Fingerprint Classification Method based on J-divergence Entropy and SVM

Gongping Yang, Yilong Yin* and Xiuyan Qi

School of Computer Science and Technology, Shandong University, Jinan 250101, China

Received: 7 May. 2013, Revised: 1 Sep. 2013, Accepted: 2 Sep. 2013
Published online: 1 Apr. 2014

Abstract: Fingerprint classification is one of the key technologies in Automatic Fingerprint Identification System (AFIS). However, the performance of most recent fingerprint classification methods is low when the quality of the fingerprint image was low. To overcome this problem, a novel method based on j-divergence entropy and SVM (Support Vector Machine) classifier is proposed in this paper. Firstly, our method transforms the fingerprint images from spatial domain to frequency domain and constructs the directional images according to frequency spectrum energy. Secondly, eigenvector around the core point is extracted. Thirdly, the dimension of eigenvector is reduced by j-divergence entropy. At last, the input image is classified by SVM classifier. Experimental results on NIST-4 database show the validity of our method, and the classification accuracy reaches 94.7% for four-class classification and 91.5% for five-class classification with zero rejection rate.

Keywords: Fingerprint identification, Fingerprint classification, Frequency-spectrum energy, J-Divergence Entropy, SVM classifier

1 Introduction

Automatic fingerprint recognition is one of the most widely used biometric technologies, and fingerprint classification is a very important step of the automatic fingerprint recognition. In large-scale fingerprint database, it is quite time consuming to match the input fingerprint against each template stored in the database. To reduce the number of matching, fingerprint classification is adopted to assign a fingerprint to one of the several pre-specified types and fingerprints belonging to different types are stored in different sub-databases. Then the input fingerprint only needs to match against templates in a certain sub-database. Therefore, fingerprint classification can be viewed as a coarse-level matching of the fingerprints and it provides an indexing mechanism for the automatic fingerprint recognition. Nowadays, most of the fingerprint classification algorithms classify fingerprints to five or four types which are respectively whorl, left loop, right loop, arch and tented arch, or whorl, left loop, right loop and arch.

Researchers have done a great number of studies about fingerprint classification algorithms. Singular points based fingerprint classification method [1,2,3,4] indentifies fingerprint type mainly by the number of core and delta point, and their topological relation. This method relies on the extracting accuracy of singular points. However, the loss detection rate and false detection rate is relatively high because of the effect of fingerprint quality [5] and area of fingerprint foreground. Structure based fingerprint classification [6] takes use of the difference of geometrical structure. This method uses the global structure of fingerprint image so it is more effective than the other methods which use only local information for fingerprint classification. Nevertheless, the computational complexity of this method is quite high and the geometrical structure of fingerprint is difficult to obtain for low quality fingerprint images. FingerCode based fingerprint classification [7,8,9] uses the texture feature around core point. This method is effective for low quality fingerprint but the classification accuracy still needs improving. Orientation field based fingerprint classification mainly uses similarity of orientation image. This method extracts fingerprint orientation image as feature vector. It is still a challenging problem to extract orientation field accurately for low quality fingerprint image.

Based on the analysis above, a fingerprint classification method based on j-divergence entropy and SVM is proposed in this paper. For the low quality
fingerprint images, this paper improves the extraction method of orientation field and the processing of feature vector in the orientation based fingerprint classification method. Some works \cite{10,11,12} use spatial field filtering or frequency field filtering, while our method uses frequency spectrum energy to acquire fingerprint orientation image. Firstly, fingerprint image is transformed to frequency domain by Fourier Transform. Then fingerprint orientation is acquired according to the energy of frequency band of fingerprint in the frequency image. Fourier Transform is an effective tool to process images that are periodic and directional. Noises can be avoided by detecting the high frequency coefficients \cite{13}. As most of the noises and fingerprint image are distributed in different frequency area of the energy image, this method reduces the effect of the noises and the orientation image is obtained accurately. At the same time, this method is relatively robust to low quality images. Because the orientation feature vectors are high dimensional, the computational complexity is high and the classification performance is correspondingly low \cite{14}. Therefore, this paper uses j-divergence entropy to reduce the high dimensional feature vector. Compared with other criteria, j-divergence entropy takes into account the posterior probability and the reduced feature vectors are more suitable for fingerprint classification. At last, features after dimensional reduction are used as input of SVM classifier. SVM can classify the fingerprints effectively as it has advantages in non-linear classification problem.

