Comparison of ROC and Likelihood Decision Methods in Automatic Fingerprint Verification

Sargur N. Srihari and Harish Srinivasan

{srihari,hs32}@cedar.buffalo.edu

Center of Excellence for Document Analysis and Recognition (CEDAR) Department of Computer Science and Engineering University at Buffalo, The State University of New York, Buffalo NY, USA

ABSTRACT

The biometric verification task is to determine whether or not an input and a template belong to the same individual. In the context of automatic fingerprint verification the task consists of three steps: feature extraction, where features (typically minutiae) are extracted from each fingerprint, scoring, where the degree of match between the two sets of features is determined, and decision, where the score is used to accept or reject the hypothesis that the input and template belong to the same individual. The paper focuses on the final decision step- which is a binary classification problem involving a single score variable. The commonly used decision method is to learn a score threshold from a labeled set of inputs and templates, by first determining the receiver operating characteristics (ROC) of the task. The ROC method works well when there is a well-registered fingerprint image. The paper shows that when there is uncertainty due to fingerprint quality, e.g., the input is a latent print or a partial print, the decision method can be improved by using the likelihood ratio of match/non-match. The likelihood ratio is obtained by modeling the distributions of same finger and different finger scores using parametric distributions. The parametric forms considered are Gaussian and Gamma distributions whose parameters are learnt from labeled training samples. The performances of the likelihood and ROC methods are compared for varying numbers of minutiae points available for verification. Using either Gaussian or Gamma parametric distributions, the likelihood method has a lower error rate than the ROC method when few minutiae points are available. Likelihood and ROC methods converge to the same accuracy as more minutiae points are available. Keywords: fingerprint verification, ROC, likelihood ratio, biometrics, forensics

1. INTRODUCTION

Fingerprints are an important class of biometrics for person identification. Person identification, regardless of the method, is ubiquitous in our daily lives. Two basic questions that often need to be answered are "who are you?" (identification) and "are you who you claim to be?" (verification). For example, one may have to be authenticated before gaining access to a bank, a protected site or to draw cash from an ATM. Biometric-based person identification is considered more reliable since the biological characteristics cannot be forgotten (like passwords) and cannot be easily shared, misplaced or duplicated. The characteristics could be a passive physical traits *e.g.*, fingerprints and hand geometry or active behavioral characteristics *e.g.*, voice and signature.

Fingerprints are a passive trait that are useful not only in biometrics but also in forensics where a latent fingerprint is lifted from a surface. The individuality of fingerprints have been studied from various viewpoints,¹ including the probability of random correspondence² and the discriminability of twins.³ These studies have dealt with well-registered fingerprint images where all features can be reliably extracted. In both biometric and forensic applications a perfect fingerprint image may not be available. This is of particular interest in forensics– where the issue of quality and quantity of the information available is of interest.^{4, 5} The amount of information available in a fingerprint may have an implication on the method of verification– which is explored in this paper.

Fingerprint verification can refer to two kinds of tasks.

1. 1:1 Verification: whether two fingerprints, one an input and other a template, are impressions of the same finger. Figure 1(a) shows an instance.



(a) 1:1 Verification.

(b) 1:N Verification.

Figure 1. Two types of verification: (a) 1:1, where a single template is compared to the input and (b) 1:N, where N enrolled templates are compared to the input.

2. 1:N Verification: Here, N impressions of a finger are enrolled (templates) and the question is whether or not an input impression belongs to the ensemble of enrolled impressions. Figure 1(b) shows an instance with N=4.

Verification accuracy for 1:N verification is usually higher than for 1:1 verification due to the presence of additional information. The focus of most of this paper is 1:1 verification with a final section on 1:N verification performance.

There is a very large literature on automatic fingerprint verification/identification [⁶]. The three main operational steps of automatic fingerprint verification are: (i) *feature extraction*— where a set of features, typically minutiae points, are determined from each fingerprint, (ii) *scoring*— where the degree of match between the two feature sets is determined, and (iii) *decision*— where the score is used to make the binary decision of whether the two fingerprints match. The robustness of the decision method with respect to the quality of the previous two steps is pertinent when one or both of the fingerprints do not contain a sufficient number of minutiae. This paper concerns the performance of two different decision methods when the numbers of available minutiae are varied.

