

The author(s) shown below used Federal funds provided by the U.S. Department of Justice and prepared the following final report:

Document Title: Quantitative Measures in Support of Latent Print Comparison

Author(s): Sargur N. Srihari

Document No.: 241288

Date Received: February 2013

Award Number: 2009-DN-BX-K208

This report has not been published by the U.S. Department of Justice. To provide better customer service, NCJRS has made this Federally-funded grant report available electronically.

Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Quantitative Measures in Support of Latent Print Comparison

FINAL TECHNICAL REPORT: NIJ Award Number: 2009-DN-BX-K208

Sargur N. Srihari
University at Buffalo, The State University of New York
Buffalo, New York 14260
Email: srihari@buffalo.edu

SUBMITTED TO:

U.S. Department of Justice, Office of Justice Programs
National Institute of Justice
810 Seventh Street N.W.
Washington, DC 20531

AWARDEE:

Research Foundation of the State University of New York

January 5, 2013

Abstract

Latent prints of friction ridge impressions have long been useful in identification, and the methodology of examining latent prints, known as ACE-V (analysis, comparison, evaluation and verification), has been well-documented. The need to quantify confidences within ACE-V has been articulated in several recent influential reports to strengthen the science of friction ridge analysis. This research addresses the evaluation of three quantitative measures: rarity of features, confidence of opinion and a probabilistic measure of similarity. The first of these, useful in the analysis phase of ACE-V, is to determine the rarity of observed features. Rarity is difficult to compute due to the large number of variables and high data requirements. The proposed solution uses probabilistic graphical models to represent spatial distributions of fingerprints represented at *level 2 detail (minutiae)*. First, the minutia coordinate system is transformed into standard position based on a point of high curvature, viz., *core point*; statistical regression (based on a Gaussian process formulation and a training set of latent prints) is used to estimate the core point. A directed probabilistic graphical model is constructed using inter-minutia dependencies and minutia confidences. The resulting model is used to determine the probability of random correspondence of the evidence in a database of n prints. The method is validated using statistical goodness-of-fit tests and illustrated using: (i) a simple configuration of minutiae, (ii) randomly selected latent fingerprints in a database, and (iii) a well-known case of erroneous identification. The second quantitative measure addressed is that of determining confidence of opinion—which is relevant to the evaluation phase of ACE-V. The proposed computation determines a likelihood ratio as a product of rarity and the probability of similarity under the identification hypothesis. The third area concerns measuring similarity probabilistically, in a manner analogous to cognition. A Markov random field is used to jointly model the minutiae in both the evidence and the known. The algorithms developed are based on statistical machine learning using several publicly available fingerprint data sets for parameter learning and testing. The developed methods can be of potential use in examiner training, presentation of opinion and validating examination procedures.

Contents

1	Executive Summary	2
2	Research narrative	7
2.1	Introduction	7
2.1.1	Current Practice	7
2.1.2	Statement of the problem	10
2.1.3	Literature review	11
2.1.4	Rationale for the research	14
2.2	Methods	14
2.2.1	Registration	14
2.2.2	Modeling Distribution of Minutiae	19
2.2.3	Rarity Evaluation	23
2.2.4	Determining the Probability of Identification	29
2.2.5	A Probabilistic Measure of Similarity	33
2.3	Discussion of findings	52
2.4	Implications for policy and practice	53
2.5	Implications for further research	53
2.6	Dissemination	54
2.6.1	Publications	54
2.6.2	Presentations	54
2.6.3	Students	54
3	References	55

Chapter 1

Executive Summary

Latent prints of friction ridge impressions have long been useful in identification. The methodology of examining latent prints, known as ACE-V (analysis, comparison, evaluation and verification), has been well-documented. Yet, the need to characterize uncertainty in latent print examination methodology has been articulated in several recent influential reports to strengthen the science of friction ridge analysis and its use in the criminal justice system. Although the evaluated probabilities are likely to point to near uniqueness of features and near certainty of individualization they will help validate existing procedures. This research addresses the evaluation of quantitative measures within the ACE-V process relating to “uniqueness” of features, and “individualization” as an opinion.

The quantitative measures developed pertain to two different stages within the ACE-V process. In the analysis phase the rarity of a given configuration of minutiae is important. In the comparison and evaluation phases, the examiner expresses his/her opinion based on consideration of both similarity and rarity. The following computational problems are explored: (i) how to model the probability distributions of fingerprint minutiae, (ii) how to determine the probability of finding a given configuration of minutiae in a database of n prints, and (iii) how to utilize these probability measures, together with probability of similarity between latent print and known, in associating a probability with an opinion.

The methods developed use machine learning approaches to several subproblems. The first problem considered is that of establishing a reference point for the spatial distribution of minutiae. A method based on regression is used to predict the core point of a latent print, where the input is an orientation map that captures information about ridge flow and target variables are the location and orientation of the core point. The particular method of regression employed is based on Gaussian processes which has the advantage of predicting a distribution for the core point rather than a point estimate. The proposed method is evaluated on a test set of latent prints and seen to provide better results than previously known non-statistical approaches used in automatic fingerprint identification systems (AFIS). The method is useful even when a core point is absent in the actual print, e.g., due to a small latent print or the friction ridge pattern has no core point.

To deal with complex distributions of minutiae, we use probabilistic graphical models which take advantage of some independencies while not ignoring strong dependencies. The resulting model is validated in several ways: statistical tests as well as with several intuitive examples. Using chi-squared goodness-of-fit tests, the resulting model is seen to be more accurate than a model which assumes complete minutiae independence. Using the model the probability of finding an input configuration of minutiae in a database of n prints can be determined. The method is illustrated with several examples including simple synthetic images, latent prints from a standard database and a latent print from a well-known case.

The likelihood ratio (LR) measure has been proposed as method to express the strength of opinion in several forensic domains. It is defined as the ratio between the joint probability that the evidence and known come from the same source, and the joint probability that the two come from two different sources. The LR has also been proposed as being suitable for comparing a latent print with a known. LR evaluation depends on the underlying probability distributions of features in both the evidence (latent print) and the known. In the friction ridge domain these distributions are difficult to compute since the input images may be partial, the features such as minutiae are uncertain, the spatial distributions of features are difficult to determine,

and the number of features are large making the computation intractable. We propose a tractable method for determining LR's using an approximation involving rarity as well as the distribution of similarity. Once the LR is determined it can be readily converted into a probability of identification/exclusion.

There are several methods for measuring similarity between a latent print and a known, including scores produced by AFIS. However there is little cognitive justification. A probabilistic measure of similarity is proposed. It models the joint distribution of both the evidence and the known using a Markov random field. The method is seen to perform better than classical fingerprint similarity measures based on geometrical configurations such as pairs and polygons of minutiae.

Several algorithms were developed in this research. They are all based on principles of machine learning and probabilistic graphical models, taking into account the computational intractability of dealing with a large number of variables. Several publicly available fingerprint data sets were used for learning parameters and testing.

The methods developed can be of potential use in examiner training, e.g., in the selection of features to be compared, presentation of opinion, e.g., opinion can be accompanied by a probability, and validating examination procedures, e.g., justification for arguments of uniqueness and individualization. The work was disseminated at several pattern recognition and forensics meetings. Several publications resulted from the work. It also supported two doctoral students both of whom developed doctoral dissertations on this topic.

List of Figures

2.1	Types of Fingerprint images used in this research: (a) rolled (NIST SD14), (b) rolled (NIST SD27), (c) plain (live-scan) (FVC2002), and (d) latent (NIST SD27).	8
2.2	Example to illustrate determining the core point of a finger print using regression: (a) fingerprint image and (b) orientation map (32×32 vector of gradient values) used as the input variable. The corresponding target variable is the core point which in this case has values $s = (253, 221)$, and $\theta = 85$	15
2.3	Core point prediction using GP regression for two latent prints (from <i>NIST27</i>): (a) image <i>g90</i> which contains a latent print within a rectangle, (b) computed orientation map containing the predicted core point (cross) and true core point (circle), (c) ten-print where the true core point is visible, and (d-f) image <i>g69</i> with corresponding images. Note that in the second case the predicted core point lies outside of the latent print.	17
2.4	Fingerprint coordinate transformation based on core point: (a) original fingerprint image with minutiae (represented by circles) and core point (dot and arrow), and (b) fingerprint image after translation and rotation of the core point to the center.	19
2.5	Gaussian mixture model of location and orientation: (a) Distribution of spatial location $s = (x_1, x_2)$ is modeled by a mixture of three bivariate Gaussians whose contours of constant density are shown (b) 3D plot of mixture model for location, (c) von Mises distributions of orientation θ for each of the three components, where the green curve corresponds to the upper cluster, blue the lower left cluster and red the lower right cluster, and (d) graphical model of mixture where \mathbf{z}_n are latent variables corresponding to mixture components and parameters are as in Eq. 2.8. [Best viewed in color] . . .	20
2.6	Sequential ordering of minutiae: (a) given minutiae $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4\}$ with centroid c , the next minutia \mathbf{x}_5 is selected by comparing the remaining minutia distances to c , thereby providing a sequencing, (b) dependency between the sorted minutiae is represented by arrows.	21
2.7	Directed probabilistic graphical model to represent the joint distribution of minutiae. This model corresponds to example in Figure 2.6(b). Minutiae locations are represented by nodes labeled \mathbf{s}_n and corresponding orientations are represented by nodes labeled θ_n . This joint distribution can be written directly from the model as $p(\mathbf{s}_1)p(\theta_1 \mathbf{s}_1)p(\mathbf{s}_2)p(\theta_2 \mathbf{s}_1, \mathbf{s}_2, \theta_1)p(\mathbf{s}_3)p(\theta_3 \mathbf{s}_1, \mathbf{s}_3, \theta_1)\dots$	21
2.8	Graphical model for rarity. Specific nPRC is the conditional probability that a known \mathbf{x} is found among at least one of $\mathbf{y}_1, \dots, \mathbf{y}_n$, where z is an indicator variable for a match.	23
2.9	Simple configurations of minutia and core points: (a) a common configuration with three minutiae m_1, m_2 and m_3 and core point c , (b) an uncommon configuration obtained by changing the orientations of m_1 and m_3 and (c) an uncommon configuration obtained by translating the three minutiae with respect to the core point. For $n = 1000$, their specific nPRC values are: (a) 1.2×10^{-2} , (b) 7.97×10^{-4} and (c) 2.3×10^{-6} respectively.	26
2.10	Brandon Mayfield case: prints used in rarity evaluation: (a) latent print <i>LFP17</i> found at crime scene with seven marked minutiae (initial annotation), (b) matching ten-print of Mayfield found in FBI database with 15 charted minutiae, (c) <i>LFP17</i> re-annotated with the same 15 minutiae as in (b), and (d) ten-print of Daoud with ten matching minutiae (from [78]). [Best Viewed in Color]	26
2.11	Two latent prints from <i>NIST27</i> : (a) <i>b115</i> is from the <i>bad</i> dataset and (b) <i>g73</i> is from the <i>good</i> dataset. In each case the left image is the print and the right its aligned version with predicted core point. Corresponding rarity values for $n = 100,000$ are given in Table 2.4. Rarity values of (a) for different values of n and different numbers of matching minutiae are plotted in Figure 2.12.	28

LIST OF FIGURES

LIST OF FIGURES

2.12 Dependence of specific n PRC of latent print $b115$ on database size n parameterized by number of corresponding minutiae $\hat{m} = 4, 8, 12, 16$ and n varying from 10^3 to 10^{14} 29

2.13 Distribution of Bozorth distance between fingerprints under the same (h^0) and different (h^1) hypotheses. 31

2.14 Determining fingerprint similarity using geometric structures: (a) minutiae pair (Bozorth), (b) a set of k -minutiae (CBFS) and (c) a polygon (Champod). 33

2.15 Neighborhood of a minutiae set: (a) fingerprint image with 39 minutiae extracted, (b) example of neighborhood r of a minutia, (c) the graph $\mathcal{G}(\mathcal{M}, \mathcal{E})$ of neighboring minutiae constructed by edging each pair of minutiae according to the neighborhood system defined by (2.34) in (b), where $r = 60$ pixels. The resulting graph contains 118 edges. 35

2.16 Correspondence process of mapping from evidence \mathcal{M} to known \mathcal{M}' 37

2.17 Feature disparity between two fingerprints: (a) eight finger-prints with extracted minutiae sets where a particular minutia (index 9) is colored, (b) the same minutia in 8 different impressions are superimposed together, and (c) 56 distances exist between all possible pairs. 37

2.18 Overview of evaluating a probabilistic measure of similarity. 38

2.19 Generative model for minutiae distortion: (a) corresponding minutiae sets from two fingerprint impressions of same finger, (b) probability density of within-finger position difference, and (c) probability density of within-finger angle difference. 39

2.20 Pairwise MRF model with node and edge potentials. 41

2.21 An example of node belief update on a MRF: (a) latent print (SD27 good 013) with 14 minutiae marked and ten print with 87 minutiae, and (b) the red nodes are matched in the MAP assignment. 43

2.22 A second example of node belief update with MAP assignments. 44

2.23 Correspondence between pairs of fingerprints—evidence on the left (minutiae in red) and known is on the right when the pairs are from: (a) same finger with low distortion (b) same finger with high distortion, and (c) different fingers. The joint probabilities and likelihood ratios for each pair using three methods are given in Table 2.6. 45

2.24 Cumulative Match Characteristics (CMC) curves: (a) good, and (b) bad 47

2.25 Distribution of joint probabilities $P(f_1, \dots, f_n)$ computed given test set of FVC2002 DB1 48

2.26 Tippett plots for LR of four different data sets. 50

List of Tables

2.1	Performance of core point prediction: <i>accuracy</i> of standard (<i>Poincare Index (PI)</i>) method and MAP estimate of proposed (<i>Gaussian Process (GP)</i>) method. Results show dramatic improvement with low-quality images.	18
2.2	Validation of Models: χ^2 test of CPDs.	23
2.3	Brandon Mayfield case: Probability of randomly matching the latent print <i>LFP17</i> with an item in FBI database (specific <i>n</i> PRC with $n = 4.7 \times 10^8$).	27
2.4	Probability of finding a match in a database. Considering the two latent prints <i>b115</i> (from the bad set) and <i>g73</i> (from the good set) in <i>NIST27</i> (shown in Figure 2.11), the probability of finding a corresponding print in a database of 100,000 prints is evaluated, using the following tolerances: minutia location $\epsilon_s = 10$ pixels and minutia direction $\epsilon_\theta = \pi/8$	28
2.5	Rate of successful correspondence (SD27)	45
2.6	Joint probability and likelihood ratio for the three pairs of prints in Fig. 2.23	47
2.7	Percentage distributions of likelihood ratios (calculated for 1500 within-finger (same) comparisons and 3,000 between-finger (different) comparisons) for FVC2002 DB1. False positive (between-finger comparison giving a value of LR greater than 1) and false negative (within-finger comparison giving a value of LR smaller than 1) rates and overall error rates are given.	49
2.8	Error rates for fingerprint identification on 4 datasets.	49
2.9	LR ranges and rate of misleading evidence in favor of prosecution (RMEP) and in favor of the defense (RMED) for four different data sets in FVC2002. The Tippett plot is shown in Fig.2.26	51
2.10	LR ranges and rate of misleading evidence in favor of prosecution (RMEP) and in favor of the defense (RMED) for different number minutiae.	52

Chapter 2

Research narrative

2.1 Introduction

Latent prints of friction ridges have been used in criminal investigations and forensic identification for over a century. The methodology used by the latent print examiner has been well-documented, e.g., the *Scientific Working Group on Friction Ridge Analysis, Study and Technology* (SWGFAST) has prepared a document entitled *Friction Ridge Examination Methodology for Latent Print Examiners* [56], which is periodically updated. Human factors in latent print examination using a systems approach is described in a recent National Institute of Standards and Technology (NIST) report [24].

The need for quantitative measures to characterize confidence in the opinion resulting from latent print examination has been articulated in several recent influential reports, particularly the NRC 2009 report [44] and the NIST Human Factors report [24]. The latter points out that the probabilities involved may point to near uniqueness of features and consequent certainty of individualization. This research addresses how to evaluate such probabilities so that the opinion of the examiner can be justified.

Latent print examiners agree that when a latent print found in a crime scene is compared to a known (inked or live scan) print, rarity of the observed configuration of features plays an important role. Determining rarity is experiential rather than quantitative. Evaluation of rarity is difficult since: (i) we are dealing with joint distributions of many variables (minutiae) and (ii) the data sets needed to determine those distributions are very large. Similarity between a pair of prints (evidence and known) is also a cognitive function of the examiner.

This research addresses: (i) how to model probability distributions of features, (ii) how they can be used to determine the rarity of evidence (as measured by the probability of random correspondence in a database of given size), (iii) how to make such evaluations computationally tractable, (iv) how rarity can be combined with similarity (between the evidence and a known) to determine the confidence of a conclusion, and (v) how to obtain a probabilistic measure of similarity. The methods used are based on machine learning [8] and probabilistic graphical models [35].

2.1.1 Current Practice

Despite being crucial evidence in individual identification, latent fingerprints comparison is not a trivial task due to their poor quality. Generally, latent fingerprints have a small surface and are distorted, smudgy, blurred or may contain artifacts. All these lead to high number of unreliable extracted features and make it hard for automatic systems to perform well. For this reason, a lights-out identification process, i.e., without human intervention, is not yet possible for latent fingerprints and a lot of human effort is required when searching a latent mark with an automatic fingerprint identification system (AFIS). Therefore, latent fingerprints are processed manually, which can be time consuming and requires prioritization of examination requests in order to avoid delayed results in high profile cases. Errors in identification from manual processing typically are a result of the limited amount of time available to the human examiner, in order for the results to have relevance in the case. A mix between fully automated and fully manual is a semi-automated (“Semi

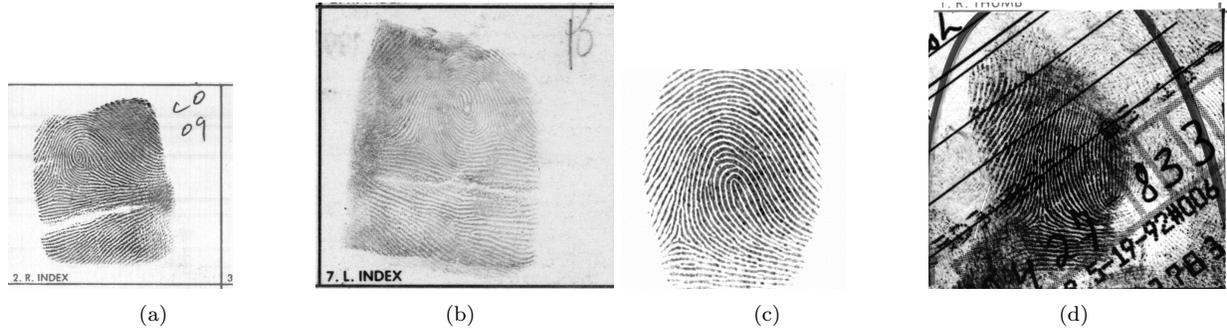


Figure 2.1: Types of Fingerprint images used in this research: (a) rolled (NIST SD14), (b) rolled (NIST SD27), (c) plain (live-scan) (FVC2002), and (d) latent (NIST SD27).

Lights-Out” [20]) approach, in which the latent print features are encoded manually by a human examiner and then compared automatically against rolled or plain known prints stored in the database. A list of 10 to 20 candidates with the highest matching scores is returned, and human intervention is needed again. The fingerprint examiner will then analyze high scoring prints in the candidate list and compare them manually to the query.

Friction ridge prints are classified in two large groups: impressions and latent prints [3]. Impressions are acquired from a cooperative subject usually by scanning the inked impression on paper or directly from the fingers with a live-scan device. It can be roughly categorized into rolled impressions and plain impressions or “flats”. Rolled impressions are obtained by carefully rolling the finger from one side to the other. Plain impressions are those in which the finger is pressed down with a moderate pressure but not rolled. Rolled impressions have a larger area including as much information as possible. Plain impressions cover fewer features but are less distorted and have clearer ridges. The other kind of prints known as latent prints are generally left at crime scenes. Latent prints are unintentional reproductions of the arrangement of ridges on the skin made by the transfer of materials (such as amino acids, proteins, polypeptides, and salts) to a surface [24]. They are not subject to retake and tend to be of considerably lower quality and information content. Examples of fingerprint images are shown in Figure 2.1.

