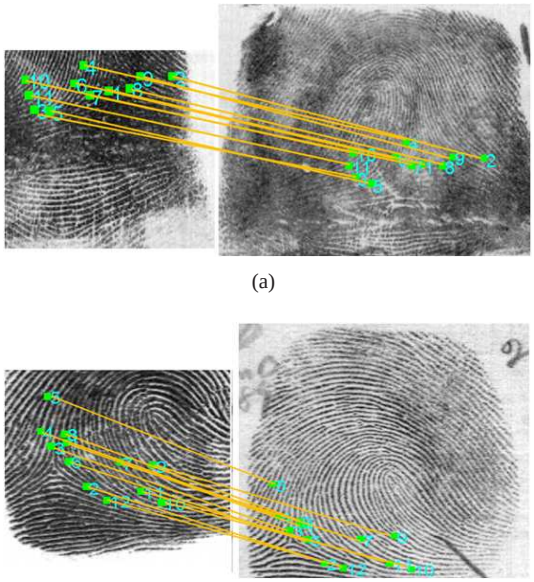
# On the Evidential Value of Fingerprints

Heeseung Choi, Abhishek Nagar and Anil K. Jain*
Dept. of Computer Science and Engineering
Michigan State University
East Lansing, MI, U.S.A.
{hschoi,nagarabh,jain}@cse.msu.edu

## Abstract

*Fingerprint evidence is routinely used by forensics and law enforcement agencies worldwide to apprehend and convict criminals, a practice in use for over 100 years. The use of fingerprints has been accepted as an infallible proof of identity based on two premises: (i) permanence or persistence, and (ii) uniqueness or individuality. However, in the absence of any theoretical results that establish the uniqueness or individuality of fingerprints, the use of fingerprints in various court proceedings is being questioned. This has raised awareness in the forensics community about the need to quantify the evidential value of fingerprint matching. A few studies that have studied this problem estimate this evidential value in one of two ways: (i) feature modeling, where a statistical (generative) model for fingerprint features, primarily minutiae, is developed which is then used to estimate the matching error and (ii) match score modeling, where a set of match scores obtained over a database is used to estimate the matching error rates. Our focus here is on match score modeling and we develop metrics to evaluate the effectiveness and reliability of the proposed evidential measure. Compared to previous approaches, the proposed measure allows explicit utilization of prior odds. Further, we also incorporate fingerprint image quality to improve the reliability of the estimated evidential value.*



(a)

(b)

Figure 1. Probabilities of non-match for (a) mated and (b) non-mated pair with differing image quality values. Both the images in (a) have the lowest quality as measured by NFIQ (NFIQ = 5) [23] whereas the two images in (b) have the best quality (NFIQ = 1). While both these pairs of fingerprints have the same matching score (7 as computed by the Verifinger matcher [14]), the non-match probabilities, after incorporating the quality value, are 0.57 and 0.90 for (a) and (b), respectively.

## 1. Introduction

The evidential value of fingerprints is a term that refers to the *value* of fingerprints as a means of person identification [10]. As stated by Galton [10], a measure of evidential value is *the probability that two fingerprints under consideration are obtained from two different persons*. Fig. 1 shows two pairs of fingerprint images where the pair in Fig. 1(a) comes from the same finger and the pair in Fig. 1(b) is from two different fingers. While both these pairs of fingerprint images have the same matching score as computed by an AFIS, they have different image quality values. This difference in their quality suggests that these two pairs be treated differently for assigning probabilities of belonging to different fingers. We call this the *probability of non-match*. The objective of this paper is to present an approach for computing the non-match probability and ways it can be evaluated.

A measure of fingerprint evidential value is crucial to arrive at a correct and conclusive forensic analysis of fingerprints. The importance of this was established by the

*A.K. Jain is also with the Dept. of Brain and Cognitive Engineering, Korea University, Anam-dong, Seongbuk-gu, Seoul 136-713, Korea.

Daubert's rule of evidence, or the Daubert standard, set by the United States Supreme Court in *Daubert v. Merrell Dow Pharmaceuticals* [1]. One of the major requirements in Daubert standard is that the admissibility of forensic evidence in court proceedings is contingent upon the availability of the error rates of the forensic evidence.

The prevalent procedure for manually matching fingerprints, which typically involves matching a latent fingerprint to a full fingerprint[1], follows a rigorous protocol, known as the ACE-V protocol. However, the lack of any thorough study measuring the error rates of the protocol has led to challenges to fingerprint evidence in a number of court cases [2]. The need for a reliable study of error rates is further compounded by the imprecise specification of the protocol as well as some inconsistencies among the implementation of the protocol [12]. Moreover, the outcomes of latent matches that are inconsequential are seldom recorded by the forensic agencies. Finally, a thorough statistical modeling of the matching process is confounded due to lack of extensive data and somewhat subjective nature of matching by latent examiners.