2. FEATURES

Friction ridge features useful in fingerprint analysis are described in several books on the subject, including the early classics $[, 7, 8^{9}]$ and more recent ones $[, 10, 11^{4}]$. They are frequently characterized as belonging to one of three levels. Level 1 features provide a broad classification of fingerprints, e.g., whorls, loops, arches, double loops, etc. Level 2 features known as minutiae are the most useful in fingerprint comparison. Level 3 features correspond to pores, cuts and imperfections, which are not considered here.

In typical AFIS systems when a latent print is to be matched against a database two types of minutiae are extracted: ridge endings and ridge bifurcations. Both minutiae are represented as triples (x, y, θ) where x and y are the two-dimensional coordinates and θ is the angle made by a short line segment representing the minutiae. The direction of the line segment is determined by the direction of the ridge in the case of a ridge ending. In the case of a bifurcation the direction is determined by the bisector of the angle between the two bifurcating ridges.

Several algorithms and software for detecting minutiae in friction ridge images are available. A program available from NIST known as MINDTCT (pronounced "min-detect" for minutiae detector) takes as input a friction ridge image file, generates image maps, binarizes the image, detects minutiae, removes false minutiae, counts neighbor ridges, assesses minutiae quality and outputs a minutiae file. The MINDTCT program performed well in a recent competitive test (known as FPVTE) conducted for the US Visit program; the best was a product from NEC. An advantage of the MINDTCT program is that the source code is freely available and therefore modifiable for research purposes. Yet for the purpose of latent print examination it may be relevant that a program such as MINDTCT may miss detecting some minutiae and detect false minutiae. When many minutiae (30 to 40) are correctly located, the fact that a few are missing or that a few are false may be inconsequential.

The features are considered representative characteristics of the sample. In order to compare two samples and to quantify their similarity, a similarity measure or a distance measure is used to compute a score that signifies the strength of match between the features of the two samples. This similarity measure can be different depending upon which biometric and what features are being considered. The similarity measure converts the data from feature space to *distance* space. Features for fingerprints are characterized by minutiae. Data in distance space for fingerprints results out of comparing two sets of minutiae points. Each minutia is characterized by a triplet $\{x, y, \theta\}$, corresponding to the (x, y) coordinates of the minutia and its angular direction θ . The following procedure can be used to compute a distance measure between two fingerprint imprints. Let T be a template image with M minutiae and I be the input image with N minutiae:

$$T = \{m_1, m_2, \dots, m_M\} \quad \text{where} \quad m_i = \{x_i, y_i, \theta_i\}, \quad i = 1, \dots, M.$$
(1)

$$I = \{m_1', m_2', \dots, m_M'\} \quad \text{where} \quad m_i' = \{x_j', y_j', \theta_j'\}, \quad j = 1, \dots, N.$$
(2)

The distance between a minutia m_j' in I and a minutia m_i in T can be calculated using spatial distance (sd) and direction difference (dd)

$$sd(m_j\prime, m_i) = \sqrt{(x_j\prime - x_i)^2 + (y_j\prime - y_i)^2}$$
(3)

$$dd(m_j\prime, m_i) = min\left(\left| \theta_j\prime - \theta_i \right|, \ 360^\circ - \left| \theta_j\prime - \theta_i \right| \right)$$

$$\tag{4}$$

3. SCORING

The result of side-by-side comparison of two fingerprints is frequently expressed by a score. By using the score of matching, the feature space representation of two fingerprints is transformed into a distance space representation of both fingerprints. The score value characterizes the strength of match between two samples.

As an example, the scoring method known as Bozorth,¹² determines the degree of match between two sets of minutiae points. This software is generally available from NIST.¹² The *Bozorth* algorithm is designed to be rotation and translation invariant with respect to the images. Details of the implementation of *Bozorth* scoring are well-known and it suffices to say that the three main steps are (i) construct an intra-fingerprint minutiae comparison table, (ii) construct an inter-fingerprint compatibility table, and (iii) traverse the inter-fingerprint compatibility table.