Methodology of ACE-V

The SWGFAST methodology document delineates (1) the principles by which examinations are conducted, (2) features to be used for friction ridge examination and (3) a method which specifies steps to be followed—which includes conclusions that may result from an examination. These are further expanded as follows:

1. *Fundamental principles* for friction ridge examinations by a latent print examiner are:
 - (a) The morphology of friction ridge skin is unique.
 - (b) The arrangement of friction ridges is permanent barring trauma to the basal layer of the epidermis.
 - (c) An impression of the unique details of friction ridge skin can be transferred during contact with a surface.
 - (d) An impression that contains sufficient quality and quantity of friction ridge detail can be individualized to, or excluded from, a source.
 - (e) Sufficiency is the examiner’s determination that adequate unique details of the friction skin source area are revealed in the impression.
2. *Features*: There are three levels of detail and a miscellaneous category of “other” features as follows:
 - (a) *Level One Detail* consists of overall ridge flow and general morphology (e.g., presence of incipient ridges, overall size). They can be used for pattern interpretation and to determine anatomical

source (i.e., finger, palm, foot, toe) and orientation. They cannot be used alone to individualize. They can be used to exclude under certain circumstances.

- (b) *Level Two Detail* describes ridge path, principally ridge path deviation such as ridge ending, bifurcation and dot. It also includes absence of ridge path deviation, e.g., continuous ridge, and ridge path morphology, e.g., size and shape. They can be used in conjunction with level one detail to either individualize or exclude.
- (c) *Level Three Detail* describes the structure of individual ridges, specifically the shape of a ridge and relative pore position. They can also include other specific friction skin morphology (i.e., secondary creases, ridge breaks, etc.) They can be used in conjunction with level one and level two detail to either individualize or exclude.
- (d) *Other features* associated with friction ridge skin (e.g., creases, scars, warts, paper cuts, blisters). They may be permanent or temporary, may exist as level one, two and three detail, and may be used in conjunction with friction ridge detail to individualize or exclude.

3. *The ACE-V Method:*

The steps involved in processing latent prints is described by the ACE-V process. It involves the recurring application of the following four steps: Analysis, Comparison, Evaluation and Verification.

- (a) *Analysis* is the assessment of a friction ridge impression to determine suitability for comparison. Factors considered include the following: Quality (clarity) and Quantity of level one, two and three detail and the anatomical source. Factors influencing quality include: residue/matrix, deposition, surface/substrate, environment, development medium, preservation method, and condition of the friction skin.
- (b) *Comparison* is the direct or side-by-side observation of friction ridge detail to determine whether the detail in two impressions is in agreement based upon similarity, sequence and spatial relationship. Features are extracted in the *comparison* phase.
- (c) *Evaluation* is the formulation of a conclusion based upon analysis and comparison of friction ridge impressions. Conclusions which can be reached are described in an updated SWGFAST document entitled *Standards for Conclusions* [57]. They are as follows:
 - i. *Individualization* (Identification): The standard for individualization is agreement of sufficient friction ridge details in sequence. Conditions that shall be satisfied: (i) determined by a competent examiner, and applied to a common area in both impressions, and (ii) based on quantity and quality of the friction ridge details, and (iii) absent any discrepancy, and (iv) reproducible conclusion. Basic principles: (i) there is no scientific basis for requiring that a predetermined number of corresponding friction ridge details be present in two impressions in order to effect individualization, (ii) individualization is supported by the theories of biological uniqueness and permanence, probability modeling, and empirical data gained through more than one hundred years of operational experience.
 - ii. *Exclusion*: The standard for exclusion is disagreement of friction ridge details. Conditions that must be satisfied are: (i) determined by a competent examiner, and (ii) applied to all comparable anatomical areas, and (iii) presence of a discrepancy, and (iv) based on sufficient quantity and quality of the friction ridge details, and (v) reproducible conclusion. Basic principles are: (i) the presence of one discrepancy is sufficient to exclude, (ii) distortion is not a discrepancy and is not a basis for exclusion, (iii) exclusion is supported by the theories of biological uniqueness and permanence, probability modeling, and empirical data gained through more than one hundred years of operational experience.
 - iii. *Inconclusive*: The standard for an inconclusive finding is the absence of sufficient friction ridge details to effect a conclusion of individualization or exclusion. Conditions that must be satisfied: (i) determined by a competent examiner, and (ii) based on quantity and quality of the friction ridge details, and (iii) insufficient agreement or disagreement in the friction ridge details, and (iv) reproducible conclusion.

- (d) *Verification*: is the independent examination by another qualified examiner resulting in the same conclusion. All individualizations (identifications) must be verified. Exclusion or inconclusive results may be verified.

2.1.2 Statement of the problem

This effort is to help quantify several elements mentioned in the SWGFAST document *Friction Ridge Examination Methodology for Latent Print Examiners* and thereby assist the examiner in testimony. In the *Fundamental principles* section of the document there is mention of “uniqueness” in 1(a), 1(b) and 1(e). Second, in the same section there is mention of “individualization” in 1(d). Third, in the section on *Method*, there is mention of “similarity” in 3(b) in the phrase “agreement based upon similarity, sequence and spatial relationship”.

There have been several influential reports that call for quantifying the terms “unique”, “individualization” and “similarity”. For instance, the NRC 2009 report *Strengthening the Forensic Sciences: A Path Forward* asks for “studies [that] would accumulate data.... to validate the ACE-V process and to attach confidence limits to individualization determinations”[44]. This essentially calls for restating individualization in probabilistic terms. The report also asks for studies on intra- and inter-variability of fingerprints which involves studies on similarity.

The more recent NIST report on *Latent Prints Examination and Human Factors: Improving the Practice through a Systems Approach* [24] discusses the probabilistic approach and states: “As the discriminating power of the latent print features used to make an identification increases, so does the probability that the matching individual is the source.” This means that as the features have been accurately ascertained, compared and found to be nearly the same, the probability of identification approaches one in the limiting case.

This research addresses the probabilistic modeling of latent print features that can be useful within the processing steps of ACE-V. Only fingerprints characterized by level-two detail are considered since: (i) they are the principal features used for identification/exclusion, and (ii) the methods developed can be eventually extended to other levels of detail and features. The models are used to make the following types of inference:

- At the end of the *analysis* phase, the examiner determines the value of the print for further comparison. The rarity of the features observed is critical to making this decision. If the examiner determines that the features are common enough to be repeated in multiple areas of friction ridge skin, then that print would not be of value for identification purposes. Since the determination of rarity is largely experience-based we address how to evaluate rarity.
- In the *comparison* phase, the examiner observes similarities and differences of the configuration of minutiae between the known and unknown impressions. Finally, in the *evaluation* phase, the examiner assesses both the rarity of the features as well as their level of similarity to reach a conclusion. We consider how to combine similarity and rarity in determining the likelihood ratio which in turn can be used for making a conclusion.

In support of these objectives we address three computational problems:

1. Registering the latent print so that spatial distributions can be defined over minutiae.
2. Characterizing the probability distributions of minutiae by taking into account minutiae dependencies; such studies have been done in the past for characterizing fingerprint individuality but minutia independence is assumed. Use the models developed for determining the rarity of given fingerprints.
3. Develop methods whereby the probability of a fingerprint (represented by a configuration of minutiae) can be combined with a probabilistic measure of similarity so that a probability of identification can be determined.

2.1.3 Literature review

We review the relevant literature in five areas: (i) the registration problem, (ii) fingerprint individuality studies, (iii) modeling minutia distribution, (iv) computing identification probability from likelihood ratios, and (v) similarity evaluation.

A. The Registration Problem

In order to compare a latent print with a known impression, it is necessary to align them properly so that the location and spatial relationships of features (e.g., minutiae) are comparable. This problem is generally referred to in the pattern recognition literature as the *registration problem*. A simple method of registration, when the prints are complete, is to determine the core point on the fingerprint. A core point refers to the center area of a fingerprint. In 1900 Henry [30] defines the core point as “the north most point of the innermost ridge line”. In practice, the core point corresponds to the center of the north most loop type singularity. For fingerprints that do not contain loop or whorl singularities, the core is usually associated with the point of maximum ridge line curvature[40].

Several algorithms have been proposed for core point detection. The most popular one, known as the Poincare Index (PI), was developed by Kawagoe and Tojo [34], and it was subsequently adopted and enhanced by others [7, 31]. The PI method is unable to detect the core point in most arch type fingerprints and performance severely deteriorates if the image quality is poor.

Another method, based on a sine map, is realized by multi-resolution analysis [32]. Methods using Fourier expansion[52], fingerprint structures [79], multi-scale analysis [39] and orientation consistency [11] have also been proposed. All of these methods, which are inspired by computational vision, require that the fingerprint is complete and that the core point is present somewhere in it. This assumption does not hold for latent fingerprints, which are usually partial and do not contain core points. So there is no way to detect them by any of the proposed computational vision approaches.

B. Fingerprint Individuality Studies

Relevant to modeling the distribution of minutiae are several classical studies on fingerprint *individuality* [67]. Dating to over a hundred years, the goal of these studies was to determine the degree to which fingerprints are unique. Galton [27] computed the probability of a fingerprint as a product of three factors: factor A is that the configuration is present/absent in each of 24 six-ridge square regions of the fingerprint, factor B is that the region is one of sixteen pattern types and factor C is that the correct number of ridges would enter and exit each region. The three factors are evaluated to have values $(1/2)^{24}$, $(1/16)$ and $(1/256)$. Their product gives the probability of any given fingerprint as 10^{-11} . This model was considered to be a gross underestimate of the variability of fingerprints and improved by Pearson [51] by replacing the $(1/2)$ term by $(1/36)$ to reflect different minutia configurations, yielding a probability of 1.09×10^{-41} . These evaluations remained untested and subsequently replaced by methods such as those based on minutia types. Stoney and Thornton [66, 65] proposed a set of desired features for a fingerprint model and attempted to meet some conditions using their survey of minutiae and subsequent data analysis. Two limitations to Stoney and Thornton were: the limited scope of the survey and lack of accommodation for prints in poor condition. Champod and Margot [12] presented a statistical model that utilized computer-generated frequencies of minutiae occurrence and minutia densities. A new variable was introduced as compound minutia length. Known weaknesses of the model include: position of the print on the finger must be known, and not allowing connective ambiguities. While efforts have been made in 10-print individuality studies, modeling of latent friction ridge prints continues to be a difficult problem due to the small finger area and poor impression quality.

C. Modeling Minutia Distribution

Several efforts have been made to characterize the distribution of fingerprint features. Since fingerprint samples can be generated from these models they may be termed as *generative*. The generative model for

level-one detail simply consists of the probabilities for each ridge flow pattern such as right loop (30%), left loop (27%), double loop (7%), arch (13%), tented arch (5%) and whorl (19%) [63].

Modeling level-two detail is more important, since fingerprint identification is largely based on features at that level, viz., minutiae. A minutia is represented by its location and direction. The direction is determined by the ridge at the location. Automatic fingerprint matching algorithms use minutiae as the salient features [80], since they are stable and are reliably extracted.

However a generative model becomes much more complex due to the large and variable number of features involved. The simplest model assumes that minutiae locations and orientations are uniformly and independently distributed [48]. An improved model assumes that although minutiae are independent of each other minutiae orientation and location were dependent. Such a model is better than mixtures of hyper-geometric and binomial distributions. A mixture model to account for the clustering tendency of minutiae was proposed [84]. A Markov point process to model minutia location has been attempted [15] with shortcomings, e.g., direction is not incorporated, being based on relative spatial relationships minutia sets from different regions will have the same rarity, and pair potential cannot be expressed in closed form making usability impractical.

There have also been efforts to model distributions beyond minutiae alone. Since latent print examiners rely on minutiae as well as ridge information, ridges can be represented discretely as ridge points and the model accounts for both minutiae and ridge points [68]. Extending to level-three detail, a model that incorporates minutiae, ridge and pore features has been recently proposed[16].

Minutiae that are spatially close tend to have similar directions with each other [58]. Moreover, friction ridges flow smoothly with very slow orientation change. The variance of the minutiae directions in different regions are dependent on both their locations and location variance [65, 14]. These observations on the dependency between minutiae need to be accounted for in eliciting reliable statistical models.

All existing models for level-two detail have the drawback of assuming independence of minutiae which leads to inaccurate probability estimates. The research described here incorporates minutiae dependencies as well as minutiae uncertainties.

D. Computing Likelihood Ratios

Forensic identification concerns whether observed evidence arose from a known source. The ACE-V method specifies three opinions: identification, exclusion and inconclusive. In order to establish confidences with opinions it is necessary to formulate a probabilistic model involving both the evidence and the known.

The generally accepted probabilistic approach is to determine the likelihood ratio (LR) [23, 1, 75] whose numerator is the joint probability of the evidence and source under the null, or *prosecution*, hypothesis that the evidence arises from the source and the denominator is the joint probability under the alternate, or *defense*, hypothesis that the evidence does not arise from the object. The evidence is deemed to have arisen from the source if $LR > 1$ and not from the source otherwise.

Determining the joint probability has high data requirements. For example, if we assume that the evidence and known are both characterized by n binary features, the joint distribution requires 2^{2n} probabilities or parameters. Even for small n this requires extremely large data sets for estimating parameters. In the case of fingerprints we may be dealing with dozens of minutiae ($n > 12$) and each minutia is characterized by three continuous values (rather than binary). Furthermore available friction ridge data sets are usually small making the estimation of joint probability distributions infeasible.

One solution is to overcome the computational and data limitations of the joint distribution is to use the distribution of distance (or similarity) between the latent print and known [45, 62]. However this involves a severe approximation in going from a high-dimensional space to a one-dimensional space. We have proposed the combination of rarity and similarity as a better approach [74].

E. Similarity Measures

A similarity measure, or its inverse which is *distance*, between two sets of features is commonly used in AFIS. It is useful in characterizing the strength of a match. In the friction ridge domain the features are representative characteristics of the fingerprint. Minutiae (level two detail) are the most widely used

feature. The representation is a feature vector whose elements are the fingerprint minutiae. Each minutiae is described by a number attributes, including its type, position and orientation in the fingerprint image. Suppose X have N minutiae, they can be represented as two sets of minutiae, $\mathcal{M}_X = \{m_1, m_2, \dots, m_N\}$, where m_i is a minutiae and contains several types of descriptions such like location, orientation and types. Let $\Lambda(m)$ denote the feature vector of minutiae $\Lambda(m) = \{l_x, l_y, \theta, t\}$. A similarity measure $s(X, Y)$ can be defined in various ways for different scenarios.

The degree of similarity in AFIS is measured by the capability of a fingerprint matcher to find true correspondences between prints of the same finger while minimizing mismatches. Due to frequent non-linear deformation in fingerprint images, directly ensuring global correspondence of minutiae is very difficult. Correctly aligning two fingerprints requires considering distortion of fingerprint image [40]. Most minutiae-based methods tend to first obtain local similarity and then perform global consolidation. Local minutiae models can be structures of pairs [80], triplets, stars, a set of minutiae or other local structures [47, 82, 29] (see Figure 2.14). Its proximity score is usually based on the frequency count of number of corresponding configurations against the total number of configurations available in the two prints. Due to the fact that matching two different pairs of fingerprints could result in the same similarity score just because the portion of matching minutiae amount are the same, neglecting the possibility of true difference (such as difference in geometrical transformation) between the similarity over spatial structure, there is clearly a severe loss of information. Evidence strength analysis based on such scores will be less reliable and accurate.

In AFIS the given fingerprint is compared against many candidates comprising a gallery (background), and zero, one, or more potential candidate matches are reported back. State-of-the-art systems can perform different types of searches: (i) *Ten-print to ten-print*: rolled or plain fingerprints of 10 fingers are searched against the database. Due to the amount of information available, lights out and human intervention is usually not needed. Feature extraction, matching and verification are done automatically. (ii) *Latents to ten-print*: this is the most critical function of an AFIS, as it helps finding the author of a crime if their fingerprints were previously registered in the AFIS. The low quality and reduced area of latent prints make it harder for the AFIS than impression to impression searches. (iii) *Latents to latents*: useful to identify two anonymous prints that have been left by the same person, even when the person has not been identified. This is a really difficult task for an AFIS, as partial fingermarks do not always have information from the same part of the finger.

According to the *Fingerprint Vendor Technology Evaluation (FpVTE)* [13] report, AFIS can achieve a rank-one identification rate of more than 99.4% with a database of 10,000 fingerprint images. Here, a rank-n identification rate measures the rate at which a correct match is present in the best n matches that are returned by the matching system. In contrast to high accuracy levels for automated matching of plain prints, the comparable rank-one identification rate for latent fingerprints is only 54%, according to a NIST report using a database of more than 40 million [20]. NIST conducted a multi-phase project on Evaluation of Latent Fingerprint Technologies (ELFT) to evaluate automatic latent feature extraction and matching techniques. In Phase I, the most accurate system showed a rank-1 accuracy of 80% (100 latents against 10,000 rolled prints). In Phase II, the rank-1 accuracy of the most accurate system was 97.2% (835 latents against 100,000 rolled prints). However, these accuracies cannot be directly compared since the Phase I and Phase II evaluations used different databases where images of latents are of very good quality.

Recent studies on latent fingerprints [25, 33, 49] consist of the following steps: (i) align two sets of minutiae; (ii) establish the correspondences between minutiae; and (iii) compute a similarity score. Improved latent matching accuracy has been reported by using extended features [33] which are manually marked. However, marking extended features (orientation field, ridge skeleton, etc.) in poor quality latents is very time-consuming and might be only feasible in rare cases. [49] developed a latent fingerprints matching algorithm that is solely based on minutiae using a descriptor-based Hough Transform and reported a rank-1 accuracy of 62.4%. Due to the poor quality, nonlinear deformation and unreliable extracted features of latent fingerprints, the uncertainties in obtaining features and aligning them favor probabilistic models that assume a distribution about the uncertainty of correspondence.

2.1.4 Rationale for the research

There are several motivations for this research: (i) to develop practical tools useful to the latent print community, (ii) provide a response to criticisms in the legal community regarding uniqueness and individualization, and (iii) model cognitive processes in ACE-V relating to rarity and similarity.

Confidence measures can be used to support the ACE-V method of latent print examination [44, 24]. Specifically, we consider the modeling of distributions of minutia configurations so that useful inferences can be made using the model. One of these is to evaluate the rarity of a given configuration which is used by latent print examiners to focus attention on the most important parts of the print. A simple analogy of the importance of rarity in identification arises when we are trying to match an individual based on height. An average value of height is much less useful than an extreme value.

Another use of rarity is in the evaluation of a measure of the strength of evidence. The strength of correspondence between the print found on a crime scene (latent print) and the known, is useful to the evaluation phase of ACE-V. When this confidence value is evaluated and found to be extremely large/small it provides a justification for the argument of individualization/exclusion which has come under criticism[54].

Many cognitive functions are performed well by human beings. Vision itself is such a function where the mammalian cortex has evolved over millions of years. Computational methods developed to understand visual processes have been referred to as computational vision[42]. Such an effort for forensic examination can be referred to as computational forensics [64]. The particular cognitive processes modeled here pertain to rarity of friction ridge patterns and similarity between patterns.

2.2 Methods

The principal effort is to model probability distributions of features found in friction ridge impressions and to use such models in answering probabilistic queries. One such query is the probability of a given input configuration. Another is the probability of finding the configuration in a database of a given size. A task of related interest is that of combining rarity information with the distribution of similarity between latent prints and known so as to provide a measure of the strength of opinion.

Modeling complex probability distributions from data sets and making inferences from them is the subject of modern methods of machine learning [8, 35]. Such methods become relevant to friction ridge analysis since the computation becomes quickly intractable with the number of features found in friction ridge impressions. So we take the machine learning approach to the tasks of image registration, modeling minutia distribution and determining similarity.

In the supervised machine learning approach, labeled data sets are needed for determining parameters and testing performance. The principal latent fingerprint dataset used in the experiments is the *NIST Special Database (SD) 27*, which contains latent prints from crime scenes and their matching rolled fingerprint mates (see Figure 2.1). Each fingerprint image is 800×768 pixels in size and has been scanned at 500 ppi. The 258 latent prints (with one duplicated) in SD27 were categorized by latent examiners into three quality levels: *good*, *bad*, and *ugly* with 88, 85, and 85 images respectively. The minutiae in NIST SD27 were manually marked by fingerprint examiners. Other NIST datasets were also used for parameter estimation, e.g., SD14 which contains 27,000 pairs of rolled fingerprint images provided by FBI. They are scanned similar to SD27 at 500 ppi and in 8-bit gray scale with size 832×768 . A live scan database, FVC 2002, was also used in some experiments: it contains four different databases (DB1, DB2, DB3 and DB4), each with 110 different fingers and 8 impressions of each finger yielding a total of 880 fingerprints.