There are two possible ways to circumvent this dilemma [11]. In the first approach, a fingerprint related testimony based on the experience of the testifier is permitted without stringent requirements on its scientific validity. The second approach is to use an automatic fingerprint identification system (AFIS) whose error rates can be easily estimated. The disadvantage of the first approach is an obvious lack of scientific validity whereas the second approach is limited due to a generally lower discriminative power of state of the art AFIS on latent matching compared to the ability of an experienced latent examiner. We believe it is indeed appropriate to adopt the second approach since the performance of the state of the art matchers is steadily improving and is approaching the ability of latent examiners [18]. Note that latent matching based on the ACE-V protocol is not infallible either, as illustrated by the infamous case involving Brandon Mayfield, who was wrongly accused and incarcerated based on latent fingerprint evidence in the Madrid bombing case [5].

There are two main approaches that have been followed in the literature to obtain the evidential value of fingerprints[2]: (i) feature modeling, and (ii) similarity score modeling. In feature modeling approach, a statistical model for the generative process of a set of fingerprint features (typically minutiae) is obtained and the evidential values for fingerprint pairs are obtained based on this model. In the similarity score modeling approach, the distribution of similarity scores between fingerprint pairs is directly modeled, separately for the mate and non-mate pairs. A measure of the evidential value can then be obtained from these two distributions.

Both the modeling approaches have their own potential strengths and limitations. In the case of feature modeling, the evidential value computed for a particular configuration of the two fingerprints being matched is expected to be reliable even for the case when only a limited number of relevant fingerprint pairs are available in the training database to estimate the value. This is because of prior knowledge used in modeling the fingerprint features. The main drawback of feature modeling, however, is that it is difficult to model a variety of fingerprint features that are typically used by state of the art matchers as well as latent examiners. This means that the feature modeling studies could use only relatively simple and often outdated features and matchers (typically minutiae location and angle) for model construction. Note that the practice of using the "12 point rule" (which states that if there are at least 12 matching minutiae between two fingerprints, they are declared as a match) was abandoned by the FBI in the early 1970s [3].

The similarity score modeling, on the other hand, allows the use of any state of the art matcher. One of the main drawbacks of similarity score modeling is that in order to estimate a tight confidence interval on the evidential value, one needs to have a large training set of fingerprint pairs (both mated and non-mated).

In this paper we follow the similarity score modeling approach to compute the evidential value of fingerprints in the form of non-match probability (NMP). We consider three different approaches to estimate the required genuine and impostor score distributions[3]: (i) histogram construction, (ii) kernel density estimation and (iii) parametric density estimation. Based on a measure of reliability of the evidential value obtained through cross-validation, we show that the non-match probability computed using the kernel density estimation performs the best.

It is well known that one factor that affects the fingerprint matching performance is the image quality [17]. To study how image quality affects the evidential value, we partition the fingerprint database into two segments based on fingerprint image quality. Based on the computed evidential values, we show that its variance is significantly reduced if the image quality is taken into account. Moreover, good quality images lead to a greater polarization of the evidential values. That is, among the good quality images the evidential values are more likely to be closer to the extremes $[0, 1]$ indicating impostor or genuine matches.

The rest of the paper is organized as follows. Section

[1]Full print to full print matching can be done effectively in "lights out" mode by AFIS unless the image quality is poor [17].

[2]We use the terms "evidential value" and "a measure for evidential value" interchangeably depending on the context.

[3]Genuine scores correspond to the match scores obtained from a pair of fingerprints when the two fingerprints belong to the same finger (mates). The impostor scores correspond to the case when the two fingerprints are obtained from different fingers (non-mates).

2 presents a summary of previous approaches to estimating fingerprint evidential value. Section 3 presents our proposed measure and methods for analyzing the evidential value and Section 4 presents experimental results. Summary and discussions are presented in Section 5.

## 2. Background

The first attempt to estimate the individuality of fingerprints was made by Galton in 1892 who proposed a statistical model of fingerprint features. His model required partitioning a fingerprint into 24 non-overlapping square regions whose width was equal to six times the inter ridge distance [10]. He argued that each of these square regions can be correctly reconstructed with a probability of $\frac{1}{2}$ if the information regarding the surrounding ridges is known. This leads to a probability of $\left(\frac{1}{2}\right)^{24}$ that the complete fingerprint can be reconstructed, given the ridge structure in the region surrounding the squares. Galton further noted that the probability that the correct number of ridges entering and exiting the 24 squares is $\frac{1}{256}$ and that the probability of occurrence of specific type of fingerprint (e.g. whorl, loop, arch, etc.) is $\frac{1}{16}$. This set of assumptions lead to a probability of $\frac{1}{256} \times \frac{1}{16} \times \left(\frac{1}{2}\right)^{24} = 1.45 \times 10^{-11}$ for correctly reconstructing a full fingerprint. This measure of fingerprint individuality is usually referred to as the *Probability of Fingerprint Configuration (PFC)* [10]. A discussion on a number of other related models is available in [19]. One of the limitations of the PFC is that it does not take into account the differences in the characteristics of a fingerprint such as the number of minutiae in the fingerprint that is being reconstructed.