Other scoring algorithms that use ridge alignment and matching, such as those discussed in,¹³ are more suited for the 1:N verification problem and when the quality of fingerprint considered is good. The *Bozorth* method is superior to that discussed in¹³ for the purpose of 1:1 verification on the FVC2002 dataset wherein the quality of the impressions is poor.

4. DECISION

The decision step is to determine from the score whether or not the two fingerprints belong to the same finger. Note that this is a task of converting a continuous-valued score into a discrete binary-valued decision. In some settings, e.g., biometrics, it is necessary to make such a hard decision whereas in other scenarios, e.g., forensics, it suffices to use scores to order a database so that a human can further evaluate the top scoring known fingerprints.

The commonly used decision method is to determine a threshold from the receiver operating characteristics (ROC) such that a score above/below the score indicates a match or a non-match.¹⁴ The ROC-based strategy learns from a large general population of ensemble of pairs of fingerprint samples (training set) and decides on an operating point by analyzing the ROC curve so that the error rate is minimized. Such a method works well when there is a complete and well-registered fingerprint image. On the other hand when there exists a partial imprint of a finger—as in the case of latent prints in forensics or due to limitations of the biometric device—other methods may be able to lower the error rate. In such situations it is useful to consider decision based on computing the likelihood ratio of match/non-match.

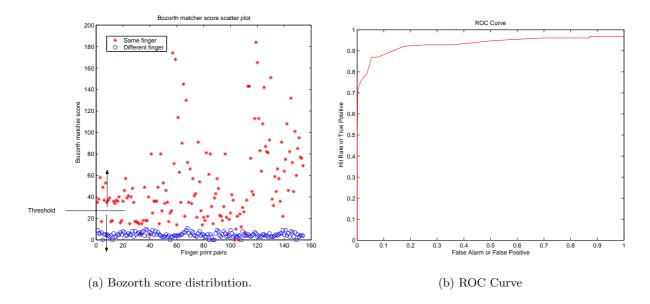


Figure 2. ROC from score distribution: (a) Distribution of Bozorth scores for a large ensemble of pairs of samples. Samples from different fingers generally score lower than those from the same finger, and (b) ROC obtained by moving a threshold in (a).

4.1. ROC-based method

The ROC-based method classifies a pair of fingerprints as belonging to the same finger or different, by learning from a large general population of ensemble of pairs. The pairs of sample either belong to the same finger or different finger. Features are extracted for each sample, and feature vectors between the two samples in a pair are compared using a scoring method. The result of such a matching is vector of distances, one value corresponding to every pair matched. The absolute value of scores help classify whether the two samples belong to the same finger. The ROC based method aims at deciding on a particular threshold in score value that separates pairs belonging to different fingers from those that belong to the same finger.

Scores are generally such that when the samples from a pair truly came from the same finger, the score is higher than if they come from different fingers. Figure 2(a) shows an example of score distribution obtained by comparing an ensemble of pairs. The optimal threshold is obtained by moving the threshold from left to right, and at every point calculate the number of pairs that were classified incorrectly. An incorrect classification can be of two types.

- 1. Pairs from different fingers classified as same(False Positive or False Alarm).
- 2. Pairs from the the same finger classified as different(False Negative).

In a similar way a correct classification can be of two types.

- 1. Pairs from the same finger classified as same(True Positive or Hit Rate).
- 2. Pairs from different fingers classified as different(True Negative).

As the threshold is moved from left to right on the distance scale, one can measure the four different variables mentioned above. A plot of hit rate(True Positive) against the false alarm (False Positive) is called the ROC curve. Figure 2(b) shows a typical ROC curve. The best threshold, or the operating point is decided by choosing the threshold that gives the least average error. The average error is defined to be the average between false

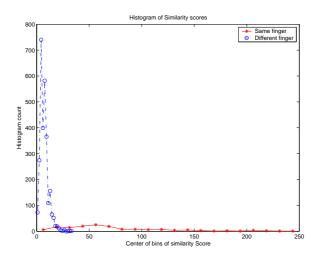


Figure 3. Histogram of scores of same and different finger pairs. Along the x-axis the centers of bins of the histogram are plotted. The y-axis is the count of how many scores fell in that bin.

negative and false positive. This definition of the average error gives equal importance to false negative and false positive. Thus the ROC method learns from a ensemble of pairs of training samples to obtain a single threshold value in score space that can be used to decide.

