# The Myth of "The Myth of Fingerprints"

Steven G. Amery
Eric Thomas Harley
Eric J. Malm
Harvey Mudd College
Claremont, CA

Advisor: Jon Jacobsen

## Summary

For over a century, fingerprints have been an undisputed personal identifier. Recent court rulings have sparked interest in verifying uniqueness of fingerprints.

We seek to determine precisely the probability of duplicate fingerprints. Our model of fingerprint structure 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.

We apply our model to the conditions encountered in forensic science, to determine the legitimacy of current methodology. We also compare the accuracies of DNA and fingerprint evidence.

Our model predicts uniqueness of prints throughout human history. Furthermore, 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 UMAP Journal* 25 (3) (2004) 215–230. 

# Introduction

## What is a Fingerprint?

A fingerprint is a two-dimensional pattern created by the friction ridges on a human finger [Beeton 2002]. Such ridges are believed to form in the embryo and to persist unchanged through life. The physical ridge structure appears to depend chaotically on factors such as genetic makeup and embryonic fluid flow [Prabhakar 2001]. When a finger is pressed onto a surface, the friction ridges transfer to it (via skin oil, ink, or blood) a representation of their structure.

Fingerprints have three levels of detail [Beeton 2002]:

1. Overall ridge flow and scarring patterns, insufficient for discrimination.
2. Bifurcations, terminations, and other discontinuities of ridges. The pairwise locations and orientations of the up to 60 such features in a full print, called *minutiae*, provide for detailed comparison [Pankanti et al. 2002].
3. The width of the ridges, the placement of pores, and other intraridge features. Such detail is frequently missing from all but the best of fingerprints.

## Fingerprints as Evidence

The first two levels have been used to match suspects with crime scenes, and fingerprint evidence was long used without major challenge in U.S. courts [OnIn.com 2003]. In 1993, however, in *Daubert v. Merrill Dow Pharmaceutical*, the U.S. Supreme Court set standards for "scientific" evidence [Wayman 2000]:

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.

Since then, there have been challenges to fingerprint evidence.

## Individuality of Fingerprints

Francis Galton [1892] divides a fingerprint into squares with a side length of six ridge periods and estimates that he can recreate the ridge structure of a missing square with probability $\frac{1}{2}$. Assuming independence of squares and introducing additional factors, he concludes that the probability of a given

fingerprint occurring is $1.45 \times 10^{-11}$. Pearson refines Galton's model and finds a probability of $1.09 \times 10^{-41}$ [Stoney and Thornton 1986].

Osterburg [1977] extends Galton's approach by dividing a fingerprint into cells that can each contain one of 12 minutia types. Based on independence among cells and observed frequencies of minutiae, he finds the probability of a configuration to be $1.33 \times 10^{-27}$. Sclove [1979] extends Osterburg's model to dependencies among cells and multiple minutiae in a single cell.

Stoney and Thornton [1986] charge that these models fail to consider key issues completely: the topological information in level-one detail; minutiae location, orientation, and type; normal variations in fingerprints; and number of positions considered. We try to correct some of these omissions.

