Combining Statistical and Structural Information for Fingerprint Identification

Qinghan Xiao and Hazem Raafat

Department of Computer Science
University of Regina
Regina, Saskatchewan
Canada S4S 0A2

Abstract

Most of the proposed fingerprint pattern recognition methods are either based on a statistical or on a structural approach. Generally, the former is used to match the fingerprint minutiae and the latter is adopted to classify the fingerprint subpatterns. In the statistical approach, the best match between two patterns can be searched, but the searching time for a given fingerprint will increase as the number of file fingerprints increases. By using the structural way, although the number of fingerprint subpatterns is small, it is still possible to reduce the amount of searching time. In order to make use of the advantageous of both approaches, they might be combined to recognize the fingerprint patterns. In this paper, the statistics of fingerprint ridges are calculated as attributes. Based on these attributes, the structural features or fingerprint subpatterns are represented. By combining the statistical and structural information extracted from fingerprint images, a multi-level algorithm is introduced to recognize the fingerprint patterns. In this way, the searching time and the computational complexity can be reduced. The proposed approach has been tested with real fingerprint data to demonstrate its effectiveness.

Keywords: Statistical and structural pattern recognition, Fingerprint classification and identification, Minutiae, Ridge attributes.

1. Introduction

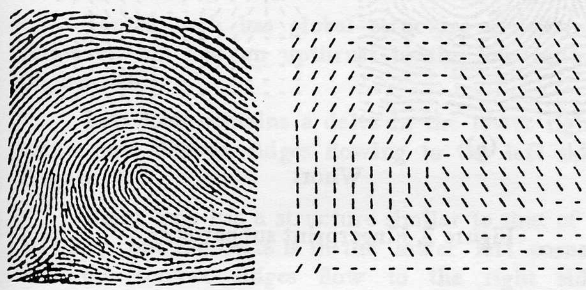
Identification of fingerprint patterns has been achieved with various approaches [1-10]. Some approaches use the statistical characteristics to match an individual print [1-5], while others use the structural information to classify the fingerprint into subpatterns [6-10]. In order to solve various application problems, different approaches for pattern recognition could be combined. For those patterns whose structural information is predominant in their description, it might be advantageous to use the structural approach. If the data involved in the pattern are easily expressed by a feature vector, the statistical approach would be used for recognition [11]. Because the fingerprint pattern falls half way between these two extreme cases, it will be efficient to make use of the statistical and structural information to solve the identification problems.

In order to reduce the searching time and the computational complexity, we develop a new approach which uses statistical as well as structural information to recognize fingerprint patterns. First, the fingerprint image is subdivided into sampling squares. Second, some statistics of the fingerprint ridges are calculated as attributes within each sampling square. Based on these statistical attributes, the global and local configurations of the fingerprint are described. Finally, the identification algorithm and experimental results are presented in terms of real fingerprint data to demonstrate the effectiveness of this approach.

2. Pattern Decomposition

According to the parallelism and continuity of the fingerprint ridges, the distribution of ridge directions can be used to describe the fingerprint subpatterns. Some articles adopt a method of a sampling matrix to classify fingerprint subpatterns [6-9]. Because of the difficulty in constructing and parsing such complex patterns as fingerprints, a directional code (4-directional codes or 8-directional codes) is used to represent the ridge directions in each matrix. Although the code makes the parsing easier, it prohibits further subdivision which has to be based on small differences in curvature of the ridges.

In this paper, an original fingerprint image consisting of 256x256 binary data points is divided into 16x16 sampling squares. Within each sampling square, there are 16x16 pixels. Unlike other methods, attributes will be used to classify and identify the fingerprint patterns. In each sampling square, the ridge count, the average direction and the direction variance of ridges are calculated to represent the fingerprint local configurations. Figure 1 shows a fingerprint image and its sampling square representation. For the reason of output on the computer, four-directional codes are selected to display and print out the sampling square schemes.



(a) (b)

Figure 1. Fingerprint image and sampling square scheme

3. Attributes of Fingerprint Ridges

In order to describe the fingerprint configurations, some statistical attributes of fingerprint ridges are introduced in this section.

The coordinates (x,y) of the fingerprint image are calculated from top left (0,0) to bottom right (255,255) with the first horizontal line of scanning as the X-axis direction.

