Summary

For over a century, fingerprints have been a nearly undisputed personal identifier, and have consequently found much practical use in criminal investigations. The 1993 *Daubert v. Merrill Dow Pharmaceutical* rulings that clarify what constitutes scientifically admissible evidence have sparked a new interest in verifying that fingerprints are truly unique from person to person. Consequently, we seek to develop a description of fingerprint structure that is capable of a precise determination of the probability of duplicate fingerprints.

To provide a complete and realistic description of fingerprint structure, our model must achieve the following objectives:

- **Topological structure** in the print determined by the overall flow of ridges and valleys should be described accurately.

- **Fine detail** in the form of ridge bifurcations and terminations must also be characterized accurately.

- **Intrinsic uncertainties** in our ability to reproduce and measure fingerprint data must be considered.

- **Definite probabilities** for specified fingerprint configurations must be calculated.

We place special emphasis on meeting the modeling criteria established by Stoney and Thornton [1986] in their assessment of prior fingerprint models.

In addition to calculating estimates of the theoretical limit of fingerprint differentiation and thereby the probability of print duplication in human history, we apply our model to the more realistic conditions encountered in forensic science to determine the legitimacy of current methodology. We also compare the accuracies of DNA and fingerprint evidence on both practical and theoretical levels.

Given the theoretical limits of fingerprint identification, our model predicts uniqueness of prints throughout human history. Furthermore, we determine that fingerprint evidence can be as valid as DNA evidence, if not more so, although both depend on the quality of the forensic material recovered.

The Myth of "The Myth of Fingerprints"

MCM Contest Question A

Team # 674

July 11, 2004

Contents

1   Introduction

Many issues have motivated the development of reliable personal identification. Historically, identification has been important in verifying individuals' claims of nationality, family membership, or socioeconomic status. Additionally, establishment of identity is crucial for the reliability of financial transactions. For example, in order to have credit at a bank, the bank must be able to uniquely associate each bank account with one person. Moreover, identity verification has applications in the legal system. In particular, reliable identification of criminals allows for greater efficiency in solving crimes, and the recognition and capture of repeat offenders.

Suppose that an investigator at a crime scene discovers a perfect imprint of a right thumb on a wall. According to popular wisdom, that thumbprint belongs to one, and only one, person in the history of the world. If a suspect is found with a thumbprint which matches the print left on the wall, then the investigator is reasonably assured of the suspect's presence at the crime scene. How correct are we to believe, however, that fingerprints are unique? Can the fingerprints humans leave behind be used reliably for identification? How do fingerprints compare with other methods of identification? In this paper we address these questions with a mathematical model of fingerprint structure.

1.1 Identity and Biometrics

Personal identity can be confirmed in modern societies in a number of ways, including secret knowledge (PIN numbers or passwords), possession of unique objects (passports or drivers' licenses), behavioral features (signature), or physical or physiological features (fingerprints). Whereas knowledge and objects may be easily transferred to someone else and behavior may be imitated with practice, certain physical body features are difficult or impossible to alter, making impersonation much more difficult. Such unalterable physical features are frequently referred to as biometrics.

Biometric measurements have been employed for identification for centuries in many different forms. Perhaps the first known usage of biometrics comes from 14th-century China, where impressions of infant hand- and footprints were taken (NCSC). In Europe, artificial biometrics such as tattoos and scars have also been used to keep track of criminals and to record what crimes they had committed. While such methods were certainly effective, humanitarian concerns render them unacceptable today. By the late 19th century, criminals were identified under the system of bertillonage, which uses body measurements such as arm and finger lengths as identifying characteristics. A 1903 incident at Fort Leavenworth demonstrated, however, that bertillonage could yield the same set of measurements to within uncertainty for two separate people and was therefore a flawed system of identity verification (NCSC).

Modern times present many possible biometrics that can be used to determine or confirm a person's identity, including retinal or iris features, DNA testing, hand geometry, vascular patterns, facial recognition, speech, and signature (NCSC). Of these, only fingerprints and DNA are recoverable from the traces humans unintentionally leave in their environment, and therefore constitute the two major biometrics of interest to the forensic sciences. As with all biometrics, each has its advantages and disadvantages. For example, fingerprints can differentiate between identical twins (Prabhakar [2001]) while DNA testing cannot. Conversely, latent fingerprints can be extremely difficult to lift accurately and often yield only partial or smudged prints, but DNA evidence contains all of the available information in those cases in which it can be recovered.

1.2 What is a Fingerprint?

A fingerprint is any two-dimensional pattern created by the friction ridges present on human fingers (Ridges and Furrows). While the human skin is

mostly smooth, the hands and feet present small ridges and valleys that increase traction. Such ridges are believed to form during the embryonic stages of human development and to persist unchanged throughout a person's lifetime. The physical ridge structure appears to depend chaotically on factors such as genetic makeup and embryonic fluid flow (Prabhakar [2001]).

When one's finger is pressed in a fluid such as ink or blood and then onto another surface, only the friction ridges touch the surface and thereby transfer to it a two-dimensional representation of their structure. Sweat and oils secreted onto the surface of the skin are also sufficient to leave behind fingerprints.

Fingerprints are distinguished from each other on three levels of detail (Beeton). The first level considers characteristics such as the overall ridge flow and scarring patterns. This level of detail is very difficult to quantify precisely and therefore is insufficient for accurate discrimination of all prints. It is, however, the easiest detail to examine qualitatively. The second level of detail considers the individual ridges and their bifurcations, terminations, and other discontinuities. These ridge features, called minutiae, afford a much more precise means of comparing fingerprints. In particular, the pairwise locations and orientations of the up to 60 minutiae typically present on a full print provide a wealth of opportunities for detailed comparison (Pankanti et al. [2002]). The third level of detail considers the width of the ridges, placement of pores, and other intra-ridge features. Such detail is frequently missing from all but the best of fingerprints, making it generally unavailable outside of a controlled environment, and is extremely difficult to quantify.

1.3   Fingerprints As Evidence

Since the late 19th century, forensic scientists have used the level one and two detail in latent fingerprints to match criminal suspects with crime scenes. For over a century, fingerprint evidence has been used without major challenge in courts of law in the United States (OnIn.com) to identify and convict countless numbers of criminals. In 1993, however, in *Daubert v. Merrill Dow Pharmaceutical*, the U.S. Supreme Court set standards that potential evidence must meet in order to be admissible as "scientific" evidence:

> "1. The theory or technique has been or can be tested.
>
> 2. The theory or technique has been subjected to peer review or publication.

3. The existence and maintenance of standards controlling use of the technique.
4. General acceptance of the technique in the scientific community.
5. A known potential rate of error."

(Wayman [2000])

In the wake of *Daubert*, six federal trial judges have concluded that handwriting analysis does not meet the Daubert criteria and thus cannot be admitted as scientific evidence (Epstein [2002]). Many defense attorneys are now trying to challenge the scientific authenticity of fingerprint evidence in similar hearings, making the individuality of fingerprints and the accuracy of fingerprint analysis particularly salient today.

