

A Statistical Evaluation Model for Minutiae-Based Automatic Fingerprint Verification Systems

J.S. Chen and Y.S. Moon

Department of Computer Science and Engineering,
The Chinese University of Hong Kong, Shatin, N. T. Hong Kong
{jschen, ysmoon}@cse.cuhk.edu.hk

Abstract. Evaluation of the reliability of an Automatic Fingerprint Verification System (AFVS) is usually performed by applying it to a fingerprint database to get the verification accuracy. However, such an evaluation process might be quite time consuming especially for big fingerprint databases. This may prolong the developing cycles of AFVSs and thus increase the cost. Also, comparison of the reliability of different AFVSs may be unfair if different fingerprint databases are used. In this paper, we propose a solution to solve these problems by creating an AFVS evaluation model which can be used for verification accuracy prediction and fair reliability comparison. Experimental results show that our model can predict the performance of a real AFVS pretty satisfactorily.

1 Introduction

Minutia-based AFVS is widely used in numerous security applications. A common practice for evaluating the reliability of an AFVS is to apply it to a fingerprint database to get the FAR and FRR. Generally speaking, the experimental result can provide sufficient confidence only if the database is big enough. As one to one matching is usually adopted in such evaluations, experiment time required will grow very fast when the database becomes bigger. As AFVSs need to be repeatedly fine tuned during development, the rise in the evaluation time will prolong the developing cycles and thus increase the cost. Also, when comparing the reliability of two AFVSs, if different databases are used, the conclusion can be essentially unfair. To solve these problems, we propose an evaluation model for AFVSs. The model can be used to predict the reliability of AFVSs as well as compare different AFVSs on a fair basis.

Actually, the accuracy of an AFVS depends on the system properties as well as the inter-class variation of fingerprints, or fingerprint individuality. Fingerprint individuality study can be traced back to more than 100 years ago [2]. From then on, most related studies have focused on minutiae based representations [1, 3, 4], among which Pankanti's model [1] has been regarded as a very simple but effective one for solving fingerprint individuality problems. This model will serve as the basis for building our AFVS evaluation model. The objective of Pankanti's model is to quantify the amount of available minutiae information to establish a correspondence between TWO fingerprints.

The rest of this paper is organized as follows. Section 2 defines some necessary symbols and terminologies. Section 3 describes the idea of our fingerprint individuality model. Section 4 gives a formal presentation of our AFVS evaluation model. Experiments are reported in Section 5 in which a real AFVS system is used to test the validity of our model. The last section is a conclusion of our work.

2 Symbols and Terminologies

The following symbols and terminologies are adopted through the rest of this paper.

Genuine minutiae: The minutiae manually (carefully) extracted by a fingerprint expert from a fingerprint image of enough image quality; *False minutiae*: Any extracted minutiae which are not genuine minutiae; *Matching score*: Number of minutiae correspondences between a master template and a live template; *Genuine matching*: The matching templates are from the same finger tip; *Imposter matching*: The matching templates are from different finger tips; *Genuine matching score*: The score of a genuine matching; *Imposter matching score*: The score of an imposter matching; *Genuine minutiae correspondence*: A declared correspondence between a genuine minutia and its counterpart; *False minutiae correspondence*: A declared minutiae correspondence which is not a genuine minutiae correspondence.

t : Matching score; $FAR(t)(FRR(t))$: False acceptance(rejection) rate; $G(t)$: Discrete Probability Density Function (PDF) of genuine matching score; $I(t)$: Discrete PDF of imposter matching score; EER : Equal error rate; $HG(x, M, K, N)$: PDF of hypergeometric distribution: $C_K^x C_{M-K}^{N-x} / C_M^N$; $b(x, n, p)$: Binomial distribution PDF: $C_n^x p^x (1-p)^{n-x}$; $chi2cdf(x, \gamma)$: Cumulative Density Function (CDF) of χ^2 distribution, where γ is the degrees of freedom; $poiss(x, \lambda)$: PDF of Poisson distribution: $\lambda^x e^{-\lambda} / x!$; $round(x)$: The integer closest to x ; $erf(x)$: Error function for Gaussian integration: $2/\sqrt{\pi} \int_0^x e^{-t^2} dt$; $N(x, \mu, \sigma)$: Normal distribution PDF: $\exp(-(x-\mu)^2 / 2\sigma^2) / \sigma\sqrt{2\pi}$.