The first rigorous and comprehensive study on feature modeling was proposed by Pankanti et al. [16]. Given a pair of fingerprints with a similarity value s, the *Probability of Random Correspondence (PRC)* refers to the probability that two randomly selected fingerprints will have similarity value the same as s. More formally, PRC is defined as

$$PRC = P(s|I) \tag{1}$$

where I refers to the impostor pair of fingerprints, i.e. they belong to different fingers.

Pankanti et al.'s approach was limited to modeling minutiae locations and directions to calculate the PRC value. They assumed a uniform distribution for minutiae location and direction. Given a query fingerprint containing n minutiae, they computed the PRC that an arbitrary template fingerprint containing m minutiae will have exactly q mated minutiae with the query as

$$PRC = P(s|I, m, n). \tag{2}$$

One limitation of this model is its relatively poor fit of uniform distribution to real minutiae distribution. Chen and Moon [6] extended this model by using von-Mises distribution to model the minutiae direction.

Figure 2. Two fingerprints from NIST SD14 belonging to the same finger that have different characteristics due to skin distortion during image acquisition. This indicates the intra-class variability in fingerprints which is one of the major obstacles in defining quantitative measures of evidential value.

Based on the observation that the minutiae tend to form clusters [20], Zhu et al. [25] used finite mixture models for modeling the distribution of minutiae. For each fingerprint, a Gaussian distribution was fit to the minutiae locations and a von-Mises distribution was used for fitting the minutiae directions in each component of mixtures. This led to a supposedly more realistic estimate of the PRC. Fang et al. [9] and Su et al. [21] extended this framework by incorporating information regarding the ridges. Chen and Jain [7] incorporated three different types of fingerprint features: level 1 (pattern type), level 2 (ridges and minutiae) and level 3 (pores) features. Su et al. [22] incorporated dependence between the neighboring minutiae using Bayesian networks. There are two main limitations of these studies: i) the matching criteria used in these techniques i.e. the number of matching minutiae, is very rudimentary and significantly biases the evidential value, and ii) intra-class variation (reflecting the variations in the matching scores between multiple impressions of the same finger) is not explicitly considered in the formulation of PRC.

The match score based approaches for computing the evidential value have also been reported in the literature. Meagher et al. [13] utilized the FBI's Integrated Automated Fingerprint Identification System (IAFIS) to compute the evidential value of fingerprints. They simulated latent images by cropping each of the $50,000$ rolled fingerprints and compared them with the original rolled images to obtain the genuine and impostor match scores. Assuming that the genuine and impostor distributions follow a Gaussian distribution, Meagher et al. estimated the probability of false correspondence, i.e the probability of finding an exact match between two unrelated fingerprints to be equal to $10^{-97}$. One major shortcoming of this study is that the intra-class variation (see Fig. 2) is not accounted for since only one image per finger was utilized; the genuine scores were computed by matching a cropped fingerprint with the full fingerprint

from which it was cropped.

Neumann et al. [15], [14] developed a fingerprint matching procedure based on different configurations of minutiae and converted the resulting similarity value $s$ into a *likelihood ratio (LR)*. The likelihood ratio is proposed as a measure of evidential value of fingerprints defined as

$$LR = \frac{P(s|G)}{P(s|I)} \tag{3}$$

where $I$ refers to impostor fingerprint pairs (non-match pairs) and $G$ refers to genuine fingerprint pairs (true-match pairs). Egli et al. [8] also used match scores acquired from an automatic fingerprint matcher to compute the evidential value in the form of likelihood ratio. The main difference among the various LR approaches is the method used for estimating the genuine and impostor densities, namely $P(s|G)$ and $P(s|I)$: in [15] kernel density estimation is used to estimate the densities whereas in [14] a mixture of Gaussian is used. Egli et al. used two different parametric distributions for fitting genuine and impostor score distributions.

## 3. Proposed Measure

In this section, we propose a new measure of fingerprint evidential value, namely the non-match probability (NMP), which overcomes certain limitations of the existing measures such as PRC and LR. One of the drawbacks of the PRC is that it does not explicitly consider the probability that the two fingerprints being matched can come from the same finger, namely the genuine match probability[4]. This affects the validity of conclusions derived as a result of the evidential analysis. While the likelihood ratio does explicitly consider the probability that the two fingerprints being considered belong to the same finger, it does not directly answer the question first posed by Galton, namely, *the probability that two fingerprints under consideration are obtained from two different persons*, whose answer is needed to elicit the evidential value of fingerprints. Furthermore, according to Taroni et al. [24], computation of likelihood ratio is just a means to obtain the probability that the suspected fingerprint is the true match which is essentially 1-NMP or the probability of a match.

