EXTRACTING FEATURES FROM FINGERPRINT

Hussein A. Lafta
Department of Computer Science, College of Sciences for Women, University of Babylon

1-Abstract

The flow pattern of ridges in a fingerprint is unique to the person in that no two people with the same fingerprints have yet been found. For automated fingerprint image matching, a machine representation of a fingerprint image is often a set of minutiae's in the print; a minimal, but fundamental, representation is just a set of ridge endings and bifurcations. However, after all the years of using minutiae, a precise definition of minutiae has never formulated.

In this work we give a defition for minutiae and derive an expression, which estimates the probability of falsely associating minute-based representations from two arbitrary fingerprints.

2-Introduction

Fingerprint based personal identification is routine used in forensic laboratories and identification units around the world and it has been accepted in the court of law for nearly a century.

The fundamental premises on which fingerprint identification is based are [1]:

1. Fingerprint details are permanent.
2. Fingerprints of an individual are unique.

The validity of the first premise has been established base on the anatomy and morphogenesis of friction ridge skin.It is the second premise, which is being challenged in recent years. In this study, we define the individuality problem as the probability of a false association: given two fingerprints from two fingerprints, determine the probability that they are "sufficiently" similar. If two fingerprints originating from two different fingerprints are examined at a very high level of detail (resolution), we may find that the fingerprints are indeed different.

However, most human experts and automatic fingerprint identification systems (AFIS) declare that fingerprints originate from the same source if they are "sufficiently" similar [2]. How much similarity is enough depends on typical (intra-class) variations observed in the multiple impressions of a finger.

3-Minutia Definition

There are very good reasons for formulating a precise definition of minutiae:

1. The ground truth of minutiae, both ridge endings and ridge bifurcations, in a fingerprint image will be well defined for manual annotation.
2. Automated minutia extraction algorithm can be compared.
3. Automated extraction algorithms can be designed with sub-pixel accuracy.
4. Vendors could compensate for biases in their minutia extraction algorithms and can construct templates that are more interoperable.
5. A good model of a minutia will allow for the definition of a well-grounded quality measure that considers how well the data fits the model.

The location and orientation of minutiae could be defined based on the result of fingerprint processing. Commonly, a pixel that has only one neighbor in the thinned image is said to be a ridge ending, while a pixel that has three neighbors is said to be a ridge bifurcation.
Clearly, the location and orientation of minutiae are greatly dependent on the various processing steps. For example, threshold a ridge at half the ridge height, or close to the top of the ridge will affect the thinning and hence the final location of the ridge ending.

In general, the minutia locations in image I(x,y) should be equal to the locations in image with transformed intensities g(I(x,y)). Hence, the positions of minutiae should not be based on some ill-defined, non linear function (like threshold) of the image function. That is, the positions should be calculated using the image function itself.

In a way, asking where the ridge of a fingerprint ends is very much likes asking, while walking off a mounting ridge in the direction parallel to the ridge, where the ridge ends and where the valley starts.

What we want to di is model the image function I (x,y) of a minutia. The perpendicular cross-section of a ridge is very much like a portion of sinusoid as the function W(x), while a parallel cross-section along the length of the ridge is a smooth step function L(x). We model these functions as

$$L(y) = 1 / (1+e^{-y/\alpha}) \quad \text{and} \quad W(x) = 1/2\ (1 + \cos(2\pi x/\beta))$$

with the minutia function m (x, y)=h L(y) W(x) defined for $-\beta/2 \le x \le \beta/2$.

Hence h is the height of the ridges, α and β are scale constants. In practice if the period of the ridge oscillations is r, the we choose β=r and α=r/10 [3]. Here r can be estimated gold-bally or locally for the ridge in question.

4-FingerPrint Processing

The purpose of fingerprint image processing is to extract a condensed representation of the image. This representation (referred as a template) is used for fingerprint matching.

Fingerprints can be represented by a large number of features, including the overall ridge flow pattern, ridge frequency, location and position of singular points(core(s) and delta(s), type, direction, and location of mintae points, ridge counts between pars of minutiae, and location of pores (see Figure 1).

Figure 1. A fingerprint image of type "right loop". The overall ridge structure, singular points, and sweat pores are shown.

From the ridge flow pattern is extracted the minutiae detail that makes a fingerprint different from other print. A first part of the detail that is usually used in fingerprint representation is the set of endings and ridge bifurcations in the flow pattern Figure 2 gives a portion of a fingerprint image.