# Our Model: Assumptions and Constraints

## Assumptions

- **Fingerprints are persistent:** they remain the same throughout a person's lifetime. Galton [1892] established this fact, in recent times verified from the processes of development of dermal tissues [Beeton 2002].
- **Fingerprints are of the highest possible quality,** without damage from abrasion and injury.
- **The pattern of ridges has some degree of continuity and flow.**
- **The ridge structure of a fingerprint is in one of five categories:** Arch, Left Loop, Right Loop, Tented Arch, or Whorl, employed in the automatic classification system of Cappelli et al. [1999] (derived from those of the FBI and Watson and Wilson [1992]). Each category has a characteristic ridge flow topology, which we break into homogeneous domains of approximately unidirectional flow. While Cappelli et al. [1999] raise the issue of "unclassifiable" prints, and they and Marcialis et al. [2001] confuse classes of ridge structures, we assume that such ambiguities stem from poor print quality.
- **Fingerprints may further be distinguished by the location and orientation of minutiae relative to local ridge flow.** Stoney and Thornton [1986] argue that the ridges define a natural coordinate system, so the location of a minutia can be specified with a ridge number and a linear measure along that ridge. Finally, minutiae have one of two equally likely orientations along a ridge.
- **Each minutia can be classified as a bifurcation, a termination, or a dot** (**Figure 1**) [Pankanti et al. 2002; Stoney and Thornton 1986]. Though Galton [1892] identifies 10 minutia structures and others find 13 [Osterburg et al. 1977], we can ignore these further structures (which are compositions of the basic three) because of their low frequency [Osterburg et al. 1977].

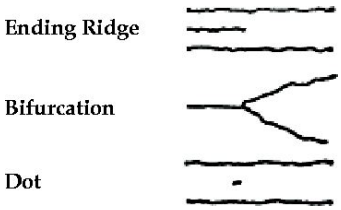


**Figure 1.** The three basic minutiae types (from Galton [1892]). We refer to ending ridges as *terminations*.

- **A ridge structure produces an unambiguous fingerprint, up to some level of resolution.** A ridge structure can vary in print representations primarily in geometric data, such as ridge spacing, curvature, and location of minutiae [Stoney and Thornton 1986]. Topological data—ridge counts, minutiae orientation, and ordering—are robust to such variation and are replicated consistently.

  A more serious consideration is connective ambiguities, such as when a given physical minutia is represented sometimes as a bifurcation and sometimes as a termination. But our highest-quality assumption dictates that such ambiguity arise only where the physical structure itself is ambiguous.

- **Location and orientation of minutiae relative to each other are independent,** though Stoney and Thornton [1986] find some dependency and Sclove [1979] model such dependency in a Markov process.

- **Ridge widths are uniform throughout the print and among different prints, and ridge detail such as pores and edge shapes is not significant.** While ridge detail is potentially useful, we have little data about types and frequencies.

- **Frequencies of ridge structure classes and configurations and minutiae types do not change appreciably with time.** We need this invariance for our model's probabilities to apply throughout human history.

## Constraints Implied by Assumptions

- Our model must consider ridge structure, relative position, orientation, and type of minutiae.

- Locations of minutiae must be specified only to within some uncertainty dependent on the inherent uncertainty in feature representation.

# Model Formulation

We examine a hierarchy of probabilities:

- that the given class of ridge structure occurs,

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

- that a region contains the specified number of minutiae,
- that the minutiae in this region follow the specified configuration, and
- that the minutiae occur with the specified types and orientations.

## Probability of Ridge Structure Class

To each of the five classes of ridge structures (Arches, Left and Right Loops, Tented Arches, and Whorls), we associate a probability of occurrence ($\nu_A$, $\nu_L$, $\nu_R$, $\nu_T$, $\nu_W$), which we estimate from observed frequency in the population.

## Probability of Ridge Structure Configuration

Each print is partitioned into regions in which the overall flow is relatively unidirectional, and the class of the print is determined from five prototypical masks characteristic of ridge-structure classes (**Figure 2**) [Cappelli et al. 1999]. The variations of flow region structure within each class then depend on parameters for the class. For example, the ridge structure of a Loop print can be determined from the locations of the triangular singularity and the core of the loop (**Figure 3**). To determine the probability of a particular region configuration, we determine the probability that the associated parameters occur.

Because of uncertainty in the parameters, we discretize the parameter space at the fundamental resolution limit $\delta_1$ (subscript indicates feature level). We use independent Gaussian distributions about the mean values of the parameters.

We now detail the parameter spaces for each ridge-structure class. The use of the prototypes requires an $X \times Y$ region within the print.