The proposed measure i.e. the non-match probability (NMP) for a similarity value (match score between a pair of fingerprints) $s$ is given by

$$NMP = P(I|s) = 1 - P(G|s) \tag{4}$$

[4]Note that, in Pankanti et al. [16], the hypothesis that two fingerprints come from same fingerprint is not rejected simply based on the fact that the two fingerprints being compared have different number of minutiae. This is one way they consider intra-user variation. The tolerance used in minutiae match also implicitly accounts for intra-user variation to some extent.

where $P(G|s)$, the posterior probability of a genuine match given a score $s$, is given as

$$P(G|s) = \frac{P(s|G)P(G)}{P(s|I)P(I) + P(s|G)P(G)}. \tag{5}$$

Here $P(s|G)$ and $P(s|I)$ denote the genuine and impostor distributions of match score $s$ and $P(I)$ and $P(G)$ denote the prior probability of an impostor or genuine match. Note that unlike previous approaches, both the genuine and impostor prior distributions are used in addition to the corresponding likelihoods in computing the NMP. Moreover, unlike the likelihood ratio values that range in $[0, \infty]$, NMP values range in $[0, 1]$, where a value close to $0$ indicates that the two fingerprints being matched are very likely to be a genuine pair and a value close to $1$ indicates that they are very likely to be an impostor pair.

### 3.1. NMP Computation

Given a large training set of fingerprint pairs that has both genuine and impostor matches, the non-match probability associated with similarity value $s$ is the fraction of pairs that are non-matches among all the pairs with $s$ as their matching value. The prior distribution can be considered as the overall proportion of genuine and impostor matches in the training set; it can also reflect any additional evidence available regarding the fingerprints in consideration being mated or not. This approach for computing NMP is effective only if a very large fingerprint database is available. In the absence of such a large database, robust techniques to estimate the relevant probability densities need to be employed.

The non-match probability is computed as:

$$NMP = P(I|s) = \frac{P(s|I)P(I)}{P(s|I)P(I) + P(s|G)P(G)}. \tag{6}$$

Here $P(s|I)$ and $P(s|G)$ can be computed from the estimated distributions. The values of priors $P(I)$ and $P(G)$ reflect additional evidence that may be available. This is one of the strengths of the proposed NMP measure as it helps to utilize any available asymmetric information towards the claim of genuine or impostor match [24]. Note that LR and PRC do not have this capability.

It is indeed possible to compute the NMP value from the PRC and LR values using the following relationships:

$$NMP = P(I|s) = \frac{P(s|I)P(I)}{P(s)} = \frac{PRC \times P(I)}{P(s)} \tag{7}$$

$$NMP = P(I|s) = \frac{1}{1 + \frac{P(s|G)P(G)}{P(s|I)P(I)}} = \frac{1}{1 + LR\frac{P(G)}{P(I)}} \tag{8}$$

Note that the above expressions require estimates of $P(s)$, $P(G)$, and $P(I)$ that in turn require some knowledge of both the genuine and impostor distribution.

Figure 3. Four different subimages from the same rolled fingerprint in the NIST SD14.

## 4. Experimental Results

There are two main criteria to evaluate a measure of the evidential value of a fingerprint match. The first is the error rate associated with the term used to compute the evidential value. The second is the confidence (variance) in the estimated evidential value. Since the error rates are dependent on the specific matcher used for computing similarity values, in this paper, we mainly focus on confidence in the evidential value as a measure of its goodness. The confidence of the evidential value is estimated using p-fold cross-validation. We also incorporate fingerprint image quality in the estimation of evidential value to enhance the confidence of the evidential value. Such an analysis is not available in the literature.

### 4.1. Protocol

We used the NIST SD14 fingerprint database in our experiments which contains two rolled impressions for each of the $27,000$ different fingers. Since the analysis of evidential value is mainly required in case of latent fingerprints[5], we simulated latent fingerprints from this database by randomly cropping four different subimages of size $400 \times 400$ from the original fingerprint (see Fig. 3). These simulated latent fingerprints were matched with the full fingerprints (not used in cropping) to obtain a set of $108,000$ genuine scores. For impostor scores, we randomly selected $500$ cropped images and matched them with randomly selected non-mated full fingerprints to obtain $1$ million impostor scores. Fingerprint feature extraction and matching were performed using Neurotechnology Verifinger software [4] which outputs match scores in the range $[0, 990]$.

