerprint. The main steps are described as follows:

Step 1. Sample a Fingerprint Image.

Step 2. Preprocess the Fingerprint Image.

Step 3. Find Principal Curves. The modified EPL principal curve algorithm is used to obtain the principal curves, which can be served as the skeleton of a fingerprint.

Step 4. Extract Minutiae. A fingerprint skeleton consists of many fingerprint ridges. A fingerprint ridge can be regarded as an undirected graph G_{VS} which consists of V and S. According to the definition of fingerprint minutiae, if the *degree* of *vertex* v_i is equal to 1 or 3, then v_i belongs to minutiae. Traversing all fingerprint ridges, we obtain the set of minutiae $A = \{v_i | d(v_i) = 1, \text{ or}, d(v_i) = 3\}$, where $d(v_i)$ denotes the degree of vertex v_i.

Step 5. Eliminate Pseudo Minutiae. In this step, three kinds of pseudo minutiae are detected and deleted from the set of minutiae A.
Step 5.1. In the set A, for each v_i whose *degree* is equal to 1, if *vertex* v_j $(i \neq j)$ satisfies the following conditions 1–5, then, v_i and v_j are pseudo Ridge endings. v_i and v_j are deleted from the set A. For example, Figs. 4(a)–4(c) belong to this kind of pseudo minutiae.

(1) $d(v_j) = 1$
(2) connect $(v_i, v_j) = \text{false}$
(3) dist $(v_i, v_j) \leq \text{maxDist1}$
(4) $\text{ABS}(\text{ABS}(\text{Angle}(v_i) - \text{Angle}(v_j)) - \pi) < \pi/9 + (\pi/6) * (1 - \text{dist}(v_i, v_j)/\text{maxDist1})$
(5) $\text{ABS}(\text{Angle}(v_i) - \text{Angle}(v_i, v_j)) < \pi/6 + (\pi/3) * (1 - \text{dist}(v_i, v_j)/\text{maxDist1})$

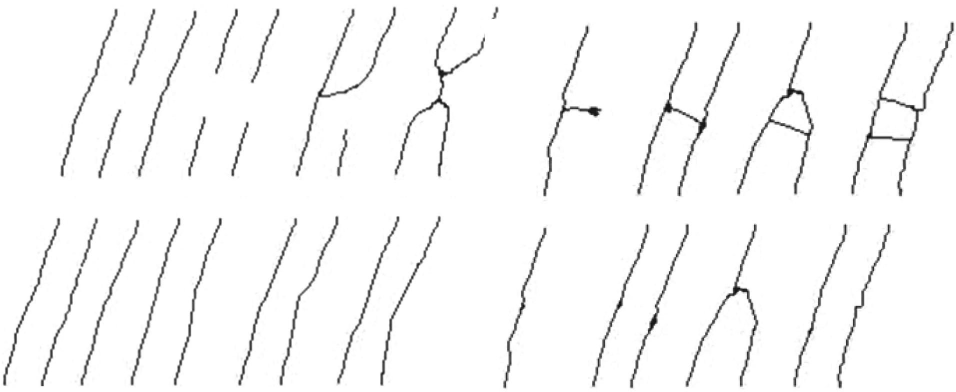
Step 5.2. In the set A, for each *vertex* v_i, if *vertex* v_j $(i \neq j)$ satisfies the following conditions (1)–(2), v_i and v_j are pseudo minutiae. v_i and v_j are deleted from the set A. For example, Figs. 4(d)–4(h) belong to this kind of pseudo minutiae.

(1) connect $(v_i, v_j) = \text{true}$
(2) dist $(v_i, v_j) \leq \text{maxDist2}$

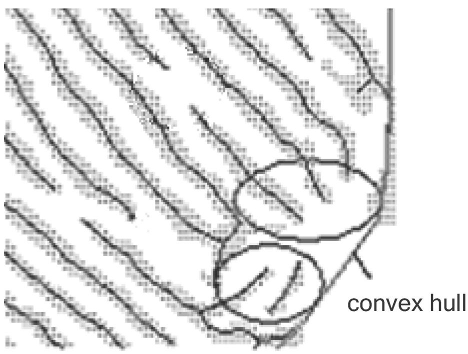
Step 5.3. For each data point $x_i (x_i \in X)$, where X is a set of boundary points of the fingerprint. The convex hull $H(X)$ is calculated as follows:

$$H(X) = \left\{ \sum_{i=1}^{k} \alpha_i x_i | x_i \in X, \alpha_i \in R, \alpha_i \geq 0, \sum_{i=1}^{k} \alpha_i = 1, k = 1, 2, \ldots \right\} \qquad (8)$$