3 Minutiae-Based Fingerprint Individuality Model

The following are the assumptions of our fingerprint individuality model.

A1) Only ridge terminations & bifurcations are considered; A2) Only locations & directions are considered for minutiae correspondence; A3) 2D fingerprint minutiae patterns follow the Complete Spatial Randomness (CSR) [5]; A4) Ridges have equal widths; A5) There is one and only one correct alignment between a master and a live template; A6) The minutiae correspondences are independent events and are equally important; A7) Only positive evidence from a minutiae correspondence is considered; A8) In an imposter matching, the minutiae direction difference between two minutiae matched in spatial position approximately follows the following distribution (PDF):

$$p_{\theta}(x) = 2/3 \left(N(x, 0, 17\sqrt{2}) + N(180-x, 0, 17\sqrt{2}) \right) + 1/(3 \times 180) \quad , (0 \leq x \leq 180) \quad (1)$$

Our model differs from Pankanti's model in assumptions A3) and A8). Assumption A3) ensures that we can describe both the spatial minutiae distribution in one single

fingerprint as well as the distribution of minutiae number among many different fingerprints. Assumption A8) is a strict mathematical description of the minutiae direction distribution. The fp383 database [6] which contains 1149 fingerprint images from 383 user finger tips was used to test the validity of these two assumptions.

For assumption A3), the hypothesis of CSR asserts: (i) the number of events (points) in any planar region A with area $|A|$ follows a Poisson distribution with mean $\lambda |A|$; (ii) given n events x_i in a region A , the x_i are independent random samples from the uniform distribution on A [5]. The test of hypothesis (i) is quite straightforward. The minutiae templates of fp383 were extracted using an AFVS which can achieve more than 95% verification accuracy on fp383 [6]. For each fingerprint, a rectangle R was placed randomly inside its effective region. The Empirical Distribution Function (EDF) of the minutiae number inside R was calculated. This EDF was then compared to a Poisson distribution with mean $\lambda |R|$, where λ was set to $54/65536\text{pixel}^2$, the average minutiae density of fp383. $|R|$ varies from 2304 to 9216pixel^2 . CSR hypothesis (i) is strongly supported by the test results. Fig. 1 shows one typical case.

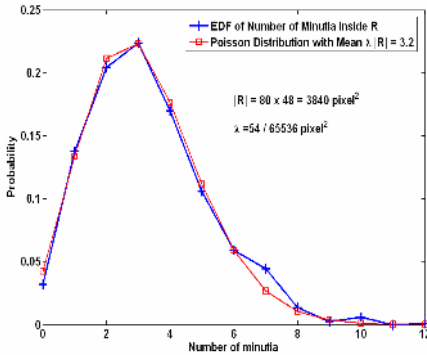


Fig. 1. Minutiae number distribution

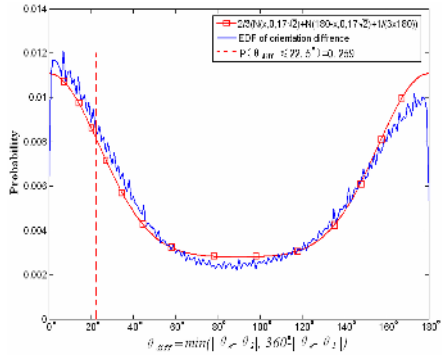


Fig. 2. Minutiae direction differences distribution

The “nearest neighbor distances” method [5] was used to test CSR hypothesis (ii). Minutiae of 40 fingerprints from fp383 were manually marked. Nearest neighbor distance test was then applied to them. Experimental results reveal that 39 fingerprints can pass the test. Boundary effect seems to be the main reason for the only fail case. In any event, for most of the test cases, uniform distribution is confirmed.