We considered three different ways to estimate the non-match probability, each differing in the way the probability density of the genuine and impostors match scores is computed: (i) histogram based estimate, (ii) kernel density based estimate, and (iii) parametric density based estimate. In this analysis, we use equal prior probability for impostor and genuine pairs, i.e. ($P(I) = P(G)$), for an easy interpretation of NMP values.

In the case of histogram based technique, the histograms of genuine and impostor scores are separately normalized so that the sum of their respective bin values is one. The non-match probability (NMP) at each score bin can be calculated as:

$$NMP = \frac{N_I^*(s)}{N_G^*(s) + N_I^*(s)} \tag{9}$$

where $N_G^*(s)$ and $N_I^*(s)$ are the normalized bin values at score $s$ for genuine and impostor matches, respectively. In the case of kernel density estimation, a Gaussian kernel with a bandwidth of $1.5$ was used to estimate the genuine and impostor distributions. In the case of parametric density estimation, Weibull distribution was used to model the genuine match scores and log-normal distribution was used to model the impostor match scores. The choice of these parametric distributions follows [8].

[5]NIST SD27 is the only public domain latent fingerprint database available. It contains 258 latents and their mated rolled impressions.

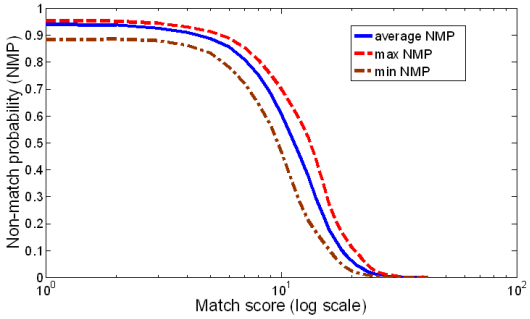



Figure 4. Non-match probabilities based on histogram density estimates. The solid line shows the mean values of NMP whereas the bounding dotted lines show the minimum and maximum values of NMP. The average variance of NMP is $0.025$.

### 4.2. Reliability of Evidential Value

The first set of experiments was conducted to determine the variation in the NMP values obtained using the three density estimators across multiple partitions of the dataset. Note that an understanding of this variability will be useful in providing a range for the most likely NMP values for a pair of fingerprints being matched.

Fig. 4 shows the NMP curves obtained using the histogram based density estimates for $10$ random partitions of the NIST SD14 database. These curves correspond to the mean, the minimum and the maximum values of the non-match probabilities over the $10$ data partitions. Figs. 5 and 6 show the corresponding curves for the kernel density based and parametric density based estimates, respectively. The average variances (averaged over different score values) of the estimates corresponding to histogram, kernel density and parametric estimates of score distributions are $0.025$, $0.025$, and $0.027$, respectively.

Considering the mean NMP values obtained using the histogram based density estimates as the ground truth, we also computed the bias in the kernel density based as well as the parametric density based NMP estimates. The average absolute difference between the histogram based NMP

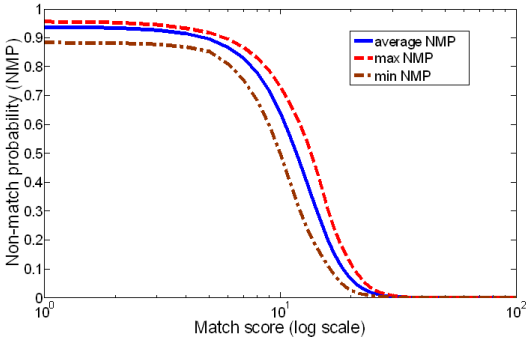



Figure 5. Non-match probabilities based on kernel density estimation. The solid line shows the mean values of NMP whereas the bounding dotted lines show the minimum and maximum values of NMP. The average variance of NMP is $0.025$.

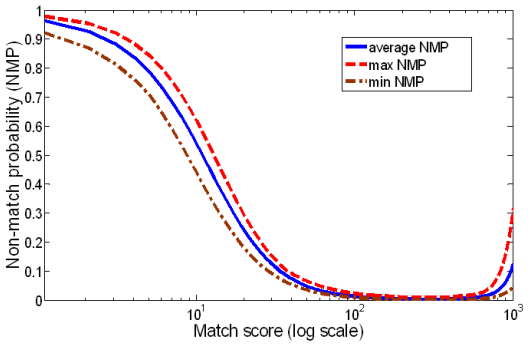



Figure 6. Non-match probabilities based on parametric density estimates. The solid line shows the mean values for of NMP whereas the bounding dotted lines show the minimum and maximum values of NMP. The corresponding average variance of NMP is $0.027$.

and kernel density based NMP is $0.006$ whereas that between histogram based NMP and parametric density based NMP is $0.106$. This shows that the parametric density based estimate has a significant bias compared to kernel density based estimates. Note that kernel density based computation of NMP is more desirable than histogram based value since it (in fact, parametric density based estimates as well) inherently extrapolates the NMP for the scores at which no matching pairs were observed in the reference database.

