Physical Sciences Research International
Vol. 2(1), pp. 1-5, January 2014
Full Length Research Paper

A modified robust and effective algorithm with mathematical calculation for minutiae detection

Arash Kalami

Department of Electrical Engineering, Urmia Branch, Islamic Azad University, Urmia, Iran.

Accepted 16 December, 2013

ABSTRACT

Fingerprint classification and recognition is still an open and very challenging problem in real world applications. To achieve good minutiae extraction in fingerprints with varying quality, preprocessing in form of binarization and skeletonization is first applied on fingerprints before they are evaluated by the neural network. Extracted minutiae are then considered as a 2D point pattern problem and an efficient algorithm is used to determine the number of matching points between two point patterns. Performance of the developed system is evaluated on a database with fingerprints from different people and experimental results are presented.

Keywords: Algorithm, back propagation pattern, features, recognitions.

E-mail: Ar.kalami@gmail.com.

INTRODUCTION

The complexity and uncertainty found in fingerprint based classification and recognition are wide in range. This leads us into difficult situation in learning, teaching and practicing the system for appropriate recognition. The Data Mining techniques are an appropriate algorithm to model the range of uncertainties found in fingerprint classification and recognition systems. Many approaches were developed for the fingerprint verification algorithm. They are model based, structure based, frequency based and syntactic. The model-based fingerprint classification technique uses the locations of singular points to classify a fingerprint into the five classes (Zhou and Gu, 2004; Chen and Lin, 2006; Nag et al., 2011; Sedghi, 2012; Eshera and Fu, 2004). In the structure-based approach, the estimated orientation field in a fingerprint image is used to classify the fingerprint into one of the five classes (Lai and Chang, 2006; Driscoll et al., 2011). A syntactic approach uses the formal grammar to represent and classify fingerprints. The variable sized minutiae based mechanism does not lend itself to indexing mechanisms. Typical graph based (Hong et al., 2008) and point pattern based approaches to match the minutiae from two fingerprints need to align the unregistered minutiae patterns of different sizes which makes them computationally expensive. Correlation based techniques (Sedghi, 2012) match the global patterns of the ridges and valleys to determine if the ridges align. The global approach to fingerprint is typically used for indexing and does not offer a very good individual discrimination. Further the indexing efficiency of the existing global representation is poor due to small number of categories that can be effectively identified and a highly skewed distribution of the population, in each category. The local texture analysis technique, is used where the fingerprint area of interest is tessellated with respect to the core point categories that can be effectively identified and a highly skewed distribution of the population, in each category. The local texture analysis technique, is used where the finger print area of interest is tessellated with respect to the core point (Karu and Jain, 2006).

PREPROCESSING

Here, the whole process of binarization, from filter mask design, orientation image estimation and smoothing to final enhancement filtering and post processing and detection and recognition are presented. All grayscale fingerprint images used in this paper and processes by this method are taken from the same database therefore

Figure 1. Advantages and disadvantages of the binarization process. Rings.

originate from the same sensor. By carefully choosing the different parameters, the process is adjusted towards that certain fingerprint sensor and no further manipulation of the process is needed. In case of sensor change or applying this method on other fingerprint images originating from different sensor, new parameters needs to be tuned to get the optimal performance for those images taken with that particular sensor. In Figure 1, it is possible to see some advantages and disadvantages of this method. The binarization method is capable to filter out some small cuts and fills small gaps or holes in the ridges. The disadvantage is that some minutiae can be interchanged like termination to bifurcation and backwards. Also, some small details in ridges can disappear.

Skeletonization is performed on so called “negative image” of the binarized fingerprint images since above specified rules uses 1 that represents the ridges and 0 for representing the background. The binarized fingerprint images uses exact the opposite signs. A negative image is simply formed by performing a logical NOT operation on the binarized fingerprint. The examination of the pixels is done in iterations where the first two rule sets are applied in turns. The pixels that can be erased are marked and first at end of each iteration are removed from the image. This process is repeated until there are no more pixels that can be removed from the image. Then the second skeletonization process is started to remove the remaining pixels to produce the 1-pixel wide lines. This process takes only one iteration. After that the image is converted back and the skeleton of the binarized fingerprint image is found. Following Figure 2 displays how the different stages between each iteration look likes. From the binary fingerprint image to the final skeleton.

