Proceedings International Conference on Pattern Recognition, Tampa, FL, December 8-11, 2008
IEEE Computer Society Press

Generative Models for Fingerprint Individuality Using Ridge Models

Chang Su, Sargur N. Srihari
Center of Excellence for Document Analysis and Recognition (CEDAR)
University at Buffalo, State University of New York
Amherst, New York 14228, U.S.A
{changsu, srihari}@cedar.buffalo.edu

Abstract

Generative models of pattern individuality attempt to learn the distribution of observed quantitative features to determine the probability of two random patterns being the same. Considering fingerprint patterns, Gaussian distributions have been previously used for minutiae location and von-Mises distributions for minutiae orientation so as to determine the probability of random correspondence (PRC) between two fingerprints. Motivated by the fact that ridges have not been modeled in generative models and the benefits from ridge points in fingerprint matching, ridge information is incorporated into the generative model by using the distribution for ridge point location and orientation. The proposed model offers a more accurate fingerprint representation from which more reliable PRCs can be computed. Based on parameters estimated from fingerprint databases, PRCs using ridge information are seen to be much smaller than PRCs computed with only minutiae.

1. Introduction

Fingerprints have been used for identification from the early 1900s. Their use for uniquely identifying a person has been based on two premises, that, (i) they do not change with time and (ii) they are unique for each individual. Until recently, fingerprints had been accepted by courts as a legitimate proof of identification. But, after the 1999 case *US vs Byron Mitchell*, fingerprint identification has been challenged under the basis that the premises stated above have not been objectively tested and the error rates have not been scientifically established.

Studies on the individuality of fingerprints date back to the late 1800s. More than twenty models have been proposed to establish the improbability of two random people having the same fingerprint. All models try to quantify the uniqueness property to be able to defend fingerprint identification as a legitimate proof of identification in the courts. These models can be classified into five different categories, namely, grid-based models, ridge-based models, fixed probability models, relative measurement models and generative models. The latest class of models, namely, the generative models aim at being flexible to represent observed distributions through different fingerprint databases and then ascertained uncertainties from models. Based on the the assumed non-independence of minutia locations and orientations, various mixture models could be used [5] [2].

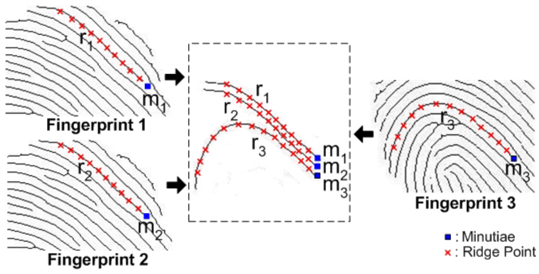

Figure 1. Ridge shape similarity infers the matching of corresponding minutiae pair and ridge point pairs for two genuine ridges (r_1 and r_2), but not for two different ridges (r_2 and r_3).

In existing generative models only minutiae have been modeled without considering ridge features. Minutiae means small details in the fingerprints, it refers to the ridge endings and ridge bifurcation. The generative models based purely on minutiae are sufficient to model biometric algorithms. However they are insufficient to model forensic scenarios where latent prints are used. Due to the poor quality, detected minutiae on latent prints usually has small quantity and low quality. Therefore ridge features such as ridge points are important in discrimination. Figure 1 illustrates the potential effectiveness of ridge features in a local view: although in some case, two different fingerprints might have the occasionally overlapped minutiae, it will be difficult for this pair to also have the overlapping ridge points at the same time. With this motivation, we further embed ridge information into existing generative models by using the distribution for ridge points. The proposed model offers more reasonable and accurate fingerprint representation and therefore a more reliable probability of random correspondence (PRC).

The following of this paper is organized as follows: Section 2 reviews the existing generative models for minutiae location and orientation [2]. Section 3 introduces a new generative model for both minutiae and ridges. Experimental results are given in Section 4. The paper concludes with a summary in Section 5.