### 4.3. Quality-based Evidential Value

It is well known that the performance of fingerprint matchers (as well as that of latent examiners) depends on the fingerprint image quality. As such, the non-match probability should depend on the quality of the fingerprint images in addition to the match scores. To investigate this, we divide the genuine and impostor match scores based on the quality of the associated fingerprint pairs. We use the NFIQ fingerprint quality measure developed by the National Institute of Standards and Technology (NIST) [23] to calculate the fingerprint image quality. The NFIQ measure assigns one of five quality levels (excellent, very good, good, fair

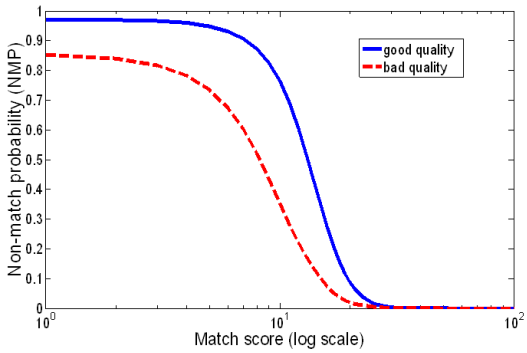



Figure 7. Non-match probabilities for the good and bad quality fingerprint pairs based on kernel density based estimates. The average variance values for the good and bad categories are $0.002$ and $0.003$, respectively. Note that these variances are much lower than the variances reported in Figures 4-6 which are about $0.025$.

and bad) to a fingerprint. For our experiments, to maintain an adequate number of fingerprint pairs of each quality type, we divide the fingerprint pairs into two quality categories: good and bad. The good category corresponds to those fingerprint pairs where both the constituent fingerprints are at least of good quality according to NFIQ (NFIQ$<=3$). The remaining fingerprint pairs are assigned to the bad category. Among the genuine pairs, there are a total of $81,527$ good quality and $26,473$ bad quality pairs while in the case of impostors, there are $836,667$ good quality and $163,333$ bad quality pairs.

Since the numbers of samples in the bad quality category is relatively small, dividing data into $10$ partitions to obtain the non-match probability could lead to large errors in estimating the densities. So, here we divide the dataset into just two partitions and estimate the average variance based on 2-fold cross validation. In order to further improve the reliability of the NMP estimate, we perform $100$ 2-fold partitioning of the dataset to obtain the average variance. The quality based non-match probability curves are shown in Fig. 7 for the case when the kernel density method is used to estimate the genuine and impostor match distributions. Clearly, the quality of the fingerprint pair significantly affects the non-match probability values. Note that the closer the NMP versus match score curve is to a step function, the more conclusive and useful are the NMP values. Based on this observation, as expected, the good quality fingerprint pairs provide more conclusive NMP values than bad quality fingerprint pairs. Further, separating the samples based on quality also reduces the variance of NMP.

We also utilized the genuine and impostor density estimates (using kernel density) computed based on the NIST SD14 database for computing the NMP values for the latent images in the NIST SD27 database which contains $258$ latent fingerprints and their mated full prints. These $258$ latent prints were classified by latent examiners into three

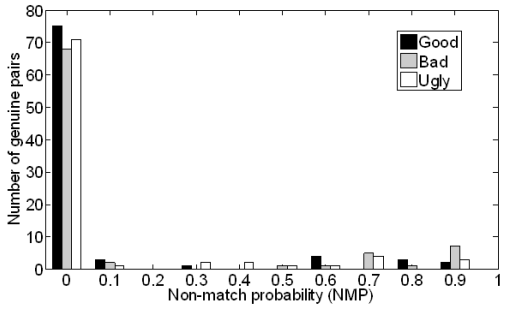
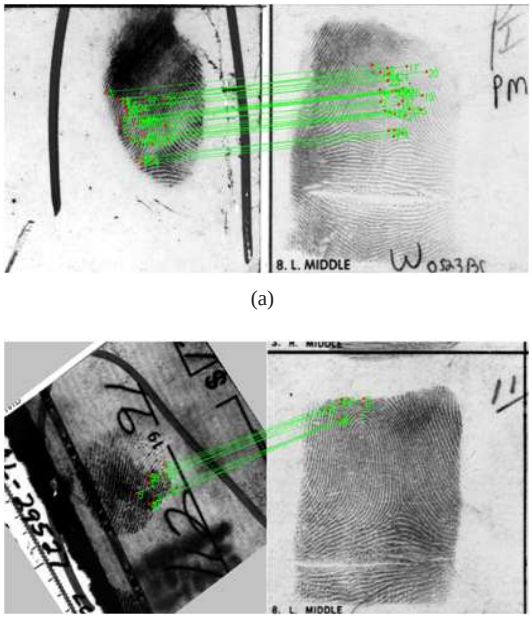