Step 5.4. For each v_i $(v_i \in A)$, if *vertex* v_i satisfies the following conditions (1)–(2), v_i is pseudo minutiae. v_i is deleted from the set A. For example, Fig. 4(i) belongs to



(a) (b) (c) (d) (e) (f) (g) (h)



(i)

Fig. 4. Removal of the most common Pseudo Minutiae.

this kind of pseudo minutiae.

(1) $(d(v_i) = 1) \cap (\text{dist}(v_i, H) \leq \text{maxDist3})$
(2) $(d(v_i) \neq 1) \cap (\text{dist}(v_i, H) \leq \text{maxDist4})$

Step 6. End. A is the fingerprint minutiae set.

Note that, in Step 4, $V = \{v_1, \ldots, v_m\}$ is a set of vertices, and $S = \{(v_{i1}, v_{j1}), \ldots, (v_{ik}, v_{jk})\} = \{s_{i1,j1}, \ldots, s_{ik,jk}\}, 1 \leq i1, j1, \ldots, ik, jk \leq m$ is a set of *edges*, such that s_{ij} is a line segment that connects v_i and v_j. We say that two vertices are *adjacent* or *neighbors* if there is an edge connecting them. The edge $s_{ij} = (v_i, v_j)$ is said to be incident with the vertices v_i and v_j. The vertices v_i and v_j are also called the *endpoints* of s_{ij} . The *degree* of a vertex is the number of edges incident with it. In Steps 5.1−5.4, $d(v_i)$ is the degree of vertex v_i; connect (v_i, v_j) denotes whether the vertex v_i and v_j are connected, connect $(v_i, v_j) = \text{true}$ implies that v_i and v_j are connected, connect $(v_i, v_j) = \text{false}$ implies that v_i and v_j are not connected; dist (v_i, v_j) is the Euclidian distance between vertex v_i and v_j; dist (v_i, H) is the minimum distance between vertex v_i and convex hull $H(X)$; Angle (v_i) is the direction angle of vertex v_i whose degree is 1; Angle (v_i, v_j) is the direction angle of arbitrary vector v_i and v_j; maxDist1, maxDist2, maxDist3, maxDist4 are threshold values, in this paper we denote aw as the average width of fingerprint lines, and by experiments we find that the effect is the best when maxDist1 $= 20 - aw$, maxDist2 $= 20 - aw$, maxDist3 $= 16$, maxDist4 $= 8$;

In Step 4, we first extract minutiae using the information of the principal curves, and then we initialize the minutiae set. In Step 5, we delete three following kinds of pseudo minutiae: (1) A pair of ridge endings caused by a broken fingerprint ridge. (2) A pair of ridge bifurcations caused by two misconnected fingerprint ridges. (3) The minutiae close to the boundary of fingerprint. Several of the experimental results are shown in Fig. 4.

In particular, during the process of minutiae extraction based on principal curves, all the parameters and functions mentioned above can be acquired from principal curves information, which is the significant advantage. Though the principal curves algorithm is more time-consuming than the thinning algorithms, the information which fingerprint principal curves contains can be used to extract minutiae more efficiently and conveniently.