The remainder of the paper is organized as follows. In Section 2, we introduce the algorithm of computing feature vector of fingerprint image. Section 3 introduces SVM classifier precisely. Section 4 gives experimental results. Finally, conclusions of this paper are given in Section 5.

2 Feature extraction for fingerprint classification

In this paper, fingerprint classification based on fingerprint orientation image is used. Fingerprint orientation image is viewed as feature vector, and the extraction process is listed as follows:

1) Conduct Fourier Transform to fingerprint image.
2) Obtain the orientation image of fingerprint image.
3) Extract the feature vector of fingerprint image.
4) Conduct dimensional reduction using j-divergence entropy.

2.1 Fourier Transform

The discrete Fourier Transform of an image with the size of $M \times N$ is given as follows:

$$F(u,v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) e^{-j2\pi(ux/M + vy/N)}$$

In Fig. 1, (a) and (c) are Fourier transformed to frequency images (b) and (d), respectively. Fast Fourier Transform is used to accelerate computational speed. Fingerprint images in NIST-4 database are used in this paper, and the image size is 512×480. As there are noises (such as some handwriting and lines) at the edge of fingerprint images, we take segmentation to the images and the image size becomes 480×480 after segmentation.

Fig. 1: Fingerprint Block and frequency images,(a) (c) are Block images,(b) (d) are frequency images of (a) (c), respectively.

Fingerprint orientation image is constructed based on block image. The method to get block image is: the eight pixels at the edge of each fingerprint image will not be partitioned, the other area is partitioned to 16×16 pixel blocks without overlap. Because the size of each block is too small, our method may be not robust to noises and block size are enlarged to 32×32. So there will be 2929 blocks after partitioning. Finally, we will get 29×29 frequency images after conduct Fast Fourier Transform to each block.

2.2 Construct Fingerprint Orientation Image

Two features of frequency image of fingerprint image are introduced:

Firstly, if the ridge orientations in an image are consistent, ridge orientation are vertical to the direction of energy distribution in frequency image. In Fig. 1, (b) and (d) are the frequency image of fingerprint block (a) and (c). It can be concluded from Fig. 1 that when the ridge
orientation in the image are consistent, the energy distribution in frequency image are following certain rules: in spatial image (a) the ridge orientation is about 45 degree and the energy distributes along the direction of 135 degree in frequency image (b). Similarly, the energy in frequency image (d) of spatial image (c) is also distributed along the direction that is vertical to the ridge orientation in image (c). The two figures in Fig. 2 are the energy image of the two fingerprint blocks in Fig. 1. It can be concluded From Fig. 2 that there are obvious peak near 135 degree and 45 degree.

![Image](43x487 to 267x574)

**Fig. 2:** (a) and (b) are the energy image of (a) and (c) in Fig. 1, respectively. The x-axis is the angle and y-axis is energy.

Secondly, the energy of periodical image is concentrated at a circle in the frequency image. According to the feature of discrete Fourier [15], the relationship between the radius of the circle and the periodical signal in the digital image is as follows:

\[ R = \frac{N}{T} \]

where T is the cycle of digital image, N is the width of image and R is radius of frequency.

Fingerprint image is periodical image and ridge orientation in fingerprint block is consistent. So the energy is certainly concentrated at the circle of with the radius R on the corresponding direction in frequency image. The size of fingerprint block in this paper is 32×32. The value of T is belong to [3,18] and so the range of R is 1 to 10. Fig. 3 shows that information of fingerprint is mainly distributed in the annular region between two circles. Noises are mainly at high-frequency and low-frequency area. Only the energy in the annular area will be considered and so this method is tolerate to low quality fingerprint in a certain degree.

Steps of computing the orientation of fingerprint block \((m, n)\) are as follows:

1) Calculating the sum energy \(L(\theta)\) at each direction in frequency image:

\[ L(\theta) = \frac{1}{N_\theta} \sum_{R(u,v) \in [R_1,R_2], \theta} |F(u,v)| \]

where \(R(u,v)\) is the distance between pixel \((u,v)\) to the center of frequency image. \(R_1\) and \(R_2\) are 1 and 10, respectively, \(r(u,v)\) is the direction of pixel \((u,v)\), where \(\theta = 1, 2, \ldots, 180\). \(N_\theta\) is the number of pixels between \(R_1\) and \(R_2\) at the direction of \(\theta\).