The ridge direction (D) is calculated within each sampling square. By tracing along a thinned ridge line from point p_1 to point p_2, the ridge direction is defined as follows:

$$D = arctg\ \frac{Y_{p_2} - Y_{p_1}}{X_{p_2} - X_{p_1}} \qquad (1)$$

The minutia direction (θ) is calculated in two ways. For ridge endings, the direction of the ridge flowing from the endpoint is defined as the minutia direction. For bifurcations, the minutia direction is defined as the average value of the direction of the two ridges that flow from this bifurcation towards the same side (Figure 2).

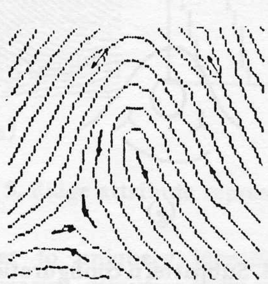


Figure 2. Minutia direction

The ridge count (C) is the number of ridges within a sampling square. It can be calculated by tracing along every fingerprint ridge in each sampling square.

The average direction and the direction variance of ridges are defined as follows:

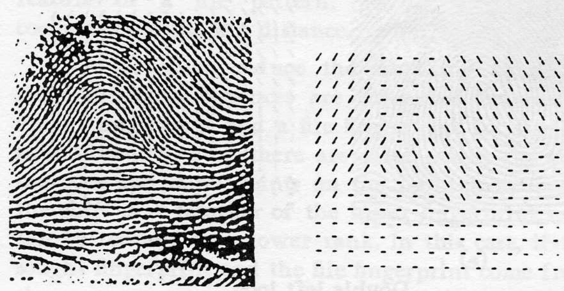
$$\bar{D} = \frac{1}{N}\sum_{i=1}^{N} D_i \qquad (2)$$

$$Var(D) = \frac{1}{N}\sum_{i=1}^{N} (D_i - \bar{D})^2 \qquad (3)$$

where N equals the number of ridges within a sampling square.

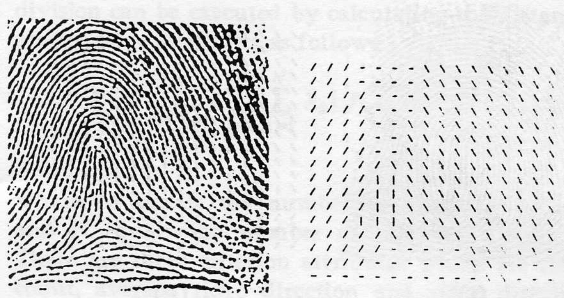
4. Pattern Representations

Based on the attributes mentioned above, the fingerprint global configuration and local features can be represented effectively. Most of the classification algorithms divide the fingerprint patterns into seven subpatterns [7-10]. They are arch, tented arch, left loop, right loop, double left loop, double right loop and whorl. Figure 3 shows the original subpattern images and their sampling square schemes. It is obvious that the sampling square schemes describe a general ridge flowing trend and give the rough positions of the core and delta points.



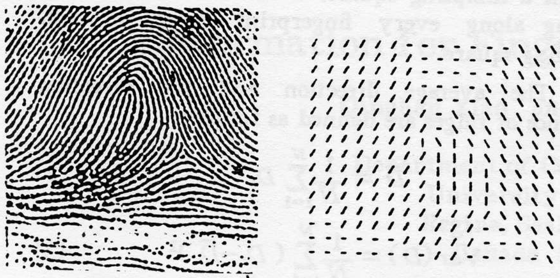
(a) (a')

Arch



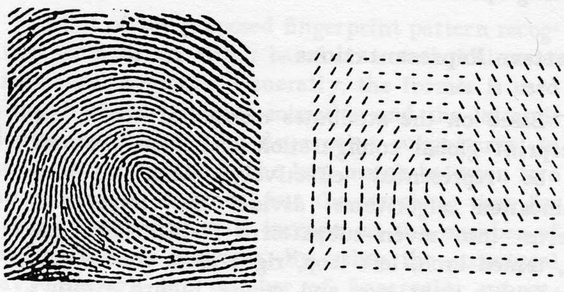
(b) (b')