The fingerprint and skeleton is laid on top of each other and the intersection of them both is highlighted with red color.

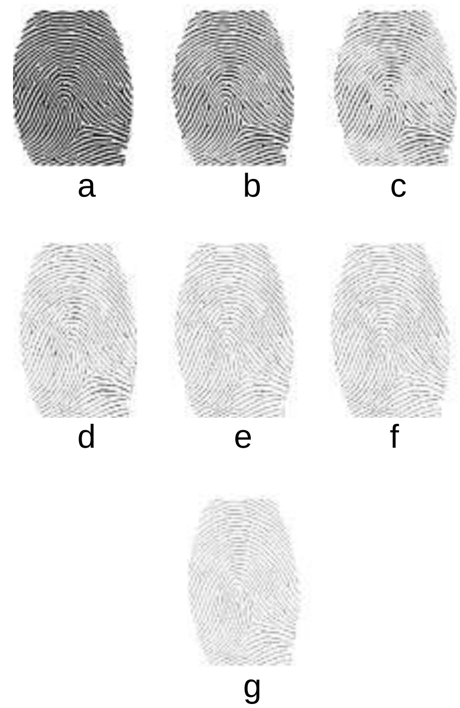

Figure 2. The steps between each iteration in the skeletonization process. (a) Binary fingerprint; (b) Iteration1; (c) Iteration 2; (d) Iteration 3; (e) Iteration 4; (f) Iteration 5; (g) Skeleton.

MATHEMATICAL CALCULATION OF NEW APPROACH

In this paper, a neural network with following characteristics was used for minutiae extraction. Input is made by stapling columns in the data portion into one column vector. Since the middle pixel is discarded, the column has 24 bit plus a 1 bit as a bias. The neural network has two hidden layers with 25 neurons in each. Output layer consists of 3 neurons, each representing one class of the patterns. The shape of this neural network is actually the same as the neural network in Figure 3b with an extent by the MAXNET. Example of how well the neural network works as a classifier is shown in Figure 3. Notice the accuracy of the finding the minutiae in the center in the zoomed section of the image in the Figure 3b. Accuracy of the neural network can be seen in Figure 3b.

Extracted minutiae from the fingerprint are together forming a point pattern in plane. Therefore matching two minutiae point patterns with each other are considered as a 2D point pattern problem. An algorithm is needed that localizes the maximum number of mutual points in the two point patterns. The algorithm described in this chapter is based on literature (Zhou and Gu, 2004) and a scientific paper (Driscoll et al., 2011). The point patterns are constructed only on positions (x, y) of minutiae in the plane. The minutiae type and orientation which provides extra information are disregarded due to possible type alternation and noise in orientation. The alternation can be caused by varying pressure between fingertip and the sensor and also by binarization process. Low pressure can cause that bifurcation minutiae appear as termination

Figure 3. Result of how well the neural network works as a classifier. Green boxes are the marked terminations. Red boxes are the marked bifurcations. (a) Extracted minutiae. (b) Zoomed section of the image.

minutiae in the fingerprint image. On other hand, high pressure can cause termination minutiae appear as bifurcation minutiae in the fingerprint image. Alternations of minutiae types by the binarization process have been pointed out. Bad quality of fingerprint image gives noisy orientation image and therefore false minutiae orientation. Alternation and false orientation of the minutiae gives higher risk that not all mutual pointes are detected. Since point patterns are based on positions of minutiae in fingerprint they form distinctive patterns.

With enough points in each pattern the positions (x, y) of the minutiae are the only information that is needed for good matching results. By using only (x, y) coordinates of minutiae yields that less memory is needed for implementation of this algorithm. The matching is performed on the two point patterns P with m number of pointes {p1, p2. . . pm} and Q with n number of points {q1, q2, . . . , qn}. The algorithm is made invariant towards possible translation and rotation among P and Q so the maximum number of matching points is found. The matching is preformed in two essential phases. First, a so called principal pair $p_{pair\_i} \leftrightarrow q_{pair\,a}$ is identified. Second, the matching pairs $p_i \leftrightarrow q_a$ are found.