4.2. Likelihood ratio method

The probabilistic approach is to model the distribution of scores obtained by comparing the ensemble of pairs of fingerprint samples. This method of learning is similar to the ROC method to start with. Here again, a large ensemble of pairs of fingerprint samples are taken. This is termed as the training set. The set is divided into two categories. Set *one* consists of pairs of finger print samples where each pair truly belongs the same finger. Set *two* consists of a pairs of finger print samples where each pair truly belongs to different persons. Features are extracted for these samples and the Bozorth matcher computes a similarity score between the samples of each pair. Let $\vec{D_S}$ denote the vector of distances between all pairs in set *one*, which represents the distribution of distances when samples truly came from the same finger. Similarly let $\vec{D_D}$ denote the vector of distances between all pairs in set *two*, which represents the distribution of distances when samples are from different fingers. Figure 3 shows the histogram of these two different distributions.

These distributions can be modeled using parametric distributions such as Gaussian or Gamma whose probability density functions (pdfs) in one dimension are given in equation (5) and (6) respectively:

Gaussian pdf:
$$P(d|\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(d-\mu)^2}{2\sigma^2}}, \quad -\infty < d < \infty$$
 (5)

where μ and σ^2 are the mean and variance of the distribution, respectively; and

Gamma pdf:
$$P(d|\alpha, \theta) = \frac{1}{\Gamma(\alpha)\theta^{-\alpha}} e^{-\theta d} d^{\alpha-1}, \quad 0 < d < \infty$$
 (6)

where α and θ can be calculated using the mean μ and variance σ of the data by using $\mu = \alpha \theta$ and $\sigma^2 = \alpha \theta^2$. The Gamma distribution was chosen since the score is a positive value. On the other hand a Gaussian pdf assigns non-zero probabilities to negative scores. Also, since the Gamma contains the Gaussian distribution as a limiting form, it is potentially a better model.

The parameters that need to be learnt for such a model are typically derived from the sufficient statistics of the distribution, and are namely μ (mean) and σ (variance) for a Gaussian, or α (shape) and θ (width) for a Gamma. These distributions are the same- and different-finger distributions. Figure 4 shows the typical

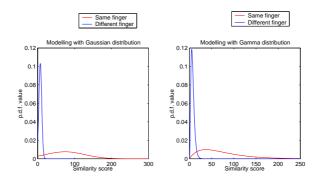
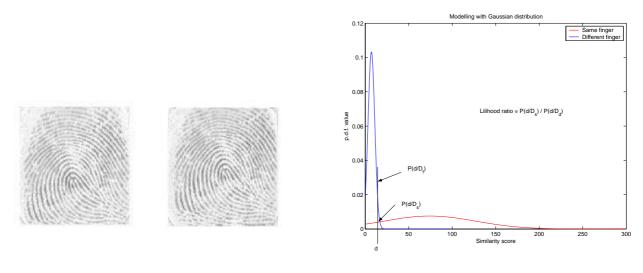


Figure 4. Two plots show modeling the histogram in figure 3 using a Gaussian(left) and a Gamma(right) distribution.

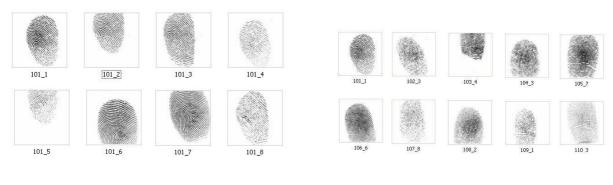


(a) Fingerprints to be compared. (b) PDF value of the Gaussian distributions at d

Figure 5. Likelihood Ration Method: (a) images with Score (Bozorth) d. The probability densities at d are evaluated to compute the likelihood ratio.

distribution p.d.f. curves obtained when the similarities are modeled using a Gaussian or a Gamma distribution.

