

# Fingerprint Analysis using Bayesian Alignment

Peter G. M. Forbes and Steffen Lauritzen

Department of Statistics, University of Oxford  
1 South Parks Road, Oxford OX1 3TG, U.K.

## Abstract

We present a model for fingerprint matching using Bayesian alignment on unlabelled point sets. An efficient Monte Carlo algorithm is developed to calculate the marginal likelihood ratio between the hypothesis that two observed prints originate from the same finger and the hypothesis that they originate from independent fingers. Our model achieves good performance on the NIST-FBI fingerprint database of 258 matched fingerprint pairs.

## 1 Motivation

Fingerprint evidence has been used for identification purposes for over one hundred years. Despite this, there has been little scientific research done on the discriminatory power and error rate associated with fingerprint identification. Within the last ten years there has been a push to give the analysis of fingerprint evidence a solid probabilistic framework, culminating in the recent paper by Neumann et al. (2012).

We develop an efficient Monte Carlo algorithm to estimate the marginal likelihood ratio between the hypothesis that two observed prints originate from the same finger and the hypothesis that they originate from independent fingers. In a courtroom presentation of fingerprint evidence the first hypothesis would be favoured by the prosecution whereas the second would be favoured by the defense. Our model could also be used as the basis of an automated fingerprint identification system (AFIS); see for example Maltoni (2009).

## 2 Fingerprint parameterization

Fingerprint evidence is based on the similarity of two or more pictures (see fig. 1a for an example). It is very difficult to put all the information from these pictures into a mathematically convenient form. Thus most fingerprint models, including Neumann et al. (2012), consider only a subset of the information: namely, the points of the image where a ridge either ends or bifurcates. These points are called *minutiae* (see fig. 1c) and contain sufficient information to uniquely identify an individual (Maltoni, 2009; Yager and Amin, 2004). A typical full fingerprint contains 100–200 minutiae, while a low quality crime scene latent fingermark may contain only one dozen (Garris and McCabe, 2000).

Each minutia consists of a position in the complex plane  $\mathbb{C}$ , an orientation on the unit circle  $\mathbb{S}^1$ , and a type (ridge ending or bifurcation) in  $\mathbb{Z}_2$ . The observed set of minutiae on a fingerprint is called a *minutia configuration* and denoted by  $M$  and represented as a point configuration on  $\mathbb{M} = \mathbb{C} \times \mathbb{S}^1 \times \mathbb{Z}_2$ . This representation is not unique. It depends on our choice of origin, orientation and scale. We must therefore ensure that our inferences are invariant under similarity transformations (see *e.g.* Dryden and Mardia 1998).

Given two minutia configurations  $A$  and  $B$ , we wish to compute the marginal likelihood for the two hypotheses

$$\begin{aligned} H_p &: A \text{ and } B \text{ originate from the same finger} \\ H_d &: A \text{ and } B \text{ originate from independent fingers.} \end{aligned} \tag{1}$$

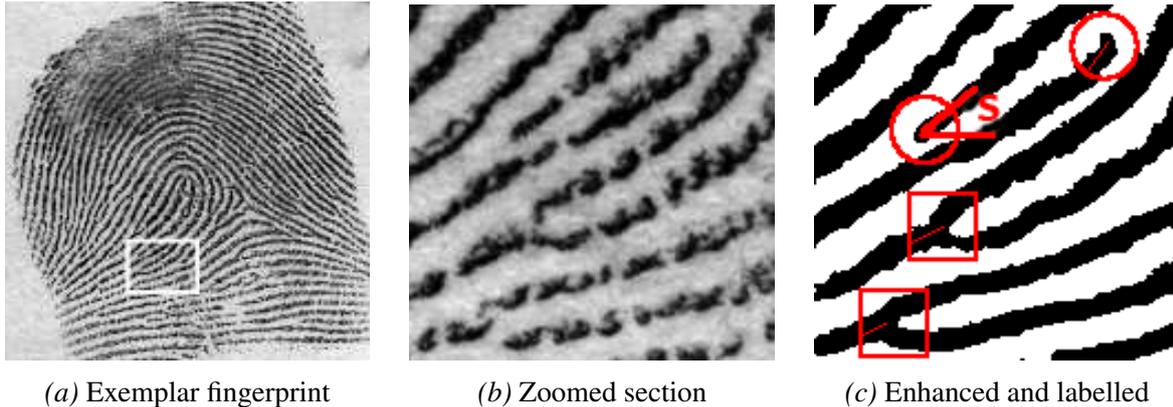


Figure 1: A typical exemplar quality fingerprint from Garris and McCabe (2000). The highlighted points in (c) are minutiae: circles are ridge endings and squares are bifurcations.

Since the fingerpad deforms whenever it is pressed against a recording surface, the apparent location and orientation of the minutiae in  $A$  and  $B$  differ. Some minutiae will be missing entirely. Any model to compare two minutia configurations must take both deformation and missing minutiae into account.