In the following four sections we discuss methods developed for (i) registration, (ii) modeling distributions of minutiae, (iii) determining rarity of a configuration, (iv) determining strength of opinion by combining rarity and similarity, and (v) determining similarity.

2.2.1 Registration

The starting point in determining the spatial distribution of minutiae is to establish the origin of a coordinate system. This problem is encountered in AFIS and in biometrics. It is typically taken to be the core point of the fingerprint. Since every fingerprint does not possess a core point a high curvature region can be

designated as the origin. Locating core points in latent fingerprints is difficult using a purely geometric approach since they are often partial images with the high curvature point left outside the print. We take the approach of regression, using an orientation map as input, to predict the location and orientation of the core point.

Latent prints usually correspond to only a small portion of the complete fingerprint. Thus feature sets, such as minutiae, extracted from the print contain only relative spatial relationships. Feature sets with the same relative spatial relationship can lead to different rarity if they come from different areas of the finger. To solve this problem, we first predict the center of the fingerprint, as defined by its core point, and then align the fingerprint by translating the coordinates so that the center is located at the core point and rotate the coordinates so that the core point orientation points north.

Since ridge flow directions reveal intrinsic features of ridge topologies, they have a critical impact on the core point. A fingerprint field *orientation map* is defined as a collection of two-dimensional direction fields. It represents the directions of ridge flows in a regularly spaced grid. The gradients of gray intensity of enhanced fingerprints are estimated to obtain reliable ridge orientation [31].

The input fingerprint image is divided into blocks of size $W \times W$ where W is the size in pixels, e.g., $W = 10$. Then compute the gradients G_x and G_y , the gradient magnitudes in the x and y directions, at each pixel in each block. Finally, estimate the local orientation of each block using:

$$\theta_o = \frac{1}{2} \tan^{-1} \left(\frac{\sum_{i=1}^W \sum_{j=1}^W 2G_x(i, j)G_y(i, j)}{\sum_{i=1}^W \sum_{j=1}^W (G_x^2(i, j) - G_y^2(i, j))} \right). \tag{2.1}$$

An example orientation map is given in Figure 2.2. In an image who field of view includes objects other than the latent print, extraneous regions are manually erased.

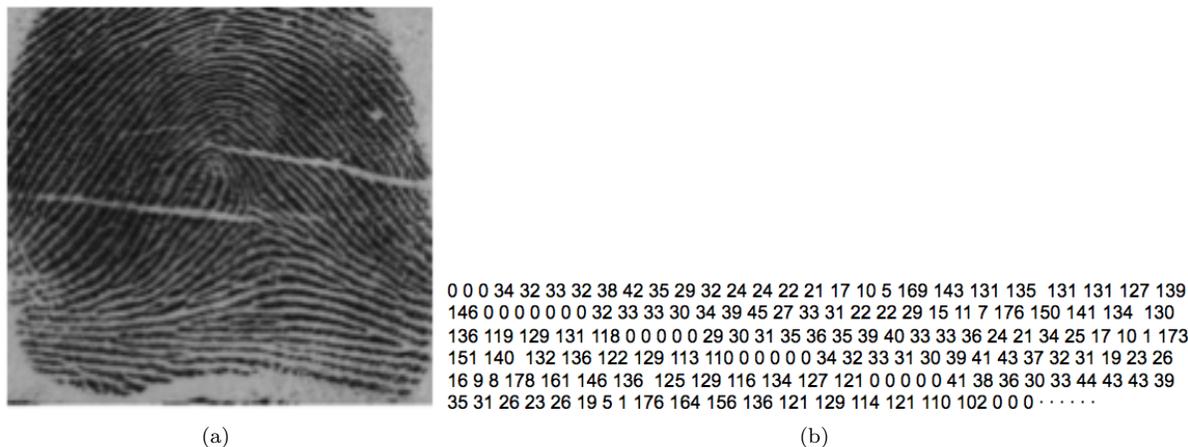


Figure 2.2: Example to illustrate determining the core point of a finger print using regression: (a) fingerprint image and (b) orientation map (32 × 32 vector of gradient values) used as the input variable. The corresponding target variable is the core point which is this case has values $s = (253, 221)$, and $\theta = 85$.

Determining Origin of Coordinate System

In order to specify the spatial location of minutiae, a coordinate system has to be specified. The origin of the coordinate system can be specified by its location and orientation. We are given an input friction ridge image that may only be partial because it is a latent print. The approach taken is to predict the core point, or a high curvature point within the image.

Regression is a machine learning task [8], where training data is used to learn a predictive model. In this problem the independent variable consists of the fingerprint field orientation map and the target variable

is the core point. Since it is a predictive model it is immaterial whether there is a core point within the observed image.

There are several possible approaches to regression, e.g., linear regression, neural networks. The particular regression method we employ is based on Gaussian processes (GPs). GPs dispense with the parametric model and instead define a probability distribution over functions directly. It provides more flexibility and is a better predictor than other approaches. The GP model also has the advantage that instead of a point estimate, the prediction is in the form of a distribution as in a fully Bayesian approach[53]. We now give some details of the GP formulation which is also described in [71].

Let the training set \mathcal{D} consist of N fingerprints, $\mathcal{D} = \{(\mathbf{g}_i, y_i) | i = 1, \dots, N\}$, where \mathbf{g}_i is an orientation map with core point y_i . The regression model with Gaussian noise is given by $y_i = f(\mathbf{g}_i) + \epsilon$ where $f(\mathbf{g}_i)$ is the value of the process or function at \mathbf{g}_i and ϵ is a random noise variable whose value is chosen independent for each observation. Assuming a noise process with Gaussian distribution $p(y_i | f(\mathbf{g}_i)) = \mathcal{N}(f(\mathbf{g}_i), \sigma^2)$, where σ^2 is the variance of noise, the likelihood function is given by

$$p(\mathbf{y} | \mathbf{f}) = \mathcal{N}(\mathbf{f}, \sigma^2 I) \quad (2.2)$$

where $\mathbf{y} = (y_1, \dots, y_N)^\top$ is the observed (learning) set of core points and $\mathbf{f} = (f(\mathbf{g}_1), \dots, f(\mathbf{g}_N))^\top$ contains the corresponding orientation maps $\mathbf{g}_i, i = 1, \dots, N$.

From the definition of a GP, its prior is given by a Gaussian whose mean is zero and covariance is defined by a covariance function $k(\mathbf{g}, \mathbf{g}')$ so that $f(\mathbf{g}) \sim \mathcal{GP}(0, k(\mathbf{g}, \mathbf{g}'))$. A GP with Gaussian kernel is used to specify the covariance between pairs of variables. $k(\mathbf{g}, \mathbf{g}') = \exp(-\|\mathbf{g} - \mathbf{g}'\|^2/2)$.

Core Point Distribution

For an input orientation map \mathbf{g}^* the predictive distribution of core point y^* can be evaluated by conditioning the joint Gaussian prior distribution on the observation (G, \mathbf{y}) , where $G = (\mathbf{g}_1, \dots, \mathbf{g}_N)^\top$. The predictive distribution is given by

$$p(y^* | \mathbf{g}^*, G, \mathbf{y}) = \mathcal{N}(m(y^*), cov(y^*)) \quad (2.3)$$

where $m(y^*) = \mathbf{k}(\mathbf{g}^*, G)[K + \sigma^2 I]^{-1} \mathbf{y}$,

$$cov(y^*) = k(\mathbf{g}^*, \mathbf{g}^*) + \sigma^2 - \mathbf{k}(\mathbf{g}^*, G)^\top [K + \sigma^2 I]^{-1} \mathbf{k}(G, \mathbf{g}^*),$$

$\mathbf{k}(\mathbf{g}^*, G) = (k(\mathbf{g}^*, \mathbf{g}_1), \dots, k(\mathbf{g}^*, \mathbf{g}_N))^\top$ and K is the Gram matrix with elements $k(\mathbf{g}_i, \mathbf{g}_j)$.

Point Estimate of Core Point

Rather than work with the distribution of the core point we can work with the maximum a posteriori probability (MAP) solution. Since \mathbf{g}^* may represent the orientation map in one of several possible locations we maximize among all m possible translations and rotations over the set $\{\mathbf{g}_i^* | i = 1, \dots, m\}$. Using Eq. (2.3), we obtain the predictive distributions $p(y^* | \mathbf{g}_i^*, G, \mathbf{y})$ for all \mathbf{g}_i^* . The core point \hat{y}^* should maximize $p(y^* | \mathbf{g}_i^*, G, \mathbf{y})$ with respect to \mathbf{g}_i^* . Thus the core point is given by

$$\hat{y}^* = \mathbf{k}(\mathbf{g}_{MAX}^*, G)[K + \sigma^2 I]^{-1} \mathbf{y} \quad (2.4)$$

where \mathbf{g}_{MAX}^* , the orientation map corresponding to the most probable core point, is given by

$$\mathbf{g}_{MAX}^* = \underset{\mathbf{g}^*}{\operatorname{argmax}} p(m(y^*) | \mathbf{g}^*, G, \mathbf{y}) \quad (2.5)$$

The point estimate is used in evaluating the performance of the GP method in Section 2.2.1 and in the coordinate transformation process described in Section 2.2.1.

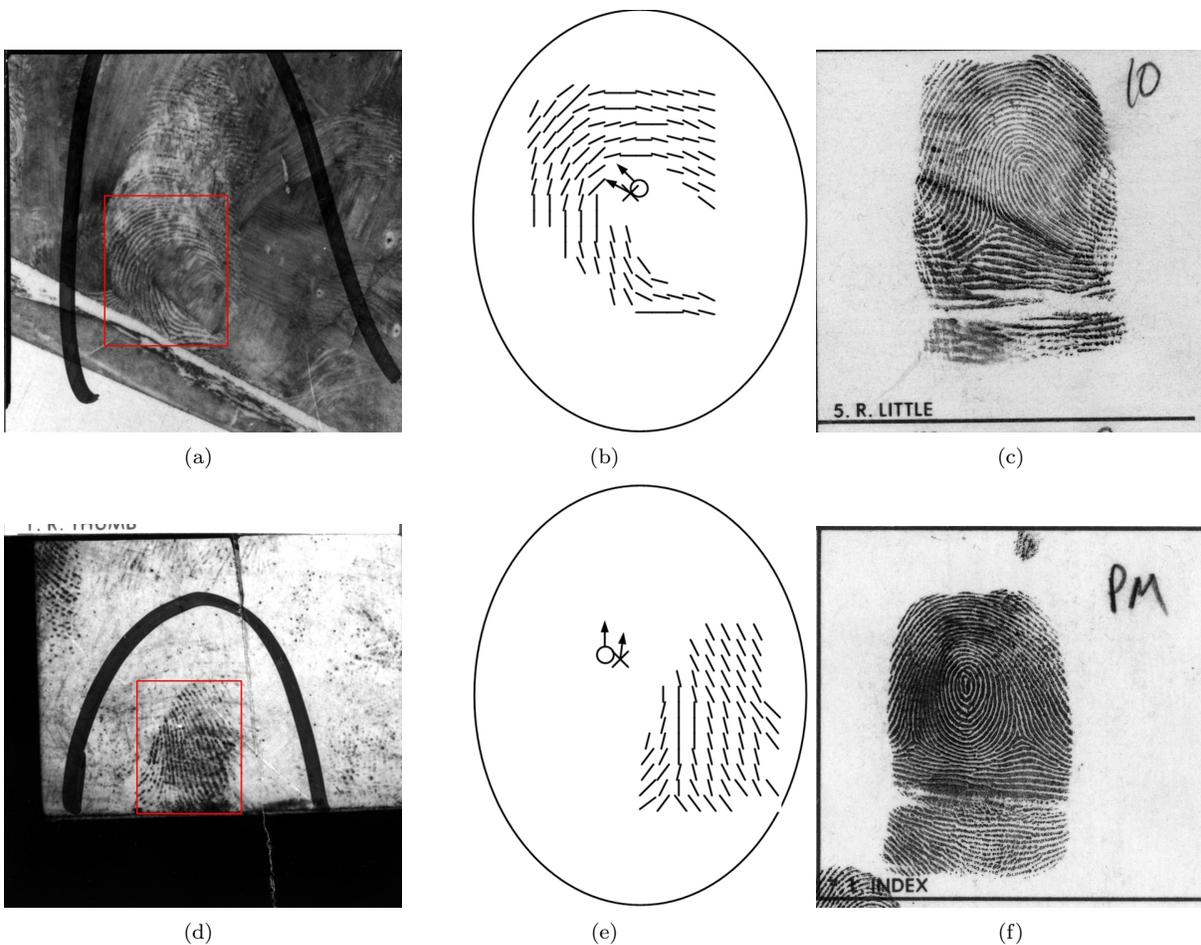


Figure 2.3: Core point prediction using GP regression for two latent prints (from *NIST27*): a) image *g90* which contains a latent print within a rectangle, (b) computed orientation map containing the predicted core point (cross) and true core point (circle), (c) ten-print where the true core point is visible, and (d-f) image *g69* with corresponding images. Note that in the second case the predicted core point lies outside of the latent print.

Table 2.1: Performance of core point prediction: *accuracy* of standard (*Poincare Index (PI)*) method and MAP estimate of proposed (*Gaussian Process (GP)*) method. Results show dramatic improvement with low-quality images.

Image Category (<i>NIST27</i>)	Standard Method (<i>PI</i>)	Proposed Method (<i>GP</i>)
Good	91%	93%
Bad	68%	87%
Ugly	47%	73%
Overall	69%	85%

Performance of GP Regression

The GP model was trained on the *NIST 4* database of fingerprint images [81] and tested on the *NIST 27* set [28]¹. Ridge orientation maps were extracted from the fingerprint images using the gradient-based approach defined by Eq. 2.1. The images were first divided into equal-sized blocks of $N \times N$ pixels, where N is the average width of a pair of ridge and valley. The value of N is 8 in *NIST 4* and varies in *NIST 27*. The gradient vectors were calculated by taking the partial derivatives of image intensity at each pixel in Cartesian coordinates. Ridge orientation is perpendicular to the dominant gradient angle in the local block. The training set consisted of orientation maps whose corresponding core points were manually marked.

The learnt GP model was applied to latent prints of varying quality: 258 prints in *NIST 27* labeled as *good* (88), *bad* (85) and *ugly* (85). Results of core point prediction for two latent fingerprints are shown in Figure 2.3. For each pixel in the smoothed orientation field the PI at pixel (i,j) is defined with respect to a digital curve which consists of a sequence of pixels that are on or within a distance of one pixel from the curve. It takes the form $PI(i,j) = \frac{1}{2\pi} \sum_{k=0}^{N_\psi-1} \Delta(k)$ where $\Delta(k)$ are orientation differences between neighboring pixels on the curve with N_ψ pixels. Assign the corresponding pixel a label 1 if its $PI = 1/2$. For each labeled connected component, if its area is larger than 7, a core point is detected at the centroid of the connected component. Further details of the PI algorithm can be found in [31].

Performance of the GP point estimate was compared to that provided by the baseline Poincare Index (PI) method [7] which is based on purely local topological considerations. To evaluate performance of both methods, test latent prints were extracted from the image database with extraneous regions manually erased. The true core point of each latent print was determined from its matching 10-print in the database. Prediction accuracy was determined by comparing the location and direction distances between predicted and true core points with the threshold parameters set at $T_s = 16$ pixels, and $T_\theta = \pi/6$.

Prediction accuracies of the PI method and the MAP estimate of the GP approach are given in Table 2.1. The good set has 88 images that mostly contain core points. Both bad and ugly sets contain 85 images of small size that usually do not include core points. For good prints, the two approaches are close. For the bad and ugly prints there is a distinct difference between the methods with GP predicting core points even when it is absent in the latent prints. The GP method also results in higher overall performance.

Since the overall error rate of the point estimate is still as high as 15%, the use of a distribution in further analysis is preferable. A disadvantage of GP core point prediction is its $O(N^3)$ complexity, where N is the number of finger prints in the training set; due to an inversion of the $N \times N$ covariance matrix. However, more efficient GP implementations are available [60, 59].

Co-ordinate transformation

After the core point is determined, with a point estimate, the Cartesian coordinate system is transformed such that the origin is the core point and the core point orientation points to $\pi/2$ (Fig. 2.4). Given a

¹*NIST 4* contains 8-bit gray scale images of randomly selected fingerprints. Each print has 512×512 pixels. The entire database contains fingerprints taken from 2000 different fingers with 2 impressions of the same finger. The database is evenly distributed over each of the five classifications with 400 fingerprint pairs from each class. *NIST 27* contains latent fingerprints from crime scenes and their matching rolled fingerprint mates. There are 258 latent cases separated into three quality categories of good, bad, and ugly.

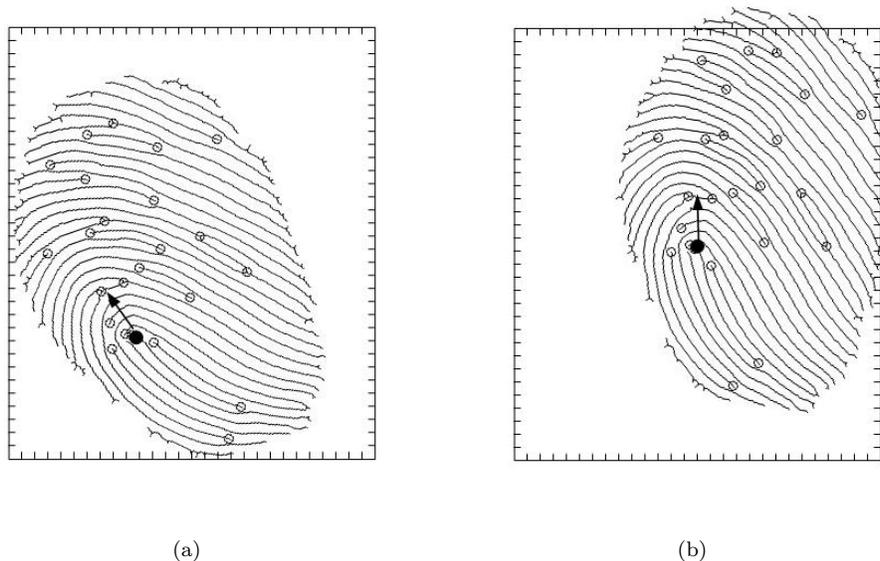


Figure 2.4: Fingerprint coordinate transformation based on core point: (a) original fingerprint image with minutiae (represented by circles) and core point (dot and arrow), and (b) fingerprint image after translation and rotation of the core point to the center.

minutia (\mathbf{s}, θ) and predicted core point (\mathbf{s}^*, θ^*) , where $\mathbf{s} = (s_x, s_y)$ represents location and θ the direction, the transformed minutia (\mathbf{s}', θ') is given by

$$s'_x = s_x - s_x^* - \sin(\theta - \theta^*) \|s - s^*\|$$

$$s'_y = s_y - s_y^* - \cos(\theta - \theta^*) \|s - s^*\|$$

$$\theta' = \theta - \theta^* + \pi/2$$

2.2.2 Modeling Distribution of Minutiae

Next is the task of modeling the joint distribution of minutiae. Each minutia is represented as $\mathbf{x} = (\mathbf{s}, \theta)$ where $\mathbf{s} = (s_x, s_y)$ is its location and θ its direction.

We begin with a model for the distribution of individual minutiae $p(\mathbf{x})$ and then consider modeling the joint distribution of a set of minutiae $\mathbf{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ where we do not wish to assume minutiae independence.

Marginal distribution of minutiae

An obvious and effective model for minutiae location is to choose a mixture of Gaussians [84, 69]. For minutiae orientation, which is an angular distribution, the circular normal or von Mises distribution [8, 41] is useful.

The distribution of minutiae location is shown in Fig. 2.5(a); the minutiae data are from 2,000 fingerprints in the *NIST4* database. This multimodal distribution is naturally modeled as a mixture of k Gaussians (Fig. 2.5(b)) with $k = 3$. Minutiae orientation θ is modeled by the von Mises distribution as shown in Fig. 2.5(c). A simple Bayesian network to represent the marginal distribution of individual points as a mixture model is shown in Figure 2.5 (d).