The process of such 1:1 verification, or side-by-side comparison of two fingerprint images, starts with feature extraction and then computing the score d between them (see Figure 5(a)). The likelihood ratio defined as $\frac{P(D_S|d)}{P(D_D|d)}$ can be calculated, where $P(D_S|d)$ is the probability density function value under the D_S distribution at the distance d and $P(D_D|d)$ is the probability density function value under the D_D distribution at the distance d. If the likelihood ratio is greater than 1, then the classification answer is that the two samples do belong the same person and if the ratio is less than 1, they belong to different persons. Figure 5(b) shows how the likelihood ratio is obtained. If we wish to do 1: N verification, there are a total of N known samples from a person(enrolled), then for one input sample, N, 1:1 verifications can be performed and the likelihood ratios multiplied. In these circumstances it is convenient to deal with log likelihood-ratios rather than with just likelihood ratios. The log likelihood-ratio (LLR) is given by $\log P(D_S|d) - \log P(D_D|d)$. The decision of same-person is favored if $\log P(D_S|d) - \log P(D_D|d) > 0$, and the decision of different-person chosen if $\log P(D_S|d) - \log P(D_D|d) < 0$. When N of these 1:1 verifications are performed these LLR's are summed and then the decision is taken.



(a) Eight samples for one finger.

(b) One sample from each of ten different fingers.

Figure 6. (a) Samples from one finger and (b) Samples from different fingers. These samples are from database 1 of the FVC2002 dataset.

5. EXPERIMENTS AND RESULTS

A database of fingerprint images available from NIST known as the FVC2002 data set was used. Experiments with a total dataset of 80 fingerprint images and later on a dataset of 800 fingerprint images were performed. For the dataset containing 80 fingerprint images, a total of 10 fingers, each providing 8 samples. Figure 6(a) and 6(b) shows few sample images from the database. For the task of verification, a number of pairs of fingers can be considered. For a given finger that has 8 samples, there are $\binom{8}{2} = 28$ same finger pairs. Hence for all the 10 fingers together there are $28 \times 10 = 280$ same finger pairs. Since there are a total of $\binom{80}{2} = 3160$ pairs, the balance of 2880 are different finger pairs. For the purpose of learning, half of the set was used as the training set and the remaining for testing, i.e. 140 samples from same finger pairs and 1440 pairs from different fingers were used for the purpose of training. Both the ROC based method and the likelihood methods used the same data set for training and testing, in order for them to be fairly compared. The first set of experiments were based on using all the minutiae points available for each sample and the experiments and results are described in section 5.1. The second set of experiments were based on carrying out fingerprint verification with reduced number of minutiae points in each sample. These are described in section 5.2.

5.1. Experiments with all minutiae

For each of the 4 databases, the training set consisted of 140 pairs of same fingers and 1440 pairs of different fingers. Each of these pairs resulted in a similarity score as a result of matching the samples in the pair using the Bozorth matcher. The remaining pairs were used for testing the accuracy of the model. The average error rate, defined as average of False positives and False negatives, is used a measure of the error rate for the model. Figure 7 shows the ROC curves learnt from the training set of the 4 different databases. Figure 8 shows the result of learning the distribution of the scores from the training set and modeling them as Gaussian distributions. After deciding on the operating point, the thresholds were used to classify the test data set. Similarly the log likelihood ratio was used to determine the classification for the likelihood method. Table 1 gives the error rates on the 4 different datasets for both the ROC based method and the likelihood method using Gaussian and Gamma distributions. The classification decision boundary is given by the sign of the log likelihood-ratio, LLR, log $P(D_S|d) - \log P(D_D|d)$.

The results presented in Table 1 show that the likelihood methods, Gaussian and Gamma, outperformed the ROC method. The Gaussian model performs slightly better the Gamma model. Although the difference is statistically insignificant, it can be explained by the fact that the amongst the two models, the Gaussian density for for the $P(D_s/d)$ as described in section 4.2, has a fatter tail towards the origin, and it accounts for greater uncertainty in that region.

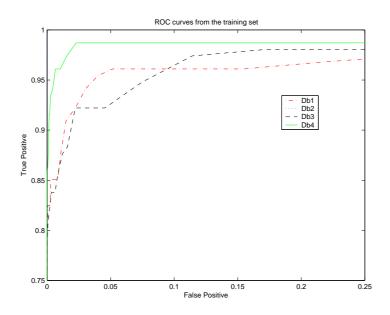


Figure 7. ROC curves obtained from the training set for each of the four databases.