1.4   Individuality of Fingerprints

Without access to the thumb of every person who has ever existed and the time and ability to compare them pairwise to each other, it is impossible to determine whether fingerprints are truly unique in human history. Nevertheless, many people have addressed the individuality issue, most frequently by estimating an upper bound for the probability that any given fingerprint configuration occurs in nature. From these probabilities and from estimates of the total human population throughout history, the probability of two thumbprints being unique up to uncertainties in measurement can be estimated. If this probability is sufficiently small, then support is lent to the claim that fingerprints are indeed unique, and thus should be considered by the courts as scientific evidence.

Francis Galton's 1892 study is the first widely known extensive scientific examination of fingerprints. Galton derives an estimate of this configuration probability by dividing the fingerprint into squares with a side length of six ridge periods. After some experimentation, he estimates he can recreate the ridge structure of any given missing square with a probability of $\frac{1}{2}$. Assuming, then, that each square is independent and introducing some additional multiplicative factors, Galton concludes that the probability of any given fingerprint occurring is 1.45×10^{-11}, which stands as one of the higher estimates in the history of such models. Pearson later refines Galton's crude model by considering minutiae more carefully and finds a probability of 1.09×10^{-41} (Stoney and Thornton [1986]).

In his 1933 model, Roxburgh employs a polar coordinate system to specify minutiae location by discrete ridge number and minutiae order

along a ridge. Roxburgh also adds in factors to account for poor image quality, minutiae ambiguity, and ridge count ambiguity to arrive at a probability of any given 35-minutiae configuration to be 5.98×10^{-46}. Amy's 1947 model, like Roxburgh's, considers minutiae ordering and orientation along ridges, but focuses on relative rather than absolute positioning of minutiae. Amy also corrects for clusters of minutiae in a line that can be treated as their own ridge (Stoney and Thornton [1986]).

In his 1977 model, Osterburg extends Galton's approach by dividing the fingerprint into cells that can each contain one of 12 minutiae types. Based on his empirical measures of minutiae type frequency and his assumption of independence among different cells, Osterburg calculates the probability of any given configuration to be 1.33×10^{-27}. In his 1979 paper, Sclove extends Osterburg's model to take into account dependencies between cells and the presence of multiple minutiae in a single cell (Stoney and Thornton [1986]).

Stoney's 1986 critique of these and other models charges that, despite the variety of techniques, assumptions, and premises employed to determine these upper bounds on the configuration probabilities, all of these models fail to consider a number of key issues completely, such as

- the topological information conferred through level-one detail,

- descriptions of minutiae location,

- minutiae orientation,

- minutiae type,

- variations in fingerprints taken from the same source, and

- number of positions considered per comparison.

In our model of fingerprint structure, we endeavor to correct at least some of these omissions of important structural data.

2 Our Model: Assumptions and Constraints

2.1 Assumptions

In order to model fingerprint structure realistically, we must make a number of assumptions. These include the following:

- **Fingerprints are persistent: they remain the same throughout a person's lifetime.** Law enforcement agencies and judicial bodies typically assume persistence of prints; in fact, their utility in criminal cases and identification systems stems directly from this temporal consistency. Francis Galton's early studies (Galton [1892]) establish this print persistence, and in recent times it has been verified by more detailed comprehension of the processes that underlie the development of dermal tissues (Ridges and Furrows).

- **Fingerprints are of the highest possible quality and are positioned consistently with respect to the global structure of the print.** Since we desire to model the structure of the fingerprint itself, free from complicating factors, we must assume we are working with representations of the best possible quality. Furthermore, because we wish that two prints differing only by an overall rotation be considered the same, we assume that we can translate and rotate a print within reasonable extents.

- **Each human fingerprint consists of a pattern of ridges with at least some degree of continuity and flow.** Empirically, all natural fingerprints have been observed to exhibit systems of parallel ridges that flow together in a distinctive pattern. While environmental factors such as abrasion and injury may affect these structures, our best-quality assumption implies that the print is taken from the physical pattern that would develop independently of such damage.

- **The ridge structure of each fingerprint unambiguously places it into one of five categories: Arch, Left Loop, Right Loop, Tented Arch, or Whorl.** These five classes are those employed in the automatic classification system developed by Cappelli et al. [1999], and derive from the systems used by the FBI and by NIST (Watson and Wilson [1992]). Furthermore, following the approach of Cappelli et al. [1999], we assume that each class of print has associated with it a characteristic ridge flow topology, which we can break into homogeneous domains in which the flow is approximately unidirectional.

  While Cappelli et al. [1999] raises the issue of prints that are "unclassifiable," even by experts, and both Cappelli et al. [1999] and Marcialis et al. [2001] demonstrate confusion between classes of ridge structures in their automatic classification, we assume that such ambiguities stem more from imperfections in print quality than in topological defects or anomalies in the print structure. Regardless, the relative

rarity of these problematic prints indicates that we should be able to cover the majority of possible prints with these five broad classes.

- **Fingerprints may further be distinguished by the location and orientation of minutiae relative to local ridge flow.** Equally as important as ridge structure in classifying prints is the location and orientation of minutiae on ridges. Stoney and Thornton [1986] argue that the ridges define a natural coordinate system across and along the ridges, and that the location of a minutia can be specified with a ridge number and a linear measure along that ridge. Finally, minutiae can have exactly one of two equally likely, fundamentally distinct orientations along a ridge.

- **Each minutia can be classified as either a bifurcation, a termination, or a dot.** While Galton [1892] identifies 10 basic minutiae structures, and while more modern authors such as Osterburg et al. [1977] extend these to up to 13 different classes of minutiae, we follow Stoney and Thornton [1986] and Pankanti et al. [2002] in arguing that the fundamental minutiae are bifurcations, terminations, and dots (see Figure 1 for illustrations), and that all other significant minutiae are compositions of these basic three. We also feel comfortable ignoring more complex minutiae because of their low rates of occurrence compared to those of simple bifurcations and terminations (Osterburg et al. [1977]).

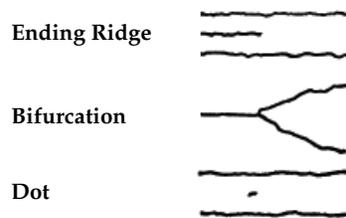Figure 1: Galton's illustrations of the three basic minutiae types (from Galton [1892]). In this paper, we refer to ending ridges as terminations.

- **Any particular physical ridge structure will produce an unambiguous fingerprint, within the limits of some level of visual resolution.** While Stoney and Thornton [1986] point out that a given finger's ridge structure in general produces a variety of print representa-

tions, this variation is present primarily in the prints' geometric data, such as ridge spacing, curvature, and location of minutiae along a ridge. This variation then justifies a degree of uncertainty in the location of minutiae within a particular ridge system. On the other hand, topological data such as ridge counts and minutiae orientation and ordering is robust to such variation, and it is therefore reasonable to assume that these structures will be replicated consistently from print to print.