### Arch

The 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 consider as fixed the width $b$ of the central corridor. The ratio of the resolution limit $\delta_1$ to the mean length of a typical segment determines the uncertainty in the angular measurement of that segment.

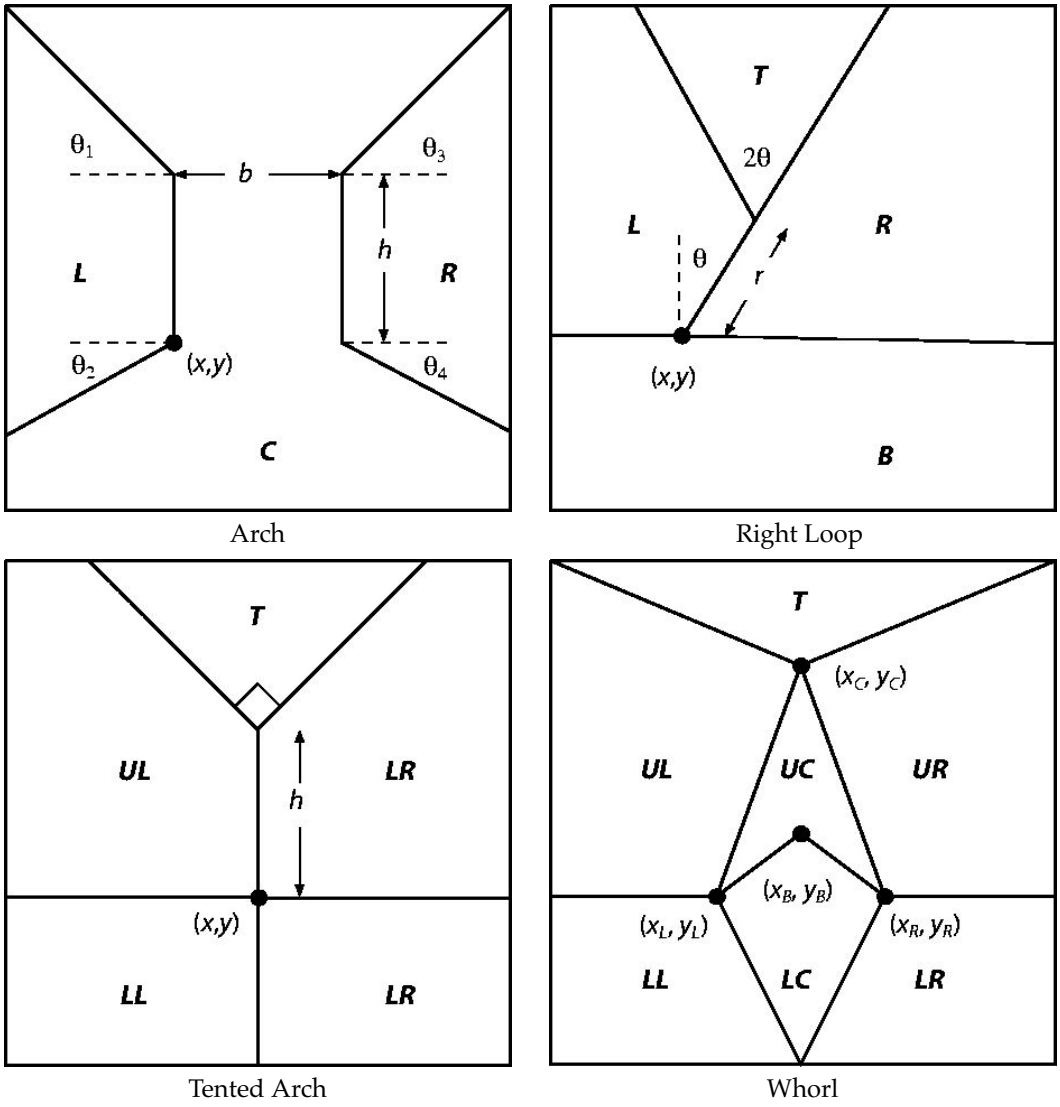


**Figure 2.** The prototypical region structures and parameters for each ridge structure class, derived from the masks in Cappelli et al. [1999].