Assumption A8) is actually based on the empirical observation that minutiae directions are NOT uniformly distributed [1]. We further observe that in certain areas (~2/3) of the fingerprints, minutiae directions tend to cluster, while uniform distribution dominates in other areas (~1/3). Let θ_m denotes the direction of a master template minutia and θ_l denotes that of a live template minutia. The direction difference between these two minutiae is defined as $\min(|\theta_m - \theta_l|, 360^\circ - |\theta_m - \theta_l|)$ [1]. We calculated the EDF of the direction differences of minutiae pairs matched in position for imposter matching in fp383. Equation (1) is obtained by fitting the observation to the experimental result, as shown in Fig. 2. Although equation (1) is

based on the experiment on fp383 only, it seems to have considerable generality. In [1], Pankanti et al claim that the possibility that the direction difference is $\leq 22.5^\circ$ is 0.267 on their database, while equation (1) suggests $0.259 (\int_0^{22.5} p_\theta(x)dx)$.

4 A Minutiae-Based AFVS Evaluation Model

In this section we will apply our fingerprint individuality model to build an AFVS evaluation model with a capability of describing the characteristics of AFVSs as well as the intra-class variation of fingerprints. We will focus on modeling the three major components of a typical AFVS: fingerprint collector, minutia extractor and matcher. The following are the assumptions for our minutiae-based AFVS evaluation model.

- E1) The minutia extractor can extract minutiae in a fingerprint image, which has “enough” image quality, with the following parameters (registration & verification):
 - a) Missing a genuine minutia is an independent event with probability p_{miss}
 - b) The extracted false minutiae form a CSR pattern with density λ_{false}
 - c) For a genuine minutia, the extracted position follows a bivariate normal distribution with equal standard deviation σ_{pos} in both dimensions; and the extracted direction follows a normal distribution with standard deviation σ_{ori} . This assumption actually tolerates the possible fingerprint intra-class variation caused by distortion.
- E2) The master template covers all areas of the corresponding finger tip. In most AFVSs, a common mechanism for ensuring high reliability is to intentionally put more control on registration to make master templates’ information more complete.
- E3) The fingerprint collector can always capture fingerprint images with “enough” image quality; in the verification phase, the effective fingerprint area is $|S|$.
- E4) The genuine minutia density of the fingerprint set to be verified is λ .
- E5) The matcher declares a correspondence between a mater template minutia and a live template minutia if and only if the following three conditions are all fulfilled:
 - a) The Euclidean distance between these two minutiae is $\leq D$
 - b) The direction difference between these two minutiae is $\leq \theta_0$
 - c) No duplicated correspondence of one minutia is allowed.
- E6) The matching score equals to the number of minutiae correspondences.

Combining the fingerprint individuality model defined in Section 3, we can formulate $G(t), I(t), FRR(t)$ and $FAR(t)$ of our AFVS evaluation model. $I(t)$ is more related to the fingerprint individuality model. Considering assumptions E1a&b) and E4), we can see that the AFVS extracted minutiae patterns still comply with our fingerprint individuality model besides the overall minutiae density is equation (2).

$$\lambda_{ovr} = \lambda(1 - p_{miss}) + \lambda_{false}, \quad p_{match}(m, n, t) = \sum_{x=t}^{\min(m,n)} HG(x, |S|/2\omega D, m, n) \times b(t, x, l) \quad (2, 3)$$

Consider an imposter matching situation X in which m minutiae exist in a master template and n minutiae in a live template. According to [1], the probability that there are exactly t minutiae correspondences between these two templates is equation (3), where ω is the ridge period and $l = \int_0^{\theta_0} p_\theta(x)dx$. According to assumption A3) and E3), the probability of the occurrence of situation X can be expressed by equation (4). Combining equations (3) and (4), we can have equation (5).

$$p_{mn} = \text{poiss}(m, \lambda_{\text{ovr}}|S|) \times \text{poiss}(n, \lambda_{\text{ovr}}|S|), \quad I(t) = \sum_{m=t}^{+\infty} \sum_{n=t}^{+\infty} p_{mn} p_{\text{match}}(m, n, t) \quad (4, 5)$$