2) Smoothing to \(\tilde{L}(\theta)\):

\[ \tilde{L}(\theta) = \frac{1}{9} \sum_{k=-4}^{4} L(\theta + K) \]

3) Calculating the orientation \(\theta\) with largest energy value:

\[ \theta = \arg\max(\tilde{L}(\theta)) \]

4) Calculating orientation of fingerprint block \(\gamma_{m,n}\) which is vertical to \(\theta\):

\[ \gamma_{m,n} = \left\{ \begin{array}{ll} \theta + \pi/2, & \theta \in [0, \pi/2], \\ \theta - \pi/2, & \theta \in (\pi/2, \pi] \end{array} \right. \]

Finally, orientation vector \(\gamma_{m,n}, (1 \leq m,n \leq 29)\) is acquired.

### 2.3 Computing Feature Vector of Fingerprint Image

Actually, fingerprint classification is coarse-level matching of the fingerprints. A fingerprint belongs to the class which the fingerprint matched. So in classification process, a reference point should be found to align fingerprint images. Because information of fingerprint types are distributed around core point, core point will be chosen as reference point and orientation information around core point is chosen as feature vector of fingerprint images. Firstly, the location of core point will be found and then feature vector of fingerprint image will be computed.

In this paper, methods in [16,17,18] are used to acquire the block at which core point is located. The steps are as follows:

1) Define the orientation vector \(d_{m,n}, (1 \leq m,n \leq 29)\) of fingerprint image:

\[ d_{m,n} = [\cos(2\gamma_{m,n}), \sin(2\gamma_{m,n})] \]
2) Calculating the sum of vector \( u_{m,n} \) in \( 3 \times 3 \) neighborhood of image block \((m,n)\):

\[
u_{m,n} = \sum_{k=-1}^{1} \sum_{l=-1}^{1} d_{m+k,n+l}
\]

(8)

the smaller the \( u_{m,n} \) is, the greater the changing of orientation around this block. And the changing of orientation in the blocks around core point is greatest.

3) Smoothing to \( u_{m,n} \) in the \( 2 \times 2 \) neighborhood:

\[
\tilde{u}_{m,n} = \sum_{k=0}^{0} \sum_{l=0}^{0} u_{m+k,n+l}
\]

(9)

4) Obtaining the block corresponding to core point:

\[
u_{i,j} = \min \tilde{u}_{p,q}, (5 \leq p \leq 17, 9 \leq q \leq 21)
\]

(10)

Core point is in the block \((i,j)\).

Information of fingerprint types distributes around core points and mainly the low-half area. So, we select an Area A which takes the block having core point as center, up to four blocks, down to eight blocks, left to eight blocks, and right to eight blocks. Fig. 5 shows that the area surrounded by the yellow dotted line is Area A.

Fig. 4: Core point is in the red rectangle and Area A is in the rectangle surrounded by yellow dotted line.

To maintain the continuity of the features, \( d_{i,j} \) (shown in Formula 7) is used as the feature of fingerprint block \((i,j)\). Therefore, the feature vector of Area A is \( D = [d_{1,1}, \ldots, d_{i,j}, \ldots, d_{13,17}] \).

2.4 Feature Extraction Based on Minimized J-Divergence Entropy

Feature extraction means that samples are represented in lower dimension by mapping (or transforming) from original high-dimension feature space. It is not suitable to use high dimensional feature to design classification considering computational complexity or classification performance. So feature extraction is needed and the dimension of samples will be reduced to design more effective classification. Feature vector D calculated in this paper is 442 dimensional which is quite high. Feature extraction is performed to classify fingerprints more effectively.

In this paper, j-divergence entropy is used to do feature extraction [14]. Entropy is a measure of uncertainty in information theory. Features with minimum uncertainty will be more useful for classification. Entropy is also used to measure the intensity of posterior probability distribution. And the best classifier is determined by posterior probability. The more intense of the posterior probability, the lower of the error probability, and the error rate of classifier will also lower. Therefore, using entropy to reduce the dimension is beneficial for fingerprint classification.