4. Experimental Results and Analysis

4.1. *Experimental results*

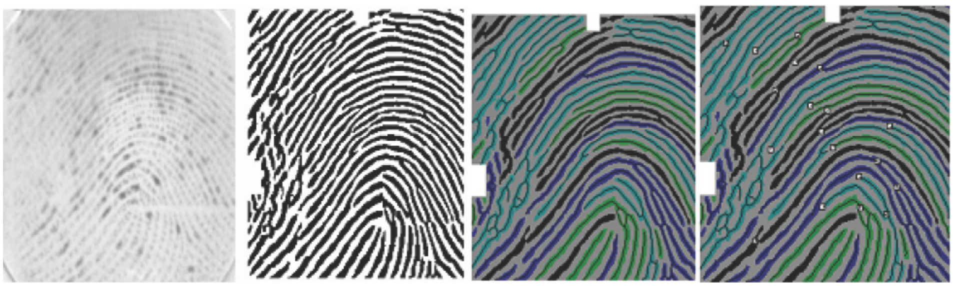
Our methodology is tested on several standard databases. The databases used include: (1) DB1 of FVC2002,[3] (2) DB4 of FVC2002,[3] (3) FERET databases,[1] (4) DB1 of FVC2000,[2] (5) DB2 of FVC2000.[2] For each database, one set of good quality images and one set of bad quality images were manually selected. Each set consists of 100 pieces of images. Good quality images have few bridges and breaks; ridges and valleys alternate and flow in a locally constant direction. On the contrary, in bad quality images, even after enhancement, ridges are not strictly continuous; parallel ridges are not well separated; a lot of creases and bruises exist. Even enhancement algorithms cannot repair them all. The Gabor algorithm was used to enhance various qualities of fingerprint images.[12,16] Fingerprint minutiae detected by algorithms were compared manually with the minutiae presented in the enhanced images. The experimental result shows the improved EPL principal curves algorithm is better in efficiency and quality than the original algorithm, and our minutiae extraction method outperforms the methods proposed by Miao. The results of intermediate stages and the detected minutiae features for a low quality fingerprint are shown in Fig. 5. The performance comparisons of this paper with other methods in several databases are given in Table 1.

4.2. *Performance analysis*

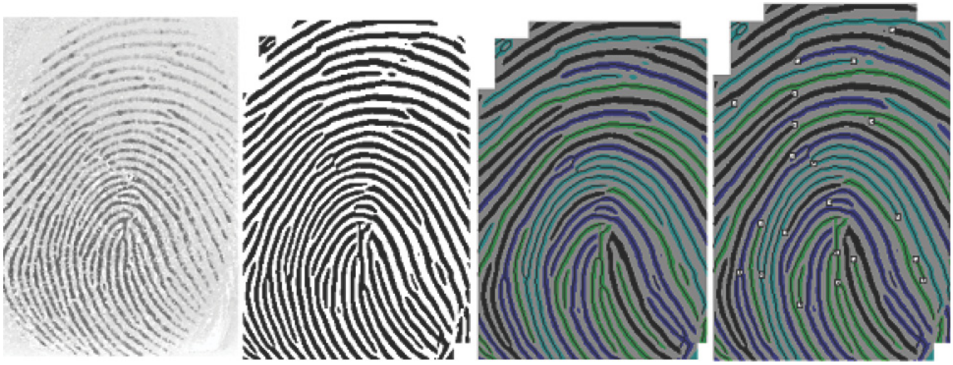
The performance of the proposed method is analyzed in the steps for obtaining fingerprint principal curves and extracting minutiae, respectively.

• The step for obtaining fingerprint principal curves

— The algorithm is implemented on DELL E7200, Intel CORE 2, WinXP, 2.47 gHz. The average running time of the improved algorithm taken for fingerprint principal curves extraction is 0.72 seconds on several standard databases, while that of



(a)



(b)

(c)

Fig. 5. Results of various stages in minutiae extraction tested on low quality images.

(d)

(e)

Fig. 5. (*Continued*)

the original EPL principal curves algorithm is about 1.87 seconds. Therefore, the algorithm proposed in this paper is more efficient than the original one.

— To compare the results of proposed algorithm, we implemented the original algorithms to get the fingerprint's skeleton which includes a set of principal curves. The results of these two algorithms are shown in Fig. 6. By a pairwise comparison, we observe that the skeletons extracted by the improved algorithm can reflect the structures of fingerprint better than the ones generated by the

Table 1. Comparison with Ref. 14 in bad quality fingerprint images.

		Performance Indicators				
DBS	Method	A(%)	A'(%)	Missing(%)	False(%)	Exchanged(%)
FVC 2002 DB1–DB2	Proposed method	85.36	87.51	4.52	7.97	2.15
	Method of Miao *et al.*,[14]	83.32	87.53	4.49	7.98	4.21
FVC 2000 DB1–DB2	Proposed method	85.41	87.53	4.58	7.89	2.12
	Method of Miao *et al.*,[14]	83.22	87.53	4.60	7.87	4.31
FERET databases	Proposed method	89.68	91.04	3.32	5.64	1.36
	Method of Miao *et al.*,[14]	87.15	91.02	3.31	5.67	3.87