A more serious consideration is possible connective ambiguities, such as when a given physical minutia is represented sometimes as a bifurcation and sometimes as a termination. Ideally, a model would account for the possibility of confusion among minutiae type. Since we assume the highest possible quality two-dimensional representation of the physical ridge structure, however, we feel safe in asserting that the majority of minutiae are consistently reproduced correctly, and that ambiguities arise only in cases where the physical structure itself is ambiguous.

- **Location and orientation of minutiae relative to each other is independent.** Stoney and Thornton [1986] indicates that some dependency exists between positions of minutiae, and Sclove [1979] explicitly considers such dependency in a Markov-type process. Nevertheless, we feel that it is enough in this model to consider the minutiae structure as an independent process, although the incorporation of such dependence constraints can be looked at as an possible extension.

- **Ridge widths are uniform throughout the print and among different prints, and ridge detail such as pores and edge shapes will not be considered significant.** Keeping ridge widths uniform simplifies the geometric consideration of the ridge structure considerably and is reasonably consistent with empirical observations. While such ridge detail is potentially useful in further differentiating prints, we have few statistics about their types and frequencies available to incorporate into a model. Furthermore, the finer the detail in these structures, the less likely is it to reproduce reliably in a print: edge and ridge widths, for example, may change appreciably with variations in pressure and ink coating.

- **Frequencies of ridge structure classes and configurations and minutiae types do not change appreciably with time.** We must assume

this invariance of print structure frequencies if the probabilities generated by our model are to apply throughout human history. Depending on the results and sensitivity of our model, however, it may be possible to relax such assumptions and still retain what conclusions we can draw from the model.

2.2 Constraints

These assumptions imply a set of constraints that our model must fulfill to give a satisfactory description of a fingerprint's structure. Such constraints include the following:

- **Our model must consider both ridge structure and relative position, orientation, and type of minutiae in order to describe a fingerprint completely.** To leave out minutiae from our model gives only a very broad characterization of the print structure, while to leave out ridge structure dismisses vital topological information about the print. Therefore, we must construct our model to replicate both levels of print information.

- **Locations of minutiae must be specified only to within some uncertainty dependent on the inherent uncertainty in feature representation.** As stated above, even under ideal circumstances, a physical characteristic of the ridge skin may be represented in a variety of ways, leading to an intrinsic uncertainty in the corresponding print structure. Our model must incorporate this fundamental uncertainty into its description of the print structure.

- **Given a particular configuration of ridge flow and minutiae on a fingerprint, our model must determine the probability that this configuration occurs.** If we are to use this model to estimate the theoretical and practical limits of the individuality of fingerprints, we must associate to any given print configuration a probability that it occurs naturally within the human population. Such probabilities will ultimately allow us to compute the probability that a given print is produced twice in a population.

3 Model Formulation

Given these assumptions and constraints, our goal is to develop a model that associates a probability with any given configuration of ridge struc-

ture and minutiae location, orientation, and type within a thumbprint. By summing these probabilities appropriately over the configuration space for prints, we can estimate the probability that two distinct humans in history will have physical ridge structures that produce indistinguishable thumbprints.

In order to determine the overall probability of a given configuration, we examine the following hierarchy of probabilities:

- the probability that the given class of ridge structure occurs,

- the probability that the ridge structure occurs in the specified configuration of ridge flow regions, and

- the probability that minutiae are distributed as specified throughout the regions.

We further break this last probability down into a composition of the following region-specific probabilities:

- the probability that a region will contain the specified number of minutiae,

- the probability that the minutiae within this region follow the specified spatial configuration, and

- the probability that the minutiae occur with the specified types and orientations.

We now analyze these probabilities in more detail.

3.1   Probability of Ridge Structure Class

We first discuss the probabilities associated with the occurrence of each class of ridge structure. As stated above, we consider five basic classes of ridge structures: Arches, Left and Right Loops, Tented Arches, and Whorls. To each class, we associate a particular probability of occurrence, which we estimate from statistical analyses of frequency among the population. We denote these probabilities ν_A, ν_L, ν_R, ν_T, and ν_W.

3.2   Probability of Ridge Structure Configuration

We base our description of ridge structure configuration on the classification developed by Cappelli et al. [1999]. In this model, each print is partitioned into regions in which the overall flow is relatively unidirectional,

and the class of the print determined from the comparison of this partition to five prototypical masks. We consider these masks to be fundamentally characteristic of each ridge structure class. The variations of flow region structure within each class then depend on a small number of parameters derived from key features characteristic to each class. For example, the ridge structure of a Loop print can typically be determined from the locations of the triangular singularity and the core of the loop. Figure 3 illustrates these locations. To determine the probability of a particular region configuration, then, we need only determine the probability that these associated parameters occur as specified.

Because there is some degree of uncertainty in the physical measurement of these configuration parameters, we discretize our parameter space based on this fundamental resolution limit. We denote this limit as δ_1, where the subscript indicates that it pertains to level-one print features. We then associate to each compartment in parameter space the probability that the parameter lies within that compartment. In general, we employ independent Gaussian distributions about the mean values for each parameter to determine these probabilities.

We now detail the parameter spaces for each ridge structure class. The use of these prototypes requires us to define a rectangular region of width X and height Y within the print. Figure 2 depicts the prototypical regions and their associated parameters. Figure 3 depicts the application of some prototype masks to specific fingerprints.

3.2.1   Arch

The variable parameters for the Arch consist of the Cartesian coordinates (x, y) of the lower corner of the left region, the height h of the central corridor, and the four angles $\theta_1, \theta_2, \theta_3, \theta_4$ at the inner corners of the left and right regions. We also consider as fixed the width b of the central corridor. We note that the ratio of the resolution limit δ_1 to the mean length of a typical segment determines the uncertainty in the angular measurement of that segment.

3.2.2   Loops, Left and Right

Since Left and Right Loops are identical except for a horizontal reflection, we can use the same parameter space to characterize both classes. The two principal features of the loop structure are the position (x, y) of the

Figure 2: The prototypical region structures and parameters for each ridge structure class, derived from the masks in Cappelli et al. [1999].

triangular singularity outside the loop and the distance r and angle θ of the core of the loop relative to this triangular singularity.

3.2.3 Tented Arch

The major structure in the Tented Arch class is the arch itself. The parameters associated with this class are therefore the position (x, y) of the base of the arch and the height of the arch h.

Arch   Right Loop   Whorl

Figure 3: The prototypical region structures applied to an Arch, a Right Loop, and a Whorl.

