Math Behind Fingerprinting

Jimmy McLaughlin

December 2018

1 Introduction

Fingerprint identification has a number of practical applications, including forensics, background checks, biosecurity, identification of bodies, and identification of amnesia sufferers. It has been in use since the early 1900s, and is now a fixture of crime scene investigation. Fingerprints can be collected and compared to a database in order to attempt to make a positive identification. Additionally, the biochemical properties of fingerprints can be examined to determine sex. No two fingerprints are the same: this applies to identical twins, and even to two fingers on the same hand. However, fingerprints can be very similar, so in order to be used as evidence a fingerprint identification must be made with a great level of certainty.

2 Collection of Fingerprints

Fingerprints can be collected from both hard and soft surfaces. When on a soft surface, such as paint or soap, fingerprints leave a visible 3D impression which is very easy for scientists to examine. On a hard surface, collecting fingerprints can be trickier. If the person had some kind of visible substance on their hand (blood, ink, dirt, etc.) then they will leave what is called a **patent** (visible) print. Otherwise, the print will be **latent** (invisible). Since latent prints cannot be seen with the naked eye, scientists will often apply a fingerprint revealing dust to surfaces which are likely to contain fingerprints (such as doorknobs or windows). Sometimes however, this dust can corrupt the fingerprint, so before even using the dust scientists will shine a bright light on a surface or cover it in super glue to reveal prints.

3 Levels of Fingerprint Detail

There are three levels of fingerprint detail. They are:

- Pattern
- Minutiae

- Pores and sweat glands

This section will explain the differences between the three.

3.1 Fingerprint Patterns

A fingerprint pattern is the general shape formed by the ridges and valleys on the fingertip. Fingerprint pattern is determined genetically, and forms after about three months in the uterus. There are six types of fingerprint patterns: Arch, Tented arch, Right loop, Left loop, Whorl, and Twinloop. About 65% of people have loops, 30% have whorls, and 5% have arches.

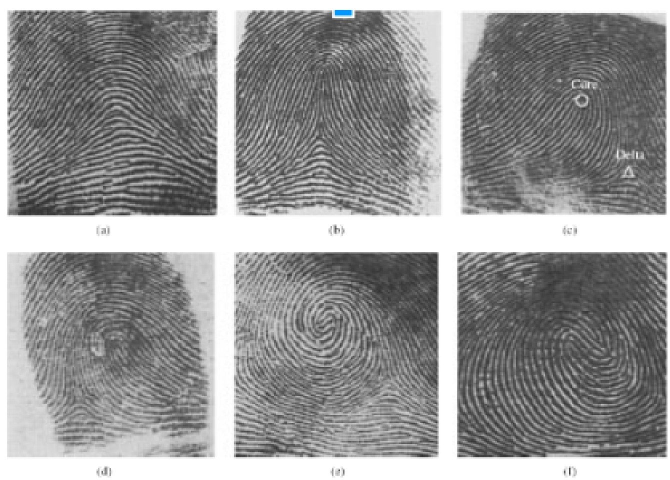

Figure 1: Types of Patterns

(a) Arch

(b) Tented arch

(c) Right loop

(d) Left loop

(e) Whorl

(f) Twinloop

Simply examining pattern type is not enough to make a fingerprint ID, since many people have each type. Thus it is the next level of detail, the minutiae, that will be key to identifying prints.

3.2 Minutiae

Minutiae are ridge-line details which form randomly and are different for every single print. The average fingerprint has 85 minutiae, with anywhere from 20-70 of these showing up in images. There are ten types of minutiae:

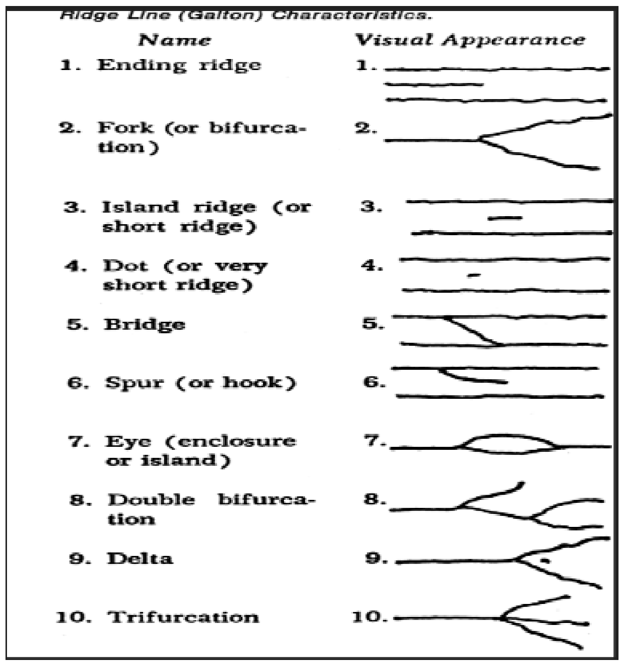

Figure 2: Types of Minutiae

Essentially, a minutiae occurs whenever there is a break in the continuous curve of the ridges. Any end or fork in the fingerprint ridges is classified as a minutiae point. There are different methods that can be used to identify location of minutiae. One is to superimpose a picture of the fingerprint over a grid composed of 1x1 millimeter squares as such:

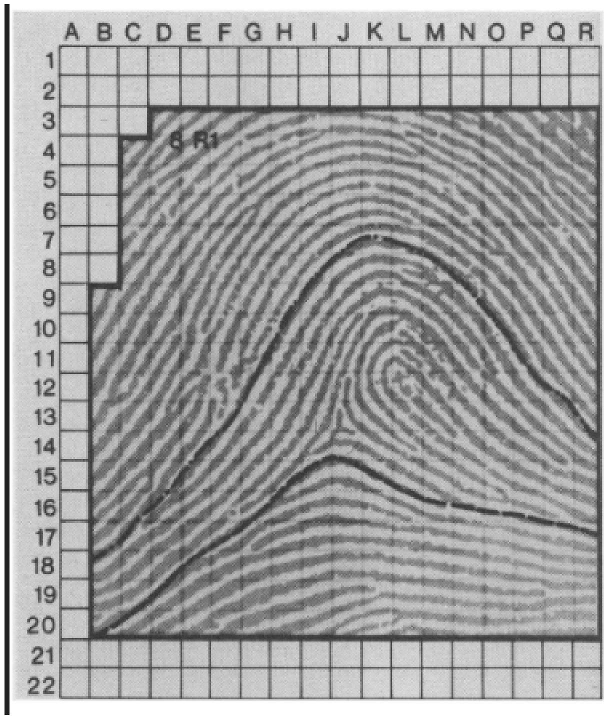

Figure 3: Fingerprint Grid

Scientists will take note of instances of minutiae in the cells. For example, F12, J13, K5, and L4 all contain minutiae. Minutiae points are classified based on their X-Y coordinates.

3.3 Pores and Sweat Glands

Oil on the hands allows for scientists to create an image of a person's sweat gland formation. In 2014, South Korean scientists developed a way to map sweat pores using a blue polymer which turns red when it interacts with any water. Currently, this method is not yet accurate enough (only about an 88% success rate) to be used as forensic evidence, but it is a revolutionary technology which could be on the horizon.

4 Using Math to Match Fingerprints

There are multiple approaches to mathematically evaluating whether two given fingerprints are the same. This section will discuss several of these methods.