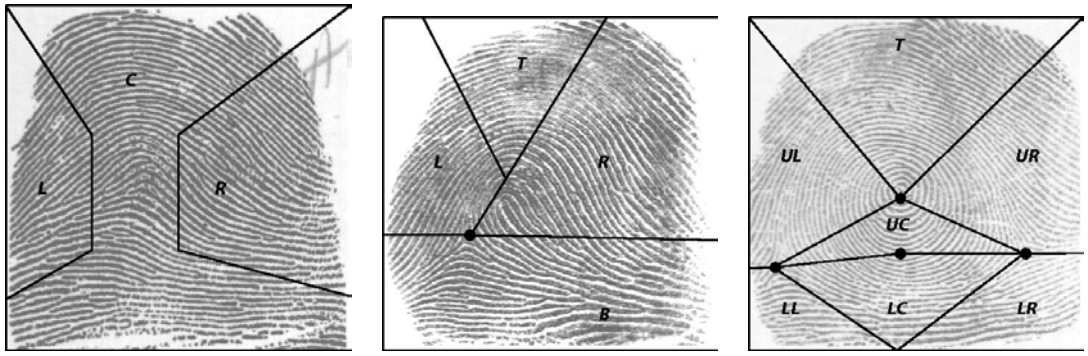


**Figure 3.** The prototypical region structures applied to an Arch, a Right Loop, and a Whorl.

### Loops, Left and Right

Since Left and Right Loops are identical except for a horizontal reflection, we use the same parameter space for both classes. The two principal features are the position $(x, y)$ of the triangular singularity outside the loop and the distance $r$ and angle $\theta$ of the core of the loop relative to this singularity.

### Tented Arch

The major structure is the arch itself; the parameters are the position $(x, y)$ of the base of the arch and the height $h$ of the arch.

### Whorl

The Whorl structure has four major features: the center of the whorl, $(x_C, y_C)$; the base of the whorl, $(x_B, y_B)$; and the triangular singularities to the left and right of the base of the whorl, at $(x_L, y_\cdot)$ and $(x_R, y_R)$. We assume that the center 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 lies above the singularities, so that $y_B \geq y_L$ and $y_B \geq y_R$.

## Probabilities of Intraregion Minutiae Distribution

Since the geometry of a region is uniquely determined by the configuration parameters, we can divide each unidirectional flow region into parallel ridges. We can represent the ridge structure of the region as a list of ridge lengths.

We assume a fundamental limit $\delta_2$ to resolution of the position of minutiae along a ridge and divide a ridge into cells of length $\delta_2$, in each of which we find at most one minutia. The probability $P_{TC}(n, l, k)$ that the $n$th ridge in the partition, with length $l$, has a particular configuration of $k$ minutiae is

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

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 $o_i$.

### Probability of Minutiae Number

Under the assumption that minutiae occur at uniform rates along a ridge, we expect a binomial distribution for the number of minutiae on the ridge. Denote the linear minutiae density on ridge $n$ by $\lambda(n)$. The probability that a minutia occurs in a given cell of length $\delta_2$ is $\delta_2 \lambda(n)$. Thus, the probability that $k$ minutiae occur is

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

### Probability of Minutiae Configuration

Assuming that all configurations of $k$ minutiae are equally likely along the ridge, the probability of the specified configuration is

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

### Probability of Minutiae Type and Orientation

The probability that minutiae occur with specified types and orientations is

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

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

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

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

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

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 that ridges develop minutiae independently.

Applying the assumption that $\lambda$ and other factors do not depend on $n$ and are hence uniform throughout the print, we can collapse these multiplicative factors to an expression for the configuration probability of the entire print:

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

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. The length $L$ 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.