Suppose a probability density function \( p(x_i) \), and the measure of the deviation of \( p(x_i) \) compared with the standard distribution \( w(x_i) \) is referred to as its relative entropy.

\[
V_{p,w} = -\sum p(x_i) \log\left[ p(x_i)/w(x_i) \right]
\]

(11)

The smaller of the relative entropy, the larger of the difference between the two probabilities distributed. When the probability distributions of the two classes are exactly the same, the relative entropy will reach the maximum (equal 0). So, we can define discrimination entropy \( W(p,q) \) to represent the difference of the two distributions \( p(x_i) \) and \( q(x_i) \):

\[
W(p,q) = V(p,q) + V(q,p)
\]

(12)

In multi-class condition, we can use \( \sum \sum W(p^{(i)}, q^{(j)}) \) to represent the deviation of different distributions, where \( i, j \) are class labels. For easier computation, we use function

\[
U(p,q) = -\sum (p_i - q_i)^2
\]

(13)

to replace \( W(p,q) \). It is to be noted that this replacement will not affect the result of choosing the best feature. When \( U \) reaches the minimum, the coordinate system is composed by the eigenvectors corresponding to the \( d \) largest eigenvalues of the matrix \( A \):

\[
A = \sum_{i,j=1}^{c} (G^{(i)} - G^{(j)})
\]

(14)

Where, \( G^{(i)} \) is the covariance matrix of the \( i \)th class, \( c \) is the classes of fingerprint. Suppose the dimension of the original feature is \( D \), and the dimension of the new feature extracted from the original feature is \( d \). Rank the eigenvalues of matrix \( A \): \( \lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \ldots \geq \lambda_d \geq 0 \), then choose the eigenvectors corresponding to the former \( d \) large eigenvalues to compose the coordinate system. In this coordinate system, the discrimination entropy is smallest, and the difference of classes probability distributions is largest. Therefore, we can use this coordinate system to classify fingerprints effectively.
3 Support Vector Machine Classifier

Support Vector Machine (SVM) researches the machine learning problem with limited samples, it has advantage to solve nonlinear pattern recognition problem and get the global optimal solution. According to the characteristic of SVM and the characteristic of eigenvector used in this paper, we use SVM as fingerprint classifier.

SVM proposed between 1992 and 1995 is still in the development stage at present. The practical risk of machine learning is composed of two parts: empirical risk and confidence interval. The traditional training methods for classification use ERM (Empirical Risk Minimization), these methods have low empirical risk but wide confidence interval, which results in over learning. SVM organizes functions to form a sequence of subsets, which makes a balance between empirical risk and confidence interval to minimize the practical risk.

The main steps of SVM are: The input space is first transformed to higher dimensional space by nonlinear transformation, and then the optimal linear classifier is found in the new space. The nonlinear transformation is realized by defining proper inner product function. There are three kinds of inner product functions at present: the polynomial inner product function, the radial-basis inner product function and the S inner product function. The aim of SVM is to solve the two-class problems, for multi-class problems, the methods are discussed in the following:

1) One-vs-all. This method trains k classifiers to solve k two-class problems, the i-th SVM uses the training samples of the i-th class as positive training samples and the training samples in other classes as negative training samples.

2) Pair wise. This method trains all the possible classifiers using samples of k classes and the time of training is $K(K - 1)/2$ totally.

3) Directed acyclic graph. To k-class problems, $K(K - 1)/2$ pair wise two-class classifiers are trained. This method is called DAGSVM according to the injection of the directed acyclic graph theory.

In this paper, we use 2,000 fingerprint images choosing from the NIST-4 as training set for obtaining the matrix. Then, we compute the eigenvalues of matrix. Rank the 442 eigenvalues in descending order, we can find that the first 20 eigenvalues account for 94.7% of the sum of all the 442 eigenvalues. Hence, we choose the 20 eigenvectors corresponding to the first 20 eigenvalues to form the projection function $U = (u_1, u_2, \ldots , u_{20})$. Then, we can reduce the original feature dimension 442 to a new space with 20-dimension by the transformation:

$$Y = XU \quad (15)$$

Where, $X$ is the original feature with 442-dimension, and $Y$ is the new feature after dimension reduction with 20-dimension.