$G(t)$ is relatively more difficult since genuine and false minutiae coexist in the templates. We simply assume that false minutiae correspondences are declared after all genuine minutiae correspondences have been declared (*). Let $\{x_m, y_m, \theta_m\}, \{x_b, y_b, \theta_b\}$ denote the occurrences of a genuine minutia existing in the master and live template respectively. According to assumptions E1c) and properties of normal distribution, independent random variables $X=(x_m- x_l)$ and $Y=(y_m- y_l)$ both follow $N(x, 0, \sqrt{2}\sigma_{\text{pos}})$; and $\Theta=(\theta_m-\theta_l)$ follows $N(x, 0, \sqrt{2}\sigma_{\text{ori}})$. Let $Z=(x_m- x_l)^2+(y_m- y_l)^2$. It can be shown that $Z/2\sigma_{\text{pos}}^2$ follows a χ^2 distribution with the degrees of freedom 2. Thus, $\text{chi2cdf}(D^2/\sigma_{\text{pos}}^2, 2)$ is the probability that the Euclidean distance between these two minutiae is $\leq D$. Also, by applying the property of normal distribution to Θ , we get $P(\Theta \leq \theta_0) = \text{erf}(\theta_0/\sqrt{2}\sigma_{\text{ori}})$. Therefore the probability that these two minutiae match is

$$p_{\text{ggm}} = \text{chi2cdf}(D^2 / 2\sigma_{\text{pos}}^2, 2) \times \text{erf}(\theta_0 / \sqrt{2}\sigma_{\text{ori}}) \quad (6)$$

Consider a genuine matching situation X , in which the number of genuine minutiae in the effective fingerprint area is α . Assume there are m_g genuine minutiae and m_f false minutiae in the master template and there are n_g genuine minutiae and n_f false minutiae in the live template. Equation (7) represents the probability that there are exactly t_g genuine minutiae correspondences and t_f false minutiae correspondences.

$$P_{\text{pm}}(\alpha, m_g, n_g, m_f, n_f, t_g, t_f) = \left(\sum_{\varphi=t_g}^{\min(m_g, n_g)} HG(\varphi, \alpha, m_g, n_g) \times b(t_g, \varphi, p_{\text{ggm}}) \right) \times p_{\text{match}}(m_g + m_f - t_g, n_g + n_f - t_g, t_g) \quad (7)$$

The probability of the occurrence of situation X can be expressed as:

$$p_{\text{cann}} = \text{poiss}(\alpha, \lambda|S|) \times b(m_g, \alpha, (1 - p_{\text{miss}})) \times b(n_g, \alpha, (1 - p_{\text{miss}})) \times \text{poiss}(m_f, \lambda_{\text{false}}|S|) \times \text{poiss}(n_f, \lambda_{\text{false}}|S|) \quad (8)$$

Combining equations (7) and (8), we have

$$G(t) = \sum_{\alpha=0}^{+\infty} \sum_{m_g=0}^{\alpha} \sum_{n_g=0}^{\alpha} \sum_{m_f=t-m_g}^{+\infty} \sum_{n_f=t-n_g}^{+\infty} \sum_{t_g=0}^t p_{\text{cann}} \times p_{\text{pm}}(\alpha, m_g, n_g, m_f, n_f, t_g, t-t_g) \quad (9)$$

Equation (9) is prohibitively complicated. Simplification can be achieved by replacing the summations with mean values for some variables. The expectation of false minutiae number is $f = \text{round}(\lambda_{\text{false}}|S|)$. The mean value of the number of genuine minutiae is $g = \alpha(1 - p_{\text{miss}})^2$. By introducing these two mean values into (9), we have

$$\hat{G}(t) = \sum_{\alpha=\max(0, t-f)}^{+\infty} \sum_{\varphi=\max(0, t-f)}^{\alpha} \sum_{t_g=\max(0, t-f)}^{\min(t, k)} \text{poiss}(\alpha, \lambda|S|) \times b(\varphi, \alpha, (1 - p_{\text{miss}})^2) \times b(t_g, \varphi, p_{\text{ggm}}) \times p_{\text{match}}(g + f - t_g, g + f - t_g, t - t_g) \quad (10)$$

Three sets of numerical simulation were performed on equation (9) and (10) with different parameters. The biggest difference between the value of $G(t)$ and $\hat{G}(t)$ is

0.004. Therefore, we can conclude that equation (10) is an accurate approximation of equation (9) in case that the error tolerance is higher than 0.01.