# Parameter Estimation

For parameters in our model, we use published values and estimates based on the NIST-4 database [Watson and Wilson 1992].

## Level-One Parameters

All lengths are in millimeters (mm); angles are in radians or in degrees.

- **Level-one spatial resolution limit** $\delta_1$: Cappelli et al. [1999] discretize images into a $28 \times 30$ grid to determine level-one detail. From these grid dimensions, the physical dimensions of fingerprints, and the assumption of an uncertainty of three blocks for any level-one feature, we estimate $\delta_1 = 1.5$.
- **Level-one angular resolution limit** $\delta_\theta$: Taking $X/2 = 5.4$ (determined below) as a typical length scale, we have $\delta_\theta = \delta_1/5.4 = 0.279$ radians.
- **Ridge structure class frequencies** $\nu_A$**,** $\nu_L$**,** $\nu_R$**,** $\nu_T$**, and** $\nu_W$: We use the estimates in Prabhakar [2001] (**Table 1**).

**Table 1.** Relative frequencies of ridge structure classes (from Prabhakar [2001]).

| $\nu_A$ | $\nu_L$ | $\nu_R$ | $\nu_T$ | $\nu_W$ |
|---|---|---|---|---|
| 0.0616 | 0.1703 | 0.3648 | 0.0779 | 0.3252 |

- **Thumbprint width** $X$ **and height** $Y$: Examining thumbprints from the NIST-4 database and comparing them with the area given by Pankanti et al. [2002], we conclude that a width that covers the majority of thumbprints is 212 pixels in the 500 dpi images, a physical length of $10.8$ mm. Similarly, $Y = 16.2$ mm.
- **Arch parameters** $(x, y)$**,** $h$**,** $b$**,** $\theta_1$**,** $\theta_2$**,** $\theta_3$**, and** $\theta_4$: We restrict the parameter space for $(x, y)$ to the lower half of the thumbprint with horizontal margins of length $b$. We estimate $b = 2.5$ from examination of Arch fingerprints in the NIST database and from Cappelli et al. [1999]. This estimate places $x \in (0, 8.3)$ and $y \in (0, 5.6)$. The mean for $(x, y)$, which we need to describe the distribution of $(x, y)$, is then $(4.2, 2.8)$. We estimate that $x$ and $y$ both have a standard deviation of $0.7$. We assume that $\theta_1, \ldots, \theta_4$ are all between $0°$ and $45°$ with mean $22.5°$ and standard deviation $5.13°$.
- **Loop parameters** $(x, y)$**,** $\theta$**, and** $r$: For a left loop (a right loop is reflection of this), $(x, y)$ must lie in the bottom left quadrant and the mean coordinate pair is $(2.7, 2.8)$. Additionally, we restrict $\theta$ to lie between $15°$ and $75°$, which allows us to estimate the mean $\theta$ as $45°$ with a standard deviation of $15°$. We estimate that $r$ must be greater than $0$ and less than $9.6$.
- **Tented arch 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 $x \in (1.8, 9)$ and $y \in (0, 8.1)$. Assuming a symmetric distribution of $(x, y)$ yields $(x, y) = (5.4, 2.8)$ with a standard deviation of $0.7$ in both directions. Logically, we place $h$ between $0$ and $Y/2 = 8.1$. Again, assuming a symmetric distribution in this parameter space and a standard deviation of one-eighth the maximum value yields $h = 4.05 \pm 1.02$.

- **Whorl 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 $x = X/4$ and $x = 3X/4$ and $y = 0$ and $y = 2Y/3$. 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.