3.2.4 Whorl

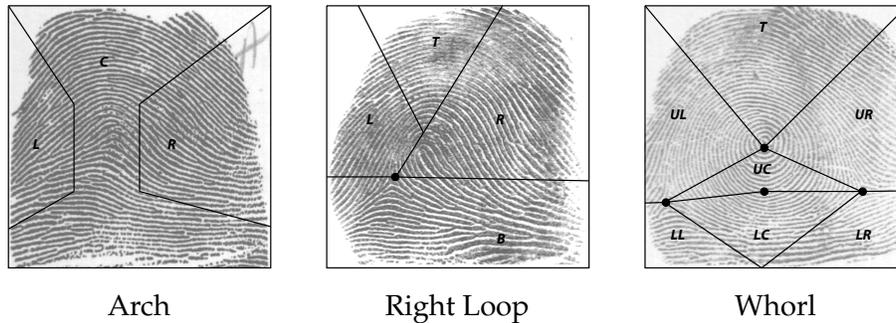The Whorl structure presents four major features that determine its structure. One of these is the center of the whorl itself, whose position is denoted by (x_C, y_C). The next is the base of the whorl, at position (x_B, y_B). The two other points are the triangular singularities to the left and right of the base of the whorl, at positions $(x_L, y_:)$ and (x_R, y_R). We further assume that the center of the whorl and the base lie between the two singularities, so that $x_L \leq x_C$ and $x_B \leq x_R$, and that the base of the whorl lies above the singularities, so that $y_B \geq y_L$ and $y_B \geq y_R$.

3.3 Probabilities of Intra-Region Minutiae Distribution

We now discuss how minutiae are distributed within each unidirectional flow region. Since the geometry of each region is uniquely determined by the configuration parameters, given parameter values we can divide each region up into a series of parallel ridges. While in practice ridges within a region bifurcate and terminate and therefore deviate from this parallel structure to some extent, assuming a consistent structure throughout confers a regular coordinate system for placement of minutiae. As a consequence of this parallel division, we can represent the ridge structure of the region as a list of ridge lengths.

As in the analysis of the ridge structure parameters, we assume we have some fundamental limit δ_2 to our resolution of the position of minutiae along a ridge. Consequently, we divide each ridge into a series of cells of length δ_2, in which we assume we find no more than one minutia. An

expression for the probability $P_{TC}(n,l,k)$ that the nth ridge in the partition, with length l, has a particular configuration of k minutiae is then

$$P_{TC}(n,l,k,\dots) = P_p(n,k,l)P_c(n,k,l)P_{to}(\{k_i,p_i,o_i\}),\tag{1}$$

where P_p is the probability that k minutiae occur on this ridge, P_c the probability that these k minutiae are configured in the specified pattern on the ridge, and P_{to} the probability that these minutiae are of the specified types and orientations, indexed by i and occurring with type probability p_i and orientation probability o_i. In general, these probabilities may depend on n, the ridge number within the region. We now discuss more specific expression for these probabilities.

3.3.1　Probability of Minutiae Number

Under the assumption that minutiae occur at uniform rates along a particular ridge, we expect a binomial distribution for the number of minutiae on the ridge. Suppose that the linear minutiae density, which may depend on the ridge number n, is denoted $\lambda(n)$. Then the probability that a minutiae occurs in a given cell of length δ_2 is given by $\delta_2\lambda(n)$. Thus, the probability that k minutiae occur is

$$P_p(n,k,l,\lambda) = \binom{l/\delta_2}{k}(\delta_2\lambda)^k(1-\delta_2\lambda)^{l/\delta_2-k}.\tag{2}$$

3.3.2　Probability of Minutiae Configuration

Assuming that all configurations of minutiae are equally likely along the ridge, the probability that these k minutiae occur in the specified configuration is the inverse of the number of possible configurations, so that

$$P_c(n,k,l) = \frac{1}{\binom{l/\delta_2}{k}}.\tag{3}$$

3.3.3　Probability of Minutiae Type and Orientation

The probability that the minutiae occur with specified types and orientations is then

$$P_{to}(\{k_i,p_i,o_i\}) = \prod_i p_i^{k_i}o_i^{k_i}.\tag{4}$$

Applying our assumption that the only level-two features are bifurcations, terminations, and dots, and that orientations are equally likely and independent along the ridge, this expression reduces to

$$P_{to} = p_b^{k_b} p_t^{k_t} p_d^{k_d} \frac{1}{2^{k_b+k_t}} \tag{5}$$

such that $k_b + k_t + k_d = k$. Then the total probability for the ridge configuration is

$$P_{TC}(n, l, k, \lambda, \{k_i, p_i, o_i\}) = (\delta_2 \lambda)^k (1 - \delta_2 \lambda)^{l/\delta_2 - k} p_b^{k_b} p_t^{k_t} p_d^{k_d} \frac{1}{2^{k_b+k_t}}. \tag{6}$$

The total probability that minutiae are configured as specified through the entire print is then product of the P_{TC}s for all ridges in all domains, since we assume ridges develop their minutiae independently.

Applying the additional assumption that λ and other factors do not depend on n and are therefore uniform throughout the print, we can collapse these multiplicative factors and develop an expression for the configuration probability of the entire print:

$$P_{TC}^{\text{global}} = (\delta_2 \lambda)^K (1 - \delta_2 \lambda)^{L/\delta_2 - K} p_b^{K_b} p_t^{K_t} p_d^{K_d} \frac{1}{2^{K_b+K_t}}. \tag{7}$$

Here, K is the total number of minutiae in the print, K_i the number of type i, and L is the total linear length of the ridge structure in the print. We note that L is determined only by the total area XY of the print and the average ridge width w and is therefore independent of the topological structure of the print.

4   Parameter Estimation

Before we can accurately calculate probabilities for fingerprint configurations, we must determine values for the many parameters in our model. Lacking the resources to investigate a number of these parameters independently, we instead elect to use values in published papers when available and to make estimates based on a limited investigation of the NIST-4 database.

4.1   Level One Parameters

We group our parameters by the level of detail that they describe. We begin with the parameters that describe the details of ridge flow.

- **Level One Spatial Resolution Limit** δ_1: Cappelli et al. [1999] discretize their images into a 28 by 30 grid to determine the level one detail. From these grid dimensions, the approximate physical dimensions of the fingerprints, and the assumption that there is an uncertainty of three blocks for any level-one feature, we estimate the natural length level-one-feature resolution limit to be $\delta_1 = 1.5$ mm.

- **Level One Angular Resolution Limit** δ_θ: Taking $X/2 = 5.4$ mm (determined below) as a typical length scale, we have $\delta_\theta = \delta_1/(5.4 \text{ mm})$, making $\delta_\theta = .279$ radians.

- **Ridge Structure Class Frequencies** ν_A, ν_L, ν_R, ν_T, **and** ν_W: We use the estimates of the relative ridge structure class frequencies presented in Prabhakar [2001], which we list in Table 1.

ν_A	ν_L	ν_R	ν_T	ν_W
0.0616	0.1703	0.3648	0.0779	0.3252

Table 1: Relative Frequencies of Ridge Structure Classes (from Prabhakar [2001]).

- **Thumbprint Width** X **and Height** Y: Examining many thumbprints from the NIST-4 database and comparing them with the area given by Pankanti et al. [2002], we conclude that a width that will cover the majority of the thumbprints is 212 pixels in the 500 dpi images, which corresponds to a physical length of 10.8 mm. Similarly, we conclude that $Y = 16.2$ mm.