$FAR(t)$ and $FRR(t)$ can then be directly deduced as (11) and (12). According to our AFVS evaluation model, the matching scores t can only take discrete values, so EER is defined as equation (13).

$$FAR(t) = 1 - \sum_{i=0}^{t-1} I(i), \quad FRR(t) = \sum_{i=0}^{t-1} G(i) \quad (11, 12)$$

$$EER = \{(FAR(t_0) + FRR(t_0)) / 2 \mid |FAR(t_0) - FRR(t_0)| = \min(|FAR(t) - FRR(t)|)\} \quad (13)$$

Equations (5), (9) ~ (13) depict the verification performance of an AFVS under our evaluation model. It is obvious that these equations are too complicated to be solved algebraically so that numerical simulations are used for all the experiments.

5 Experimental Results and Discussions

To test the validity of our model, the AFVS mentioned in Section 3 was used. The AFVS was first applied to fp383 to get the practical verification performances ($G_I(t)$, $I_I(t)$, $FAR_I(t)$, $FRR_I(t)$ and EER_I). Then, model parameters were evaluated for this AFVS and numerical simulation was performed to achieve its theoretical verification performance ($G'(t)$, $I'(t)$, $FAR'(t)$, $FRR'(t)$ and EER') under our evaluation model.

Kingston's estimation on the genuine minutiae density of $0.246\text{minutiae}/\text{mm}^2$ [1] was adopted here so that $\lambda = 51/65536\text{pixel}^2$. $\omega = 8.2\text{pixels}/\text{ridge}$ for 450 dpi images [1]. D and θ_0 were set to 20pixels and 22.5° respectively. Core points were used as reference points. During the matching process, only the minutiae whose distances from the core point lie between 80pixels and 16pixels were considered. This leads to $|S| = 19300\text{pixels}^2$. The automatic minutiae extraction results of 40 fingerprints were compared to their manually extracted templates which gives out $P_{miss} = 0.3$, and $\lambda_{false} = 18/65536\text{pixel}^2$. σ_{pos} and σ_{ori} were estimated by fitting $Z/2\sigma_{pos}^2$ to a χ^2 distribution and θ to a normal distribution respectively which leads to $\sigma_{pos} = 2.5$ and $\sigma_{ori} = 5.0$.

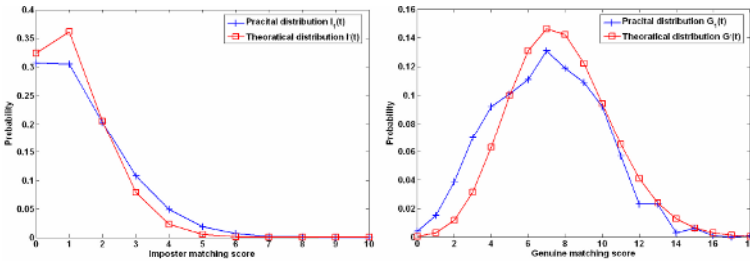


Fig. 3. Comparisons of theoretical and practical distributions of $G(t)$ and $I(t)$

Fig. 3 compares the practical and theoretical distribution of $I(t)$ and $G(t)$. There are mainly three reasons for the overestimation of $G(t)$: a) The core points of around 2.7% fingerprints in fp383 could not be consistently extracted [6]. Deviation in the reference point locations will surely degrade the genuine matching score. b) The

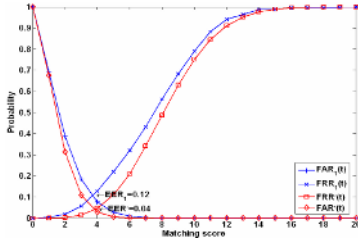


Fig. 4. Comparison of the ROC curves

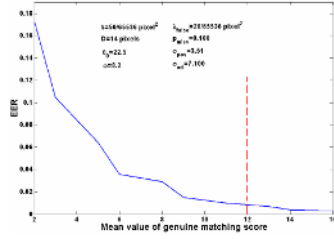


Fig. 5. EER values under different |S| values

overestimate of the effective fingerprint area as different fingerprints has different core point location. c) The assumption (*) made in section 4 is not always true.