**Table 2** summarizes the estimates for these four classes of ridge structures.

**Table 2.** Parameter range estimates for the ridge structure classes. All lengths in millimeters (mm), angles in degrees.

| Arch parameter ranges | |
|---|---|
| $(x, y)$ | $(4.2, 2.8) \pm (0.7, 0.7)$ |
| $h$ | $4.05 \pm 0.7$ |
| $b$ | $2.5 \pm 0$ |
| $\theta_1$–$\theta_4$ | $22.5° \pm 5.13°$ |
| **Loop parameter ranges** | |
| $(x, y)$ | $(2.7, 2.8) \pm (0.7, 0.7)$ |
| $\theta$ | $45° \pm 15°$ |
| $r$ | $4.58 \pm 0.7$ |
| **Tented Arch Parameter Ranges** | |
| $(x, y)$ | $(5.4, 2.8) \pm (0.7, 0.7)$ |
| $h$ | $4.05 \pm 1.02$ |
| **Whorl parameter ranges** | |
| $(x_L, y_L)$ | $(2.7, 4.1) \pm (0.7, 0.7)$ |
| $(x_C, y_C)$ | $(5.4, 12.2) \pm (0.7, 0.7)$ |
| $(x_R, y_R)$ | $(8.1, 4.1) \pm (0.7, 0.7)$ |
| $(x_B, y_B)$ | $(5.4, 4.1) \pm (0.7, 0.7)$ |

## Level-Two Parameters

- **Level-two spatial resolution limit $\delta_2$:** We estimate $\delta_2$ by $r_0$, the spatial uncertainty of minutiae location in two dimensions [Pankanti et al. 2002], and propose $\delta_2 = 1$ for best-case calculations.

- **Relative minutiae type frequencies $p_d$, $p_b$, and $p_t$:** Almost every compound minutia can be broken into a combination of bifurcations and terminations separated spatially. Counting these compound minutiae appropriately, we determine the relative minutiae frequencies in **Table 3**.

- **Ridge period $w$:** We use $0.463$ mm/ridge for the ridge period, the distance from the middle of one ridge to the middle of an adjacent one [Stoney and Thornton 1986].

**Table 3.** Frequencies of simple minutiae types (from Osterburg et al. [1977]).

| $p_b$ | $p_t$ | $p_d$ |
|---|---|---|
| 0.356 | 0.581 | 0.0629 |

- **Mean number of minutiae per print** $\mu$: Under ideal circumstances, we discern 40 to 60 minutiae on a print [Pankanti et al. 2002]; we take $\mu = 50 \pm 10$.
- **Linear minutiae density** $\lambda$: We calculate $\lambda$ by dividing the average number of minutiae per a thumbprint $\mu$ 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 that there have been 100 billion humans [Haub 1995].

# Model Analysis and Testing

Let the probability that a print has a configuration $x$ be $p_c(x)$. Assuming that fingerprint patterns are distributed independently, the probability that two prints match is $p_c^2(x)$. The sum of these probabilities over the configuration space is the total probability that some match occurs.

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

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

## Level-One Detail Matching

We restrict each parameter to a region of parameter space in which we expect to find it and assume that it is uniformly distributed there. This approximation is enough to estimate order of magnitude, which suffices for our analysis. Then

$$p_{c1}(i) = \frac{\nu_i}{\left( \prod_{j \in V(i)} \frac{L_j}{\delta_1} \right)}, \qquad \textbf{(1)}$$

where $L_j$ is the range of parameter $j$ in $V(i)$, the set of parameters for a type-$i$ ridge structure. For **(1)** 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. The product is simply the total number of compartments in the

parameter space, since we assume a uniform distribution in that range. Calculating $p_{c1}(i)$ for each ridge structure type, and summing squares, we find the probability that two thumbprints have the same overall ridge structure:

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

## Level-Two Detail Matching

If we disregard the infrequent dot minutiae, we obtain the probability

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

for a configuration $j$ with $k$ minutiae, $k_b$ of which are bifurcations (and the rest ridges), placed in $C = XY/w\delta_2$ cells. If we simplify minutia-type frequencies to $p_b = p_t = 1/2$, 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

$$\begin{aligned} p_2 &= \sum_{k=0}^{C}\sum_{k_b=0}^{k}\left((\delta_2\lambda)^k(1-\delta_2\lambda)^{C-K}\frac{1}{4^k}\right)\binom{C}{k}\binom{k}{k_b}2^k \\ &= \sum_{k=0}^{C}(\delta_2\lambda)^{2k}(1-\delta_2\lambda)^{2(C-k)}\frac{1}{4^k}\binom{C}{k} \\ &= \left(\frac{5(\delta_2\lambda)^2 - 8\delta_2\lambda + 4}{4}\right)^C. \end{aligned}$$

Match probabilities for $\lambda = 0.13 \pm 0.03/\text{mm}$, $\delta_2 = 1$ mm, and $C = 250$ to $400$ cells range from $2.9 \times 10^{-23}$ to $9.8 \times 10^{-60}$; probabilities for the more realistic values $\lambda = 0.05 \pm 0.03/\text{mm}$, $\delta_2 = 2$–$3$ mm, and $C = 100$ to $250$ cells range from $3.7 \times 10^{-5}$ to $1.7 \times 10^{-47}$.

## Historical Uniqueness of Fingerprints

Denote the probability of a match of any two left thumbprints in the history of the human race by $p$ and the world total population by $N$. The probability of at least one match among $\binom{N}{2}$ thumbprints is

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

**Figure 4a** illustrates the probability of at least one match for $N = 10^{11}$, while **Figure 4b** shows a log-log plot of the probability for very small $p$-values. Since even conservative parameter values in the ideal case give $p \ll 10^{-30}$, our model solidly establishes uniqueness of fingerprints through history.

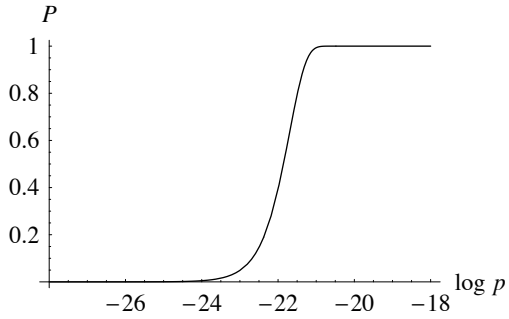


**Figure 4a.** For $N = 10^{11}$, probability of at least one thumbprint match through history.

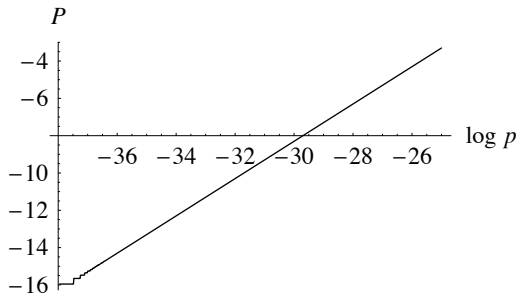


**Figure 4b.** Log-log plot of probability.

# Strengths and Weaknesses of the Model

## Strengths

- **Topological coordinate system**: We take topological considerations into account, as demanded by Stoney and Thornton [1986].
- **Incorporation of ridge structure detail**: We use this in addition to the minutiae detail that is the primary focus of most other models.
- **Integration of nonuniform distributions**: 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.
- **Accurate representation of minutia 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 [Osterburg et al. 1977; Pankanti et al. 2002] neglect some of this information.
- **Flexibility in parameter ranges**: We test a range of parameters in both ideal and practical scenarios and find that the model behaves as expected.

## Weaknesses

- **Ambiguous prints, smearing, or partial matches**: We assume that ambiguities in prints reflect ambiguities in physical structure and are not introduced by the printing. This is certainly not the case for actual fingerprints.
- **Domain discontinuities**: We have no guarantee of continuity between regions of flow; continuity requirements may affect the level-one matching probabilities significantly.