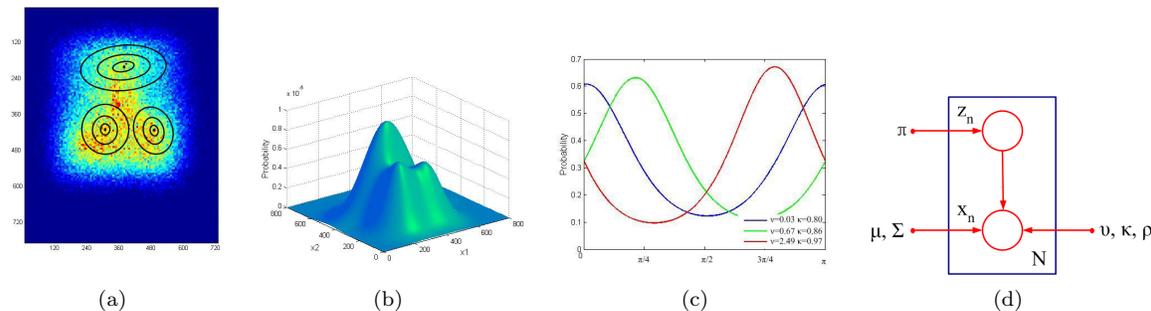


Figure 2.5: Gaussian mixture model of location and orientation: (a) Distribution of spatial location $s = (x_1, x_2)$ is modeled by a mixture of three bivariate Gaussians whose contours of constant density are shown (b) 3D plot of mixture model for location, (c) von Mises distributions of orientation θ for each of the three components, where the green curve corresponds to the upper cluster, blue the lower left cluster and red the lower right cluster, and (d) graphical model of mixture where \mathbf{z}_n are latent variables corresponding to mixture components and parameters are as in Eq. 2.8. [Best viewed in color]

Joint distribution of minutiae

It is unsatisfactory to model the joint distribution of minutiae as a product of their marginal distributions particularly when significant dependences exist. In the case of fingerprints, studies indicate that minutiae direction is related to its own location as well as the location of neighboring minutiae [58, 65, 14]. These known dependences can be readily incorporated into a causal Bayesian network. Conditional dependences of minutiae can be incorporated by linearizing them by defining a unique sequence of points. Linearization is used in agglomerative clustering of N points where a point is assigned to a cluster whose mean is nearest; where nearest can be defined in terms of distance to the mean (centroid), closest or furthest point of the cluster[19].

The sequence starts with the point \mathbf{x}_1 whose location is closest to the center (core point). Each remaining point \mathbf{x}_n is the spatially closest to the centroid defined by the arithmetic mean of the location coordinates of all the previous points $\mathbf{x}_1, \dots, \mathbf{x}_{n-1}$. Given this sequence, the fingerprint can be represented by a minutiae sequence $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_N)$. The sequence is robust to the variance of the minutiae because the next minutia is decided by the all the previous minutiae. Given the observation that spatially closer minutiae are more strongly related, we only model the dependence between \mathbf{x}_n and its nearest minutiae among $\{\mathbf{x}_1, \dots, \mathbf{x}_{n-1}\}$. Although not all dependences are taken into account, this is a good trade-off between model accuracy and computational complexity. Figure 2.6(a) presents an example where \mathbf{x}_5 is determined because its distance to the centroid of $\{\mathbf{x}_1, \dots, \mathbf{x}_4\}$ is minimal. Figure 2.6(b) shows the minutiae sequence and the minutiae dependency (arrows) for the same configuration of minutiae.

A directed probabilistic graphical model, also known as a Bayesian network, can be used to represent correlations between minutiae². Examples of such relationships are: minutia orientation is dependent on near minutiae (both location and orientation), minutia location is conditionally dependent on the location of neighboring minutiae given their directions. A graphical model for the minutiae set in Figure 2.6 is given in Figure 2.7, where nodes \mathbf{s}_n and θ_n represent the locations and directions of minutia \mathbf{x}_n . For each conditional distribution, a directed link is added from nodes corresponding to the variables on which the distribution is

²It should be noted that a Bayesian network does not necessarily imply a full Bayesian approach involving prior and posterior distributions of parameters. The main implication is the explicit characterization of conditional probabilities in the distributions[8, 35].

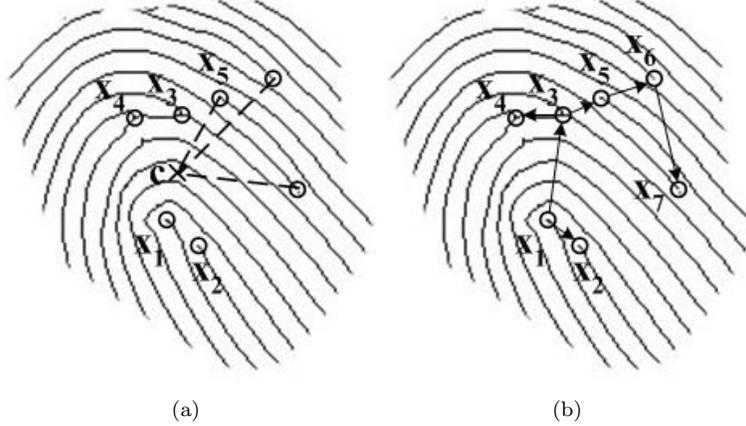


Figure 2.6: Sequential ordering of minutiae: (a) given minutiae $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4\}$ with centroid c , the next minutia \mathbf{x}_5 is selected by comparing the remaining minutia distances to c , thereby providing a sequencing, (b) dependency between the sorted minutiae is represented by arrows.

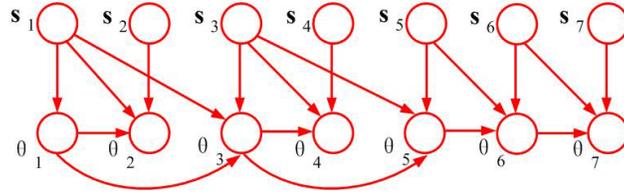


Figure 2.7: Directed probabilistic graphical model to represent the joint distribution of minutiae. This model corresponds to example in Figure 2.6(b). Minutiae locations are represented by nodes labeled \mathbf{s}_n and corresponding orientations are represented by nodes labeled θ_n . This joint distribution can be written directly from the model as $p(\mathbf{s}_1)p(\theta_1|\mathbf{s}_1)p(\mathbf{s}_2)p(\theta_2|\mathbf{s}_1, \mathbf{s}_2, \theta_1)p(\mathbf{s}_3)p(\theta_3|\mathbf{s}_1, \mathbf{s}_3, \theta_1)\dots$

conditioned. A general expression for the joint distribution of any minutiae set \mathbf{X} is

$$\begin{aligned}
 p(\mathbf{X}) &= p(\mathbf{s}_1)p(\theta_1|\mathbf{s}_1) \prod_{n=2}^N p(\mathbf{s}_n)p(\theta_n|\mathbf{s}_n, \mathbf{s}_{\psi(n)}, \theta_{\psi(n)}) \\
 &= p(\mathbf{s}_1, \theta_1) \prod_{n=2}^N p(\mathbf{s}_n)p(\theta_n|\mathbf{s}_n, \mathbf{s}_{\psi(n)}, \theta_{\psi(n)})
 \end{aligned} \tag{2.6}$$

where $\mathbf{s}_{\psi(n)}$ and $\theta_{\psi(n)}$ are the location and direction of minutiae \mathbf{x}_i which have the minimal spatial distance to minutia \mathbf{x}_n , and $\psi(n) = \operatorname{argmin}_{i \in [1, n-1]} \|\mathbf{x}_n - \mathbf{x}_i\|$. To compute the joint probability $p(\mathbf{X})$, three distributions are needed: $p(\mathbf{s})$, distribution of minutia location, $p(\mathbf{s}, \theta)$, the distribution of minutia location and direction, $p(\theta_n|\mathbf{s}_n, \mathbf{s}_{\psi(n)}, \theta_{\psi(n)})$, the conditional distribution of minutiae direction given its location, and the location and direction of the nearest minutiae. Each of these are addressed next.

Minutia location. Since minutiae tend to form clusters [58] a mixture of Gaussians, with K_1 components, is used:

$$p(\mathbf{s}) = \sum_{k_1=1}^{K_1} \pi_{k_1} \mathcal{N}(\mathbf{s}|\mu_{k_1}, \Sigma_{k_1}). \tag{2.7}$$

Minutiae location and direction. Since minutiae in different regions are associated with different region-

specific directions, a mixture of joint Gaussian and von-Mises distributions with K_2 components is used:

$$p(\mathbf{s}, \theta) = \sum_{k_2=1}^{K_2} \pi_{k_2} \mathcal{N}(\mathbf{s} | \mu_{k_2}, \Sigma_{k_2}) \mathcal{V}(\theta | \nu_{k_2}, \kappa_{k_2}). \quad (2.8)$$

While we can get $p(\mathbf{s})$ by marginalizing $p(\mathbf{s}, \theta)$ for each continuous value of \mathbf{s} , it is time consuming and unnecessary and thus Eq 2.7 is used.

Conditional minutia direction. Minutiae direction, given its location and the the nearest minutiae among \mathbf{x}_1 to \mathbf{x}_{n-1} , is a mixture of von-Mises densities with K_3 components:

$$p(\theta_n | \mathbf{s}_n, \mathbf{s}_{\psi(n)}, \theta_{\psi(n)}) = \sum_{k_3=1}^{K_3} \pi_{k_3} \mathcal{V}(\theta_n | \nu_{k_3}, \kappa_{k_3}). \quad (2.9)$$

where π_{k_i} are non-negative component weights that sum to one, $\mathcal{N}(s | \mu_k, \Sigma_k)$ is the bivariate Gaussian probability density function of minutiae with mean μ_k and covariance matrix Σ_k , and $\mathcal{V}(\theta | \nu_k, \kappa_k)$ is the von-Mises probability density function of minutiae orientation with mean angle ν_k and precision (inverse variance) κ_{k_3} [8]

$$\mathcal{V}(\theta | \nu_k, \kappa_k) = \frac{1}{2\pi I_0(\kappa_k)} \exp[\kappa_k \cos(\theta - \nu_k)]. \quad (2.10)$$

The number of components K_i in each of the mixture distributions can be determined using the Bayes information criterion (BIC). Other parameters are determined using the EM algorithm (see Appendix (1)).

Model Validation

The next step is to validate the probabilistic graphical model. In machine learning where the goal is prediction, the usual method of validation of a model is to determine error rates on a test set. Here we are not making a decision. Instead, validation implies evaluating a measure as to how well the model fits the data. A standard goodness-of-fit measure is the Pearson chi-squared (χ^2) test. The test determines whether observed sample frequencies are consistent with expected frequencies specified in the null hypothesis. The test, applied to binned data, requires a sufficient sample size in each bin in order for the chi-square approximation to be valid [18].

Three different tests were conducted for each of the distributions in Eqs. 2.7, 2.8, and 2.9. For minutiae location, the space was partitioned into 16 non-overlapping blocks. For minutiae location and orientation, there were 16×4 non-overlapping blocks.

For minutia dependency, the orientation space was divided into 9 non-overlapping blocks. The blocks were combined with adjacent blocks until both observed and expected numbers of minutiae in the block were greater than or equal to 5. The test statistic used was the chi-square random variable $\chi^2 = \sum_i \frac{(O_i - E_i)^2}{E_i}$ where O_i is the observed minutiae count for the i th block, and E_i is the expected minutiae count for the i th block. The p -value, the probability of observing a sample statistic as extreme as the test statistic, associated with each test statistic χ^2 is then calculated based on the chi-square distribution and compared to the significance level. For the *NIST 4* dataset, we chose significance level equal to 0.01. The generative models were trained using 4000 fingerprints. BIC yielded $K_1 = K_2 = 3$. For K_3 there are 4096 different values for different condition settings.

To test the three models, the numbers of fingerprints with p -values above (corresponding to accept the model) and below (corresponding to reject the model) the significance level were computed. The results are shown in Table 2.2.

Of the 4,000 fingerprints, 3,387 were accepted and 613 rejected for minutia location model, and 3,216 were accepted and 784 rejected for minutia location and orientation model. To test the model for minutia dependency, we first collected all the linked minutia pairs in the minutia sequences produced from 4,000 fingerprints. Then these minutia pairs were separated by the binned locations of both minutiae (32×32) and orientation of leading minutiae (4). Finally, the minutia dependency models were tested on corresponding minutia pair sets. Of the 4,096 dependency models, 3,558 were accepted and 538 rejected. The results

Table 2.2: Validation of Models: χ^2 test of CPDs.

Distribution	Data size	Accept	Reject
Location $p(\mathbf{s})$	4000	3387	613
Location & orientation $p(\mathbf{s}, \theta)$	4000	3216	784
Conditional orientation $p(\theta_n \mathbf{s}_n, \mathbf{s}_{\psi(n)}, \theta_{\psi(n)})$	4096	3558	538

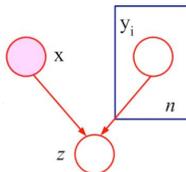


Figure 2.8: Graphical model for rarity. Specific nPRC is the conditional probability that a known \mathbf{x} is found among at least one of $\mathbf{y}_1, \dots, \mathbf{y}_n$, where z is an indicator variable for a match.

imply that both generative models have a reasonable and accurate fit for fingerprints.

2.2.3 Rarity Evaluation

The probability that an input \mathbf{x} coincides with one of n samples, within specified tolerance, is defined as the *specific nPRC* [69]. Since it is conditional to the known input it can also be referred to as the *conditional nPRC*. Specific nPRC is different from the probability of random correspondence (PRC), which is the probability that a random pair of samples will have the same value. It is also different from nPRC which is the probability that some pair of samples among n have the same value [72].

We make this idea precise using the graphical model of Figure 2.8, where \mathbf{x} represents the input in a feature space, and $\mathbf{Y} = \{\mathbf{y}_1, \dots, \mathbf{y}_n\}$, in plate notation, represents the set of n database entries. The binary-valued indicator random variable z takes one of two values $\{z^0, z^1\}$. It has value z^0 when at least one database element \mathbf{y}_i has the value \mathbf{x} , and z^1 otherwise. The conditional probabilities can be written as $P(z^0 | \mathbf{x} \in \mathbf{Y}) = 1, P(z^0 | \mathbf{x} \notin \mathbf{Y}) = 0, P(z^1 | \mathbf{x} \in \mathbf{Y}) = 0, P(z^1 | \mathbf{x} \notin \mathbf{Y}) = 1$. By marginalizing over \mathbf{Y} , the specific or conditional nPRC is given by

$$P(z = z^0 | \mathbf{x}) = \sum_{\mathbf{Y}} P(z = z^0 | \mathbf{x}, \mathbf{Y}) P(\mathbf{Y}) \quad (2.11)$$

where $P(\mathbf{Y})$ is the joint probability, or likelihood, of \mathbf{Y} . Since the space of \mathbf{Y} consists of all possible entries in the database, which is huge, we use an alternative method of determining the conditional nPRC. Assume that database entries are independent and identically distributed. Since the probability of \mathbf{x} not matching each element in the database is $1 - P(\mathbf{x})$ we have

$$P(z = z^0 | \mathbf{x}) = 1 - (1 - P(\mathbf{x}))^n \quad (2.12)$$

If \mathbf{x} is continuous with distribution $p(\mathbf{x})$ the conditional nPRC is the probability of finding an element among n within tolerance ϵ is

$$P(\mathbf{x}, \mathbf{Y} : \epsilon, n) = 1 - (1 - p_\epsilon(\mathbf{x}))^n \quad (2.13)$$

where p_ϵ is the probability of \mathbf{x} in the interval $\mathbf{x} \pm \epsilon$.

Minutia tolerance is defined as follows. Minutia pair $\mathbf{x}_a = (\mathbf{s}_a, \theta_a)$ and $\mathbf{x}_b = (\mathbf{s}_b, \theta_b)$ correspond if for tolerance $\epsilon = [\epsilon_s, \epsilon_\theta]$

$$\| \mathbf{s}_a - \mathbf{s}_b \| \leq \epsilon_s \wedge |\theta_a - \theta_b| \leq \epsilon_\theta \quad (2.14)$$

where $\| \mathbf{s}_a - \mathbf{s}_b \|$ is the Euclidean distance between their locations. A match implies the existence of at least

\hat{m} corresponding pairs for specified ϵ and \hat{m} .

Since \mathbf{s} and θ are continuous-valued, and tolerance is considered, the evaluation of specific n PRC involves integrating over the relevant distributions. Also, since minutiae quality varies greatly in latent images, it is useful to take into account minutia confidence in specific n PRC evaluation. We describe next these evaluations with and without consideration of uncertainty in minutiae and core point.

Minutiae and core point with point estimates

Here we assume that the core point and minutiae have no associated uncertainty, i.e., they are point estimates. For the core point the MAP estimate, as provided by the GP in Eq. 2.4, is used and for detected minutiae the confidence values are ignored.

Let \mathbf{X} be a fingerprint with N minutiae. Let \mathbf{X}' be another fingerprint with M minutiae. Let $\tilde{\mathbf{X}}$ be the set of \hat{m} minutiae common to both \mathbf{X} and \mathbf{X}' where the matching minutiae are within ϵ of each other. The distribution of the common points can be written from Eq 2.6 as

$$p_\epsilon(\tilde{\mathbf{X}}) = p_\epsilon(\mathbf{s}_1, \theta_1) \prod_{n=2}^{\hat{m}} p_\epsilon(\mathbf{s}_n) p_\epsilon(\theta_n | \mathbf{s}_n, \mathbf{s}_{\psi(n)}, \theta_{\psi(n)}) \quad (2.15)$$

where the three product terms are expanded as

$$p_\epsilon(\mathbf{s}_n, \theta_n) = \int_{|\mathbf{s}-\mathbf{s}'_n| \leq \epsilon_s} \int_{|\theta-\theta'_n| \leq \epsilon_\theta} p(\mathbf{s}, \theta) ds d\theta,$$

$$p_\epsilon(\mathbf{s}_n) = \int_{|\mathbf{s}-\mathbf{s}'_n| \leq \epsilon_s} p(\mathbf{s}) ds, \text{ and}$$

$$p_\epsilon(\theta_n | \mathbf{s}_n, \mathbf{s}_{\psi(n)}, \theta_{\psi(n)}) = \int_{|\theta-\theta'_n| \leq \epsilon_\theta} p(\theta | \mathbf{s}_n, \mathbf{s}_{\psi(n)}, \theta_{\psi(n)}) d\theta$$

The specific n PRC is computed using Eq. (2.13) by

$$p_\epsilon(\mathbf{X}, \hat{m}, n) = 1 - (1 - p_\epsilon(\mathbf{X}, \hat{m}))^n \quad (2.16)$$

where $p_\epsilon(\mathbf{X}, \hat{m})$, the probability that \hat{m} pairs of minutiae correspond, is given by

$$p_\epsilon(\mathbf{X}, \hat{m}) = \sum_{m' \in M} p(m') \binom{m'}{\hat{m}} p_\epsilon(\tilde{\mathbf{X}}_i) \quad (2.17)$$

where M contains all possible numbers of minutiae in one fingerprint among n fingerprints, $p(m')$ is the probability of a random fingerprint having m' minutiae, minutiae set $\tilde{\mathbf{X}}_i = (\mathbf{x}_{i1}, \mathbf{x}_{i2}, \dots, \mathbf{x}_{i\hat{m}})$ is the subset of \mathbf{X} and $p_\epsilon(\tilde{\mathbf{X}}_i)$ is the joint probability of minutiae set $\tilde{\mathbf{X}}_i$ given by Eq. (2.15).

Minutiae uncertainty

Next we consider a model for uncertainty associated with minutiae. These arise from confidence values assigned either by a human examiner or by an AFIS system.

Assume that confidence of minutia \mathbf{x}_n is given by (d_{s_n}, d_{θ_n}) , where d_{s_n} is location confidence and d_{θ_n} is direction confidence. Given minutiae $\mathbf{x}_n = (s_n, \theta_n)$ and their confidences, the distributions of location s' and direction θ' can be modeled by Gaussian and von-Mises distributions

$$c(\mathbf{s}' | \mathbf{s}_n, d_{s_n}) \sim \mathcal{N}(\mathbf{s}' | \mathbf{s}_n, d_{s_n}^{-1}) \quad (2.18)$$

$$c(\theta' | \theta_n, d_{\theta_n}) \sim \mathcal{V}(\theta' | \theta_n, d_{\theta_n}) \quad (2.19)$$

where the precision (inverse variance) of location distribution d_{s_n} represents location confidence and the concentration parameter of direction distribution d_{θ_n} represents direction confidence.