- **Arch Characteristic Parameters** (x, y), h, b, θ_1, θ_2, θ_3, **and** θ_4: We restrict the parameter space for (x, y) to the lower half of the thumbprint with horizontal margins of length b. We estimate b as 2.5 mm from examination of Arch fingerprints in the NIST database and from Cappelli et al. [1999]. This estimate places $x \in (0 \text{ mm}, 8.3 \text{ mm})$ and $y \in (0 \text{ mm}, 5.6 \text{ mm})$. The average value for (x, y), which we need to describe the distribution of (x, y) is then $(4.2 \text{ mm}, 2.8 \text{ mm})$. We estimate that x and y both have a standard deviation of 0.7 mm. We assume that θ_1–θ_4 are all restricted between 0 and 45 degrees with average value 22.5 degrees with a standard deviation 5.13 degrees.

- **Loop Characteristic Parameters** (x, y), θ, **and** r: We determine parameters only for left loop, since the right loop is the horizontal reflec-

Arch Parameter Ranges	
(x, y)/mm	$(4.2, 2.8) \pm (0.7, 0.7)$
h/mm	4.05 ± 0.7
b/mm	2.5 ± 0
θ_1–θ_4/°	22.5 ± 5.13
Loop Parameter Ranges	
(x, y)/mm	$(2.7, 2.8) \pm (0.7, 0.7)$
θ/°	45 ± 15
r/mm	4.58 ± 0.7
Tented Arch Parameter Ranges	
(x, y)/mm	$(5.4, 2.8) \pm (0.7, 0.7)$
h/mm	4.05 ± 1.02
Whorl Parameter Ranges	
(x_L, y_L)/mm	$(2.7, 4.1) \pm (0.7, 0.7)$
(x_C, y_C)/mm	$(5.4, 12.2) \pm (0.7, 0.7)$
(x_R, y_R)/mm	$(8.1, 4.1) \pm (0.7, 0.7)$
(x_B, y_B)/mm	$(5.4, 4.1) \pm (0.7, 0.7)$

Table 2: Parameter range estimates for the ridge structure classes.

tion of the left loop. We reason that (x, y) must lie in the bottom left quadrant and that the average coordinate pair is $(2.7 \text{ mm}, 2.8 \text{ mm})$. Additionally, we restrict θ to lie between 15 degrees and 75 degrees, which allows us to estimate the average θ as 45 degrees with a standard deviation of 15 degrees. We estimate that r must be greater than 0 mm and less than 9.6 mm.

- **Tented Arch Characteristic Parameters** (x, y) **and** h: Along the y direction we restrict the bottom of the arch (x, y) to lie in the bottom half of the thumbprint.. We further estimate that x lies in the middle two-thirds of X. These assumptions yield $x \in (1.8 \text{ mm}, 9 \text{ mm})$ and $y \in (0 \text{ mm}, 8.1 \text{ mm})$. Assuming a symmetric distribution of (x, y) yields $(x, y) = (5.4 \text{ mm}, 2.8 \text{ mm})$ with a standard deviation of 0.7 mm in both directions. Logically, we place h between 0 mm and $Y/2 = 8.1$ mm. Again, assuming a symmetric distribution in this pa-

rameter space and a standard deviation of one-eighth the maximum value yields $h = 4.05$ mm ± 1.02 mm.

- **Whorl Characteristic Parameters** (x_L, y_L), (x_C, y_C), (x_R, y_R), **and** (x_B, y_B): We expect (x_L, y_L) to be in the bottom left quadrant for all but the most extreme examples and similarly (x_R, y_R) to lie in the bottom right quadrant. We place (x_B, y_B) between $x = \frac{X}{4}$ and $x = \frac{3X}{4}$ and $y = 0$ and $y = \frac{2Y}{3}$. The topmost point, (x_C, y_C), we place in the top half of the thumbprint. We again put the average values in the center of their restricted areas.

The parameter range estimates for these four classes of ridge structures are summarized in Table 2.

4.2 Level Two Parameters

Parameters associated with level-two details include the following.

- **Level Two Spatial Resolution Limit** δ_2: With only minimal experience in comparing fingerprints, it is difficult to arrive at a reasonable value for the spatial resolution limit. As a result, we use the value given by Pankanti et al. [2002] for the spatial uncertainty of minutiae location in 2-dimensions, r_0, as an estimate for our δ_2. We propose $\delta_2 = 1$ mm as a value for best-case calculations. In practice, this value may be as high as 3 mm because of the low quality of latent prints.

- **Relative Minutiae Type Frequencies** p_d, p_b, **and** p_t: Osterburg publishes the relative frequencies of 12 different types of minutiae. Almost all of the the compound minutiae can be broken into a combination of bifurcations and terminations separated spatially. Counting these compound minutiae appropriately, we determine these relative minutiae frequencies, which are listed in Table 3.

p_b	p_t	p_d
0.356	0.581	0.0629

Table 3: Frequencies of Simple Minutiae Types (from Osterburg et al. [1977])

- **Ridge Period** w: We use Stoney's value of 0.463 mm/ridge for the ridge period, the distance from the middle of one ridge to the middle of an adjacent ridge (Pankanti et al. [2002]).

- **Mean Number of Minutiae per Print** μ: Under ideal circumstances, we expect to be able to discern between 40 and 60 minutiae on an average print (Pankanti et al. [2002]), so we take $\mu = 50 \pm 10$. In practice, forensic specialists may be able to discern as few as 10 minutiae from a latent print because of poor quality, so we take $\mu = 20 \pm 10$ when considering these conditions.

- **Linear Minutiae Density** λ: We calculate λ by dividing the average number of minutiae per a thumbprint μ by the total ridge length of a thumbprint XY/w. Under ideal conditions, this gives $\lambda = 0.13 \pm 0.03$ minutiae/mm. In practice, we may have $\lambda = 0.05 \pm 0.03$ minutiae/mm (Pankanti et al. [2002]).

Finally, we estimate the total number of humans in the history of the planet. We use the value given in Haub [2002], approximately 100 billion. Given the number of approximations Haub makes in calculating this number, no greater accuracy is warranted.

5 Model Analysis and Testing

Given two prints taken at random from the population, we wish to determine the probability that they present the same print structure. For each point x in our print configuration space, the probability that each print has that configuration is the probability of occurrence $p_c(x)$ for that configuration. Assuming that the fingerprint patterns are distributed independently, then, the probability the prints match is $p_c^2(x)$. Then the sum of these probabilities as x ranges over the entire configuration space is the total probability of a match.

Operating under the assumptions described above, the probabilities associated with the two levels of detail are determined independently, so the total occurrence probability factors into $p_{c1}(x_1)p_{c2}(x_2)$. Denoting the level-one configuration subspace as C_1 and the level-two subspace as C_2, the total probability of the prints matching is

$$p = \sum_{i \in C_1} \sum_{j \in C_2} (p_{c1}(i)p_{c2}(j))^2 = \left(\sum_{i \in C_1} p_{c1}^2(i) \right) \left(\sum_{j \in C_2} p_{c2}^2(j) \right) = p_1 p_2. \quad (8)$$

Thus, we can determine the probabilities that the prints' level-one and level-two structures match separately. This separation greatly increases both the efficiency and accuracy of our computations.