Tented arch



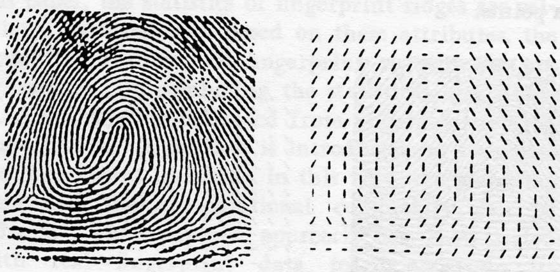
(c) (c')

Left loop



(d) (d')

Right loop



(e) (e')

Double left loop

(f) (f')

Double right loop

(g) (g')

Whorl

Figure 3. Fingerprint subpatterns

Because of their very different structures, these seven subpatterns can be classified by their outlines shown in Figure 4.

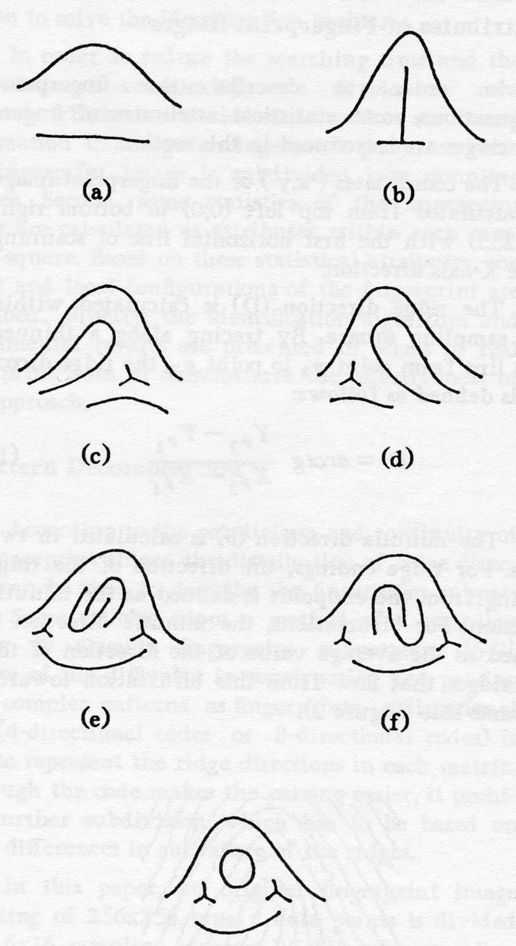


Figure 4. Subpattern configurations

By analyzing the directional relations among the sampling squares, we can describe the subpatterns as follows.

Arch is a pattern of convex ridges with a peak in the middle (Figure 4(a)).

Tented arch has global structure of convex ridges on top and an upthrust in the low center (Figure 4(b)).

Left loop contains a delta in the lower right corner and convex ridges flowing to the left side (Figure 4(c)).

Right loop has a structure similar to that of a left loop, but the delta is in the lower left corner and the convex ridges flow to the right side (Figure 4(d)).

Double left loop has two opposite loops and two deltas. If the upside down loop is on the left side, it is called a double left loop (Figure 4(e)).

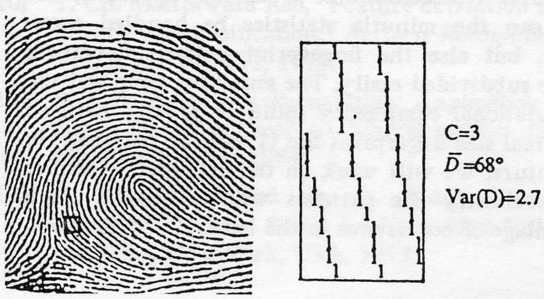
Double right loop has a similar structure as double left loop, but the location of upside down loop is on the right side (Figure 4(f)).

Whorl consists of two deltas and global convex ridges, which is shown as Figure 4(g).

In this paper, the attributes of ridges in the sampling square are considered for further subdivision. The set of attributes is denoted by A_i.

$$A_i = [a_1(i), a_2(i), ..., a_n(i)] \qquad (4)$$

where $a_1(i)$, $a_2(i)$, ..., $a_n(i)$ are the attributes of the ith sampling square. They are ridge count, average direction and direction variance of ridges within the sampling square. Figure 5(a) shows a thinned fingerprint image. Figure 5(b) shows the ridge attributes within the lower left hand sampling square marked in Figure 5(a).