### 3 Description of our model

Lauritzen et al. (2012) note the similarity between minutiae matching and the alignment problems often studied in bioinformatics. We adapt the hierarchical Bayesian model for unlabelled point set matching of Green and Mardia (2006) to the problem of fingerprint matching.

We model the latent minutia configuration  $M$  as a Poisson point process (see *e.g.* Møller and Waagepetersen 2004) on  $\mathbb{M}$ . The observed minutia configurations  $A$  and  $B$  are partial, deformed copies of  $M$ .

We account for deformations and observation errors via the deformation functions  $D_A$  and  $D_B : \mathbb{M} \rightarrow \mathbb{M}$ . To allow for unobserved minutiae, we suppose  $A$  and  $B$  are independent thinnings of  $D_A(M)$  and  $D_B(M)$  respectively.

### 4 Results

The marginal likelihood ratio is estimated by a reversible jump Markov chain Monte Carlo algorithm (Green, 1995). Our algorithm was implemented in C# and generates approximately  $10^5$  samples per thread per second on a 3GHz Intel Xeon processor.

To test the algorithm on real world data, we use a small database provided by the National Institute for Standards and Technology and the Federal Bureau of Investigation (Garris and McCabe, 2000). This database consists of 258 fingermarks and their corresponding exemplar fingerprints. All marks and prints have their minutiae labelled by expert fingerprint examiners. The likelihood ratios are shown in fig. 2.

### 5 Conclusion and Future Work

We have described a simple model for quantifying the evidence two fingerprints originate from the same finger. We can efficiently sample from the posterior distribution of this model using MCMC methods. Preliminary tests show that the model is often capable of distinguishing true fingerprint matches from false fingerprint matches.

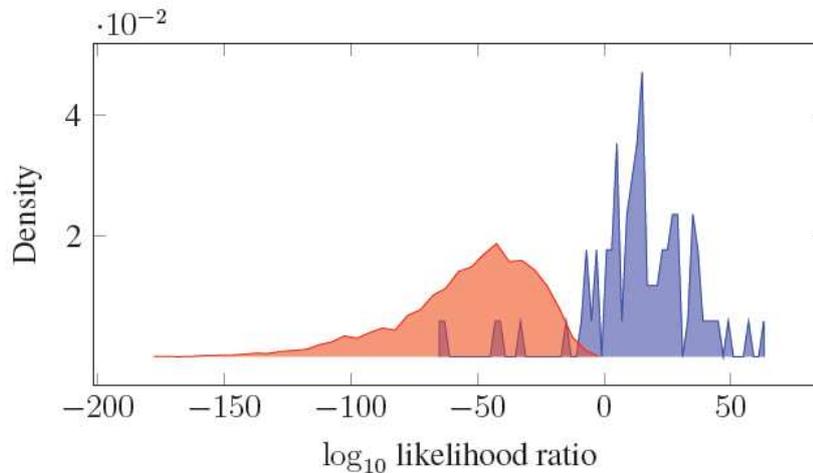


Figure 2: Estimated likelihood ratios. Red corresponds to prints from distinct fingers while blue corresponds to prints from the same finger.

In future work we will enhance our model to better capture the complicated spatial distribution of the minutiae. We also plan to conduct a thorough validation of our model assumptions against large databases of fingerprints.

## References

- Dryden, I. and K. Mardia (1998). *Statistical Shape Analysis*. Chichester: Wiley.
- Garris, M. and R. McCabe (2000). NIST special database 27: Fingerprint minutiae from latent and matching tenprint images. Technical report, NIST, Gaithersburg, MD, USA.
- Green, P. J. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika* 82(4), 711-732.
- Green, P. J. and K. V. Mardia (2006). Bayesian alignment using hierarchical models, with applications in protein bioinformatics. *Biometrika* 93(2), 235-254.
- Lauritzen, S., R.G. Cowell, and T.Graversen (2012). Discussion on the paper by Neumann et al. (2012). *Journal of the Royal Statistical Society: Series A* 175(2), 405-406.
- Maltoni, D. (2009). *Handbook of Fingerprint Recognition* (2nd ed.). New York: Springer-Verlag.
- Møller, J. and R. P Waagepetersen (2004). *Statistical Inference and Simulation for Spatial Point Processes*. Boca Raton: Chapman and Hall/CRC.
- Neumann, C., I. W. Evett, and J.E. Skerrett (2012). Quantifying the weight of evidence from a forensic fingerprint comparison: a new paradigm (with discussion). *Journal of the Royal Statistical Society: Series A* 175(2), 371-415.

Yager, N. and A. Amin (2004). Fingerprint verification based on minutiae features: a review.  
*Pattern Analysis and Applications* 7(1), 94-113.