5.1   Level-One Detail Matching

To determine the level-one probabilities $p_{c1}(i)$ we wish to use Gaussian distributions about the mean of each parameter. While we have crude estimates of the ridge structure parameters, they are not sufficiently refined to provide an accurate estimate of these Gaussian distributions. Thus, we make a simplifying assumption: we restrict each parameter to a region of parameter space in which we would reasonably expect to find the parameter and assume that the parameter is uniformly distributed in that region. This crude approximation does not provide the most precise determination of the probability, but it does give us enough accuracy to get at least an estimate of the order of magnitude, which suffices for our analysis. Making this assumption, we see that,

$$p_{c1}(i) = \frac{\nu_i}{\left(\prod_{j \in V(i)} \frac{L_j}{\delta_1} \right)} \tag{9}$$

where L_j is the range of parameter j in $V(i)$, the set of parameters for a type i ridge structure. In order for Equation 9 to be accurate, we should make any L_j corresponding to angular parameters the product of the angle range with our typical length of 5.4 mm. Note that the product is simply the total number of compartments in our parameter space, since we are assuming a uniform distribution in that range. Calculating this probability, $p_{c1}(i)$, for each ridge structure type, and summing their squares, we find,

$$p_1 = \sum_{i \in C_1} p_{c1}^2(i) = .00044 \tag{10}$$

This is the probability that any two thumbprints have the same overall ridge structure.

5.2   Level-Two Detail Matching

Recalling that Equation 7 gives the probability of a particular minutiae configuration, we note that if we disregard the relatively infrequent dot minutiae and focus only on the more fundamental bifurcations and ridges, we obtain the probability

$$p_{c2}(j) = (\delta_2 \lambda)^k (1 - \delta_2 \lambda)^{C-K} p_b^{k_b} p_t^{k-k_b} \frac{1}{2^k} \tag{11}$$

$\delta_2\lambda$	0.10	0.13	0.16
$C = 250$	2.9×10^{-23}	2.3×10^{-30}	1.3×10^{-37}
$C = 300$	8.9×10^{-28}	2.7×10^{-36}	5.5×10^{-45}
$C = 350$	2.7×10^{-32}	3.2×10^{-42}	2.3×10^{-52}
$C = 400$	8.5×10^{-37}	3.8×10^{-48}	9.8×10^{-60}

Table 4: Second-level match probabilities for $C = 250$ to 400 cells, $\lambda = 0.13 \pm 0.03/\text{mm}$, and $\delta_2 = 1$ mm.

$\delta_2\lambda$	0.05	0.10	0.15	0.20
$C = 100$	3.7×10^{-5}	9.6×10^{-10}	1.6×10^{-14}	2.0×10^{-19}
$C = 150$	2.3×10^{-7}	3.0×10^{-14}	2.1×10^{-21}	8.6×10^{-29}
$C = 200$	1.4×10^{-9}	9.2×10^{-19}	2.8×10^{-28}	3.8×10^{-38}
$C = 250$	8.7×10^{-12}	2.9×10^{-23}	3.6×10^{-35}	1.7×10^{-47}

Table 5: Second-level match probabilities for $C = 100$ to 250 cells, $\lambda = 0.05 \pm 0.03/\text{mm}$, and $\delta_2 = 2\text{--}3$ mm.

for a configuration j corresponding to k minutiae, k_b of which are bifurcations, placed in $C = XY/w\delta_2$ cells. If we further simplify our minutiae type frequencies to be $p_b = p_t = 1/2$, and note that there are $\binom{C}{k}\binom{k}{k_b}2^k$ ways to configure j given k and k_b, the total probability of a match becomes

$$p_2 = \sum_{k=0}^{C} \sum_{k_b=0}^{k} \left((\delta_2\lambda)^k (1 - \delta_2\lambda)^{C-K} \frac{1}{4^k} \right)^2 \binom{C}{k}\binom{k}{k_b}2^k \tag{12}$$

$$= \sum_{k=0}^{C} (\delta_2\lambda)^{2k}(1 - \delta_2\lambda)^{2(C-k)} \frac{1}{4^k} \binom{C}{k} \tag{13}$$

$$= \left(\frac{5(\delta_2\lambda)^2 - 8\delta_2\lambda + 4}{4} \right)^C . \tag{14}$$

Table 4 displays the match probabilities obtained by taking $\lambda = 0.13 \pm 0.03/\text{mm}$ and $\delta_2 = 1$ mm and estimating C between 250 and 400 cells. Table 5 displays the probabilities for the more realistic parameter values of $\lambda = 0.05 \pm 0.03/\text{mm}$, $\delta_2 = 2\text{--}3$ mm, and $C = 100\text{--}250$ cells.

5.3   Historical Uniqueness of Fingerprints

Now that we have estimates of upper bounds for the probability of two prints at random matching, we can estimate the probability that any two left thumbprints in the history of the human race match. We focus specifically on only the thumbprint of a particular hand in order to justify the claim of independence between prints.

Suppose the probability of a match is denoted p and the total population of the world N. Then the probability that at least one match occurs among the $\binom{N}{2}$ thumbprint comparisons that must be performed is given by

$$P = 1 - (1 - p)^{\binom{N}{2}}. \tag{15}$$

Figure 4 illustrates the probability of at least one match for $N = 10^{11}$, while Figure 5 shows a log-log plot of the probability for very small p-values. For p below 10^{-30}, the chance of a match between two thumbprints falls below $P = 10^{-8}$, at which point we can reasonably conclude that fingerprints are unique throughout history. Since even conservative parameter values in the ideal case give $p \ll 10^{-30}$, our model solidly establishes such uniqueness.
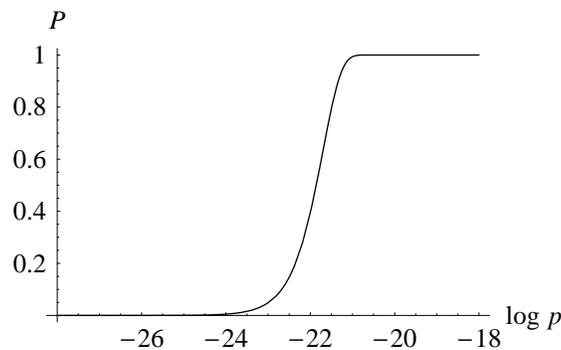

Figure 4: Probability of at least one thumbprint match through history, with $N = 10^{11}$.
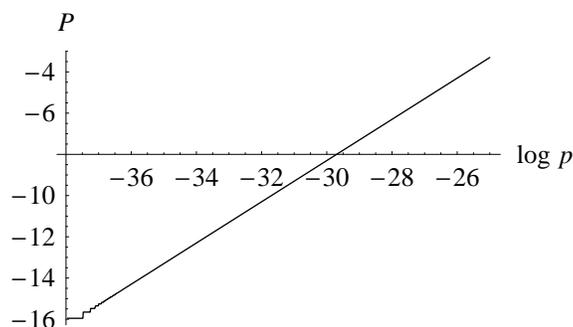
Figure 5: Log-log plot of the probability of a historical thumbprint match, with $N = 10^{11}$.