4.1 Matching Fingerprints using Probability

Images of two fingerprints are observed to determine whether or not they are the same. This is done by examining the minutiae points. When two minutiae points of the same type are observed in the same cell, this is considered to be a minutiae match. The more matches that are observed, the higher the likelihood that the prints are the same. In a sample of 50,000 prints, the mean number of minutiae matches between two non-identical prints is 0.5, and the standard deviation is 1.83. So if a high number of minutiae matches is observed, there is an extremely high likelihood that the prints are identical. This information can be used to derive a formula expressing the probability of getting at least n number of matches between two prints: $P(X \geq n) = P(Z \geq \frac{(n-\mu)}{\sigma})$ where X is the number of observed minutiae matches, μ is the mean, σ is the standard deviation, and Z is the Z-score. So for example the probability of getting at least 27 matches is: $P(Z \geq \frac{(27-0.5)}{1.83}) = P(Z \geq 14.5) = 2.03 * 10^{-46}$. In order for a positive legal identification to be made, the probability of the observed minutiae configuration must be 10^{-20} or less. Another way to say this is that the probability of a false positive must be 10^{-20} or less for a fingerprint match to be considered evidence. Since some types of minutiae are more common than others, a system of weights has been developed to express how much an individual observation of a certain type contributes towards making a positive ID. For each type, the weight is the negative log probability of it being observed in a given cell. The weights for each minutiae type are shown in the third column of this chart:

5. Weights for the Characteristics

Cell configuration	*Estimate of probability, $\hat{p}_i$*	*Weight, $-\log_{10}\hat{p}_i$*
Delta	.00198	2.70
Island	.0177	1.75
Bridge	.0122	1.91
Spur	.00745	2.13
Dot	.0151	1.82
Ending ridge	.0832	1.08
Fork	.0382	1.42
Lake	.00640	2.19
Trifurcation	.000582	3.24
Double bifurcation	.00140	2.85
Broken ridge	.0139	1.86
Other multiple occurrence	.0355	1.45
No characteristic (empty cell)	.766	0.116

Figure 4: Minutiae Type Weights

Since the probability of a false positive must be $\leq 10^{-20}$, the equation $2.70(D) + 1.75(I) + ... + 0.116(E) \geq 20$ must be satisfied to make a positive identification, where D is the number of cells containing a delta, I is the number of cells containing an island, E is the number of empty cells, and so on. Thus observing one cell of any type is not enough to make an ID: there needs to be multiple cells matching between two images before they can be declared identical.
The second column on the above chart shows the probability of observing each type. This can be expanded to create a multinomial formula expressing the probability of getting a given type of minutiae distribution: $P = (p_0)^{k_0}(p_1)^{k_1}....(p_{12})^{k_{12}}$ where p_i is the probability of a minutiae type occurring and k_i is the number of observations of that type. So for example, to calculate the probability of observing 5 deltas and 4 islands in a sample of 9 cells, the p-values corresponding to delta (0.00198) and island (.0177) can be inserted into the multinomial formula to obtain $P = (.00198^5)(.0177^4) = 2.9869 * 10^{-21}$ (note that 5 and 4 are the k-values corresponding to each minutiae type). Thus, probability can be used to determine whether or not two fingerprints are identical.

4.2 Matching Fingerprints using Similarity Metrics

Metrics are a mathematical tool whereby a function assigns a distance between every two elements of a set. The notation for a metric is $D(A, B) = x$, where x is the distance and A and B are set elements. They are a commonly used tool in fingerprint comparison, and can be defined in one of two ways:

- Discrete Fingerprint Metric: If two fingerprints, A and B, are identical, then $D(A, B) = 0$. If they are not identical, then $D(A, B) = 1$.

- Continuous Fingerprint Metric: For two fingerprints, A and B, let r be a real number representing the **dissimilarity** between them, such that $D(A, B) = r$. (It is at the discretion of scientists how to define this dissimilarity value).

These two definitions can be combined to create a new metric that evaluates whether or not two fingerprints are the same based on the size of their dissimilarity number, r, using the following method:

- Begin by taking many samples of the same fingerprint.
- Let S be the set of all dissimilarity numbers ($r = D(A, B)$) between any two elements A and B in the sample.
- Create a constant c to be the maximum element of this set S. Essentially, c represents the maximum amount of dissimilarity that can occur between two **identical** fingerprints.

- Define the metric as such: For any two fingerprints, A and B, if $D(A, B) \leq c$, then they are the same print. If $D(A, B) > c$, then they are not the same print.

This metric provides a way of mathematically evaluating whether or not two fingerprints are identical.