The principal pair is the two points $p_{pair\_i} \leftrightarrow q_{pair\,a}$ corresponding to each other in the P and Q. It is the pair of points under which the maximum number of pairs $p_i \leftrightarrow q_a$ is found. Identification of $p_{pair\_i} \leftrightarrow q_{pair\,a}$ a is done by examining the scale s and rotation _ differences of the vectors. The values are calculated from points pi, pj , qa and qb where i $\neq$ j and a $\neq$ b.

$$\left|\overrightarrow{p_i p_j}\right| = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2} \quad (1)$$

The search for principal pair $p_{pair\_i} \leftrightarrow q_{pair\,a}$ is

$$\theta_{\overline{p_i p_j}} = \begin{cases} \arctan\dfrac{y_j - y_i}{x_j - x_i} & (x_j - x_i) > 0 \\ \arctan\dfrac{y_j - y_i}{x_j - x_i} + \pi & (x_j - x_i) < 0 \\ \dfrac{\pi}{2} & (x_j - x_i) = 0 \quad y_j - y_i > 0 \\ -\dfrac{\pi}{2} & (x_j - x_i) = 0 \quad y_j - y_i < 0 \end{cases} \quad (2)$$

$$\left|\overrightarrow{q_i q_j}\right| = \sqrt{(x_b - x_a)^2 + (y_b - y_a)^2} \quad (3)$$

$$\theta_{\overline{q_a q_b}} = \begin{cases} \arctan\dfrac{y_b - y_a}{x_b - x_a} & (x_b - x_a) > 0 \\ \arctan\dfrac{y_b - y_a}{x_b - x_a} + \pi & (x_b - x_a) < 0 \\ \dfrac{\pi}{2} & (x_b - x_a) = 0 \quad y_b - y_a > 0 \\ -\dfrac{\pi}{2} & (x_b - x_a) = 0 \quad y_b - y_a < 0 \end{cases} \quad (4)$$

$$S = \frac{\left|\overrightarrow{q_a q_b}\right|}{\left|\overrightarrow{p_i p_j}\right|} \quad (5)$$

$$\theta = \theta_{\overline{q_i q_j}} - \theta_{\overline{p_i p_j}} \quad (6)$$

conducted by testing each point p_i toward all points q_a. For every pair $p_i \leftrightarrow q_a$ the so called Matching Pairs Support (MPS) is calculated. The MPS value w_{ia} is the number of most common θ between the vectors with $S_{MIN} < s < S_{MAX}$ where $S_{MIN} = 0.98$ and $S_{MAX} = 1.02$. The value s is chosen around 1 because if the point patterns P and Q are originating from the same person and sensor the scale should be 1. Due to plasticity of the skin the points can shift the position to some extent and introduce noise to the coordinates. Therefore some variation in s is needed to be taken in a consideration. After each calculation towards pair $p_i \leftrightarrow q_a$ the cumulative sum M(θ)is updated. The M(θ) is array where θ denots the position in the array that is increased with 1. To make things easier the θ is converted from radians to degrees and quantized to 0.25 precision and is denoted θ^+. The original θ is in the interval [$-\pi$, π], the converted and quantized θ^+ is in the interval [0, 359.75] degrees with steps of 0.25 in-between. When all pairs $p_i \leftrightarrow q_a$ has been exploited, the accumulator sum is searched for the peak. The θ^+ that has the biggest peak in the M(θ) is the corresponding rotation between P and Q for the $p_i \leftrightarrow q_a$ pair. The following pseudo code shows how the above described MPS works. Set the accumulator sum M(θ^+) = 0 for all θ^+

S_{MIN} = 0.98

S_{MAX} = 1.02

for j = 1, ...,m, i $\neq$ j

for b = 1, ..., n, a $\neq$ b

$$S = \frac{|\overline{q_a q_b}|}{|\overline{p_i p_j}|}$$

if $S_{MIN} < s < S_{MAX}$

$$\theta = \theta_{\overline{q_i q_j}} - \theta_{\overline{p_i p_j}}$$

$$\theta^+ = quantize(\theta \frac{180}{\pi})$$

M(θ^+ _+) = M(θ^+) + 1

end

end

end

Search the M(θ^+) for which θ^+ it has the biggest peak. Return the biggest value labeled as w_{ia} together with the according θ^+.