The ROC curves are shown in Fig. 4. We can see that our model can predict the distribution of $I(t)$ and $G(t)$ satisfactorily. The overestimation of $G(t)$ which directly leads to an obvious underestimate of EER is probably caused by inconsistency between the model assumptions and the experimental settings as discussed above. In addition, the quaternion $\{p_{miss}, \lambda_{false}, \sigma_{pos}, \sigma_{ori}\}$ actually decides the intrinsic reliability of a extraction process, making it possible to separate the extractor and the matcher when evaluating an AFVS. Clearly, our model can help AFVS developers to improve their systems by analyzing how different parameters can affect the system reliability. Fig. 5 and Fig. 6 show the relationship between EER and $|S|$, D and θ_0 respectively. The conclusion made in [6] that “when $|S|$ is big enough, the increasing of $|S|$ will not lead to an obvious improvement in EER ” can be easily observed from Fig. 5. Fig. 6 shows that best system accuracy that can be achieve when $D \approx 3\sigma_{pos}$ and $\theta_0 \approx 3\sigma_{ori}$.

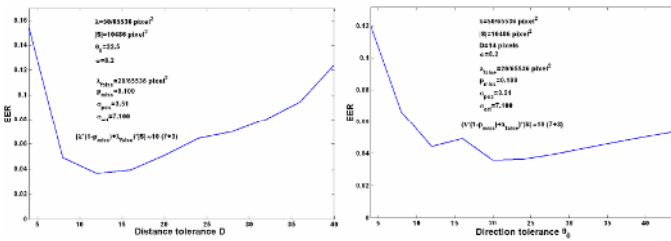


Fig. 6. The relationship between EER and distance/direction tolerance

6 Conclusion and Acknowledgement

We have proposed an evaluation model for minutiae-based AFVSs. We first adopt Pankanti’s model with some strengthening assumptions to describe the fingerprint individuality. Then we parameterize the three major components of an AFVS. Equations are then derived to describe the verification performance under the model assumptions. Experimental results show that our model can predict the distribution of the $G(t)$ and $I(t)$ of an AFVS satisfactorily. Furthermore, our model can serve as an assistant for AFVS developers to improve their system reliability since (a) our model

makes it possible to analyze different components in an AFVS separately; (b) how different model parameters will affect the system reliability can be used as a guidance for the developers to fine tune their systems.

This work was partially supported by the Hong Kong Research Grants Council Project 2300011, "Towards Multi-Modal Human-Computer Dialog Interactions with Minimally Intrusive Biometric Security Functions".

References

- [1] S. Pankanti, S. Prabhakar, A. K. Jain, On the Individuality of Fingerprints, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, pp. 1010-1025, vol. 24, no. 8, August 2002
- [2] F. Galton, *Finger Prints*, London: McMillan, 1892
- [3] M. Trauring, Automatic Comparison of Finger Ridge Patterns, *Nature*, pp. 938-940, 1963
- [4] D. A. Stoney, J. I. Thornton, A Critical Analysis of Quantitative Fingerprint Individuality Models, *J. Forensic Sciences*, pp. 1187-1216, vol.31, no. 4, October 1986
- [5] P. J. Diggle, *Statistical Analysis of Spatial Point Patterns*, Oxford University Press, 2003
- [6] K. C. Chan, Y. S. Moon, P. S. Cheng, Fast Fingerprint Verification Using Sub-regions of Fingerprint Images, *IEEE Trans. On Circuits and Systems for Video Technology*, pp. 95-101, vol. 14, issue 1, January 2004