4.3 Matching Fingerprints using Minutiae Graphs

Identifying the distribution of minutiae can help evaluate whether or not two prints are identical. Topological displays of minutiae, known as Minutiae Voronoi Diagrams, can be generated by the following rules:

- Locate all the minutiae on a given fingerprint image.
- Assign to every minutia a planar sub-region known as a cell, which contains all points that are closer to the given minutia than to any other minutia in the image.

A fingerprint image and its corresponding Minutiae Voronoi Diagram are shown:

Figure 5: Fingerprint Image

Note that each black dot represents a minutiae point.

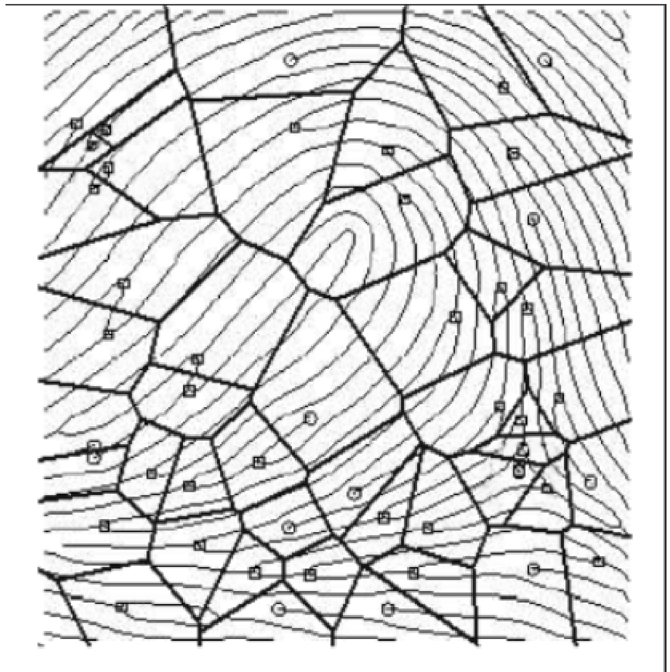

Figure 6: Minutiae Voronoi Diagram

The dots represent the minutiae points, and the black lines around them represent the cells. This diagram can be used to generate a minutiae graph. The goal of the graph is to represent possible configurations of minutiae while ensuring that identical prints are matched correctly (satisfying the condition $D(A,B) \leq c$ as discussed earlier). Graphs will be analyzed to determine whether or not they are isomorphic, using the Fingerprint Similarity Metric, which states the following:

- Two fingerprints, A and B are equivalent (D(A,B) = 0) if their graphs are isomorphic and distinct (D(A,B) = 1) if their graphs are not isomorphic.

The advantage of using isomorphism as a way of testing minutiae graphs is that if a fingerprint image is accidentally rotated, scaled, or translated incorrectly it will not affect the results since isomorphism only examines structure. In the minutiae graph, the cells are represented as black lines, just as in the Minutiae Voronoi Diagram, but the dots are represented as plus signs:

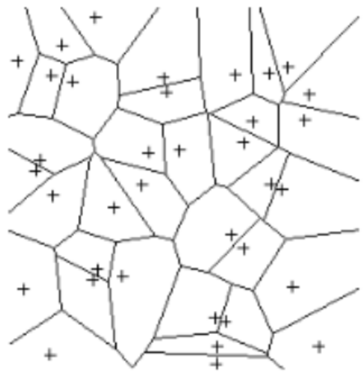

Figure 7: Minutiae Graph

(Note that the fingerprint shown in this graph is not the same as the one shown in the Minutiae Voronoi Diagram). Comparing this graph to another minutiae graph, shown in figure 8, it can be concluded that the two fingerprints are not the same since their graphs are not isomorphic:

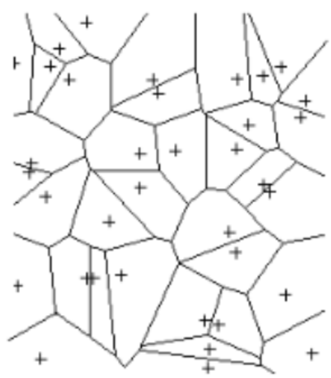

Figure 8: Minutiae Graph

The number of possible isomorphism classes N corresponding to a given number of minutiae points n can be given by the following equation:
$N = 14e^{0.8(n-6)}$
Observe that even a small number of minutiae points will have a very high number of corresponding isomorphism classes (ex: $n = 20$ leads to $N = 1,023,826$). This value N represents the number of distinct fingerprint possibilities based on a sample of n minutiae points.