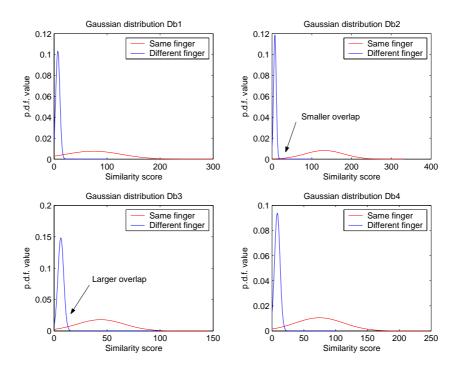


Figure 8. Distribution of similarity scores for each of four databases modeled with Gaussian distribution.

	ROC	Log-likelihood under		
Database	method	Gaussian	Gamma	
1	4.16%	3.93%	3.93%	
2	6.45%	2.57%	3.65%	
3	14.43%	13.96%	13.96%	
4	6.14%	6.14%	7.24%	
overall	7.79%	6.65	7.19%	

Table 1. Error rates for 1:1 fingerprint verification on the four different databases.

5.2. Experiments with reduced minutiae

The number of available minutiae in the two fingerprints directly affects the extent of match between two fingerprints. Partial fingerprints always account for fewer minutiae and we present an approach to simulate the reduced minutiae scenario, which are intended to simulate the effects arising from noisy image acquisition techniques. In this approach, minutiae points are randomly removed from the set of total available. An alternative set of experiments, to consider only minutiae points from a region as those available, was also performed but briefly discussed due to page limitations.

5.2.1. Randomly removing minutiae

In this experiment, the process of selecting the minutiae points to be removed from the sample was done at random. Since the process of removing the points were at random, the experiments discussed in this section were averaged over 5 different random seeds. The number of minutiae points removed from each sample was proportional to the total amount of minutiae points present in the sample. The 80 fingerprints in each of the four databases were examined by reducing the number of minutiae points in each gradually. When the number of minutiae points are reduced in the sample, the minutiae matches between a pair reduces significantly. As a result, there is more uncertainty in the classification. It is seen in our experiments that modeling uncertainty statistically proved better than the traditional ROC-based method. As the number of minutiae points were reduced, the experiments mentioned in section 5.1 were carried out and the average error rate of the two methods were compared with the average number of minutiae points per sample. 5 different experimental setups were created, each of which had on an average 5 minutiae points less than the previous setup. Figure 9 shows the average error rate for each of the databases varying with the average number of minutiae points present per sample. The size of the training and test set was fixed to be the same as discussed in the previous section. It is interesting to note from the figures that the likelihood methods perform better when there are fewer number of minutiae points available on certain databases. The average difference between the error rates between the methods decreased non-linearly as minutiae were gradually removed. If more minutiae were included, the difference between the error rates of the two methods became smaller. In all cases, modeling the distribution statistically was never worse than the ROC-based method.

The above experiment was performed on a database of all available **800 fingerprint images**. This gives a better idea about the performance of the statistical methods against the ROC based method. The results for the same are shown in Figure 11(a). The error rates and the standard error of the mean are shown in the corresponding Table 2.

5.2.2. Minutiae from a region

This experiment was based on the idea of simulating a partial fingerprint. One minutiae point was chosen at random and N nearest minutiae points to this chosen one (based on Euclidean distance on location) were said to be the available minutiae. Now when the first chosen minutiae point is at the edge of the fingerprint, then only minutiae closest to that will be available, simulating the effect of having only edge of the fingerprint available for matching. Again here, the experimental results were averaged over 5 different positions of the chosen minutiae to simulate different regions of partial fingerprint. Here again, the likelihood based methods performed consistently better than the ROC based methods. Also similar to the previous experiment the error rate increases non-linearly as the number of available minutiae points are reduced. The experimental results on the 800 fingerprint database is shown in Figure 11(b). The experiment was carried over 6 different positions of the chosen minutiae. The error rates and the standard error of the mean are shown in Table 2.