To locate the best suitable pair $p_i \leftrightarrow q_a$ as principal pair the MPS value w_{ia} is compared to its previous value. If the new w_{ia} is bigger, then the new pair $p_i \leftrightarrow q_a$ is chosen as a new principal pair. The following pseudo code shows how the most suitable $p_i \leftrightarrow q_a$ is selected.

Set p_{max} = 0

for i = 1, ..., m

for a = 1, ..., n

Calculate the MPS value w_{ia} for $p_i \leftrightarrow q_a$.

If $p_{max} < w_{ia}$

$p_{max} = w_{ia}$, $\theta_0 = \theta^+, \quad pair\_i = i, \quad pair\_a = a$

end

end

end

Pair $P_{pair\_i} \leftrightarrow q_{pair\_a}$ is taken as the principal pair and indicates the rotation difference θ_0 between P and Q.

Localization of the matching pairs $p_i \leftrightarrow q_a$ among P and Q is executed in three stages. The first stage is based on the principal pair and the parameters S_{NORM} and θ_0 where S_{NORM} = 1.

This method is actually very similar to the principal pair determination method. Calculations of s and θ^* are now performed for vectors $\overrightarrow{p_{pair\_i} p_j}$ and $\overrightarrow{q_{pair\_a} q_b}$. Those values are then compared toward S_{NORM} and θ_0 respectively. If there is only one vector in P and Q that satisfy the following rules $|s - S_{NORM}| \leq \Delta s$ and $|\theta^* - \theta_0| \leq \Delta\theta$ then the pair pj $\leftrightarrow q_{ib}$ is matching pair. The matched pair can be collected to the point set called G. However, sometimes there are two or more points closer together and more then one vector satisfies the above rules. In that case, identification of the matching pair is impossible to do by above mentioned method. Those pointes have to be stored in another point set called F and localization of the matching pairs is done in the second stage. Following pseudo code demonstrates the categorization of the different points to set G and F.

S_{NORM} = 1

Put $p_{pair\_i} \leftrightarrow q_{pair\_a}$ to the set G

Set matching _flag[b] = 1 for all b.

for j = 1, ...,m, j $\neq$ pair_i

Set ib = 0, count = 0, Temp = []

for b = 1, ..., n, b $\neq$ pair_a

if matching _flag[b] = 1

$$S = \left| \frac{\overrightarrow{q_{pair\_a} q_b}}{\overrightarrow{p_{pair\_i} p_j}} \right|$$

$$\theta^* = (\theta_{\overline{q_{pair\_a} q_b}} - \theta_{\overline{p_{pair\_i} p_j}}) \frac{180}{\pi}$$

If $|s - S_{NORM}| \leq \Delta_s$ & $|\theta^* - \theta_0| \leq \Delta\theta$

ib = b, count = count + 1

Collect the pair $p_j \leftrightarrow q_a$ as a possible matching pair into

temporally set Temp.

end

end

end

if count = 1

The $p_j \leftrightarrow q_a$ is taken as a matching pair and is collected in set G.

Matching_ flag[I $_b$] = 0

elseif count > 1

Collect the points in set Temp to the set F.

end

end

Observe that the θ^* is not quantized to get more accurate difference measurement between θ^* and θ_0. However, it is recalculated to the interval of [0, 359.75]. The matching flag[b] serves as an indicator for points q_b that has yet not been uniquely matched. After the categorization of the points, set G holds the pairs that have been uniquely matched.

Those pairs in set G are taken as the matching pairs. The pairs that are stored in set F are the pairs $p_j \leftrightarrow q_b$ where p_j have more then one q_b points to form a possible matching pair with. An important aspect when categorizing the pairs into set G and F are the values of thresholds Δs and $\Delta\theta$. If those values are chosen too small the risk is that above mentions rules will not be fulfilled. Therefore high possibility that no pairs will be chosen to set G and F, even if the P is corresponding pattern to Q. If the threshold values are chosen to large there will be no unique pairs and all points will be collected to the set F. With empty set G the localization of matching pairs in set F can't be done. This means that the threshold values Δs and $\Delta\theta$ has to be carefully chosen. The second stage of localizing the matching pairs is done by finding the unique pairs $p_i \leftrightarrow q_b$ in the set F. To do that the points pj in set F are first transformed:

$$p_{transf\_i} = T_r(p_j) \tag{7}$$

to align with points q_b. Then the matching is done by measuring the distance between the point p_{transf_i} and q_b. The point q_b that is closest to the point p_{transf_i} with distance smaller then threshold d_1 as:

$$\left\| T_r(p_j) - q_b \right\| \leq d_1 \tag{8}$$

is a matching pair $p_j \leftrightarrow q_b$. The matched pair $p_j \leftrightarrow q_b$ is taken away from set F and is collected to the set G.