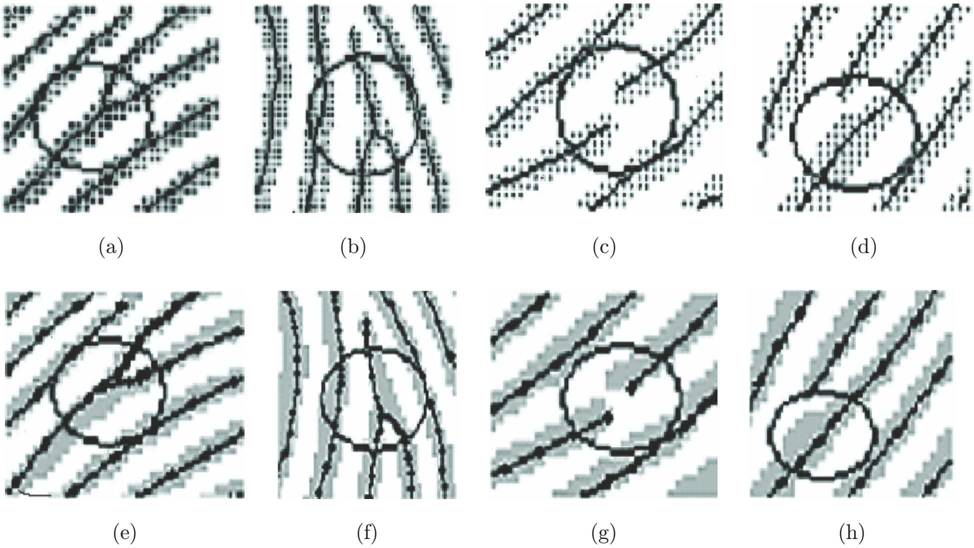


Fig. 6. (a)−(d) are the skeletons extracted by the modified EPL principal curves algorithm while (e)−(h) are the skeletons extracted by the original algorithm.

original algorithm. The principal curves in set (a)–(d) have smoother lines, higher accuracy, and are less error-prone than the ones in set (e)–(h).

- The step of the minutiae extraction and pseudo minutiae elimination

  In the experiment, we can obtain more accurate results of the minutiae extraction based on principal curves within a short running time. Three advantages over other methods based on original principal curves, gray-level image and thinning image are as follows:

— The average running time for our minutiae extraction algorithm is 41.76 ms, far better than 0.18 s,[23] 0.177 s[11] and 47.87 ms.[14] The reason is that the produced principal curves contain more information that can benefit the extracting minutiae, and from which pseudo minutiae is deleted. Additionally, searching for the original fingerprint image repeatedly which is time-consuming is avoided.

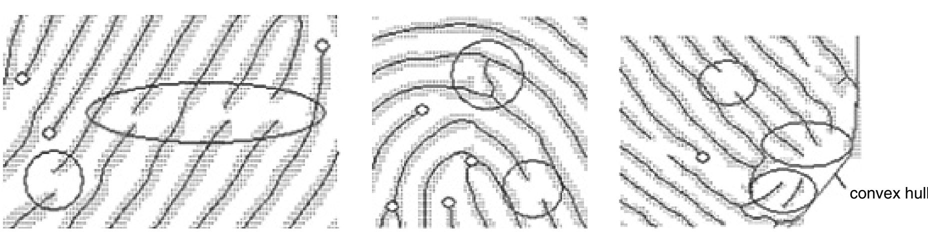
— As for the evaluation of minutiae extraction, we can use the accuracy rate (A) and (A') described below