Ranges for these values can be assigned based on image resolution, e.g., since ridges are ten pixels wide in the *NIST* dataset, location confidence d_{s_n} is in the interval $[0.01, 1]$ and orientation confidence d_{θ_n} is in the interval $[1, 10]$, where a high confidence value implies a high quality minutia.

By application of the sum rule of uncertainty of minutiae, the conditional distributions involved in Eq. (2.15) are given by

$$p_\epsilon(\mathbf{s}_n, \theta_n) = \int \int \int \int_{s' \theta' |\mathbf{x} - \mathbf{x}'| \leq \epsilon} c(\mathbf{s}' | \mathbf{s}_n, d_{s_n}) c(\theta' | \theta_n, d_{\theta_n}) p(\mathbf{s}, \theta) d\mathbf{s}' d\theta' d\mathbf{s} d\theta \quad (2.20)$$

$$p_\epsilon(\mathbf{s}_n) = \int \int_{s' |\mathbf{s} - \mathbf{s}'| \leq \epsilon_s} c(\mathbf{s}' | \mathbf{s}_n, d_{s_n}) p(\mathbf{s}) d\mathbf{s}' d\mathbf{s} \quad (2.21)$$

$$p_\epsilon(\theta_n | \mathbf{s}_n, \mathbf{s}_{\psi(n)}, \theta_{\psi(n)}) = \int \int_{\theta' |\theta - \theta'| \leq \epsilon_\theta} c(\theta' | \theta_n, d_{\theta_n}) p(\theta | \mathbf{s}_n, \mathbf{s}_{\psi(n)}, \theta_{\psi(n)}) d\theta' d\theta \quad (2.22)$$

Specific n PRCs can again be computed by Eq. (2.16) and Eq. (2.17). Since confidence distributions are sharply peaked and the definite integral intervals are small, numerical integration can be used in probability calculation.

The complexity of joint probability of a print with \hat{m} matching minutiae is $O(\hat{m})$. The computational cost of specific n PRC for a certain minutia set is $O(M\hat{m})$, where M is the maximum number of minutiae in a fingerprint.

Core point uncertainty

Here we take the Bayesian approach where the core point does not have a fixed value but has a distribution instead. Let y be the core point whose distribution is $p(y)$ as given in Eq. 2.3. Thus the distribution of $\tilde{\mathbf{X}}$, which is the set of minutiae common to \mathbf{X} and \mathbf{X}' , is obtained by integrating out the core point parameter as

$$p_\epsilon(\tilde{\mathbf{X}}) = \int p_\epsilon(\tilde{\mathbf{X}} | y) p(y) dy \quad (2.23)$$

where $p_\epsilon(\tilde{\mathbf{X}} | y)$ is given by Eq.2.15 and the minutia set are subjected to coordinate transformation specified by core point y .

Examples of rarity evaluation

Evaluation of fingerprint rarity is demonstrated with three sets of examples: (i) a few simple minutia configurations where we assume point estimates of core point and minutiae, (ii) the Madrid train bombing case where we use a point estimate of the core point and minutiae with/without confidence values, and (iii) latent prints from a standard data set where we use a fixed core point and an uncertain core point.

As in any machine learning scenario, the evaluation of rarity depends on the data set from which the parameters are determined. The database should be representative enough. We used the largest publicly available database for the learning phase, viz., *NIST special database 4*, which contains 2,000 8-bit gray scale fingerprint image pairs.

A. Simple minutia configurations Determining the rarity of configurations of few minutiae is useful to the latent print examiner who needs to decide whether to proceed further. A minutia configuration involving only three minutiae, with no uncertainty in either minutiae or core point, is shown in Figure 2.9. It has a high specific n PRC of 0.012 in a database of 1,000 entries. When the common minutia structure is perturbed, by changing a few minutiae orientations, a much lower probability of finding a match is observed.

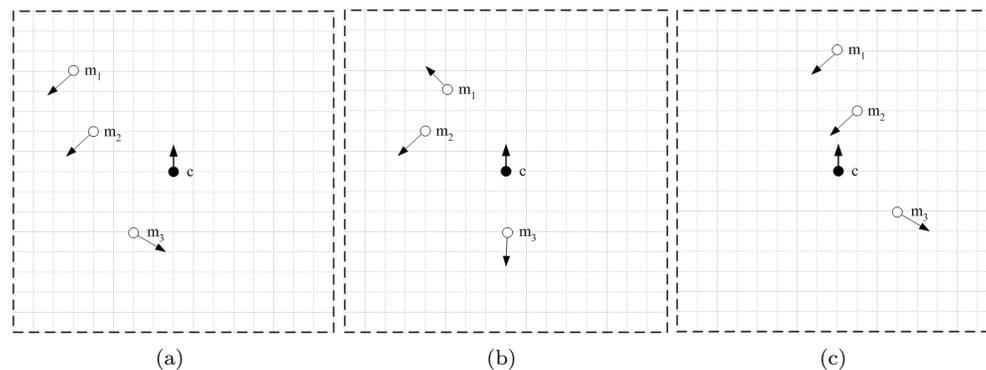


Figure 2.9: Simple configurations of minutiae and core points: (a) a common configuration with three minutiae m_1 , m_2 and m_3 and core point c , (b) an uncommon configuration obtained by changing the orientations of m_1 and m_3 and (c) an uncommon configuration obtained by translating the three minutiae with respect to the core point. For $n = 1000$, their specific n PRC values are: (a) 1.2×10^{-2} , (b) 7.97×10^{-4} and (c) 2.3×10^{-6} respectively.

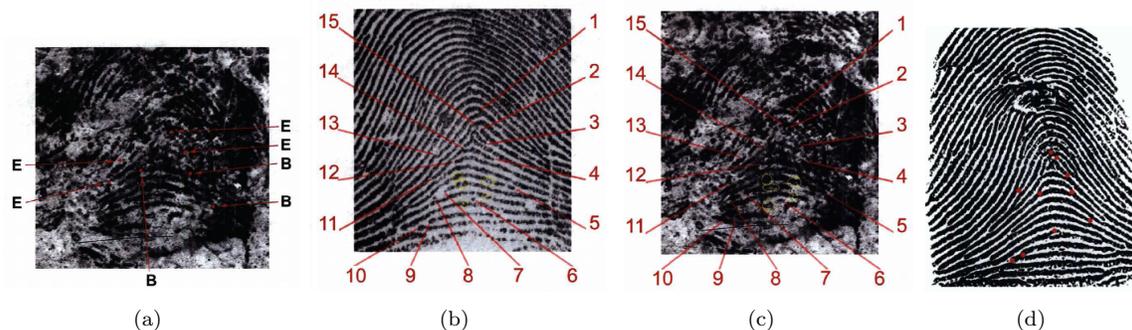


Figure 2.10: Brandon Mayfield case: prints used in rarity evaluation: (a) latent print $LFP17$ found at crime scene with seven marked minutiae (initial annotation), (b) matching ten-print of Mayfield found in FBI database with 15 charted minutiae, (c) $LFP17$ re-annotated with the same 15 minutiae as in (b), and (d) ten-print of Daoud with ten matching minutiae (from [78]). [Best Viewed in Color]

B. The Brandon Mayfield Case

The Brandon Mayfield case [78] provides a useful test scenario since both the latent print and minutiae annotations of the image are available. Particulars of the case are relevant since they illustrate the value of rarity evaluation.

In 2004 there was a terrorist bombing in the Madrid train system leading to the death of nearly 200 individuals and injury of 2,000 more people. A latent print was found on a plastic bag of detonators in a nearby van. This was erroneously identified as corresponding to the inked ten-print of Brandon Mayfield, an attorney in Oregon. The identification of Mayfield was effected through an AFIS search of the Federal Bureau of Investigation (FBI) criminal files consisting of 470 million fingerprints.

The latent print, tagged as $LFP17$, was initially marked as having seven identifiable minutiae (Figure 2.10(a)). Upon observing the Mayfield ten-print (Fig. 2.10(b)), $LFP17$ was re-annotated as having 15 minutiae that correspond to the Mayfield ten-print (Figure 2.10(c)). Subsequently Spanish National Police (SNP) identified the true perpetrator as an Algerian national, Ouhmane Daoud, whose ten-print is shown in Fig. 2.10(d). Rarities of the minutiae sets in this problem, evaluated under two scenarios: full confidence of minutiae and minutiae with uncertainty, are given in Table 2.3. In both cases a point estimate for the core point was assumed.

Case 1: Point estimates of minutiae (Row 1 of Table 2.3): The probability of finding the seven minutiae in the FBI database (specific n PRC with $n = 4.7 \times 10^8$), is 0.78, which is a very high probability. For the

Table 2.3: Brandon Mayfield case: Probability of randomly matching the latent print *LFP17* with an item in FBI database (specific *n*PRC with $n = 4.7 \times 10^8$).

Minutiae Assumption	Uncertainty	Seven Minutiae in Fig 2.10(a)	Fifteen Minutiae in Fig 2.10(c)
1. Point Estimates		0.78	1.2×10^{-6}
2. Uncertain Minutiae		0.93	7.8×10^{-7}

15 minutiae in the re-annotated print, the specific *n*PRC is 1.2×10^{-6} which is much lower than with seven minutiae, but in absolute terms it is about 1 in a million, which is quite high.

Case 2: Uncertain minutiae (Row 2 of Table 2.3) We manually assigned a confidence to each minutia in each set as described in Section 2.2.3. For the seven minutiae set in Figure 2.10(a), confidences (d_s, d_θ) were assigned clockwise starting from 1 o'clock, as follows : (0.1, 7), (0.6, 8), (0.2, 8), (0.7, 6), (0.9, 9), (0.6, 8), (0.9, 9). Specific *n*PRC evaluates to 0.93, which is a high probability. For the 15 minutiae set in Figure 2.10(c) the confidences for (d_s, d_θ) were: (0.3, 7), (0.1, 7), (0.6, 8), (0.2, 8), (0.6, 7), (0.8, 6), (0.9, 9), (0.9, 9), (0.6, 5), (0.5, 5), (0.6, 8), (0.8, 9), (0.9, 9), (0.7, 5), (0.6, 5). The specific *n*PRC is again close to 1 in a million.

Consider now the effect of incorporating minutia uncertainties into the model. In both cases there is a much higher probability of match with fewer minutiae (7) than with more minutiae (15). With only seven minutiae, the probability of random correspondence increases to 0.93, a significant increase of 0.15. With 15 minutiae, the specific *n*PRC is 7.8×10^{-7} , which is a small decrease of 0.42×10^{-6} . The relative increase/decrease of specific *n*PRC with minutia uncertainty, particularly when many minutes are considered, can be attributed to the configuration of minutiae—since the overall distribution is a mixture of the individual distributions.

Consider next the rarity of the ten minutiae that are common between the re-annotated latent print and the ten-print of Daoud shown in Fig. 2.10(d). The specific *n*PRC of the ten minutiae, assuming point estimates, is 0.014, again a high probability. In general, configurations of ten minutiae are quite rare, e.g., when 200 ten minutiae sets were randomly chosen from fingerprints in the *NIST4* dataset, which contains 4,000 images, with $n = 470 \times 10^6$, the average specific *n*PRC was 7.1×10^{-8} ; this average value is by definition known as *n*PRC [72]. We conclude that while ten minutiae have significant discriminatory power in general, the particular configuration of ten minutiae common to the latent print and the Daoud print were those that are more common.

Evaluation of rarity in this case illustrates the potential use of such a measure in practice. The seven or ten matching minutiae initially found in both the evidence and known are relatively common (equivalently, they have low rarity). Thus a higher degree of match should be sought in subsequent examination before issuing an opinion.

C. Latent prints from Standard Dataset

Next we consider the effect of core point uncertainty on the evaluation of rarity. Two latent prints from the *NIST 27* data set are shown in Figure 2.11: print *b115* is from the *bad* quality set and print *g73* is from the *good* quality set. Latent print *b115*, contains $N = 16$ minutiae and *g73* contains $N = 39$ minutiae. Minutiae confidences were manually assigned by visual inspection; the values assigned for the 55 minutiae are analogous to the smaller minutiae set of minutiae in the example of Section 2.2.3.

Specific *n*PRCs for the two prints, computed for varying numbers of matching minutiae pairs \hat{m} , assuming fingerprint database size $n = 100,000$, are given in Table 2.4. The tolerance is set at $\epsilon_s = 10$ pixels and $\epsilon_\theta = \pi/8$. In each case the first column shows the specific *n*PRC with point estimates for the core point and the second column shows the specific *n*PRC with core point uncertainty as determined by GP.

The cases show that specific *n*PRC depends on the given latent fingerprint. As the number of minutiae to be matched, \hat{m} , increases the probability of finding a random match decreases. When all the minutiae are considered an extremely low value of specific *n*PRC is observed. Thus the values of specific *n*PRC provide a measure for the strength of latent fingerprint evidence.

Whether to use for the core point, a point estimate or a distribution of it, in rarity evaluation is an

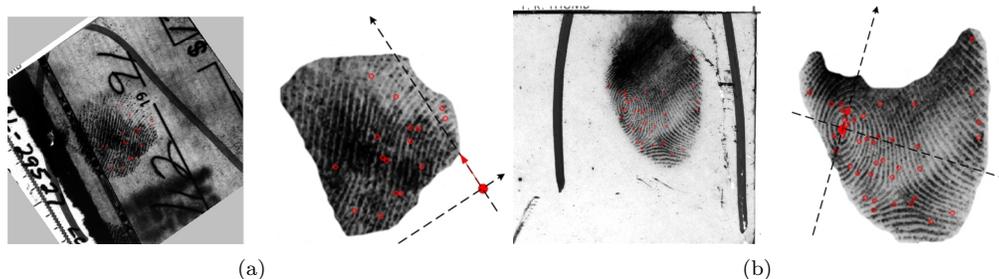


Figure 2.11: Two latent prints from *NIST27*: (a) *b115* is from the *bad* dataset and (b) *g73* is from the *good* dataset. In each case the left image is the print and the right its aligned version with predicted core point. Corresponding rarity values for $n = 100,000$ are given in Table 2.4. Rarity values of (a) for different values of n and different numbers of matching minutiae are plotted in Figure 2.12.

Table 2.4: Probability of finding a match in a database. Considering the two latent prints *b115* (from the bad set) and *g73* (from the good set) in *NIST27* (shown in Figure 2.11), the probability of finding a corresponding print in a database of 100,000 prints is evaluated, using the following tolerances: minutia location $\epsilon_s = 10$ pixels and minutia direction $\epsilon_\theta = \pi/8$.

X = Latent Print <i>b115</i>				X = Latent Print <i>g73</i>			
N = No. of minutiae in X	No. matching \hat{m}	Specific n PRC = $p_\epsilon(\mathbf{X}, \hat{m}, n = 10^5)$		N = No. of minutiae in X	No. matching \hat{m}	Specific n PRC = $p_\epsilon(\mathbf{X}, \hat{m}, n)$	
		Core-point is Point-Estimated	Core-point has a distribution			Core-point is Point-Estimated	Core-point has a distribution
16	2	0.73	0.79	39	4	1	1
	4	9.04×10^{-6}	1.81×10^{-5}		8	3.11×10^{-14}	1.50×10^{-15}
	8	2.46×10^{-19}	4.54×10^{-17}		12	2.56×10^{-25}	1.07×10^{-26}
	12	6.13×10^{-31}	6.05×10^{-28}		24	3.10×10^{-52}	9.93×10^{-55}
	16	1.82×10^{-46}	2.93×10^{-41}		39	7.51×10^{-79}	6.16×10^{-82}

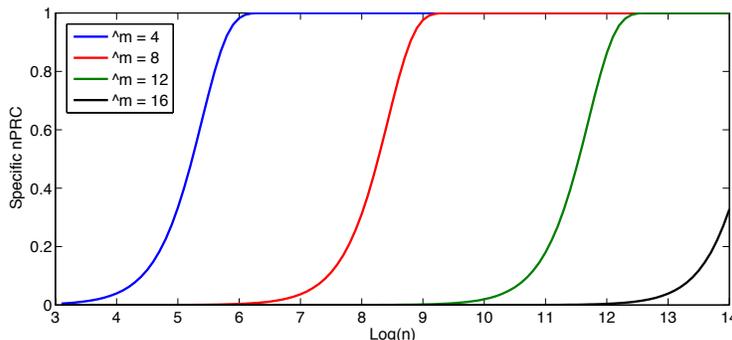


Figure 2.12: Dependence of specific n PRC of latent print $b115$ on database size n parameterized by number of corresponding minutiae $\hat{m} = 4, 8, 12, 16$ and n varying from 10^3 to 10^{14} .

interesting question. It is analogous to the question of whether the maximum likelihood estimate is better than a Bayesian estimate in general parameter estimation. The answer to that question is that when there is uncertainty about the parameter, as when the sample size is limited, the Bayesian estimate is better. Otherwise both give the same results. In the case of fingerprint comparison, if the core point is not visible, we estimated it. As was seen in Table 2.1 the MAP estimate of the core point has a 15% error rate. Thus it is better to use the distribution of the core point. Note that in the Brandon Mayfield case the core point was visible in the latent print, thus a distribution was not needed.

It is also interesting to observe how rarity varies with database size. As can be seen in Figure 2.12 for latent print $b115$, finding a match for four minutiae, or $\hat{m} = 4$ (blue curve), in a database of a million entries, i.e., $\log n = 6$, is guaranteed (specific n PRC=1). Whereas for matching twelve minutiae (green curve) a trillion database entries ($\log n = 12$) are needed for a random match.

2.2.4 Determining the Probability of Identification

Finally we discuss here as to how the rarity computation can be used in determining the probability of identification/exclusion. It is based on formulating the likelihood ratio method which is gaining acceptance in forensics in general and fingerprints in particular.

Let $S = \{s_i\}$ be a set of sources. They correspond to, say, fingers. Let \mathbf{o} be a random variable representing an object drawn from a source s_i , e.g., impression of a known finger. Let \mathbf{e} be a random variable representing evidence drawn from a source s_j , e.g., a crime scene impression. The task is to determine the probability of whether \mathbf{o} and \mathbf{e} came from the same.

We can state two opposing hypotheses:

h^0 : \mathbf{o} and \mathbf{e} are from the *same* source ($i = j$); and

h^1 : \mathbf{o} and \mathbf{e} are from *different* sources ($i \neq j$), which are the *identification* and *exclusion* hypotheses of forensics; some forensic statistics literature also refers to them as *prosecution* and *defense* hypotheses[1].

We can define two joint probability distributions $P(\mathbf{o}, \mathbf{e}|h^0)$ and $P(\mathbf{o}, \mathbf{e}|h^1)$ which specify as to how often each instance of the object and evidence occur together when they belong to the same source or to different sources. The relative strengths of evidence supporting the two hypotheses is quantified by the likelihood ratio

$$LR_J = LR(\mathbf{o}, \mathbf{e}) = \frac{P(\mathbf{o}, \mathbf{e}|h^0)}{P(\mathbf{o}, \mathbf{e}|h^1)}. \quad (2.24)$$

The corresponding log-likelihood ratio, $LLR(\mathbf{o}, \mathbf{e}) = \ln P(\mathbf{o}, \mathbf{e}|h^0) - \ln P(\mathbf{o}, \mathbf{e}|h^1)$, has representational advantages: its sign is indicative of same or different source, it has a smaller range than LR, and additivity of contributions of independent features³.

³In performing LLR additions, since LR values in the interval $(1, \infty)$ convert to positive LLRs and LR values in the interval

It is useful to convert LR into probabilities of identification and exclusion using a Bayesian formulation. Let the *prior* probabilities of the hypotheses be $P(h^0)$ and $P(h^1)$ with $P(h^0) + P(h^1) = 1$. Defining the prior odds as $O_{prior} = \frac{P(h^0)}{P(h^1)}$, we can express the prior probability of the same source as $P(h^0) = O_{prior}/(1+O_{prior})$. The prior odds can be converted into posterior odds as $O_{posterior} = \frac{P(h^0|\mathbf{o},\mathbf{e})}{P(h^1|\mathbf{o},\mathbf{e})} = O_{prior} \times LR(\mathbf{o}, \mathbf{e})$. Thus we can write the posterior probability of the same source as $P(h^0|\mathbf{o}, \mathbf{e}) = O_{posterior}/(1 + O_{posterior})$. The particular case of equal priors is of interest in forensics, as opinion without prior bias. In this case we get a simple form for the probability of identification as

$$P(h^0|\mathbf{o}, \mathbf{e}) = \frac{LR(\mathbf{o}, \mathbf{e})}{1 + LR(\mathbf{o}, \mathbf{e})} = \frac{\exp(LLR(\mathbf{o}, \mathbf{e}))}{1 + \exp(LLR(\mathbf{o}, \mathbf{e}))}. \quad (2.25)$$

The probability of exclusion is $P(h^1|\mathbf{o}, \mathbf{e}) = 1 - P(h^0|\mathbf{o}, \mathbf{e}) = 1/[1 + LR(\mathbf{o}, \mathbf{e})] = 1/[1 + e^{LLR(\mathbf{o}, \mathbf{e})}]$.