6 Strengths and Weaknesses of the Model

6.1 Strengths

In comparison to existing models of fingerprint structure and probability, our model has several special advantages, which we detail below.

- **Topological Coordinate System**: One strength of our model is that we take topological considerations into account in defining the placement of minutiae within the print. Stoney and Thornton [1986] notes that this natural ridge-based coordinate system should be a major consideration of any proper model of fingerprint structure because of the topological order it confers to the print. By explicitly considering the division of the print region into ridges, we build this order into our model from the ground up.

- **Incorporation of Ridge Structure Detail**: Another key feature of our model is that it uses the detail associated with the flow of ridges on the print in addition to the minutiae detail, which is the primary focus of most other models that evaluate the individuality of prints. We feel that these minutiae-driven models neglect important structural elements in their analyses and thereby in many cases incorrectly estimate the probability of print correspondence.

- **Integration of Non-Uniform Distributions**: We incorporate nonuniform distributions into our model in two key places: we allow for

more complex distributions of the ridge structure parameters, such as Gaussian distributions for singularity locations, and we consider that distribution of minutiae along ridges may depend on the location of the ridge in the overall structure. In addition, applying uniform distributions in their place allows us to recover the salient features of other, less sophisticated models.

- **Accurate Representation of Minutiae Type and Orientation**: We follow models such as those developed by Roxburgh and Stoney in emphasizing the bidirectional orientation of minutiae along ridges, and we further consider the type of minutiae present as well as their location and orientation. Cruder models of minutiae structure, such as those developed by Pankanti et al. [2002] or Osterburg et al. [1977], neglect at least some of this key information.

- **Flexibility in Parameter Ranges Considered**: Our model is sufficiently robust to consider parameter ranges for both ideal and practical cases. Furthermore, we explicitly test a range of parameters in both of these scenarios and find that the model behaves as expected under these variations. We further expect stable behavior under variation of other parameter sets, including minutiae distributions and ridge flow parameters.

6.2 Weaknesses

As is true of every model, our model has a number weak points, which we address below. In particular, we emphasize the complicating features we have elected not to incorporate into our model.

- **Ambiguous Prints, Smearing, or Partial Matches**: We assume that any ambiguities in our prints reflect ambiguities in the physical structure of the finger and not those introduced by the creation of the two-dimensional representation. As we discuss in our assumptions, this is certainly not the case for actual fingerprints. Furthermore, we have no explicit mechanism to account for smearing or other obscuring effects that arise in forensic situations, although adjustment of the effective print area and resolution limit parameter may suffice to replicate these considerations. In the theoretical limit, however, we assume no such ambiguities, so we can safely neglect their effects.

- **Domain Discontinuities**: One flaw in our representation of the ridge flow structure is that we have no guarantee of continuity between re-

gions of flow. Such continuity requirements places an additional constraint on the probabilities of ridge-flow configurations, which may affect the level-one matching probabilities significantly.

- **Nonuniform Minutiae Distribution**: We explicitly assume that the distribution of minutiae along a ridge is uniform. Stoney and Thornton [1986] points out, however, that models should account for variations in minutiae density across the print and that clustering of minutiae occurs and may depend on the relative location of minutiae within a print. Although we have a mechanism for varying this distribution across ridge numbers, we currently have no data on what this distribution should be.

- **Left/Right Orientation Distribution**: We automatically assume that the distribution of minutiae orientation will be independent and uniform throughout the print. Amy notes, however, that the preferential divergence or convergence of ridges in a particular direction can lead to an excess of minutiae with a particular orientation (Stoney and Thornton [1986]). Ideally, a model would derive its orientation distribution from analysis of the topology of the ridge flow.

- **Level Three Information**: As stated above, we neglect level three information such as pores and edge shapes because of uncertainty regarding its reproducibility in prints. It is conceivable that particularly significant level-three information might remain consistent from print to print, in which case it would be desirable to incorporate these distinguishing details into the model.

- **Continuous Distributions**: We have chosen to discretize all of the parameter spaces associated with the model, in part to incorporate uncertainties naturally and in part to simplify computations. A more accurate formulation may involve keeping some of these parameter spaces continuous and integrating rather than summing over uncertainty regions. Such calculations will almost certainly be more complicated to carry out, however.

7   Comparison with DNA Methods

7.1   DNA Fingerprinting

The genetic material in all living organisms consists of deoxyribonucleic acid (DNA), a macromolecule in the shape of a double helix with nitrogen-base "rungs" connecting the two helices. The different configurations of these nitrogen bases encode the genetic information for each organism, and, except for identical twins and other cases in which organisms effectively split into multiple separate organisms, are unique from organism to organism. Applied to humans, this uniqueness provides another method of distinguishing one person from another. A naive approach to such genetic comparison might involve direct comparison of base-pair sequences. The approximately three billion base pairs in each person's genome make this process infeasible, however. Instead, scientists sequence patterns in a person's DNA called variable number tandem repeats (VNTR), which are sections of the human genome that have no apparent genetic function. Since VNTRs vary widely between individuals, they too provide a way to differentiate between people through genetic material. This VNTR process was first developed as an identification technique in England by Jeffreys et al. [1995].

7.2   Constructing a DNA Fingerprint

A DNA "fingerprint" is constructed by first extracting a DNA sample from body tissue, hair, or bodily fluids such as blood, semen, or saliva. The sample is then amplified by polymerase chain reaction and restriction enzymes used cut the samples into smaller fragments. The segments are then sorted by size using a process called electrophoresis. The VNTRs are marked with radioactive probes and exposed on X-ray film, where they form a characteristic pattern of black bars. This pattern constitutes the DNA fingerprint. A sample DNA fingerprint from a rape case is displayed in Figure 6. The highlighted marks are, from left to right, a VNTR sample from the defendant, a VNTR ruler, and the VNTRs found in semen taken from the rape victim. The matching VNTRs in the suspect's sample and the semen sample support the claim that the semen originated from the defendant.

7.3   The Probability of a DNA Fingerprint

If the arrangement of VNTRs produced from two different samples match, the two samples probably came from the same person. The probability
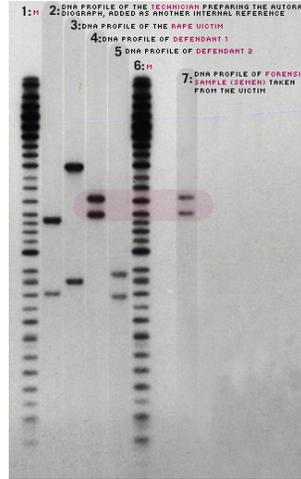
Figure 6: Sample DNA fingerprint. Adams [2002]

of two different patterns exhibiting the same VNTR by chance varies between 10^{-2} to 10^{-4}, depending on the VNTR (Roeder [1994], Woodworth [2001]). The total probability of an individual's DNA fingerprint occurring by chance is computed under the assumption that the VNTRs are independent. This independence has been verified for the ten most commonly used VNTRs by Lambert et al. [1995]. By testing multiple VNTRs, the probability of the tested configuration occurring by chance decreases dramatically. For instance, the probability of the DNA profile obtained from the stain on White House intern Monica Lewinsky's dress occurring by chance is reported to be 1 in 7.9 trillion (Adams [2002]).