(a) (b)

Figure 5. Ridge attribute representations

In order to uniquely identify or verify a suspect's fingerprint, minutiae have to be used as the identification features [3,5]. In our algorithm, minutia location and minutia direction are selected as the features to match the fingerprint patterns.

5. Recognition Algorithm

By combining both statistical and structural information, a multi-level algorithm has been developed to recognize the fingerprint patterns. The algorithm is explained in the following steps:

Step 1: This step is a classifier which analyzes a global configuration of fingerprint on the basis of sampling square scheme. A 3x3 window is used to move along the rows and the columns of the image to recognize the corner (/‾ , ‾\ , /\ , /| , |\ ), the delta (—< , /_\ , >—) and the ridge flowing tendency on the scheme. According to the subpattern characteristics mentioned above, a fingerprint image can be classified into one of the seven subpatterns shown in Figure 4. If some fingerprint impressions could not be classified into any of those subpatterns in this step, they will be labelled as accidental then turned to the next level for further subdivision.

Step 2: According to the distance criterion [12], the distance of features can be calculated as follows:

$$R_k(i,j) = \lambda_k | f_k(i) - f'_k(j) | \qquad (5)$$

where R_k (i,j) is the kth distance between the ith feature of an under searching pattern and the jth feature of a file pattern. λ_k is a weighting coefficient for the kth distance.

In order to reduce the searching time, the ridge attribute distances are computed between a sought fingerprint and a file fingerprint for further subdivision. Because there are some marks for aiming core and delta points on the input instrument, the shift and rotation of the input fingerprint card can be reduced to a lower rank. In this case, if the sought fingerprint and the file fingerprint come from the same person, there should be no significant differences among their ridge attributes within each sampling square. Based on these properties, the subdivision can be executed by calculating the distances among ridge attributes as follows:

$$R = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{k=1}^{N} \lambda_k | f_k(i) - f'_k(i) | \qquad (6)$$

where M equals the number of sampling squares and N equals the number of criterion attributes. There are three criterion attributes which are ridge count, average ridge direction and ridge direction variance. If R is less than a threshold T (selected by experiments), the matching score will be calculated between sought and file fingerprints in the next step.

Step 3: Because there are many fingerprints which have the similar ridge flowing tendency but dissimilar minutiae, the minutia similarity has to be calculated for uniquely identifying a fingerprint. The dissimilarity between two patterns can be measured by distance criterion and can be calculated as follows:

$$J(F_i) = \underset{j}{Min}\left\{ \sum_{k=1}^{N} R_k(i,j) \right\} \quad (7)$$

where the minimum value is selected as the real dissimilarity measure of the feature F to correspond with the best registration. If the sought fingerprint has M_1 minutiae and the file fingerprint has M_2 minutiae, the matching score between the sought fingerprint and file fingerprint can be calculated by using the following expression:

$$S = \frac{1}{M_1} \sum_{i=1}^{M_1} \frac{B}{1 + \underset{j}{Min}\left\{ \sum_{k=1}^{N} R_k(i,j) \right\}} \quad (8)$$

$$j=1, ..., M_2$$

where B is a constant which is selected to normalize S from 0 to 100. The criteria are the minutia location and the minutia direction. If a pair of fingerprints have a low similarity in step 2, it cannot get a high score in step 3 since step 2 represents the dissimilarity of the ridge attributes and step 3 is used to measure the similarity of the minutiae.

Step 4: After comparing the matching scores of all selected file fingerprints, a list of five candidates is printed out in the order of matching scores.

6. Experimental Results and Discussions

The presented recognition algorithm has been programmed in FORTRAN 77 running on a VAX-11/780 computer. Each rolled fingerprint is scanned by a TV camera that sweeps horizontally across the fingerprint from top left to bottom right to form a representative 256-level digitized output matrix of 256x256 pixels over a 2.0 cm square format. The database consists of 500 fingerprints — 15 arches, 10 tented arches, 125 left loops, 100 right loops, 40 double left loops, 40 double right loops and 170 whorls. They are selected on the basis of the natural distribution.