Thus the key to determining the probability of identification is to determine LR defined by Eq. 2.24, which in turn requires the distributions $P(\mathbf{o}, \mathbf{e}|h^i)(i = 0, 1)$, defined over all possible values of objects and their evidential forms. If \mathbf{o} and \mathbf{e} are n -dimensional binary vectors with each feature taking K possible values, then $2K^{2n}$ parameters are needed to specify the joint distribution. Determining these distributions is computationally and statistically infeasible. Computationally, kernel density estimation [2] and finite mixture models [43] have been proposed, but they have limitations as well⁴. More important is the statistical limitation of having a sufficient number of samples for so many parameters. Today, objects and evidence can be represented by ever finer features due to higher camera resolution and automatic feature extraction methods and their possible evidential forms is infinite.

Similarity Approximation

One method of simplification is to use a (dis)similarity function between object and evidence. The approach is to define $d(\mathbf{o}, \mathbf{e})$ as a scalar *distance* between object and evidence and define another likelihood ratio as follows

$$LR_D = \frac{P(d(\mathbf{o}, \mathbf{e})|h^0)}{P(d(\mathbf{o}, \mathbf{e})|h^1)}. \quad (2.26)$$

The number of parameters needed to evaluate LR_D is constant, or $O(1)$, and is independent of the number of features n . Due to its simplicity, this method has been widely used in fingerprint identification [45], handwriting analysis [62], pharmaceutical tablet comparison [9], etc. The probability distributions of dissimilarity using the Bozorth measure [80] (see Fig. 2.14) are given in Figure 2.13.

For certain feature spaces and distance measures, e.g., continuous features with Euclidean distance, this approach is equivalent to a *kernel* method [55]. The scalar distance d is just the magnitude of the vector difference \mathbf{d} . However, because it maps two distributions of $2n$ variables each into two scalar distributions there is severe loss of information (many pairs of \mathbf{o} and \mathbf{e} can have the same distance). A natural extension is to use vector difference \mathbf{d} , which quantifies the distribution of both the magnitude and the orientation of the difference between \mathbf{o} and \mathbf{e} , giving a much fine-grained characterization of the difference between \mathbf{o} and \mathbf{e} . While this likelihood ratio LR_{VD} , provides the simplification of mapping two distributions of $2n$ variables each into two distributions of n variables each, there is still a loss of information in the many to one mappings.

Similarity and Rarity Approximation

If the object and evidence are continuous scalar random variables drawn from the same or different sources; samples are normally distributed about its source mean with a known constant variance σ^2 ; the source mean is normally distributed with mean μ and variance τ^2 with $\tau \gg \sigma$, and there are p object samples with mean

(0, 1) convert to negative LLRs, the precisions of LR < 1 must be high, otherwise the ranges of positive and negative LLRs will not be symmetric.

⁴Kernel density estimation is expensive in memory as it needs to store the entire training data set, and the cost of evaluating the density grows linearly with size of data set. For mixture models, an important issue is to select the number of components, also, the training algorithm such as EM may converge to a local optimum or the boundary of the parameter space [26]

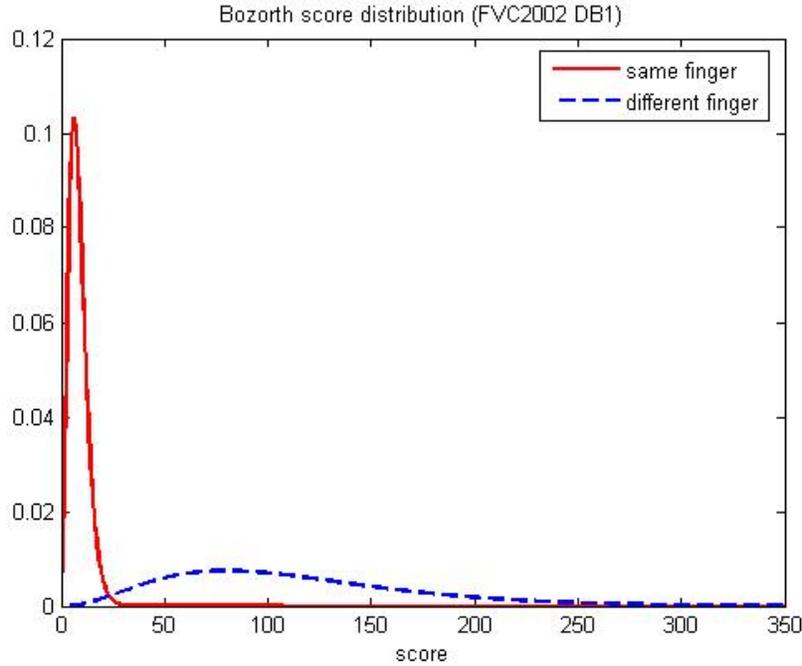


Figure 2.13: Distribution of Bozorth distance between fingerprints under the same (h^0) and different (h^1) hypotheses.

o , and q evidence samples with mean e with $p = q$, then the likelihood ratio can be approximated by

$$LR(o, e) = \frac{\tau}{\sigma\sqrt{2/p}} \exp\left\{-\frac{(o-e)^2}{4\sigma^2/p}\right\} \exp\left\{-\frac{(m-\mu)^2}{2\tau^2}\right\}, \quad (2.27)$$

where $m = (o + e)/2$ is the mean of o and e . This result is due to Lindley[38].

Lindley's result, which was defined for the univariate Gaussian case, can be generalized to vectors and graphs[74]. We begin by rewriting Eq. (2.27) for the univariate Gaussian case as the product of two factors, the first as the density of the difference (or dissimilarity) and the second as the reciprocal of the density of mean of object and evidence (or rarity), as follows.

$$\begin{aligned} LR &= \frac{\frac{1}{\sqrt{2}\sigma} \exp\left\{-\frac{(o-e)^2}{4\sigma^2}\right\}}{\frac{1}{\tau} \exp\left\{-\frac{(m-\mu)^2}{2\tau^2}\right\}} \\ &= \frac{\frac{1}{\sqrt{2\pi}\sqrt{2}\sigma} \exp\left\{-\frac{(o-e)^2}{2(\sqrt{2}\sigma)^2}\right\}}{\frac{1}{\sqrt{2\pi}\tau} \exp\left\{-\frac{(m-\mu)^2}{2\tau^2}\right\}} \\ &= \frac{\mathcal{N}(o-e|0, (\sqrt{2}\sigma)^2)}{\mathcal{N}(m|\mu, \tau^2)} \\ &= P(o - e|h^0) * \frac{1}{P(m)}. \end{aligned} \quad (2.28)$$

We generalize Eq. 2.28 from the univariate case to more general definition of difference. LR is *approximated* as the product of two factors:

$$LR_{DR} = P(\mathbf{d}(\mathbf{o}, \mathbf{e})|h^0) * \frac{1}{P(\mathbf{m}(\mathbf{o}, \mathbf{e}))}, \quad (2.29)$$

where $\mathbf{d}(\mathbf{o}, \mathbf{e})$ is the *difference* between \mathbf{o} and \mathbf{e} , and $\mathbf{m}(\mathbf{o}, \mathbf{e})$ is the *mean* of \mathbf{o} and \mathbf{e} . This approximation performs much better than a likelihood ratio LR_D that is based on distance (or equivalently, dissimilarity)

alone [74].

Similarity Term in LR computation

The first product term in Eq. 2.29 can be expanded to explicitly include dependence on \mathbf{o} and \mathbf{e} as

$$P(\mathbf{d}(\mathbf{o}, \mathbf{e})|h^0) = \sum_{\mathbf{o}, \mathbf{e}} P(\mathbf{d}(\mathbf{o}, \mathbf{e})|\mathbf{o}, \mathbf{e}, h^0)P(\mathbf{o}, \mathbf{e}|h^0). \quad (2.30)$$

According to Eq. 2.30 the difference distribution can be determined from pairs of (object, evidence) samples, i.e., knowns and latent prints, from the same source. The dependence on the distribution of sources (fingerprints) can be explicitly written as

$$P(\mathbf{o}, \mathbf{e}|h^0) = \sum_S P(\mathbf{o}, \mathbf{e}|h^0, s_i)P(s_i).$$

Rarity Term in LR computation

The second product term in Eq. 2.29 can be expanded as

$$P(\mathbf{m}(\mathbf{o}, \mathbf{e})) = P(h^0)P(\mathbf{m}(\mathbf{o}, \mathbf{e})|h^0) + P(h^1)P(\mathbf{m}(\mathbf{o}, \mathbf{e})|h^1).$$

It can be estimated from representative pairs of known and latents irrespective of whether they arise from the same finger; to keep the distribution *unbiased*, the number of sample pairs sampled under h^0 (same finger) should almost *equal* the number of sample pairs that are drawn under h^1 (different finger).

The weighted “mean” of the known and latent can be interpreted as the set of *common features* or *matched characteristics*, as they are the *nearest* candidate to represent the true mean. Rarity is the reciprocal of the probability (or in the continuous case, probability density function) of observing these common features in a general population. For categorical discrete features or graphs, $\mathbf{m}(\mathbf{o}, \mathbf{e})$ is the *common features* between \mathbf{e} and \mathbf{o} . The difference between categorical features is regarded as another categorical variable, which can be mapped to numbers for easy implementation, although those numbers do *not* indicate any ordering. So both $\mathbf{d}(\mathbf{o}, \mathbf{e})$ and $\mathbf{m}(\mathbf{o}, \mathbf{e})$ have the *same* dimension.

Since rarity is the reciprocal of probability its definition follows from that of probability. We can formally rarity in discrete and continuous spaces as follows.

Def. 1 (discrete) Given a probability space (Ω, \mathcal{F}, P) , the *rarity* of a random event $\xi \in \mathcal{F}$ is defined by

$$R(\xi) = \frac{1}{P(\xi)}, \quad (2.31)$$

where Ω is the sample space, $\mathcal{F} \subseteq 2^\Omega$ is the set of events, P is the probability measure, and $P(\xi) (\neq 0)$ is the probability of the event ξ .

Def.2(continuous) Let $\mathbf{x} = (x_1, \dots, x_n)^T$ be a *continuous* n -dimensional random vector with the p.d.f. $p(\mathbf{x})$ defined on a domain S . Suppose for every assignment of $\mathbf{x} \in S$, there is a *confidence interval* $(\mathbf{x} - \epsilon/2, \mathbf{x} + \epsilon/2)$ associated with \mathbf{x} at a given confidence level $1 - \alpha$, where $\epsilon = (\epsilon_1, \dots, \epsilon_n)^T$, and α is a small positive value less than 1. This interval $(\mathbf{x} - \epsilon/2, \mathbf{x} + \epsilon/2)$ is a n -dimensional *region* whose volume is $\prod_{i=1}^n \epsilon_i$. Then the rarity of the event that \mathbf{x} takes value \mathbf{x}_0 is given by

$$R(\mathbf{x} = \mathbf{x}_0) = \frac{1}{\int_{\mathbf{x}_0 - \epsilon/2}^{\mathbf{x}_0 + \epsilon/2} p(\mathbf{x}) d\mathbf{x}}. \quad (2.32)$$

Since the magnitude of ϵ is usually small, the density in the region $(\mathbf{x}_0 - \epsilon/2, \mathbf{x}_0 + \epsilon/2)$ can be considered as constant, therefore (2.32) can be approximated by

$$R(\mathbf{x} = \mathbf{x}_0) = \frac{1}{p(\mathbf{x}_0) \prod_{i=1}^n \epsilon_i}, \forall \mathbf{x}_0 \in S. \quad (2.33)$$

Experimental results using the rarity-similarity formulation in general forensics is discussed in [73].

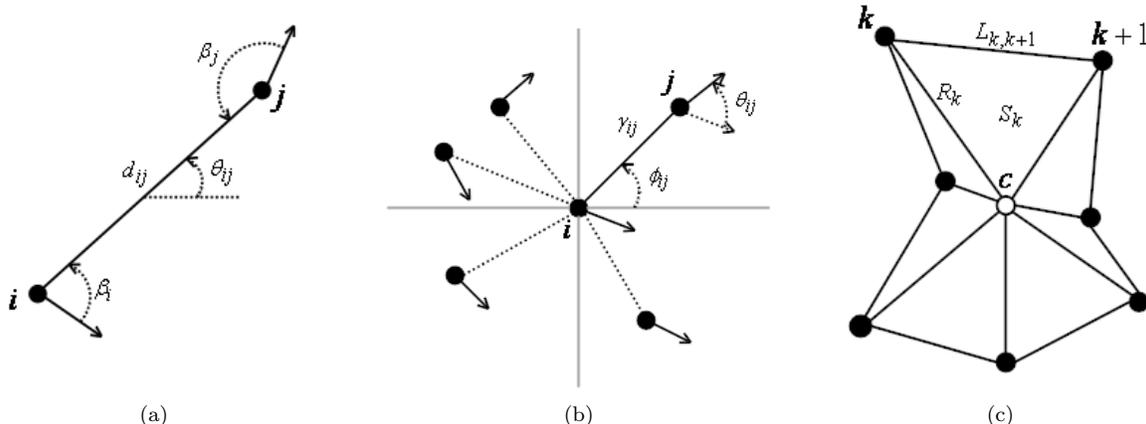


Figure 2.14: Determining fingerprint similarity using geometric structures: (a) minutiae pair (Bozorth), (b) a set of k -minutiae (CBFS) and (c) a polygon (Champod).

Computational Considerations

The distributions of difference and mean each involve n variables. The complexity of computing the two distributions exactly for non-Gaussian distributed data is still exponential with the number of features, with the number of parameters required being $2K^n$. However, we are dealing with two separate n -dimensional distributions rather than a single $2n$ -dimensional distribution as in the case of the joint distribution method. For determining the n -dimensional distributions, approximation techniques based on mixture models or probabilistic graphical models, can be used. Finally, the explicit factorization into similarity and rarity matches human intuition thereby having explanatory power which is essential in forensics.

2.2.5 A Probabilistic Measure of Similarity

The comparison phase of the ACE-V process is the side-by-side observation of friction ridge detail in two prints to determine agreement in the detail. Unlike human observation, which is a complex process involving comparison variety of perceived information, AFIS matchers generally use local minutiae structures [25] to quickly find a coarse correspondences between two sets of minutiae and then consolidate the local matching results at a global level. Based on local neighborhood structures properties invariant to rotation and translation are defined.

Existing methods use predefined feature sets that represent N annotated minutiae in “agreement” with each other. The structures can be pairs of minutiae [80], a set of k minutiae [17], stars, or other local structures [21] (see Figure 2.14). In one such method based on polygons [46, 45], the goal is to find the most optimal geometrical transformation that relates the two sets of minutiae without establishing local correspondence. Similarity is determined as the Euclidean distance between feature vectors. There are several shortcomings: not finding optimal correspondences while minimizing mismatches, dependency between minutiae points depends on the defined structure, structure is fixed once the polygon is built, inconsistent geometric structures for distorted fingerprints, and validation of discriminative power.

Cognitive Motivation

In the past few decades, psychologists have developed theories that unify similarity and recognition [5]. It is assumed that the greater the similarity between a pair of stimuli, the more likely one will be confused with the other in recognition. Thus, similarity is defined as a function of perceptual distributions. Many probabilistic similarity models have been proposed [6, 4]. These models assume that the perceptual effect of a stimulus is random and on any single trial it is a multivariate variable. Assuming that the percept

is probabilistic fundamentally changes the predictions of these models. Similarity models differ according to whether they assume the percept is deterministic or probabilistic and whether they assume that the judgement process through which a response is selected is deterministic or probabilistic [4]. The judgement process is deterministic if the same response is always made given the same information and it is probabilistic if, for each information state, a response is selected randomly by sampling from some probability distribution.

Models that assume a probabilistic percept can account for violations in some distance axioms[5]. Let two vectors of some dimensions S_A and S_B represent two stimuli, the perceived similarity of S_A to S_B is $s(S_A, S_B)$. Suppose the dissimilarity and similarity are inversely related, the perceived dissimilarity is measured by the psychological distance between the two $d(S_A, S_B)$. Let $\delta(S_A, S_B)$ be the judged similarity of S_A to S_B which is determined experimentally. These two similarities are related by a monotonic function g in the following assumption,

$$\delta(S_A, S_B) = g[d(S_A, S_B)]$$

If d is a metric function which satisfies following distance axioms:

- **Constance self-dissimilarity:**

$$d(S_A, S_A) = d(S_B, S_B)$$

This axiom is potentially testable because if judged dissimilarity is monotonically related to perceived dissimilarity, it implies that $\delta(S_A, S_A) = \delta(S_B, S_B)$ for all S_A and S_B . By reviewing empirical evidence against this assumption [36], it is argued that stimuli having few features in common with other objects in the stimulus domain, have a greater perceived self-similarity and so a smaller perceived self-dissimilarity. There is absolute distinction between the similarity between two matched fingerprint and two non-matched ones.

- **Minimality:**

$$d(S_A, S_A) < d(S_A, S_B)$$

Two different stimuli are always at least as dissimilar as either stimulus is to itself. This axiom is also potentially testable because it implies $\delta(S_A, S_A) < \delta(S_A, S_B)$ for all S_A and S_B .

- **Symmetry:**

$$d(S_A, S_B) = d(S_B, S_A)$$

This implies similarity is a symmetric relation and therefore $\delta(S_A, S_B) = \delta(S_B, S_A)$. This property can also be tested. The validity of this assumption may depend critically on the experimenter's instructions. For example, violations may be more likely if subjects are asked to judge the similarity of S_A to S_B than if they are asked to judge the similarity of S_B to S_A .

- **Triangle inequality:**

$$d(S_A, S_C) \leq d(S_A, S_B) + d(S_B, S_C)$$

This property cannot be tested experimentally. Empirical testing of the triangle inequality is problematic when perceived and judged dissimilarity are only monotonically related. In this case the fact that the perceived dissimilarities satisfy (or violate) the triangle inequality places no logical constraints on the judged dissimilarities. Tversky and Gati proposed a substitute for the triangle inequality, called corner inequality [77] and it is verified that this does not hold for probabilistic measures.

To sum up, a probabilistic similarity model can violate three distance axioms: non constant self-similarity, asymmetry, violation of the corner inequality. Even if d satisfies distance axioms, δ may not or vice versa. Existing fingerprint similarity models are all deterministic in representing perceived features and the decision process. However, in reality the information that forms a percept varies over time, the uncertainties reside in obtaining fingerprint itself favor similarity models that assume probabilistic percepts. Local interaction or structure among minutiae should be maintained despite geometric transformation and distortion.