$$A = 1 - \frac{d+f+x}{m+d-f} \tag{9}$$

$$A' = 1 - \frac{d+f}{m+d-f} \tag{10}$$

where A is the accuracy rate, d is the number of missing minutiae, f is the number of false minutiae, x is the number of minutiae owing to exchange of ridge ending and bifurcation minutiae (called error of type), and m is the number of minutiae obtained by our algorithm. $(m + d - f)$ is the number of minutiae that are contained in the fingerprint, $(d + f + x)$ is the number of all misidentified minutiae (including error of type), $d/(m + d - f)$ is the missing rate, $f/(m + d - f)$ is false acceptance rate, $x/(m + d - f)$ is exchanged rate. A' is the accuracy rate without x, $(d + f)$ is the number of all misidentified minutiae (not including error of type). The statistics in Table 1 shows that even for bad quality fingerprinting images, our total error is within acceptable limits. The reason is that the extracted skeletons are the result of fingerprint data optimization, and it can describe the minutiae more accurately than that extracted from thinning images and gray-level images. And based on the vertex out-degree direction and convex hull of the fingerprint, the broken and border pseudo minutiae can be respectively differentiated, as shown in Fig. 7. Our minutiae extraction is based on principal curves, and the minutiae information is a part of the skeleton topology information. Thus, it does not need to judge the type of the minutiae, which accelerates the speed of minutiae extraction. Furthermore, we can extract the minutiae efficiently unless the skeleton topology is severely sabotaged. Though the minutiae extraction[14] is also based on principal curves, the original principal curves algorithm results in the skeleton's deviation and being misconnected, which causes a lower accuracy rate of minutiae.

— Concerning the capability of anti-noise, the minutiae results of the proposed algorithm cannot be affected by the huge noisy fingerprint images which do not have large amount of connected components after filtering, while for those which have large amount of connected components after filtering, the number of extracted minutiae decreases relatively, and the number of lost minutiae increases. The reason is that when there is a large scale of connected components in the filtered image, the extracted skeletons have a corresponding net structure, so the similar minutiae in the net structure are deleted based on the judging condition of the connected pseudo minutiae. However, the connected components are centralized around the core area of fingerprint where a limited number of minutiae exists. As a result, the influence on authentic minutiae extraction is limited.



(a) (b) (c)

Fig. 7. The minutiae extraction based on principal curves. Small circles are minutiae.

5. Conclusions

The original EPL principal curve algorithm is modified based on the features of fingerprint data, and is applied for fingerprint minutiae extraction. Experimental results indicate that the improved algorithm can obtain smoother, higher accuracy and less error-prone fingerprint skeletons within a short running time, and the minutiae extraction method based on principal curves is obtained. The improved EPL principal curves is more accurate than the method.[14] So the proposed algorithm is more feasible in practical applications. Our future work will focus on the following three points: (1) The time complexity of algorithms needs to be further reduced. The proposed algorithm takes more computational cost in finding principal curves than in extracting minutiae. For instance, we will try to reduce the complexity by improving our algorithm to parallel algorithm. (2) The methods for differentiating pseudo minutiae need to be improved. The proposed method for the classification and the detection of pseudo minutiae is similar to traditional methods, and consequently the improvement of minutiae accuracy rate is small. If the information implied in skeletons, especially the skeleton topology and the relation between skeletons, can be fully used, then the improvement of accuracy would be greater. (3) Using the fingerprint skeletons as a feature of fingerprint for fingerprint matching will also be considered.

Acknowledgment

This work is supported by the National Natural Science Foundation of China (Grant No. 61075056, 60970061), and the Research Fund for the Doctoral Program of Higher Education (Grant No. 20060247039).

References

1. FERET web site:<http://hi.baidu.com/shjyy1983/blog/item /93a89513d04751045baf53c5.html>.
2. FVC2000 web site:<http://bias.csr.unibo.it/fvc2000/download.asp>.
3. FVC2002 web site:<http://bias.csr.unibo.it/fvc2002/download.asp>.
4. T. Hastie and W. Stuetzle, Principal curves, *J. Amer. Statist. Assoc.* **84** (1989) 502–516.
5. C. Y. Huang, L. M. Liu and D. C. D. Hung, Fingerprint analysis and singular point detection, *Patt. Recogn. Lett.* **28**(4) (2008) 1937–1945.
6. T. Y. Jea, Minutiae-based partical fingerprint recognition, Ph.D. Thesis, University at Buffalo, the State University of New York, 2005.
7. E. Jochen, T. Gerhard and E. Ludger, Data compression and regression based on local principal curves, *Advances in Data Analysis, Data Handling and Business Intelligence* (Springer Berlin Heidelberg, 2009).
8. B. Kegl, Principal curves: Learning, design, and applications, Ph.D. Thesis, Concordia University, Canada (1999).
9. B. Kegl and A. Krzyzak, Piecewise linear skeletonization using principal curves, *IEEE Trans. Patt. Anal. Mach. Intell.* **24** (2002) 59–74.
10. P. W. Kwan, J. B. Gao, Y. Guo and K. Kameyama, A learning framework for adaptive fingerprint identification using relevance feedback, *Int. J. Patt. Recogn. Artif. Intell.* **24**(1) (2010) 15–38.