2. Generative Models for Minutiae

In [2], a minutiae is denoted by (X, D) where X denotes the location of a minutiae and D denotes the orientation. A joint distribution of (X, D), assuming that minutiae location and orientation are independent, can be represented by the mixture density function as

$$f(s, \theta|\Theta_G) = \sum_{g=1}^{G} \tau_g f_g^X(s|\mu_g, \Sigma_g).f_g^D(\theta|\nu_g, \kappa_g, \rho_g) \tag{1}$$

where G is the number of mixture models, τ_g is the non-negative weight for that model. $f_g^X(s|\mu_g, \Sigma_g)$ is the probability density function of a bivariate Gaussian distribution over the position of minutiae. $f_g^D(\theta|\nu_g, \kappa_g, \rho_g)$ is the probability density function over the orientations and is given by

$$\rho_g \upsilon(\theta).I\{0 \leq \theta < \pi\} + (1-\rho_g)\upsilon(\theta-\pi).I\{\pi \leq \theta < 2\pi\} \tag{2}$$

where $\upsilon(\theta)$ is the von-mises distribution modeling the angular random variables in $[0, \pi)$ with the von-mises parameters ν_g and κ_g.

Given a template T with n minutiae and an input/query Q with m minutiae and w out of them match, the probability of Random Correspondence is given by

$$\begin{aligned} &PRC_0 = p^*(w; Q, T) = \\ = &\begin{pmatrix} n \\ w \end{pmatrix} .(p_m(Q, T))^w (1 - p_m(Q, T))^{n-w} \end{aligned} \tag{3}$$

The probability is a binomial probability whose parameters are n and $p_m(Q, T)$. The latter is the probability that a random minutia from Q will match a minutia from T. Since most of the matchers try to maximize the number of matchings, $p_m(Q, T)$ can be calculated by the conditional expectation which is equivalent to the number of minutia matches between Q and T. The estimation can be written as

$$\frac{n.p_m(Q, T)}{(1 - (1 - p_m(Q, T))^n)} = w_0 \tag{4}$$

where w_0 is the number of matches by the matching algorithm [1] with the proposed models fit into Q and T.

3 Generative Model for Minutiae and Ridges

Until now, only minutiae was used in the framework of generative models for fingerprints. Ridge details in fingerprints provide vital information about fingerprints and it is intuitive to see that any generative model that can utilize ridge details as well into its framework can only be a better representation of the generative model for fingerprints. In [3], we proposed an algorithm to utilizes ridge information more effectively– by choosing representative points along the ridges. Our extensive experimental results on both full-size fingerprint matching and partial fingerprint matching demonstrate that fingerprint verification using minutiae together with general ridge information is more accurate than using minutiae only. These observations, meanwhile, implicate that the fingerprint individuality would be more accurately studied if general ridges would be considered together with minutiae when designing a generative model. Motivated by this work, we formulate below way to introduce ridge information to the generative model for minutiae discussed in section 2. The parameter set presenting ridge information $\mathcal{T}$ is appended to the three existing parameters x, y, θ of the minutiae and a generative model is built for these parameters.

In order to take ridge information into account, the ridge is represented as a set of ridge points sampled at equal interval of inter ridge width. Three types of ridges are defined as (i) short ridges: $l(r) \leq L/3$, (ii) medium ridges: $L/3 < l(r) < 2L/3$ and (iii) long ridges: $2L/3 \leq l(r) \leq L$, where L is the maxima ridge length. These three possible ridge length types can be associated with any minutiae. Without loss of generality, we can assume that there exist only three possible ridge length types corresponding to a minutiae. For the generative model, the ridge length type is modeled as a uniform distribution $F^l(l_r|a, b)$, where $[a, b]$ is the interval of the uniform distribution.

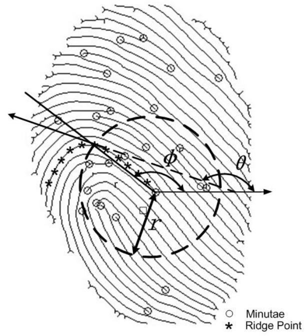

Figure 2. Presentation of ridge points in polar coordinates.