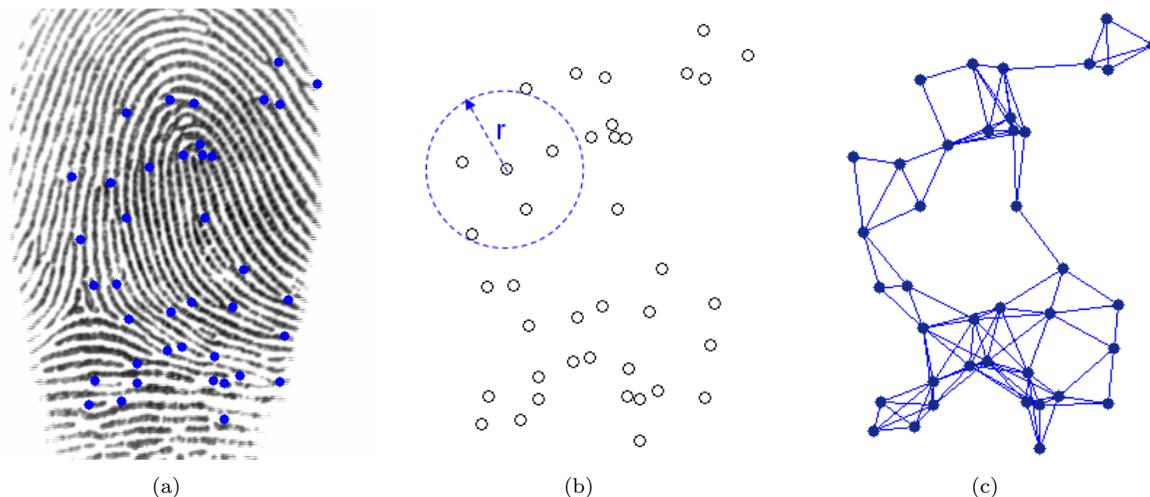


Figure 2.15: Neighborhood of a minutiae set: (a) fingerprint image with 39 minutiae extracted, (b) example of neighborhood r of a minutia, (c) the graph $\mathcal{G}(\mathcal{M}, \mathcal{E})$ of neighboring minutiae constructed by edging each pair of minutiae according to the neighborhood system defined by (2.34) in (b), where $r = 60$ pixels. The resulting graph contains 118 edges.

Markov Random Field Representation

We propose a probabilistic measure, which calculates a probability based on the most optimal correspondences between minutiae in two fingerprints, using a Markov Random Field (MRF) representation. MRFs are PGMs but unlike Bayesian networks there is no directionality associated with edges between variable nodes [8, 35].

In the proposed method, a mapping field that associates one set of minutiae to the other is defined for a neighborhood system. We are looking for the most likely assignments of minutiae alignment via maximizing the likelihood of associating the two sets of minutiae. Features are defined locally within the neighborhood to capture spatially relationships such like local interaction or dependence among minutiae. The uncertainties of fingerprints similarity are described by a generative model. Information of percepts are well encoded into the potentials of MRF.

Probabilistic inference is then performed to achieve a consistent global optimization. Similarity between two fingerprints is measured as the joint probability over the MRF. Distributions of the similarities given the alternative hypotheses are used to model the uncertainty in fingerprint.

Consider a set of minutiae $\mathcal{M} = \{m_i\}$ extracted from a fingerprint image, a neighborhood system for \mathcal{M} is defined as $\mathcal{N} = \{\mathcal{N}_i | m_i \in \mathcal{M}\}$, where \mathcal{N}_i is the set of neighbor minutiae m_i , defined as the set of minutiae within a radius of r from m_i .

$$\mathcal{N}_i = \{m_j \in \mathcal{M} \mid \text{dist}(m_i, m_j)^2 < r^2, j \neq i\} \quad (2.34)$$

Where the $\text{dist}(\bullet, \bullet)$ denotes the Euclidean Distance and ϵ takes an integer value. If there is no any minutiae within in the radius, then assign the nearest minutiae of m_i to be its neighborhood in order to guarantee the connectivity among all the minutiae. Let graph $G(\mathcal{M}, \mathcal{E})$ denotes the N minutiae network for a given fingerprint, a small r leads to a very densely connected graph while a large enough r will result in a fully connected graph. Neither the case is our desired. A proper r should be chosen in order to capture local relationship among local minutiae within a certain degree of complexity of the graph. The edges \mathcal{E} in \mathcal{G} represent links between sets of neighboring minutiae pair $\mathcal{E} = \{(m_i, m_j) | m_j \in \mathcal{N}_i\}$. An illustration of constructing the graph given a set of minutiae is shown in Figure 2.15. The neighborhood \mathcal{N}_i is a subset of \mathcal{M} and has the following properties:

1. A minutiae is not a neighbor to itself: $m_i \notin \mathcal{N}_i$.
2. the neighbor relationship is mutual:
 $m_i \in \mathcal{N}_j \iff m_j \in \mathcal{N}_i$.

The variables with a neighborhood system \mathcal{N} satisfies the Markov properties:

$$\begin{cases} P(m) > 0, \forall m \in \mathcal{M} \\ P(m_i | m_j, m_j \in \mathcal{M} \setminus m_i) = P(m_i | m_s, m_s \in \mathcal{N}_i) \end{cases} \quad (2.35)$$

The first constrain indicate the positivity condition and the second constrain implies the local relationship within the neighborhood. It is equivalent to the requirement that the conditional probability $P(m_i | m_j)$ for each i and j depends only on $\{m_s\}$ for $m_s \in \{m_i\} \cup \mathcal{N}_i$. \mathcal{M} is an MRF with joint probability $P(m_1, \dots, m_n)$ determined by its local conditional probabilities. Only the neighboring minutiae have direct interactions with each other.

Before modeling a joint distribution of minutiae $p(m_1, \dots, m_n)$ in a fingerprint, we begin with modeling distribution for single minutiae $p(m)$ with its feature vector $\Lambda(m)$. In [84, 70], generative models for minutiae locations and orientations have been used for fingerprint individuality analysis. An effective choice for minutiae location modeling is to use mixture of Gaussians and use von Mises distribution for minutiae orientation. For relative distance between pair of minutiae, a obvious model is to choose a mixture of Gammas.

A naive model is to assume minutiae independence and multiply marginal distributions together. However, such model does not account for highly correlated spacial features. In the model of Section 2.2.2, minutiae dependence was modeled by a causal Bayesian network where conditional dependencies of minutiae is defined by a unique sequence of points. MRFs can express a wider variety of spatial relationships. It has been widely employed in computer vision takes given the power of modeling contextual dependence patterns or and characterizing mutual influences/constrains among objects as image pixels and correlated features [37, 83, 10, 50].

Given a neighborhood \mathcal{N} , the joint distribution of minutiae obeys a Gibbs distribution [35] parameterized by a set of clique potentials:

$$P_G(m_1, \dots, m_n) = \frac{1}{Z} \tilde{P}_G(m_1, \dots, m_n) \quad (2.36)$$

where

$$\tilde{P}_G(m_1, \dots, m_n) = \prod_{i \in \mathcal{M}} \Phi_i(m_i) \prod_{i \in \mathcal{M}} \prod_{j \in \mathcal{N}_i} \Psi_{i,j}(m_i, m_j) \quad (2.37)$$

is the unnormalized distribution and $Z = \sum_{m_1, \dots, m_n} \tilde{P}_G(m_1, \dots, m_n)$ is the normalization constant called the partition function, Φ_i and $\Psi_{i,j}$ are single-node potentials and pairwise (edge) potentials.

Minutiae Correspondence

A correspondence relationship between two fingerprints is established before measuring similarity between them. The process is to align the two sets of minutiae via a most likely manner. Given a set of minutiae $\mathcal{M}_X = \{m_{X_i}, i = 1, \dots, N_X\}$ in the mark fingerprint, each minutiae is associated with another one from the set $\mathcal{M}_Y = \{m_{Y_s}, s = 1, \dots, N_Y\}$ in a known print via an unique mapping $f(m_{X_i}) : m_{X_i} \rightarrow m_{Y_s}$.

Let set $\mathbf{f} = \{f_1, \dots, f_N\}$ be the mapping for all the minutiae in \mathcal{M} , where f_i denotes $f(m_{X_i})$ for simplicity. Each f_n is a multinomial random variable taking N_Y values from \mathcal{M}_Y which is the set for all possible mappings. The overall configuration space of \mathcal{F} is the Cartesian product $\mathcal{M}_Y \times \mathcal{M}_Y \cdots \times \mathcal{M}_Y = \mathcal{M}_Y^{N_Y}$. The correspondence process of minutiae sets is a mapping as shown in Figure 2.16.

The goal is to perform minutiae-based latent fingerprints comparison based on probabilistic approach. We propose a probabilistic model, which calculates a probability for establishing an optimal correspondences between two fingerprints.

Each fingerprint is represented by a set of minutiae. Each minutia is described by a number attributes, including its position, orientation and type. The process of minutiae correspondence is defined via an set of

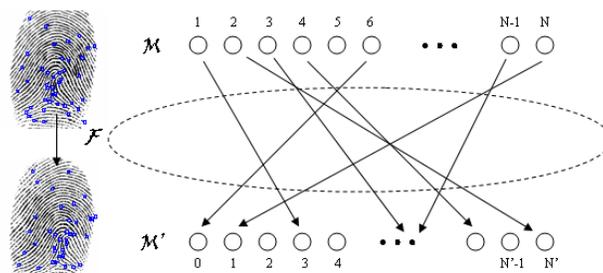


Figure 2.16: Correspondence process of mapping from evidence \mathcal{M} to known $\tilde{\mathcal{M}}$.

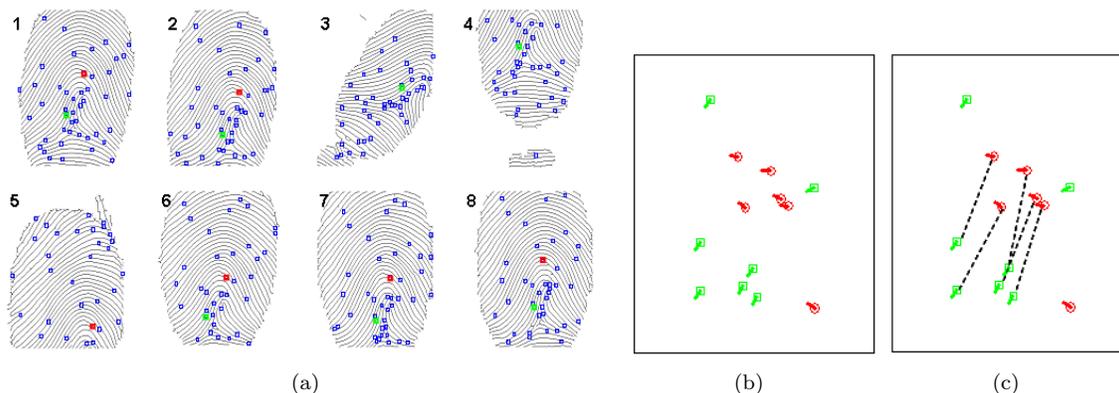


Figure 2.17: Feature disparity between two fingerprints: (a) eight finger-prints with extracted minutiae sets where a particular minutia (index 9) is colored, (b) the same minutia in 8 different impressions are superimposed together, and (c) 56 distances exist between all possible pairs.

unique mapping $f(x_i) : x_i \rightarrow y_j$, that associate each minutiae in a latent set $X = \{x_i, i = 1, \dots, N_X\}$ with another one from the known set $Y = \{y_j, j = 1, \dots, N_Y\}$.

Let set $\mathbf{f} = \{f_1, \dots, f_N\}$ be the mapping set for all the minutiae in X , where f_i denotes $f(x_i)$ for simplicity. Each f_n is a multinomial random variable taking N_Y values from Y which contains the set for all the possible mates. The mapping process over the entire minutiae set defines a set of random variables over a neighborhood system \mathcal{N} defined as follow: a neighborhood of f_i is defined as the set of mapping w.r.t the minutiae within a radius of r from x_i , $\mathcal{N}_i = \{f_s \in \mathbf{f} | \text{dist}(x_i, x_s)^2 < r^2, s \neq i\}$. The $\text{dist}(\cdot, \cdot)$ denotes the Euclidean Distance and r is a empirical constant chosen in order to capture local relationship among local minutiae within a certain degree of complexity of the graph. In our experiment, r is set to 75, which is about 6-10 times of the ridge width. If there is no any minutiae within in the radius, then assign the nearest minutiae of f_i to be its neighborhood in order to guarantee the connectivity among all variables. Let graph $G(\mathbf{f}, \mathcal{E})$ denotes the mapping network for the latent set. The edges \mathcal{E} represent links between sets of neighboring pair $\mathcal{E} = \{(f_i, f_s) | f_s \in \mathcal{N}_i\}$. An illustration of constructing the graph given a set of 39 minutiae is shown in Figure 2.15.

Mapping Field (Prior)

Denote the probability that random variable f_i takes the value y_j to be $P(f_i) = P(f_i = y_j)$ and the joint probability for the whole set $P(\mathbf{f}) = P(f_1, \dots, f_n)$ that defined over the neighborhood system \mathcal{N}_i satisfies the Markov properties. Variables f_i follows a multinomial distribution on N_Y states $P(f_i | \mu_i) = \prod_{j=1}^{N_Y} \mu_j^{\mathbf{I}_j(f_i)}$, where μ_j is the model parameter and $\mathbf{I}_j(f_i)$ is the indicator function for f_i takes the value from the mapping $f_i \rightarrow y_j$.

We encode this individual node preferences as node potentials by defining feature functions with a set of

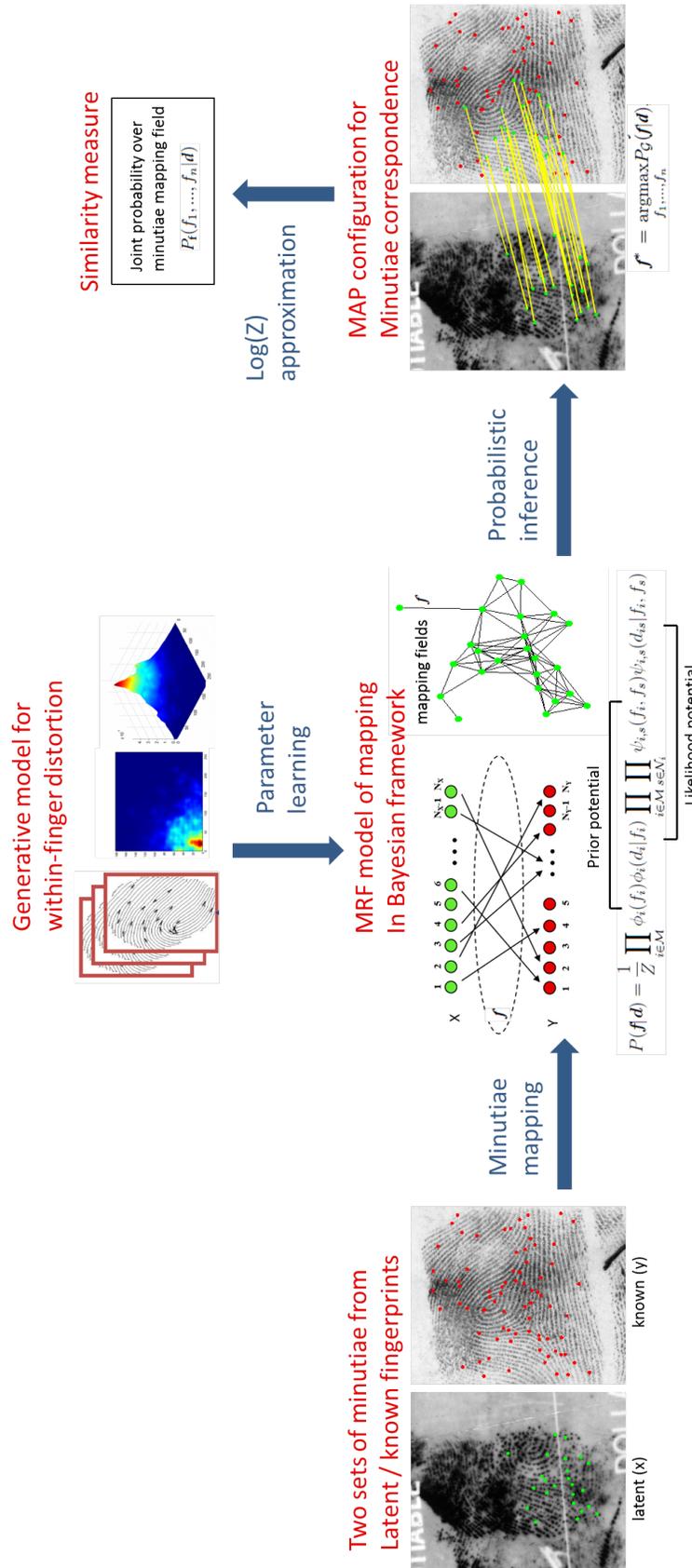


Figure 2.18: Overview of evaluating a probabilistic measure of similarity.

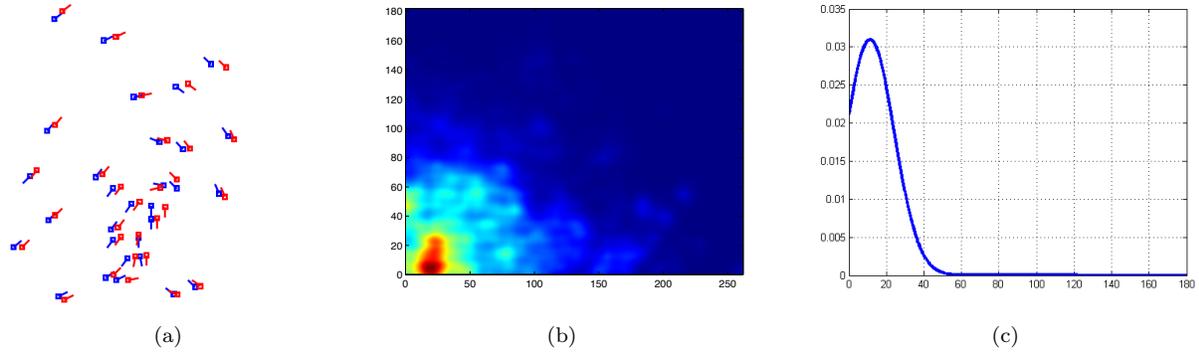


Figure 2.19: Generative model for minutiae distortion: (a) corresponding minutiae sets from two fingerprint impressions of same finger, (b) probability density of within-finger position difference, and (c) probability density of within-finger angle difference.

weights η_i that measures the effort to paid for associating each minutiae in X with the other one in Y . Since a Markov network structure in dose not generally reveal all of the structure in a Gibbs parameterization and parameters are coupled in potentials, for clarity, we define node potential

$$\phi(f_i) = \frac{e^{\sum_j \eta_j \mathbf{I}_j(f_i)}}{\sum_{j'} e^{\eta_{j'}}} \quad (2.38)$$

where $\mu_j = \frac{e^{\eta_j}}{\sum_{j'} e^{\eta_{j'}}$ to ensure the multinomial parameter constraint $\sum_{i=1}^{N_Y} \mu_j = 1$. Similarly, we encode the pairwise preferences as edge potentials.

$$\psi(f_i, f_s) = \frac{e^{\sum_{is} \eta_{is} \mathbf{I}_{jt}(f_i, f_s)}}{\sum_{i's'} e^{\eta_{i's'}}} \quad (2.39)$$

The variables with a neighborhood system \mathcal{N} satisfies the Markov properties: $P(f) > 0, \forall f \in \mathbf{f}$ indicate the positivity condition and $P(f_i|f_k, f_k \in \mathbf{f} \setminus f_i) = P(f_i|f_s, f_s \in \mathcal{N}_i)$ implies the local relationship within the neighborhood. It is equivalent to the requirement that the conditional probability $P(f_i|f_k)$ for each i and s depends only on $\{f_s\}$ for $f_s \in \{f_i\} \cup \mathcal{N}_i$. \mathbf{f} is a multinomial Markov Random Fields with the joint probability $P(f_1, \dots, f_n)$ determined by its local conditional probabilities.

$P(\mathbf{f}; \eta) =$

$$\exp\left\{\sum_{i \in X} \sum_{j \in Y} \eta_j \mathbf{I}_j(f_i) + \sum_{i \in X, s \in \mathcal{N}_i} \sum_{j, t \in Y} \eta_{is} \mathbf{I}_{jt}(f_i, f_s) - C(\eta)\right\} \quad (2.40)$$

where $\mathbf{I}_{jt}(f_i, f_s)$ is indicator function for the pair $f_i \rightarrow y_j$ and $f_s \rightarrow y_t$, $C(\eta)$ is the log partition function. By Hammersley-Clifford theorem[61] the mapping field (2.40) is equivalently a Gibbs distribution parameterized by a set of clique potentials.

$$P(\mathbf{f}) = \frac{1}{Z} \prod_{i \in X} \phi_i(f_i) \prod_{i \in X, s \in \mathcal{N}_i} \psi_{i,s}(f_i, f_s) \quad (2.41)$$

where Z is the partition function. In this paper we restrict our attention to single and pairwise potentials.