Because there are no two fingerprints with the same details, a concordance method, that is designing and checking the classifier by using the same set of data, is used to evaluate the algorithm. In the experiments, all the fingerprints are scanned two times. In the first time they are stored to establish a fingerprint database. Then they are scanned under various conditions to examine the recognition algorithm. From the first level classifier, we found out that there are eight samples which could not be classified into any one of the seven subpatterns. Unlike other classification algorithms, there are no rejection subpatterns in our classifier. Therefore, those eight samples are sent to the second level classifier to compare them with all other fingerprints in the database. In order to reduce the matching time for next step, the threshold T is experimentally selected to get rid of 90% of the file fingerprints in the second level classifier. In the matching step, the highest matching scores varied from 32 to 91 with different candidates. To meet the demands of law enforcement agencies, an automatic fingerprint identification method has to narrow the range of manual verification and reduce the false dismissal rate. In our experiments, a list of five candidates is always submitted to the law enforcement agencies. There were no false dismissals as confirmed by the experts. Although we cannot conclude from the results that two fingerprints are exactly the same, the proposed method provides a technique which can significantly improve the present cumbersome manual searching process.

Since a multi-level identification algorithm has been used by combining statistical and structural information, the searching time is reduced for the subpatterns which have a low percentage in the natural distribution. It is necessary to point out that the computational complexity for the recognition algorithm is less than $O[M_1 \times M_2]$ which is the computational complexity of the matching formula (8).

From the experimental results we can see the effectiveness of combining the statistical and structural representations. With this approach, not only can the minutia statistics be handled accurately, but also the fingerprint pattern structures can be subdivided easily. The searching time and the computational complexity could be greatly reduced for a real size fingerprint file (1 million persons). In the future, we will work on the first-level classifier to subclassify the patterns which have a high percentage of occurrence in the natural distribution.

7. Acknowledgements

This research is supported in part by the K. C. Wong Education Foundation of Hong Kong and the Natural Sciences and Engineering Research Council of Canada operating grant A8154.

8. References

[1] Q. Xiao and Z. Bian, "An approach to fingerprint identification by using the attributes of feature lines of fingerprint," *Proc. 8th Int. Conf. on Pattern Recognition*, Paris, France, pp. 663-665, 1986.

[2] Q. Xiao, X. Dong and Z. Li, "A recognition method for fingerprint," *Acta Automatica Sinica*, Vol. 10, No 1, pp. 69-73, 1984 (in Chinese).

[3] M. Kawashima and K. Kiji, "Personal identification by fingerprint or palm-print," *Information Processing*, Vol. 25, No. 6, pp. 599-605, 1984.

[4] F. Pernus, S. Kovacic and L. Gyergyek, "Minutiae based fingerprint recognition," *Proc. 5th Int. Conf. on Pattern Recognition*, Florida, USA, pp. 1380-1382, 1980.

[5] M. Eleccion, "Automatic fingerprint identification," *IEEE Spectrum*, Vol. 10, No.9, pp. 36-45, 1973.

[6] B. Moayer and K. S. Fu, "A tree system approach for fingerprint pattern recognition," *IEEE Trans. Comput.*, Vol. C-25, pp. 262-274, 1976.

[7] B. Moayer and K. S. Fu, "An application of stochastic languages to fingerprint pattern recognition," *Pattern Recognition*, Vol. 8, pp. 173-179, 1976.

[8] B. Moayer and K. S. Fu, "A syntactic approach to fingerprint pattern recognition," *Pattern Recognition*, Vol. 7, pp. 1-23, 1975.

[9] O. Nakamura, K. Goto and T. Minami, "Fingerprint classification by directional distribution patterns", *Syst. Comput. Controls*, Vol. 13, No. 5, pp. 81-89, 1982.

[10] T. Ch. Malleswara Rao, "Feature extraction for fingerprint classification," *Pattern Recognition*, Vol. 8, pp. 181-192, 1976.

[11] S. T. Bow, *Pattern Recognition: Application to Large Data-Set Problems*, Marcel Dekker, Inc., New York, USA, 1984.

[12] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*, John Wiley and Sons, New York, USA, 1973.