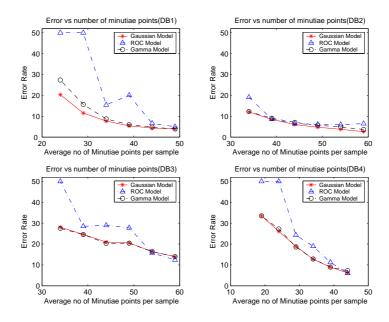


Figure 9. Variation in the error rates for 1:1 verification for each of the 4 databases with the average number of minutiae points available per sample (after random removal). The last data point in each of the database corresponds to the case where no minutiae points were removed.

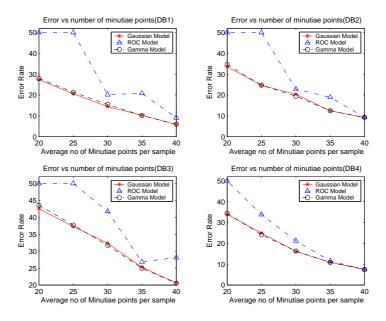
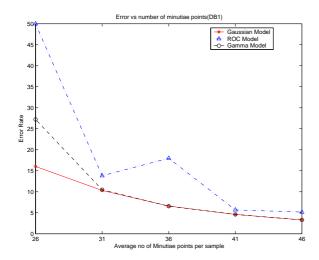
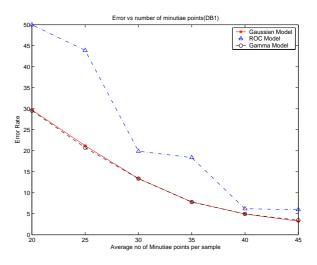


Figure 10. Variation in the error rates for 1:1 verification for each of the four databases with the average number of minutiae points available per sample (corresponding to a region).





(a) Variation in error rates for all 800 images after randomly removing minutiae points.

(b) Variation in error rates for for all 800 images after considering minutiae from a region.

Figure 11. Error rates for a database of 800 fingerprint images.

Randomly removed minutiae			Minutiae available in a region				
Available	ROC	Gamma	Gaussian	Available	ROC	Gamma	Gaussian
Minutiae	Error Rate	Error Rate	Error Rate	Minutiae	Error Rate	Error Rate	Error Rate
46	5.17 ± 0.02	$3.29 {\pm} 0.07$	$3.28 {\pm} 0.08$	45	$5.98 {\pm} 0.05$	$3.45 {\pm} 0.06$	$3.24{\pm}0.03$
41	$5.68 {\pm} 0.02$	$4.61 {\pm} 0.04$	$4.60 {\pm} 0.05$	40	$6.19{\pm}0.20$	$4.96 {\pm} 0.18$	$4.93 {\pm} 0.20$
36	17.97 ± 0.11	$6.56 {\pm} 0.11$	$6.59 {\pm} 0.09$	35	$18.36 {\pm} 0.19$	$7.77 {\pm} 0.16$	$7.82 {\pm} 0.15$
31	$13.81{\pm}1.10$	$10.48 {\pm} 0.19$	$10.31 {\pm} 0.21$	30	$19.88 {\pm} 0.79$	$13.36 {\pm} 0.13$	$13.37 {\pm} 0.20$
26	50.00 ± 0.00	27.22 ± 0.17	$16.02 {\pm} 0.24$	30	$19.88 {\pm} 0.79$	$13.36 {\pm} 0.13$	$13.37 {\pm} 0.20$
				20	$50.00 {\pm} 0.00$	$29.55 {\pm} 0.30$	$29.72 {\pm} 0.35$

 Table 2. Mean and standard deviation of error rates for all available 800 fingerprint images with randomly removed minutiae.