Figure 8. Histograms of the NMP values for 258 latent prints in NIST SD27 categorized as good, bad, and ugly when matched to their corresponding (mated) rolled prints.

quality types, namely: good, bad and ugly. There are 88 good, 85 bad and 85 ugly latent images in the database. For consistency with two-quality level partitioning of the NIST SD14 database, we combine the ugly and bad quality fingerprints into a single bad category and assume the quality category of the latent image as the quality category of the latent-fingerprint pair under consideration. See Fig. 8 for a histogram of the NMP values for 258 latent prints when matched with their mated rolled prints. Note that most of the NMP values are close to zero indicating a genuine match with high confidence; the high NMP values observed for some genuine matches indicate that those match decisions are suspect. The match scores were computed based on manually marked minutiae provided in the NIST SD27 database using the Verifinger matcher[6].

Fig. 9 shows two latent-full print pairs; one is from the good quality category and the other is from the bad quality category (these two specific pairs were considered in [22]). The NMP value for the first pair (Fig. 9(a)) without considering the quality information is $4.47 \times 10^{-56}$ but the quality-based NMP value increases to $1.44 \times 10^{-52}$. For the second pair of bad quality (Fig. 9(b)), the NMP value is decreased from $0.784$ to $0.521$ as a result of considering the quality information. This means that for a poor quality latent-rolled fingerprint pair, the genuine pairs are more likely to have low match scores thereby reducing the NMP value. We also computed the PRC as well as LR values corresponding to these two latent-full print pairs as reported in Table 1. Note that an NMP value of about $10^{-52}$ for the g73 latent-rolled print pair means that out of $10^{52}$ pairs of fingerprints that have the same matching score as that between the g73 latent-rolled print, only one of them is expected to be an impostor pair while remaining pairs are expected to be genuine. On the other hand, the corresponding likelihood ratio of $10^{51}$ simply means that the chances of the same score value being obtained from a genuine pair is $10^{51}$ times more likely than the chances of the same score value being obtained from an impostor pair. The PRC value can, however be understood using a simple counting experiment. The PRC value of about $10^{-55}$ for the g73 latent-rolled print pair with a match score of $65$ means that out of the $10^{55}$ impostor pairs observed in the past, only one of them had a matching score of $65$. The drawback of this interpretation is that it does not explicitly consider the probability that a genuine pair has a match score of $65$. Consider a hypothetical scenario where the probability that a genuine pair has a match score $65$ is $10^{-56}$ . Then the pair is likely to be a genuine match despite a seemingly very small PRC value. While we agree that such scenarios are rare, it does point to the fact that certain aspects of the evidential value cannot be incorporated in PRC.

[6]Since there is no SDK available to us for latent to full print matcher, we rely on the matcher for full to full print to compute the match score.




(a)

(b)

Figure 9. Two sample latent fingerprints and their corresponding full prints from NIST SD27. (a) A good quality latent-rolled pair (g73), and (b) a bad quality latent-rolled pair (b115). The match scores for (a) and (b) are 65 and 9, respectively.

| Latent case | Match score | Quality considered | Estimation of evidential value | | |
|---|---|---|---|---|---|
| | | | NMP (kernel density) | LR (kernel density) | PRC |
| "g73" | 65 | Yes | $1.44\times10^{-52}$ | $6.94\times10^{51}$ | $7.76\times10^{-55}$ |
| | | No | $4.47\times10^{-56}$ | $2.23\times10^{55}$ | $2.35\times10^{-58}$ |
| "b115" | 9 | Yes | 0.524 | 0.910 | 0.018 |
| | | No | 0.784 | 0.276 | 0.024 |

Table 1. Likelihood Ratio (LR) and PRC values for the two latent-rolled print pairs in Fig. 9. Note that due to the definition of LR, the effect of considering quality information on the LR value is opposite of that on PRC and NMP values.

## 5. Summary

In this paper, we have presented a comprehensive framework to analyze the evidential value of fingerprints and proposed a new measure, called the non-match probability (NMP). This measure is more intuitive than the existing measures based on the probability of random correspondence (PRC) and likelihood ratio (LR). Further, it is easier to empirically analyze the reliability of this measure. We show that incorporating the image quality can lead to improvement in the confidence of evidential value of fingerprints. In future, we plan to develop techniques to combine evidence from multiple sources (e.g., multiple matchers and multiple latent prints of the same finger) and increase the size of the database to estimate the genuine and impostor distributions. We also plan to develop relationship between the match score-NMP graph and the corresponding Receiver Operating Characteristics (ROC) curve in order to elicit the pros and cons of using NMP values as a threshold in defining the operating point of a biometric recognition system rather than the False Accept Rate that is commonly used.