The minutiae extraction process typically consists of ridge extraction, followed by thinning and minutiae extraction. Ridge extraction, or ridge segmentation, is essentially the step of finalizing the fingerprint image. That is, somehow the fingerprint image.

$I: (x,y) \longrightarrow [0,255]$ is converted to $B: (x,y) \longrightarrow \{0,1\}$, where the value 0 corresponds to valleys and 1 to ridges.

One way to accomplish this is to select a global threshold T and converting the image I(x,y) to a binary image as

$$B(x,y)= \begin{cases} 1 , & I(x,y) > T \\ 0 , & I(x,y) <= T \end{cases}$$

Due to the poor quality of many fingerprint images, this approach is most often adequate for extracting minutiae. In areas of the fingerprint that are dry, no ridges may be detected, while in areas where the finger is wet, no valley may be detected.

The typical solution is to use a threshold T(x,y), which is a function of the spatial location. Virtually every published method of feature extraction [4] computes the orientation field of the fingerprint image, which reflects the local ridge direction at every pixel.

At this point, a binary image has been computed which more or less faithfully represents the original image. The ridges will have width that will vary over the fingerprint images.

The next processing steps are typically composed of a sequence of image operations: Directional smoothing, Thinning [6], Morphological filtering and minutia pruning. (Post processing). These types of operations may be performed in different order.

5-Metric of Similarity

In order to solve the individuality problem, we need to first define a prior the representation of fingerprint (pattern) and the metric of similarity.

In this study, we have chosen minutiae representation of fingerprints because it is utiljzed by forensic experts, has been demonstrated to be relatively stable, and has been adopted by most of the automatic fingerprint matching systems.

The similarity metric is the number of corresponding minutiae between two minutiae sets. Given a representation scheme and a similarity metric, there are two approaches for determining the individuality of the fingerprints. In the empirical approach, representative samples of fingerprints are collected and using a typical fingerprint matcher.

The accuracy of the matcher on the samples provides an indication of the uniqueness of the fingerprint with respect to the matcher. There are known problems (and costs) associated with the collection of a large number of reprehensive samples. In the theoretical approach to individuality estimation, one model all realistic phenomenon affecting inter-class and intra-class pattern variations. Given the similarity metric, one could then, estimate the probability of a false association. In this work, we propose a theoretical formulation of the fingerprint individuality.

5-Fingerprint Probability Model

We have developed a model to obtain a realistic and accurate probability of correspondence between fingerprints.

The probability obtained using this model can be estimated by making the following assumptions:

1. We consider only two types of minutiae features: ridge endings and ridge purification. Additionally, we do not distinguish between the two types of minutiae because the can not be accurately discriminated. Since minutiae can reside only on ridges, which follow a "flow" pattern, we implicitly model the statistical dependence between minutiae directions and locations.
2. We assume a uniform distribution of minutiae in a fingerprint with the restriction that two minutiae can not be very close to each other. this assumption approximates the slightly overviewed uniform distribution of minutiae found by Stony[5].
3. Correspondence of a minutiae pair is an independence event and each correspondence is equally important.

4. We do not take into account fingerprint image quality in individuality determination since it is very difficult to reliable assign a quality index to a fingerprint.

The fingerprint correspondence problem involves matching a template with the input. we assume that a reasonable alignment has been established between the template and the input. The alignment of the input minutiae set with the template minutiae is done so that the minutiae correspondences can be determined within a small tolerance. given a input fingerprint containing n minutiae, our goal is to compute the probability that any arbitrary fingerprint(template in a database of fingerprints) containing m minutiae will have exactly q corresponding minutiae with the input see Figure 3.

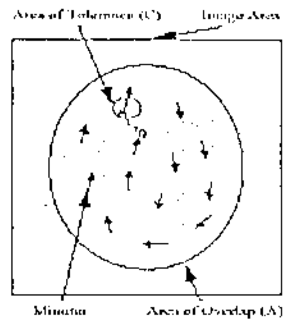

Figure 3. Fingerprint and minutiae.