Transformation of the p_j to p_{transf_i} is done as following:

$$p_{transf\_i} = T_r(p_j) \Rightarrow \begin{pmatrix} x_{ptransf\_j} \\ y_{ptransf\_j} \end{pmatrix} = \begin{pmatrix} t_x \\ t_y \end{pmatrix} + \begin{pmatrix} s\cos\theta & -s\sin\theta \\ s\sin\theta & s\cos\theta \end{pmatrix} \begin{pmatrix} x_{pj} \\ y_{pj} \end{pmatrix} \tag{9}$$

where the tx and ty denotes the translation in x and y coordinates between P and Q point patterns. Estimation of the transformation parameters r = [tx, ty, s $\cos\theta$, s $\sin\theta]^T$ for the function $T_r(p_j)$ are based on matching pairs in set G. The following calculations are done to estimate the parameters:

$$r = \frac{1}{\det} \begin{pmatrix} l_p & 0 & -\mu_{xp} & \mu_{yp} \\ 0 & l_p & -\mu_{yp} & -\mu_{xp} \\ -\mu_{xp} & -\mu_{yp} & k & 0 \\ \mu_{yp} & -\mu_{xp} & 0 & k \end{pmatrix} \begin{pmatrix} \mu_{xQ} \\ \mu_{xQ} \\ l_{P+Q} \\ l_{P\_Q} \end{pmatrix} \tag{10}$$

CONCLUSION

The goal of this paper is to develop a complete system for fingerprint verification through extracting and matching minutiae. A neural network is trained using the back-propagation algorithm and works as a classifier to locate various minutiae. The performance of the developed system was evaluated on database with 2 fingerprints from 20 different people. The database was assembled from pre-stored fingerprints in (Zhou and Gu, 2004). The test showed that the system and algorithm is fully capable of distinguishing the related fingerprints apart from the non-related fingerprints. The system has proved to be robust towards translation, rotation and/or missing minutiae between the matched fingerprints.

REFERENCES

Chen P, Lin H, 2006. A DWT based approach for image steganography. Int J Appl Sci Eng. 4(3):275-290.

Driscoll EC, Martin CO, Ruby K, Russel JJ, Watson JG, 2011. Method and Apparatus for verifying Identity using Image Correlation. U.S. Patent 5067162.

Eshera M, Fu KS, 2004. A Similarity Measure between Attributed Relational Graphs for Image Analysis. Proceedings of 7th International Conference on Pattern Recognition, Montreal, Canada. pp. 75-77.

Hong L, Wang Y, Jain AK, 2008. Fingerprint image enhancement algorithm and performance evaluation. IEEE Trans Pattern Analysis Machine Intel, 21(4):777–789.

Karu K, Jain AK, 2006. Finger print classification. Pattern Recognition, 29(3):389-404.

Lai B, Chang L, 2006. Adaptive Data Hiding for Images Based on Harr Discrete Wavelet Transform. Lecture Notes in Computer Science, 4319:1085-1093.

Nag A, Biswas S, Sarkar D, Sarkar PP, 2011. A novel technique for image steganography based on DWT and Huffman encoding. Int J Comp Sci Secur, 4(6):561-565.

Sedghi T, 2012. A Fast and Effective Model for cyclic Analysis and its application in classification" Arab J Sci Eng, 38(4):927-935.

Sedghi T, 2012. Non linear transform for retrieval system in consideration of feature combination technique. Int J Nat Eng Sci, 6(3):21-23.

Zhou J, Gu J, 2004. A model-based method for the computation of fingerprints orientation fField. IEEE Trans Imag Proc, 13(6):885-888.