For ridges with different lengths, different ridge points are picked as anchors. For medium ridges, $(L/3)^{th}$ ridge point is picked and for long ridges, both $(L/3)^{th}$ and $(2L/3)^{th}$ are picked. None ridge point will be chosen for short ridges. The reason for choosing such ridge points is described in [3]. For the generative model, the ridge points are modeled as a joint distribution of the ridge point location and orientation. The proposed model for fingerprint presentation is based on a mixture consisting of G components. Each components is distributed according the density of the minutiae and the ridge points. The equations of the generative model $f(\cdot|\Theta_G)$ for short, medium and long ridges are given in Eq.5, Eq.6 and Eq.7 respectively.

$$F^l(l_r) \cdot \sum_{g=1}^{G_1} \pi_g F_g^m(s_m, \theta_m|\Theta_G) \tag{5}$$

$$F^l(l_r) \cdot \sum_{g=1}^{G_2} \pi_g F_g^m(s_m, \theta_m|\Theta_G) \cdot F_g^{\frac{L}{3}}([r, \phi, \theta]_{\frac{L}{3}}|\Theta_G) \tag{6}$$

$$F^l(l_r) \cdot \sum_{g=1}^{G_3} \pi_g F_g^m(s_m, \theta_m|\Theta_G) \cdot F_g^{\frac{L}{3}}(r, \phi, \theta]_{\frac{L}{3}}|\Theta_G) \cdot F_g^{\frac{2L}{3}}([r, \phi, \theta]_{\frac{2L}{3}}|\Theta_G) \tag{7}$$

where $F_g^m(\cdot)$ represents the distribution of the minutiae location s_m and direction θ_m. $F_g^i(\cdot)$ presents the distribution of the i^{th} ridge points. They are defined as

$$F_g^m(s_m, \theta_m|\Theta_G) = f_g^X(s_m|\mu_{gm}, \Sigma_{gm}) \cdot f_g^D(\theta_m|\nu_{gm}, \kappa_{gm}, \rho_{gm}) \tag{8}$$

$$F_g^i(r_i, \phi_i, \theta_i|\Theta_G) = f_g^P(r_i, \phi_i|\mu_g^i, \Sigma_g^i, \nu_{g\phi}^i, \kappa_{g\phi}^i, \rho_{g\phi}^i) \cdot f_g^D(\theta_i|\nu_{g\theta}^i, \kappa_{g\theta}^i, \rho_{g\theta}^i) \tag{9}$$

Table 1. Results for testing the goodness of fit of the mixture models with and without ridge information.

p-value	Minutiae and Ridge Information	Only Minutiae
$p - value > 0.01$ (Model Accepted)	679	574
$p - value \leq 0.01$ (Model Rejected)	121	226

In Eq.8, $f_g^X(\cdot)$ and $f_g^D(\cdot)$ are defined as in Eq.1, where s_m and θ_m present the minutiae location and direction. The distribution of ridge points takes into account their location and direction as well, but we present them in polar coordinates. Figure 2 depicts this idea. In Eq.9, $f_g^D(\cdot)$ presents the distribution of the ridge point direction, θ_i is the direction of the i^{th} ridge point and $f_g^P(\cdot)$ is the distribution of ridge point location given by

$$f_g^P(r_i, \phi_i|\mu_g^i, \sigma_g^i, \nu_{g\phi}^i, \kappa_{g\phi}^i, \rho_{g\phi}^i) = f_g^R(r_i|\mu_g^i, \sigma_g^i) \cdot f_g^D(\phi_i|\nu_{g\phi}^i, \kappa_{g\phi}^i, \rho_{g\phi}^i) \tag{10}$$

where r_i is the distance from the i^{th} ridge point to the minutiae, ϕ_i is the positive angle required to reach the i^{th} ridge point from the polar axis and $f_g^R(r_i|\mu_g^i, \sigma_g^i)$ is a one dimension Gaussian distribution with mean μ_g^i and variance σ_g^i, which are learned from the FVC2002 and fixed in the model.

To estimate the unknown parameters in the generative model, we develop an algorithm based on the EM algorithm. The number of components G for the mixture model was found after validation using k-means clustering.

4 Experiments and Results

Generative models without ridge information introduced in Section 2 and with the ridge information introduced in Section 3 have been implemented and experiments have been conducted on FVC2002 DB1 [4]. The database has 100 different fingerprints with 8 impressions of the same finger.