Since the fingerprint minutiae are defined by their location, (x, y), and by angle of the ridge on which they reside, θ, the input and the template mintiae sets, I and T respectively, are defined :

$$I = \{\{x'_1, y'_1, \theta'_1\},\{x'_2, y'_2, \theta'_2\},\ldots,\{x'_n, y'_n, \theta'_n\}\} \qquad (1)$$

$$T = \{\{x_1, y_1, \theta_1\},\{x_2, y_2, \theta_2\},\ldots,\{x_m, y_m, \theta_m\}\} \qquad (2)$$

A minutiae j in the input fingerprint is consider as "corresponding" or "matching" to the minutiae i in the template, if and only if

$$\sqrt{(x'_1 - x_1)^2 + (y'_1 - y_1)^2} \le r_0 \quad \text{and} \qquad (3)$$

$$\mathrm{Min}\left(|\theta'_i - \theta_j|, 360 - |\theta'_i - \theta_j|\right) \le \theta_0, \qquad (4)$$

Where r_0 is the tolerance in the distance and θ_0 is the tolerance in the angle. Let A be the total area of overlap between the input and the template fingerprints after a reasonable alignment has been achieved. the probability that any arbitrary minutiae in the input will match any arbitrary minutiae in the template , only in terms of location, and only in terms of direction, are given by Eqs. (5), and (6) respectively. Eq. (5) assumes that (x_1, y_1) and (x'_1, y'_1) are independent and Eq.(6) assumes that θ, θ' are independent.

Let $\delta_x = x' - x$, $\delta_y = y' - y$ and $d = \sqrt{\delta_x^2 + \delta_y^2}$

$$P(d \le r_0) = \frac{\pi r^2}{A} = \frac{C}{A} \qquad (5)$$

$$P(\mathrm{Min}(|\theta'_i - \theta_j|, 360 - |\theta'_i - \theta_j|) \le \theta_0) = \frac{2\theta_0}{360} \qquad (6)$$

If the template contains m minutiae, the probability that only one minutiae in the input will correspond to any of the template minutiae is given by mC/A. Now, given two input minutiae, the probability that only "first" one corresponds to any of the m template minutiae is the product of the probabilities that the first input minutiae has a correspondence (mC/A) and the second minutiae dose not have a correspondence ((A-mC)/(A-C)). Thus, the probability that exactly one of the two input minutiae matches any of the m template minutiae is 2 × mC/A × (A-mC)/(A-C), since either the first input minutiae alone may have a correspondence or the second input minutiae alone may have a correspondence. if the input fingerprint has n minutiae , the probability that exactly one input minutiae matches one of the m template minutiae is

$$P(A,C,m,n,1) = \binom{n}{1} \left(\frac{mC}{A}\right) \left(\frac{A-mC}{A-C}\right) \qquad (7)$$

The probability that there are exactly P corresponding minutiae between the n input and m template minutiae is given by :

$$P(A,C,m,n,p) = \binom{n}{p} \left(\frac{mC}{A}\right) \left(\frac{(m-1)C}{A-C}\right) \qquad (8)$$

Letting M A/C, assuming M to be integer (A >> C), and rearranging, we obtain

$$P(A,C,m,n,p) = \frac{\binom{n}{p} \left(\frac{mC}{A}\right) \left(\frac{(m-1)C}{A-C}\right)}{\binom{M}{n}} \qquad (9)$$

The above analysis considers a minutiae correspondence based solely on the minutiae location. Next we consider a minutiae correspondence that depends on minutiae directions as well as minutiae positions. For the sake of this analysis, let as assume that the minutiae directions are completely independent of the minutiae positions and matching minutiae position and minutiae direction are therefore independent events.

Let L be such that

$P(\mathrm{Min}(|\theta'_i - \theta_j|, 360 - |\theta'_i - \theta_j|) \leq \theta_0) = L$ in Eq. (6).

Given n input and m template minutiae, the probability of p minutiae falling into the similar positions can be estimated by Eq. (9). once p minutiae positions are matched , the probability that $q \leq p$ minutiae among them have similar direction is given by :

$$\binom{p}{q} (L)^q (1-L)^{p-q}, \qquad (10)$$

where L is the probability of two position-matched minutiae having similar direction and 1- L is the probability of two position-matched minutiae taking difference directions. Therefore, probability of matching q minutiae in both position as well as direction is given by

$$P(A,C,m,n,p) = \sum_{P-q}^{\min(m,n)} \left(\frac{\binom{n}{p}\left(\frac{M-m}{n-p}\right)}{\binom{M}{n}} \times \binom{p}{q} (L)^{q} (1-L)^{p-q} \right) \qquad (11)$$

6- Discussion and Parameter Estimation

Until now, we have assumed that locations are uniformly distributed within the entire fingerprint area. However, the number (or the area) of ridge arose all fingerprint types is approximately the same. Since A is the area of the overlap between the template and the input fingerprints, the ridges occupy approximately A/2 of the area, with the other half being occupied by the valley. Since the minutiae can lie only on ridges, i.e., along a curve of length A/w, where w is the ridge period, he value of M in Eq. (11) should therefore be changed from M−A/C to M−(A/w) /(2 r_0) , Where $2r_0$ is the length tolerance in minutiae location. this analysis assumes that the ridge direction information/ uncertainty is completely captured by Eq. (6).