7.4 Comparison of Traditional and DNA Fingerprinting

In order to compare the efficacy of DNA fingerprinting and traditional fingerprinting, we estimate reasonable probabilities of fingerprint pattern uniqueness in a forensic setting. While level-two data is often severely limited by the print quality, we expect level-one detail to remain relatively unchanged unless significant sections of the print are obscured or absent. As such, we use a value of 10^{-3} for p_1 from Equation 10, allowing for a very conservative loss of seven-eighths of the level-one information. Multiplying the entries in Table 5 by this level-one factor estimate of 10^{-3}, we observe that all but the three worst probabilities are less than one in a bil-

lion.

Traditional fingerprinting relies on the expert ability of forensic scientists at the crime scene to locate and lift prints from a variety of surfaces. The quality and duration of a latent print depends upon the surface the print was left on, as well as the humidity and temperature of the environment. While latent prints can last up to 40 years, this represents an extreme value and a latent print may degrade after less than a week. Most fingerprints are also extremely fragile and as a result are destroyed by contact with any other surface. In most cases, only partial fingerprints can be recovered for analysis (OnIn.com).

DNA, on the other hand, is a very resilient molecule and consequently can be quickly and completely recovered from a drop of dried blood, semen, or sweat. In older crimes where DNA from bodily fluids has degraded or is not available, it can still be extracted from human teeth and bones.

DNA fingerprinting has its flaws, however. Most significantly, false positives can arise by chance and by the mishandling of DNA samples. Thompson [2003] reports one such error which led to the false conviction of Timothy Durham for the rape of an 11-year-old girl. A DNA test showed that Durham's genotype matched that of the semen donor. The laboratory had failed to completely separate male from female DNA during extraction of the semen stain, however, and the combination of the VNTRs from the victim and the rapist produced a fingerprint that matched Durham's. Although false positives can occur (Thompson [2003]), the rate at which they occur is difficult to estimate on the basis of existing data.

8   Results and Conclusions

Until recently, fingerprint evidence had been used without major challenge in courts of law in the United States. In 1993, however, scientific evidence came under fire by the courts in the case of *Daubert v. Merrill Dow Pharmaceutical*. Recent court cases specifically threaten the validity of fingerprint evidence and have renewed interest in whether fingerprint individuality has scientific merit.

In this paper we present a model which determines whether fingerprints are unique. We consider both the topological structure of a fingerprint and the fine detail present in the individual ridges. Furthermore, we incorporate the uncertainties intrinsic in recording fingerprints and compute definite probabilities which suggest that fingerprints are reasonably

unique among all humans which have lived on Earth.

Fingerprint evidence compares well with DNA evidence in forensic settings. Our model predicts that with even a reasonably small fingerprint area and number of features, the probability that a match between a latent print and a suspect's print occurs by chance will be less than 1×10^{-9}. As with DNA evidence with few VNTRs, however, fingerprints of poor quality with few features can give inconclusive results. Both sets of evidence can result in uncertainties of false association that are too high to convict a suspect beyond reasonable doubt.

Although it addresses many vital criteria, our model can nevertheless be improved in several ways. For example, we assumed that minutiae are distributed uniformly and are independent of the placement and orientation of neighboring minutiae. We could increase the accuracy of our model by relaxing this assumption and including empirically observed distributions of minutiae. We assume that fingerprints are always of the highest quality. To accurately predict whether a latent print matches a suspect's print, this model should consider more explicitly smearing, smudging, and other image ambiguity and quality issues.

References

Julian Adams. DNA fingerprinting, genetics and crime: Dna testing and the courtroom. Found at `http://www.fathom.com/courese/21701758/`, 2002.

Mary Beeton. Scientific methodology and the friction rdige identification process. Found at `http://www.ridgesandfurrows.homestead.com/files/Scientific_Methodology.pdf`.

Raffaele Cappelli, Alessandra Lumini, Dario Maio, and Davide Maltoni. Fingerprint classification by directional image partitioning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(5):402–421, 1999.

Robert Epstein. Fingerprints meet *Daubert*: The myth of fingerprint "science" is revealed. *Southern California Law Review*, 75:605–658, 2002.

Francis Galton. *Finger Prints*. Macmillan, London, 1892.

Carl Haub. How many people have ever lived on earth? *Population Today*, November 2002.

A.J. Jeffreys, V. Wilson, and S.L. Thein. Individual-specific fingerprints of human DNA. *Nature*, 316:76–79, 1995.

J.A. Lambert, J.K. Scranage, and I.W. Evett. Large scale database experiments to assess the significance of matching DNA profiles. *International Journal of Legal Medicine*, 108:8–13, 1995.

Gian Luca Marcialis, Fabio Roli, and Paolo Frasconi. Fingerprint classification and of flat and structural approaches. *Lecture Notes in Computer Science*, 2091:241–246, 2001.

NCSC. National center for state courts. Found at `http://www.ncsconline.org/`.

OnIn.com. The history of fingerprints. Found at `http://onin.com/fp/fphistory.html`.

James W. Osterburg, T. Parthasarathy, T.E.S. Ragvahan, and Stanley L. Sclove. Development of a mathematical formula for the calculation of fingerprint probabilities based on individual characteristics. *Journal of the American Statistical Association*, 72:772–778, December 1977.

S. Pankanti, S. Prabhakar, and A. Jain. On the individuality of fingerprints. In *IEEE Transaction on Pattern Analysis and Machine Intelligence*, pages 1010–1025, August 2002.

Salil Prabhakar. *Fingerprint Matching and Classification Using a Filterbank*. PhD thesis, Michigan State University, 2001.

Ridges and Furrows. Ridges and furrows. Found at `http://www.ridgesandfurrows.homestead.com/`.

K. Roeder. DNA fingerprinting: a review of the controversy. *Statistical Science*, 9:222–247, 1994.

Stanley L. Sclove. The occurence of fingerprint characteristics as a two-dimensional process. *Journal of the American Staistical Association*, 74:588–595, September 1979.

David A. Stoney and John I. Thornton. A critical analysis of quantitative fingerprint individuality models. *Journal of Forensic Sciences*, 31(4):1187–1216, October 1986.

William Thompson. Examiner bias. Found at `http://www.bioforensics.com/conference/ExaminerBias/`, 2003.

C.I. Watson and C.L. Wilson. *NIST Special Database 4, Fingerprint Database*. U.S. National Institute of Standards and Technology, 1992.

James L. Wayman. Daubert hearing on fingerprinting: When bad science leads to good law: The disturbing irony of the *Daubert* hearing in the case of *U.S. v. Byron C. Mitchell*. Found at `http://www.engr.sjsu.edu/biometrics/publications_daubert.html`, 2000.

George Woodworth. Probability in court - DNA fingerprinting. Found at `http://www.stat.uiowa.edu/~gwoodwor/statsoc`, 2001.