5.3. Experiments with 1:N Verification

The 1 : N Verification process consists of two steps (i) Enrollment: Here a known set of N impressions for a finger are enrolled and labeled so as to belong to that finger, and (ii) Authentification: The input impression is compared against all the N enrolled(templates), one by one to result in N different scores and the mean of the N score is taken to be the final score. Once again, the distribution of these scores can be used to model a Gaussian/Gamma distributions. The 80 fingerprint database consisting of 10 fingers, 8 impressions each was used as the data set and N = 4 samples were enrolled for each finger. The input fingerprint could be of same or different finger. All the remaining fingerprints that were not used for enrollment were used for testing and using the *Bozorth Matcher* score and all minutiae preserved a verification accuracy of **99.8**% was acheived on DB1 of FVC2002. The accuracy was consistent across all models(ROC, Gaussian, Gamma) since all minutiae were considered. Using the *Alignment matching* method discussed in¹³ along with a few modifications, the accuracy was found to be 98.5%. The presene of N templates instead of 1 makes the task of fingerprint verification that much easier, as indicated by the high verification accuracy rates. For the same DB1 FVC2002 dataset, in the 1 : 1 verification, an accuracy of 96.07% was achieved(reported in table 1 in section 5.1), whereas the 1 : N verification yielded 99.8%(reported above in this section).

6. CONCLUSIONS AND DISCUSSION

This paper has presented a comparison of alternative methods for the decision step of fingerprint comparison in a biometric setting. The standard ROC method of comparison was compared with a method based on likelihood ratios. The likelihood ratios of same and different finger scores were computed by modeling the score distributions parametrically. The parametric forms considered were the Gaussian and Gamma distributions.

When a large number of minutiae points are available in both the input and the template, the error rates of both the ROC and likelihood methods are similar. However when the number of available minutiae are reduced, the likelihood methods perform significantly better than ROC method. This suggests that the likelihood decision is especially useful in the regime of small numbers of available minutiae. It is to be noted that the results are with available numbers of minutiae rather than the number of matching minutiae which is expected to be a fraction of the available minutiae.

That the likelihood method has a lower error rate than the ROC is not surprising since it is a version of the minimum error rate Bayes classifier for the two-class case.¹⁴ Since the ROC uses a single threshold it is equivalent to a linear decision boundary. On the other hand the likelihood method allows for non-linear decision boundaries which amount to multiple decision points along the score line. Further decrease in the error rate can be expected by using more accurate models of score distributions, e.g., mixture models and more robust parametric distributions such as the Student-*t* distribution.

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

REFERENCES

- 1. D.A.Stoney. Measurement of fingerprint individuality. Advances in Fingerprint Technology, H. Lee and R.E. Gaensslen (eds.), 2001.
- 2. S. Srihari, H. Srinivasan, and C. Su. Generative model of fingerprint individuality using ridge information. CEDAR Technical Report, page 20 pages, 2008.
- S. Srihari, H. Srinivasan, and G. Fang. Discriminability of the fingerprints of twins. Journal of Forensic Identification, 58:109–127, 2008.
- 4. D.R.Ashbaugh. Quantitative-Qualitative Friction Ridge Analysis. CRC Press, Boca Raton, FL, 1999.
- C.Champod and P. A. Margot. Computer assisted analysis of minutiae occurrences in fingerprints. Proceedings Int. Symp. Fingerprint Detection and Identification, 1996.

- D. Maltoni, D. Maio, A.K. Jain, and S. Prabhakar. Handbook of Fingerprint Recognition. Springer, New York, 2003.
- 7. F.Galton. Finger Prints. McMillan, London, 1892.
- 8. E.R. Henry. Classification and Uses of Fingerprints. London: Routledge, 1900.
- 9. B.Wentworth and H. H. Wilder. Personal Identification. R. G. Badger, Boston.
- 10. A.A. Moenssens. Fingerprint Techniques. Chilton Book Company, Radnor, PA, 1971.
- 11. D.Z. Loesch. Quantitative Dermatoglyphics. Oxford University Press, New York, 1983.
- M. D. Garris, C. I. Watson, R. M. McCabe, and C. L. Wilson. User's guide to nist fingerprint image software (nfis). NISTIR 6813, National Institute of Standards and Technology, U.S. Department of Commerce: Gaithersburg, M. D., 2004.
- R Jain, A. Lin Hong Bolle. Online fingerprint verification. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 19:302–314, April 1997.
- 14. R.O. Duda, P.E. Hart, and R. Stork. Pattern Classification. Wiley, New York, 2001.