Our model has several parameters, namely r_0, L, w, A, m, n and q. The value of L further depends on θ_0. The values of r_0, L, and w are estimated in this section for a given sensor resolution.

The value of r_0 should be determined to account for the variation in the different impressions of the sane finger. However, since the spatial tolerance is dependent upon the scale at which the fingerprint images are scanned, we need to calculate it for the specific sensor resolution.During aligning, we give a matted pair of fingerprints using the overall transformation, (δ_x,δ_y) for each corresponding minutiae pair was computed; distance offset $(d=\sqrt{\delta_x^2 + \delta_y^2})$ estimates for all minutiae in all mated fingerprint pars were pooled to obtain a distribution for distance between the corresponding minutiae. we are seeking that value of r_0 for which $P(d \leq r_0) \geq 0.975$, the value of r_0 which counts for at least 97.5% of variation in the minutiae positions of genuine fingerprint matching. The value of r0 is found be 15 pixels for fingerprint scanned at 500 dpi resolution[5].

To estimate the value of L, we first estimate the value of θ_0. the value of θ_0 can also be estimated thruogh expriments so that it counts for 97.5% variation in the minutiae angles in genuine fingerprint matching.

The value of w was taken as reported in [4]. It estimated the value of ridge period as 0.463 mm/ridge from a database of 412 fingerprints.

7-Conclusions

The model proposed here is relatively simple. It ignores most of the known (weak) dependence among the features and does not directly include features such as ridge counts, fingerprint classes , ridge frequencies, permanent scars, etc.

For these reasons, we suspect that the proposed model does no yet compete in predicting the performance of a human fingerprint expert matcher. By additionally considering a more detailed fingerprint representation (e.g., different minutiae types, sweat pore information), the confidence genuine mates can be reinforced and the spurious associations among the impostors can be ruled out. The extension of our proposed model to include additional features is a topic for our future research.

References

[1] Sharath P., Sail P. and Anil K. J., "On the Individuality Of Fingerprint", sharat@ watson.ibm.com .

[2] A.K. Jain, L. Hong, S. Pankatti and R.Boll, " An identity Authentication System Using Fingerprint" Proc. IEEE Vlo.85, No. 9, pp 1365-1388, 1997.

[3] D.A Stoney and J.I Thornton, "A Critical Analysis of Quantative Fingerprint Individuality Models", journal of Forensic Science, Vol.31, NO. 4, Oct 1986, pp 1187-1216.

[4] S.L. Sclove, "The Occurrence of Fingerprint Characteristics as a Two Dimensional Process", Journal of American Statistical Association, Vol. 74, No.367, pp. 588-595, 1979.

[5[Nikolaos G.B.,"A Parallel Symmetric Thinning Algorithim", 1988

[6] Y.Sun and M.A. Abidi, "Surface Matching By 2D Points Fingerprint", yiyong@iristown.engr.utk.edu, 2000.

[7] Ruud M. Bolle,"Fingerprint Minutiae: A constructive definition", bole @ us.ibm.com, 2002.

الخلاصة

تكون بصمة الإبهام لكل شخص فريدة, و لا يوجد لحد الآن شخصين متشابهين في بصمة الإبهام. يكون التمثيل الآلي لبصمة الإبهام غالبا على شكل تفصيلات في بصمة الإبهام, خاص بالحد الأدنى, لكن أساسي. يكون التمثيل بالضبط مجموعة من نهايات مرتفعة و متشعبة.على أي حال بعد كل السنين التي كان يستخدم فيها تفصيلات الإبهام لا يمكن قط صياغة تعريف دقيق لتفصيلات الإبهام.

في هذا العمل تم إعطاء تعريف لتفصيلات الإبهام و اشتقاق تعبير لتخمين احتمالية فشل ربط تمثيل تفاصيل الإبهام الناتجة من بصمتي أبهام اعتباطية.