4 Experimental Results

In this paper, we use the NIST-4 fingerprint database to do our experiments. NIST-4 fingerprint database includes 4000 fingerprint images, which captures from 2000 persons, i.e., each person provides 2 fingerprint images. And the image size is 520480. The fingerprint labels range from f0001 to f2000 and from s0001 to s2000. Images with the same number of f and s belong to the same fingerprint, e.g., the images labeled f0001 and s0001 are impressions of the same fingerprint. In our experiment, we use 2000 fingerprint images (ranging from f0001 to f1000 and from s0001 to s1000) as training set and use the remaining 2000 images (ranging from f1001 to f2000 and from s1001 to s2000) as testing set. In training phase, train the fingerprints according to the first class they belong to; in testing phase, consider every classification result is correct, no matter whatever class label is determined by the classifier. Table 1 is our classification result with no false rejection.

<table>
<thead>
<tr>
<th></th>
<th>left loop</th>
<th>right loop</th>
<th>whorl</th>
<th>arch</th>
<th>tented arch</th>
</tr>
</thead>
<tbody>
<tr>
<td>left loop</td>
<td>377</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>right loop</td>
<td>6</td>
<td>359</td>
<td>1</td>
<td>3</td>
<td>30</td>
</tr>
<tr>
<td>whorl</td>
<td>8</td>
<td>8</td>
<td>378</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>arch</td>
<td>8</td>
<td>6</td>
<td>0</td>
<td>399</td>
<td>22</td>
</tr>
<tr>
<td>tented arch</td>
<td>10</td>
<td>10</td>
<td>1</td>
<td>35</td>
<td>317</td>
</tr>
</tbody>
</table>

In Fig. 5, we compare our experimental result with the experimental results of Qinzhi Zhang et al. [25], Cheong Hee Park et al. [16] and Jin-Hyuk Hong et al. [26], on the condition that all these experiment use the same training set and test set. And, on the condition that the false rejection rate is 0%, 6%, 9% and 20%, our method can achieve 91.5%, 93%, 93.4%, 95% and 94.7%, 95.8%, 96.2%, 96.8%, respectively, corresponding to the two kinds of fingerprint classes - 5 classes or 4 classes.
5 Conclusions

Currently, the performance of fingerprint classification algorithms is still unsatisfactory in the case of low quality fingerprint images. Therefore, we propose a fingerprint classification algorithm based on discrimination entropy and SVM. Firstly, we compute the energy spectrum of fingerprint image to obtain fingerprint directional graph. Secondly, we extract the orientation field around the core point. Then, we use the discrimination entropy to reduce the dimension of the feature vector. Finally, SVM classifier is applied for classification. Experimental results show that our algorithm is very effective.

From the distribution character of energy in fingerprint spectrum graph, we can see that our algorithm has certain robustness. The feature extraction method based on discrimination entropy reduces the dimension of feature vector. It does not only improve the classification performance, but also improve the classification speed. SVM has many advantages in dealing with no-linear problems, and it can obtain the globally optimal solution. Therefore, it is beneficial for fingerprint classification.

However, our method uses directional graph as feature vector, the direction tendencies of some fingerprint images are not obvious. So, our method has certain limitation for this type of fingerprints. In addition, our method uses core point as reference point, but it is a difficult work to extract the core point exactly. All these problems affect the performance of our algorithm to some extent.

Acknowledgement

This work is supported by National Natural Science Foundation of China under Grant No. 61173069, 61070097, Program for New Century Excellent Talents in University of Ministry of Education of China under Grant No. NCET-11-0315 and Shandong Natural Science Funds for Distinguished Young Scholar. The authors would like to thank Yanbin Ning and Chunxiao Ren for their helpful comments and constructive advice on structuring the paper.

References


[25] Qinzheng Zhang, Hong Yan; Fingerprint classification based on extraction and analysis of singularities and pseudo ridges [J]. Pattern Recognition, 11, 2233-2243 (2004).


**Gongping Yang** received his Ph.D. degree in computer software and theory from Shandong University, China, in 2007. Now he is an associate professor in School of Computer Science and Technology, Shandong University. His research interests are biometrics, medical image processing, and so forth.

**Yilong Yin** is the Director of MLA Group and a Professor of Shandong University. He received his Ph.D. degree in 2000 from Jilin University. From 2000 to 2002, he worked as a post-doctoral fellow in the Department of Electronic Science and Engineering, Nanjing University. His research interests include machine learning, data mining, and biometrics.

**Xiuyan Qi** received her B.S. degree from Shandong University in 2005. Now she is a master student at School of Computer Science and Technology, Shandong University. Her research interests include biometrics, image processing and pattern recognition.