11. X. Liang, A. Bishnu and T. Asano, A combinatorial approach to fingerprint binarization and minutiae extraction using Euclidean distance transform, *Int. J. Patt. Recogn. Artif. Intell.* **21**(7) (2007) 1141–1158.
12. H. Lin, Y. F. Wan and J. Anil, Fingerprint image enhancement: Algorithm and performance evaluation, *IEEE Trans. Patt. Anal. Mach. Intell.* **20**(8) (1998) 777–789.
13. D. Maio and D. Maltoni, Direct gray-scale minutiae diction in fingerprints, *IEEE Trans. Patt. Anal. Mach. Intell.* **19**(1) (1997) 27–40.
14. D. Q. Miao, Q. S. Tang and W. J. Fu, Fingerprint minutiae extraction based on principal curves, *Patt. Recogn. Lett.* **28**(6) (2007) 2184–2189.
15. W. Qing and Y. L. Yin, A novel method using videos for fingerprint verification, *Int. J. Patt. Recogn. Artif. Intell.* (2011), to appear.
16. G. Shlomo, A. Mayer and K. Daniel, Fingerprint image enhancement using filtering techniques, *Real-Time Imag.* **8** (2002) 227–236.
17. S. N. Srihari and H. Srinivasan, Comparison of ROC and likelihood decision methods in automatic fingerprint verification, *Int. J. Patt. Recogn. Artif. Intell.* **22**(3) (2008) 535–553.
18. R. Tibshirani, Principal curves revisited, *Stat. Comput.* **2** (1992) 183–190.
19. F. R. Wang, P. Cheng and X. Zou, A novel improved approach of direct minutia extraction from gray-level fingerprint image, *J. Imag. Graph.* **9**(5) (2004) 557–564.
20. H. N. Wang, Lee and C. M. Thomas, Extraction of curvilinear features from noisy point patterns using principal curves, *Patt. Recogn. Lett.* **29**(7) (2008) 2078–2084.
21. J. F. Wang, C. L. Lin, C. W. Yen, Y. H. Chang, T. Y. Chen, K. P. Su and M. L. Nagurka, Determining the association between dermatoglyphics and schizophrenia by using fingerprint asymmetry measures, *Int. J. Patt. Recogn. Artif. Intell.* **22**(3) (2008) 601–616.
22. J. H. Wegstein, An automated fingerprint identification system, US Government Publication. Washington, 1982.
23. Y. L. Yin, X. B. Ning and X. M. Zhang, An improved algorithm for minutiae extraction in fingerprint images, *J. Imag. Graph.* **12**(7) (2002) 1302–1306.
24. S. S. Yu and W. H. Tsai, A new thinning algorithm for gray-scale images by the relaxation technique, *Patt. Recogn.* **23**(10) (1990) 1067–1076.
25. J. P. Zhang and J. Wang, An overview of principal curves, *Chinese J. Comput.* **26**(2) (2003) 129–146.
26. H. Y. Zhang and D. Q. Miao, The research of off-line handwritten character recognition based on principal curves, Ph.D Thesis, Tongji University, 2005.
27. H. Y. Zhang, D. Q. Miao and D. X. Zhang, Analysis and extraction of structural features of off-line handwritten digits based on principal curves, *J. Comput. Res. Develop.* **42**(8) (2005) 1344–1349.
28. J. P. Zhang, D. W. Chen and U. Kruger, Adaptive constraint K-segment principal curves for intelligent transportation systems, *IEEE Trans. Intell. Transp. Syst.* **9**(4) (2008) 666–677.

Hongyun Zhang received the Ph.D. degree in pattern recognition and intelligence system from Tongji University, Shanghai, China, in 2005. She is currently a lecturer at Tongji University.

Her research interests include pattern recognition, data mining, rough set theory and granular computing.

Duoqian Miao received the Ph.D. degree in pattern recognition and intelligence system from Chinese Academy of Engineering, Beijing, China, in 1997. He is a professor at Tongji University.

His research interests include rough set theory, granular computing, pattern recognition and data mining.

Caiming Zhong is currently pursuing his Ph.D. in pattern recognition and intelligence system at Tongji University, Shanghai, China.

His research interests include cluster analysis, manifold learning and image segmentation.