5 Connections to Other Aspects of Math

Fingerprint analysis has a surprising amount of math behind it, and the topics I explored drew on concepts from Linear Algebra, Abstract Algebra, Probability and Statistics, and Elementary Analysis. Additionally, there are a lot of differential equations used in fingerprint analysis, although I did not look into these very much because they are pretty complex and would require a lot of time to understand and explain. There are several other student talks which could be connected to mine in various ways. Xinyao talked about reaction diffusion systems and their role in the formation of animal fur patterns. Since minutiae form randomly in the womb, their origins likely follow a similar sort of method. Charlie's presentation on predicting and solving crimes using math focused more on social and population predictors, but fingerprinting ties into the general category of using math to save lives and catch criminals. Chengwu and Hanmi's talk on matching students and advisors used a similar system of weights to what I discussed in my section on fingerprinting using probability. Obviously this topic isn't something most people will use direcly unless they go into forensics, but it provides some interesting mathematical approaches to a field that is a mixture of science and guesswork.

References

[1] "Fe-Fi-Fo-Thumb." *Math.washington.edu*, University of Washington, 9 Feb. 2004. sites.math.washington.edu/~morrow/mcm/uw28_04.pdf.

[2] Osterburg, James W, et al. "Development of a Mathematical Formula for the Calculation of Fingerprint Probabilities Based on Individual Characteristics." *Taylor and Francis Group, Journal of the American Statistical Association, Vol. 72, No. 360*, Dec. 1977. www.jstor.org/stable/pdf/2286458.pdf?refreqid=excelsior%3Aa849393ae3bde882407d905a7a33ff15.

[3] Huppenthal, John, et al. "Mathematical Modeling: Fingerprint Analysis." Semanticscholar.org, 10 May 2013. pdfs.semanticscholar.org/4d5a/3c599a03b938b0092874fbf24ec55ebf0500.pdf.

[4] "Fingerprint Analysis." *DNA Evidence: How It's Done*, NFSTC. www.forensicsciencesimplified.org/prints/how.html.

[5] "Fingerprint Analysis." *DNA Evidence: How It's Done*. www.forensicsciencesimplified.org/prints/why.html.

[6] Serle, Jack. "What Determines Your Fingerprint Pattern?" *Science Focus - BBC Focus Magazine*, BBC Focus Magazine. www.sciencefocus.com/the-human-body/what-determines-your-fingerprint-pattern/.

[7] Fierro, Pamela Prindle. "A Closer Look at the Fingerprint Patterns of Twins." *Verywell Family*, Verywellfamily, 4 Sept. 2018. www.verywellfamily.com/twins-and-fingerprints-2446684.

[8] Thakkar, Danny. "Minutiae Based Extraction in Fingerprint Recognition." *Bayometric*, Bayometric, 8 Aug. 2018. www.bayometric.com/minutiae-based-extraction-fingerprint-recognition/.

[9] Wogan, Tim. "Fingertip Sweat Pore Maps to Catch Criminals." *Chemistry World*, The Royal Society of Chemistry, 14 May 2014. www.chemistryworld.com/news/fingertip-sweat-pore-maps-to-catch-criminals/7361.article.

[10] Walvisch, Jamie. "Fingerprinting to Solve Crimes: Not as Robust as You Think." *The Conversation*, The Conversation, 23 Oct. 2017. theconversation.com/fingerprinting-to-solve-crimes-not-as-robust-as-you-think-85534.

[11] "Metric (Mathematics)." *Wikipedia*, Wikimedia Foundation, 7 Dec. 2018. en.wikipedia.org/wiki/Metric_(mathematics).

[12] "Fingerprint." Wikipedia, Wikimedia Foundation, 30 Nov. 2018. en.wikipedia.org/wiki/Fingerprint.