- **Nonuniform minutia distribution**: We assume that the distribution of minutiae along a ridge is uniform. However, models should account for variations in minutiae density and clustering of minutiae [Stoney and Thornton 1986]. Although we have a mechanism for varying this distribution, we have no data on what the distribution should be.
- **Left/right orientation distribution**: We assume that the distribution of minutiae orientation is 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].
- **Level-three information**: We neglect level-three information, such as pores and edge shapes, because of uncertainty about its reproducibility in prints.

# Comparison with DNA Methods

## DNA Fingerprinting

The genetic material in 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 configurations of these nitrogen bases encode the genetic information for each organism and are unique to the organism (except for identical twins and other cases in which an organism splits into multiple separate organisms).

Direct comparison of base-pair sequences for two individuals is infeasible, so scientists sequence patterns in a person's DNA called *variable number tandem repeats (VNTR)*, sections of the genome with no apparent genetic function.

## Comparison of Traditional and DNA Fingerprinting

While level-two data is often limited by print quality, we expect level-one detail to remain relatively unchanged unless significant sections of the print are obscured or absent. We use $p_1 = 10^{-3}$ from **(2)**, allowing for a conservative loss of seven-eighths of the level-one information. Multiplying by this level-one factor $10^{-3}$, all but the three worst probabilities are less than $10^{-9}$.

DNA fingerprinting has its flaws: False positives can arise from mishandling samples, but the frequency is difficult to estimate. The probability of two different patterns exhibiting the same VNTR by chance varies between $10^{-2}$ and $10^{-4}$, depending on the VNTR [Roeder 1994; Woodworth 2001]. The total probability of an individual's DNA pattern occurring by chance is computed under the assumption that the VNTRs are independent, which has been verified for the ten most commonly used VNTRs [Lambert et al. 1995].

# Results and Conclusions

We present a model that determines whether fingerprints are unique. We consider both the topological structure of a fingerprint and the fine detail present in the individual ridges. We compute probabilities that suggest that fingerprints are reasonably unique among all humans who have lived.

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 is less than $10^{-9}$. Both DNA evidence with few VNTRs and fingerprints of poor quality with few features can give inconclusive results, resulting in uncertainty beyond a reasonable doubt.

# References

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

Beeton, Mary. 2002. Scientific methodology and the friction ridge identification process. *Identification Canada* 25 (3) (September 2002). Reprinted in *Georgia Forensic News* 32 (3) (November 2002) 1, 6–8. `http://www.ridgesandfurrows.homestead.com/files/Scientific_Methodology.pdf` .

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

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

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

Haub, Carl. 1995. How many people have ever lived on Earth? *Population Today* 23 (2) (February) 4–5. Reprinted. 30 (8) (November/December 2002) 4–5. `http://www.prb.org/Content/ContentGroups/PTarticle/0ct-Dec02/How_Many_People_Have_Ever_Lived_on_Earth_.htm` .

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

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

Marcialis, Gian Luca, Fabio Roli, and Paolo Frasconi. 2001. Fingerprint classification by combination of flat and structural approaches. In *International*

*Conference on Audio- and Video-Based Biometric Person Authentication AVBPA01* (June 2001, Halmstad, Sweden), edited by J. Bigun and F. Smeraldi. *Springer Lecture Notes in Computer Science* 2091: 241–246.

National Center for State Courts (NCSC). 2004. `http://www.ncsconline.org/` .

German, Ed. 2003. The history of fingerprints. `http://onin.com/fp/fphistory.html` .

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

Pankanti, S., S. Prabhakar, and A. Jain. 2002. On the individuality of fingerprints. *IEEE Transaction on Pattern Analysis and Machine Intelligence*: 1010–1025. `http://citeseer.ist.psu.edu/472091.html` .

Prabhakar, Salil. 2001. Fingerprint matching and classification using a filterbank. Ph.D. thesis. East Lansing, Michigan: Michigan State University.

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

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

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

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

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

Wayman, James L. 2000. 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*. `http://www.engr.sjsu.edu/biometrics/publications_daubert.html` .

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