We compare the results to that of [2]. First we fit the mixture model on the data set and compare the adequateness of the generative models. To test the goodness of fit of the mixture models to the observed minutiae and ridge features, the Chi-square statistical hypothesis test is used. The results are given in Table 1. For the FVC2002 DB1, we computed the number of

Table 2. PRC for different fingerprint matches with varying m ,n and w

m	n	w	PRC_0 Minutiae and Ridge Information	PRC_0 Only Minutiae
26	26	12	6.9×10^{-8}	3.6×10^{-4}
		20	2.3×10^{-18}	2.3×10^{-11}
36	36	12	1.1×10^{-6}	1.5×10^{-3}
		20	3.1×10^{-15}	3.1×10^{-9}
		32	2.4×10^{-34}	5.5×10^{-24}
46	46	12	4.2×10^{-6}	3.8×10^{-3}
		20	7.8×10^{-14}	5.2×10^{-8}
		32	4.8×10^{-30}	6.6×10^{-20}
		42	1.4×10^{-48}	7.6×10^{-35}

Table 3. The probability of at least two fingerprints matching among k fingerprints with average number of minutiae 39 and average number of matching minutiae pairs 27

k	Minutiae and Ridge Information p(k)	only Minutiae p(k)
100	7.64×10^{-21}	5.84×10^{-12}
10^5	7.72×10^{-15}	5.90×10^{-6}
10^{10}	7.72×10^{-5}	0.702
10^{20}	0.9999999999999946	1

fingerprints with p-values above (corresponding to accept the model) and below (corresponding to reject the model) the threshold 0.01. Of the 800 fingerprints, 679 are accepted with ridge model which is higher than 574 and 121 are rejected which is smaller than 226. The results imply that the mixture model with ridge information offers a better fit to the observed fingerprints compared to the model without ridge information.

Then random fingerprints are generated from the model. Values of PRC_0 are calculated using the formulae introduced in Section 3. The results are presented in Table 2. The PRCs are calculated through varying number of minutiae in template(m), input(n) and the number of ridges matched(w). Our highlight is that the PRC values estimated by the model embedding ridge information are never greater than PRC values without ridge information, which indicates that ridge information strengthens individuality of fingerprints.

Based on this model, the probability of at least two fingerprints matching among k fingerprints, $p(k)$, is computed. The Table 3 shows the probability $p(k)$ for some values of k. For example, in $100,000$ randomly chosen fingerprints, there is only 7.72×10^{-15} probability that some pair of them will match if we consider both minutiae and ridge in matching. This probability is much smaller than previous minutiae only model which is 5.90×10^{-6}.

5 Summary

Generative models of individuality attempt to model the distribution of features and then use the models to determine the probability of random correspondence. While models have been proposed for minutiae, ridges have not been considered. The paper proposes a model for both minutiae and ridge information. We modified the previous generative model with a mixture distribution to model ridge information. The new generative model is compared with the generative model without ridge information on the FVC2002 DB1. The experiments show that the proposed model offers a more reasonable and more accurate fingerprint representation. PRCs with ridge information are much smaller than PRCs without ridge information. The results provide a much stronger argument for the individuality of fingerprints in forensics than previous generative models.

Acknowledgement This work was funded by Grant Number 2005-DD-BX-K012 awarded by the National Institute of Justice, Office of Justice Programs, U. S. Department of Justic. Points of view are those of the authors and do not necessarily represent the official position or policies of the U. S. Department of Justice.

References

[1] S. Chikkerur, A. N. Cartwright, and V. Govindaraju. K-plet and cbfs: A graph based fingerprint representation and matching algorithm. *International Conference on Biometrics*, pages 309–315, 2006.

[2] S. Dass, Y. Zhu, and A. K. Jain. Statistical models for assessing the individuality of fingerprints. *Fourth IEEE Workshop on Automatic Identification Advanced Technologies*, pages 3–9, 2005.

[3] G. Fang, S. N. Srihari, H. Srinivasan, and P. Phatak. Use of ridge points in partial fingerprint matching. *Proc. of SPIE: Biometric Technology for Human Identification IV*, pages 65390D1–65390D9, 2007.

[4] D. Maio, D. Maltoni, and R. Cappelli. Fingerprint verification competition. *http://bias.csr.unibo.it/fvc2002/*, 2002.

[5] S. Pankanti, S. Prabhakar, and A. K. Jain. On the individuality of fingerprints. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(8), 2002.