## Acknowledgement

This work is supported by the FBI Biometric Center of Excellence. Anil Jain's research was partially supported by WCU (World Class University) program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (R31-2008-000-10008-0).

## References

[1] Daubert v. Merrell Dow Pharmaceuticals. 113 S. Ct. 2786, 1993. 2

[2] http://www.onin.com/fp/daubert_links2.html. 2

[3] http://what-when-how.com/forensic-sciences/standards-of-proof/. 2

[4] http://www.neurotechnology.com. 5

[5] A Review of the FBIś handling of the Brandon Mayfield Case, 2006. Office of the Inspector General, Oversight and Review Division, U.S. Department of Justice. 2

[6] J. Chen and Y. Moon. A minutiae-based fingerprint individuality model. In *Proc. IEEE Computer Vision and Pattern Recognition*, June 2007. 3

[7] Y. Chen and A. K. Jain. Beyond minutiae: A fingerprint individuality model with pattern, ridge and pore features. In *Proc. 2nd International Conference on Biometrics*, pages 523–533, 2009. 3

[8] N. Egli, C. Champod, and P. Margot. Evidence evaluation in fingerprint comparison and automated fingerprint identification systems - modelling within finger variability. Technical Report 167, Forensic Science International, 2007. 4, 5

[9] G. Fang, S. N. Srihari, and H. Srinivasan. Generative models for fingerprint individuality using ridge types. In *Proc. 3rd International Symposium on Information Assurance and Security*, pages 423–428, August 2007. 3

[10] F. Galton. *Finger Prints*. Macmillan, London, 1892. 1, 3

[11] L. Haber and R. N. Haber. Experiential or scientific expertise. *Law, Probability and Risk*, 7(1):143–150, 2008. 2

[12] L. Haber and R. N. Haber. Scientific validation of fingerprint evidence under daubert. *Law, Probability and Risk*, 7(1):87–109, 2008. 2

[13] S. B. Meagher, B. Buldowle, and D. Ziesig. 50k fingerprint comparison test. USA v. Byron Mitchell, US District Court Eastern District of Philadelphia. Government Exhibits 6-8 and 6-9 in Daubert Hearing before Judge J. Curtis Joyner, July 1999. 3

[14] C. Neumann, C. Champod, R. Puch-Solis, N. Egli, A. Anthonioz, and A. Bromage-Griffiths. Computation of likelihood ratios in fingerprint identification for configurations of any number of minutiae. *Journal of Forensic Sciences*, 52(1):54–63, 2007. 4

[15] C. Neumann, C. Champod, R. Puch-Solis, N. Egli, A. Anthonioz, D. Meuwly, and A. Bromage-Griffiths. Computation of likelihood ratios in fingerprint identification for configurations of three minutiae. *Journal of Forensic Sciences*, 51(6):1255–1266, 2006. 4

[16] S. Pankanti, S. Prabhakar, and A. K. Jain. On the individuality of fingerprints. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 24(8):1010–1025, August 2002. 3, 4

[17] C. Wilson et al. Fingerprint vendor technology evaluation 2003: Summary of results and analysis report. Technical Report NISTIR 7123, NIST, June 2004. 2

[18] M. Indovina et al. ELFT Phase II - an evaluation of automated latent fingerprint identification technologies. Technical Report NISTIR 7577, NIST, 2009. 2

[19] D. A. Stoney. *Measurement of Fingerprint Individuality, in Advances in Fingerprint Technology, H. C. Lee and R. E. Gaensslen (eds.)*. CRC Press, Boca Raton, 2001. 3

[20] D. A. Stoney and J. I. Thornton. A critical analysis of quantitative fingerprint individuality models. *Journal of Forensic Sciences*, 31(4):1187–1216, October 1986. 3

[21] C. Su and S. N. Srihari. Probability of random correspondence for fingerprints. In *Proc. 3rd International Workshop on Computational Forensics*, pages 55–66, August 2009. 3

[22] C. Su and S. N. Srihari. Evaluation of rarity of fingerprints in forensics. In *Proc. Neural Information Processing Systems*, December 2010. 3, 7

[23] E. Tabassi, C. Wilson, and C. Watson. Fingerprint image quality. Technical Report NISTIR7151, NIST, 2004. 1, 6

[24] F. Taroni, S. Bozza, A. Biedermann, P. Garbolino, and C. Aitken. *Data Analysis in Forensic Science: A Bayesian Decision Perspective*. Wiley, 2010. 4

[25] Y. Zhu, S. Dass, and A. K. Jain. Statistical models for assessing the individuality of fingerprints. *IEEE Trans. on Information Forensics and Security*, 2(3):391–401, 2007. 3