Generative models for Minutiae distortion

To decide the values of weights η_i , we consider spatial distortion constraint over the graph, in that truly corresponding minutiae tend to distributed closely. A minutiae shows up in different locations and orienta-

tions in a fingerprint and its distorted ones due to some transformation and distortion. Assuming disparities among these corresponding minutiae forms some noise distribution density, more effort need to be paid when associating a minutiae to the one that is unlikely to occur within the noise tolerance. A generative model of minutiae distortion is built to compute the probability of a possible match. The probability density is based on the differences among distorted minutiae and distorted edges in a background database. Let $\xi^{(1)}(\Lambda_i, \Lambda_{f_i}) = \{\Delta l_x \Delta l_y \Delta \theta\}$ denote the disparity vector between x_i and $f(x_i)$. Let Ξ_{i_s} and Ξ_{f_i, f_s} denotes the vector that contains length and angle for the edge between (x_i, x_s) and $(f(x_i), f(x_s))$ respectively, then disparity vector for edge difference is $\xi^{(2)}(\Xi_{i_s}, \Xi_{f_i, f_s}) = \{\Delta \mathcal{L}, \Delta \Theta\}$. These differences are nonnegative and thus mixtures of gamma distribution is proposed

$$p(\xi^{(1)}(\Lambda_i, \Lambda_{f_i})) = \sum_k \mathcal{G}(\pi_k^{(1)}, \alpha_k^{(1)}, \beta_k^{(1)}) \quad (2.42)$$

$$p(\xi^{(2)}(\Xi_{i_s}, \Xi_{f_i, f_s})) = \sum_k \mathcal{G}(\pi_k^{(2)}, \alpha_k^{(2)}, \beta_k^{(2)}) \quad (2.43)$$

where

$$\mathcal{G}(\xi; \pi_k, \alpha_k, \beta_k) = \pi_k \frac{\xi^{\alpha_k - 1}}{\Gamma(\alpha_k) \beta_k^{\alpha_k}} \exp\left\{-\frac{\xi}{\beta_k}\right\}$$

k is number of mixtures and π_k is the weight for each component. $\Gamma(\alpha)$ is Gamma function which exists for all possible value of α_k . α_k and β_k are estimated using EM method. (2.42) and (2.43) are generative model for the feature difference between a pair of corresponding minutiae.

The distortion model is trained on a fusion of rolled and plain databases containing 200 fingers with 8 impressions each from FVC2002 and 7000 fingers with two images each from NIST SD14. Distances were normalized w.r.t ridge width due to variation on different fingerprint image sizes. The average ridge width of FVC2002 database is roughly 9 pixels and 12 pixels for NIST SD14 and NIST 27. Figure 2.19 shows plots of probability density of difference between paired minutiae.

The parameters η_i, η_{ij} in (2.40) are calculated as the log probabilities decided from evaluating mixtures of gamma distributions at given distance. They are negative with small values that increase the effort need to pay for a mapping when the prior of minutiae assignment $P(f_i)$ is low.

Observation Model (Likelihood)

$P(\mathbf{f})$ can be considered as prior that represents any information carried in the association process of minutiae. If we assume some observation \mathbf{d} that encompasses the information of dissimilarity between the mark fingerprint X and known Y , given a particular set of minutiae association $\{f_i\}$, d_i is considered as the normalized distance between a given minutiae $x_i \in X$ and the candidate minutiae $x_s \in Y$. Let Λ_i denote the feature vector of location, orientation and type of minutiae m_i and $\Xi(x_i, x_s)$ denotes the vector that contains length and angle for the edge between x_i and x_s . Assuming the features of mapped minutiae and observed minutiae are related via a noise model. The noise is additive white Gaussian with unknown variances

$$\phi(d_i | f_i) = \exp\left(-\sum_{\kappa_1} \left(\frac{\Lambda_i^{(\kappa_1)} - \Lambda_{f_i}^{(\kappa_1)}}{\sqrt{2}\sigma_1^{(\kappa_1)}}\right)^2\right) \quad (2.44)$$

where the Gaussian noise covariance is estimated by maximizing the likelihood function $p(\mathbf{d} | \mathbf{f})$. Similarly, observation model for the pairwise relationship is defined as

$$\psi(d_{i_s} | f_i, f_s) = \exp\left(-\sum_{\kappa_2} \left(\frac{\Xi_{i_s}^{(\kappa_2)} - \Xi_{f_i, f_s}^{(\kappa_2)}}{\sqrt{2}\sigma_2^{(\kappa_2)}}\right)^2\right) \quad (2.45)$$

where κ_1 and κ_2 represents the types of feature attribute $\kappa_1 = \{1, 2, 3, 4\}$ represents the x, y location, orientation and type for a single minutiae. $\kappa_2 = \{1, 2\}$ represents the length and angle of an edge between

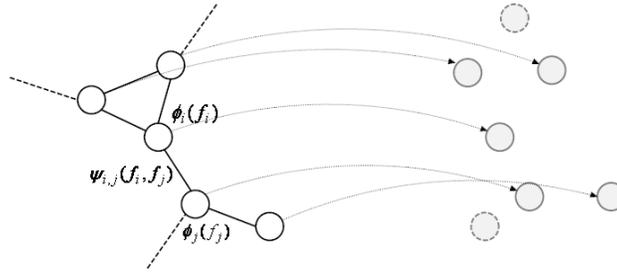


Figure 2.20: Pairwise MRF model with node and edge potentials.

two neighboring minutiae.

MAP Assignments of Correspondence

The problem of finding corresponding two sets of minutiae is to find the most likely consistent global assignments to all the variables of the MRF by maximizing a posterior estimation $\mathbf{f}^* = \operatorname{argmax}_{f_1, \dots, f_n} P_{\mathcal{G}}(\mathbf{f}|\mathbf{d})$, where $P(\mathbf{f}|\mathbf{d}) \propto P(\mathbf{f}, \mathbf{d}) = P(\mathbf{f})P(\mathbf{d}|\mathbf{f})$ by Baye's rule. Combine prior (2.40) and observation potentials (2.44) (2.45) and rearrange the single node and pairwise node terms for

$$P(\mathbf{f}|\mathbf{d}) = \frac{1}{Z} \prod_{i \in X} \Phi_i(f_i) \prod_{i \in X, s \in \mathcal{N}_i} \Psi_{i,s}(f_i, f_s) \quad (2.46)$$

where $\Phi_i(f_i) = \phi_i(f_i)\phi_i(d_i|f_i)$ is the single-node potential and $\Psi_{i,s}(f_i, f_s) = \psi_{i,s}(f_i, f_s)\psi_{i,s}(d_{is}|f_i, f_s)$ corresponds to the pairwise potential (edge potential). The pairwise MRF is illustrated in Figure 2.20. Our problem is to find the assignments for the unobserved variables that is most probable given the observation. To find the MAP configuration of each variables, we use the max-product algorithm of belief propagation on the MRF.

To find the MAP configuration of each variables, max-product algorithm of loopy belief propagation is used. The algorithm iteratively computes local beliefs of each mapping $b_i(f_i)$ that could possibly assign one of minutiae in Y and update probabilities via message passing.

Belief Propagation

Given the MRF $G(V, E)$, each node sends out a message to each of its neighbors and receives a message from each neighbor. Let x_i and x_j be two neighboring nodes in \mathcal{G} . We denote $m_{i,j}(f_j)$ the message that node f_i sends to node f_j . At every iteration, the max-product update rules of messages is

$$m_{i,j}(f_j) \leftarrow \alpha \operatorname{arg max}_{f_i} \Phi_i(f_i) \Psi_{i,j}(f_i, f_j) \prod_{k \in \mathcal{N}_i \setminus j} m_{k,i}(f_i) \quad (2.47)$$

where α is the normalization constant which does not really influence the final beliefs but affects the numerical stability of the algorithm. The message $m_{i,j}(f_j)$ is a vector of dimensionality of number of configurations/states of the variables, or the length of the minutiae set N_Y , which each component being proportional to how likely this node f_i believes that node f_j is being in corresponding to the state. On the right-hand-side, we take product over all incoming messages of node f_i except for the one coming from node f_j . Then wrap up with all the local information together and send it out to node f_j . The belief at a node f_i is proportional to the product of all the messages coming to it and the local evidence at this node ($\Phi_i(f_i)$).

$$b_i(f_i) = \alpha \Phi_i(f_i) \prod_{j \in \mathcal{N}_i} m_{j,i}(f_i) \quad (2.48)$$

The node belief is an approximation of the unnormalized marginal probability at this node $\tilde{p}_i(f_i) = \sum_{\forall j \in \mathcal{M} \setminus i} P(\mathbf{f})$.

It is also a vector of dimensionality of number of states and must sum to 1 ($\sum_{f_i} b_i(f_i) = 1$) at each node.

We can also define the edge belief to be the pairwise nodes belief similarly

$$b_{ij}(f_i, f_j) = \alpha \Psi_{ij}(f_i, f_j) \Phi_i(f_i) \Phi_j(f_j) \prod_{k \in \mathcal{N}_i \setminus j} m_{k,i}(f_i) \prod_{l \in \mathcal{N}_j \setminus i} m_{l,j}(f_j) \quad (2.49)$$

This edge belief is an approximation of the unnormalized marginal probability at the tow node clique by marginalizing the joint probability over every other node except i and j $\tilde{p}_{i,j}(f_i, f_j) = \sum_{\forall k \in \mathcal{M} \setminus \{i,j\}} P(\mathbf{f})$. It is

a square matrix of dimensionality of $N_Y \times N_Y$ and must sum to 1 ($\sum_{f_i, f_j} b_{ij}(f_i, f_j) = 1$) for every edge. On every iteration, each edge carries a message containing information of how it thinks about the values of two nodes on both ends.

The algorithm is initialized with all message vectors set to $1/N'$. The message update rule is then applied iteratively in a synchronous schedule in which all node simultaneously sends their messages in parallel. The algorithm converges to a fixed point regardless of initial assignments.

Figures 2.21 and 2.22 show the update of two node beliefs and the algorithm converges to a fixed point after 12 iterations. At convergence, the belief for any node is the maximum posterior conditioned on the observation $b_i(f_i) = \alpha \max_{f_i} P(f_i | d_i)$ and the belief for any edge (f_i, f_s) is the maximum posterior $b_{i,j}(f_i, f_s) = \alpha \max_{f_i, f_s} P(f_i, f_j | d_{ij})$ that must agree with the node belief of f_i and f_s .

The minutiae correspondence is achieved as $\mathbf{f}^* = \arg \max_{f_i} b_i(f_i)$ when the algorithm converges. It is possible that multiple minutiae are matched to the same one. When such a conflict happens, only $x_m = \arg \max_{i_m} b_{i_m}(f_{i_m})$ is considered to be matched successfully. Figure 2.23 shows three examples of minutiae matching results.

Similarity Measure

The similarity score calculated by existing matching algorithms is a proximity measure of capability of the matcher to find true correspondences, e.g. the fraction of number of matched minutiae over average number of minutiae in the two fingerprints. These formula-based score is deterministic and information about uncertainties in matching process could be lost due to frequent nonlinear deformation and in fingerprint images, where exact match is not possible. We define a probabilistic measure as the joint probability over the minutiae mapping set given the observation $P_{\mathbf{f}}(f_1, \dots, f_n | \mathbf{d})$ (2.46). When optimal correspondence is achieved, this probability reflects the likelihood that they are matched. Computing this joint probability requires approximation of the partition function Z . We use the Bethe free energy approximation [35].

The method for calculating overall similarity between two fingerprints is based on an approximation of the joint probability of $P(f_1, f_2, \dots, f_n)$. We could assume that this joint probability can be viewed as the probability of evidence, that is defined as a relationship between the mark and known as the similarity measure in Eq. (2.2.5). This relationship can be modeled by the minutiae corresponding process $\mathbf{f}: X \rightarrow Y$, that leads for compute the joint probability given the observation $P_{\mathbf{f}}(f_1, \dots, f_n | \mathbf{d})$ of all the variables in the MRF as in Eq. (2.46). Computing this joint probability requires approximation of the partition function Z . We use the Bethe free energy approximation.

The Kullback-Leibler distance (or mutual information) between the joint probability $p(\mathbf{f})$ and the overall belief $b(\mathbf{f})$

$$D(b(\mathbf{f}) || p(\mathbf{f})) = \sum_{\mathbf{f}} b(\mathbf{f}) \ln \frac{b(\mathbf{f})}{p(\mathbf{f})} \quad (2.50)$$

The Kullback-Leibler distance is non-negative and it is zero if and only if the two probability functions are equal. Since $p(\mathbf{f})$ is a Gibbs distribution $p(\mathbf{f}) = \frac{1}{Z} e^{-E\mathbf{f}}$, thus

$$D(b(\mathbf{f}) || p(\mathbf{f})) = \sum_{\mathbf{f}} b(\mathbf{f}) E(\mathbf{f}) + \sum_{\mathbf{f}} b(\mathbf{f}) \ln b(\mathbf{f}) + \ln Z \quad (2.51)$$

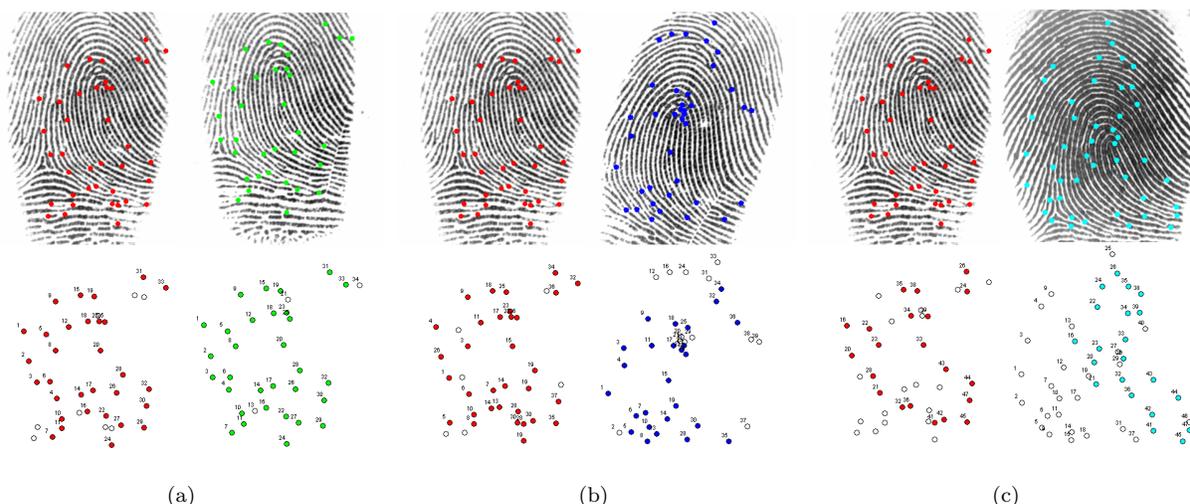


Figure 2.23: Correspondence between pairs of fingerprints– evidence on the left (minutiae in red) and known is on the right when the pairs are from: (a) same finger with low distortion (b) same finger with high distortion, and (c) different fingers. The joint probabilities and likelihood ratios for each pair using three methods are given in Table 2.6.

The Bethe free energy of the joint distribution given a belief

$$E_{Bethe} = - \sum_{ij} \sum_{f_i, f_j} b(f_i, f_j) \ln \psi(f_i, f_j) \phi(f_i) \phi(f_j) + \sum_{ij} \sum_{f_i, f_j} b(f_i, f_j) \ln b(f_i, f_j). \quad (2.52)$$

The beliefs represent the pseudo-marginals of the nodes taking a particular value. Then our approximation of log partition function is $-\min_b E_{Bethe}(b)$.

Correspondence Experiments

We conducted experiments on latent-to-ten print search on livescan dataset FVC 2002, FVC 2004, rolled dataset NIST SD14 and latent data set NIST SD27. To evaluate the approach, we first compared each latent fingerprint in SD27 against its rolled mate. For every latent, the corresponding minutiae set found by our approach was compared with the ground-truth (provided with SD27 validated by professional latent examiners). The result of this small-scale test is shown in Table 2.2.5. The successful correspondence rate (SCR) is calculated as the total number of correctly mated minutiae (NCM) divided by the number of available minutiae (NM) in the latent for all 257 tests.

	Good	Bad	Ugly	All
No. of Available minutiae (NM)	2776	1502	1025	5303
No. correctly mated (CM)	2532	1322	711	4565
Successful Correspondence Rate (SCR)	91.2%	88.0%	69.4%	86.1%

Table 2.5: Rate of successful correspondence (SD27)

To make the latent-to-ten print search more realistic, as in [33, 49], we expanded the background database by adding fingerprints from the NIST Special Database 14. The SD14 database contains 27,000 pairs of rolled fingerprint images provided by FBI. They are all scanned in a similar way as SD27 at 500 ppi and in 8-bit

gray scale with size of 832×768 . The latent fingerprint is matched against the combined gallery of 20,257 rolled impressions with 20000 in SD14 and 257 in SD27. 7000 pairs of fingerprints from SD14 are left out for training the distortion model together with FVC2002 DB1 and DB2. We used the NIST MINDTCT to extract minutiae automatically from NIST SD 14 rolled fingerprints and FVC 2002 plain fingerprints.

The results were evaluated for three fingerprint qualities since latent fingerprint image quality critically affect the match performance. Also the number of minutiae is another indicator of fingerprint quality [33, 49]. In [49] latents in SD27 is classified into three groups: large ($n \geq 22$), medium ($13 < n < 21$), and small ($n \leq 13$), containing 79, 96, an 83 prints, respectively. Experimental results are presented for all these six quality groups. The baseline algorithm is a non-probabilistic matching algorithm used in [33] that also defines a minutiae descriptor based on neighborhood and adopts a greedy strategy for global matching.

A cumulative match characteristic (CMC) is a method of showing measured accuracy performance of a biometric system operating in the closed-set identification task. Templates are compared and ranked based on their similarity. The CMC shows how often the individual's template appears in the ranks (1, 5, 10, 100, etc.), based on the match rate. A CMC compares the rank (1, 5, 10, 100, etc.) versus identification rate. CMCs for the SD27 latent prints shown in in Figure 2.24 for the proposed method and the baseline method indicate that we are able to perform strong latent print comparison with large intraclass variations and distortion.

Likelihood Ratio

To assess the value of the evidence, two generic questions need to be answered: (1) what is the probability of getting this degree of similarity if the two come from the same source (prosecution hypothesis h^0)? (2) what is the probability of observing this degree of similarity if the two come from different sources (defense hypothesis h^1)? The answers to the two questions refer to the numerator and the denominator of the likelihood ratio (LR). Computing the numerator requires evaluating all the possibilities of comparison between the fingerprint and its distorted ones at the observed evidence (similarity). For the denominator of the likelihood ratio, the evidence is positioned among the possibilities offered by fingerprints from different sources.

We conducted experiments with different methods for calculating LRs for the comparison of two fingerprints. The LR is computed in three different ways. One is according to (2.2.5) in which the probabilistic similarity measure is used as approximation:

$$LR_S = \frac{p(P(f_1, \dots, f_n)|h^0)}{p(P(f_1, \dots, f_n)|h^1)} \quad (2.53)$$

The other is to consider the joint probability of the evidence and known approximately as the joint probability over the distance of their common features:

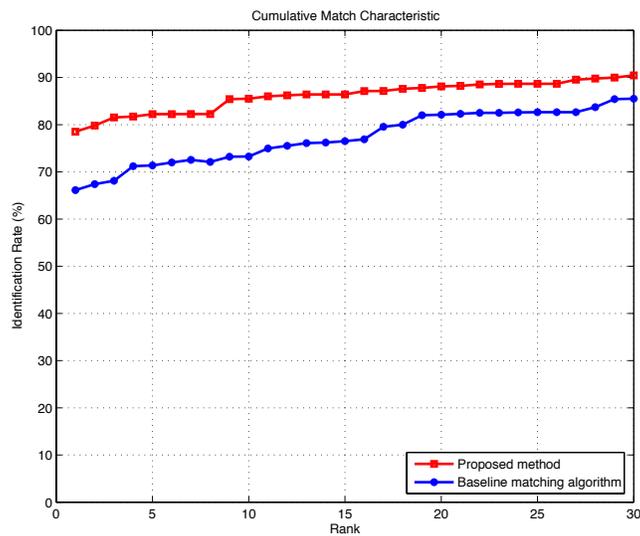
$$LR_J = \frac{p(d_1, d_2, \dots, d_k|h^0)}{p(d_1, d_2, \dots, d_k|h^1)} \quad (2.54)$$

where k is number of corresponding pairs of minutiae between two fingerprints and d_k is the distance between the pairs. These common features are modeled using a MRF. Potentials of the MRF are learned from our generative model for distance. Another method for calculating LR is by adopting rarity into the equation:

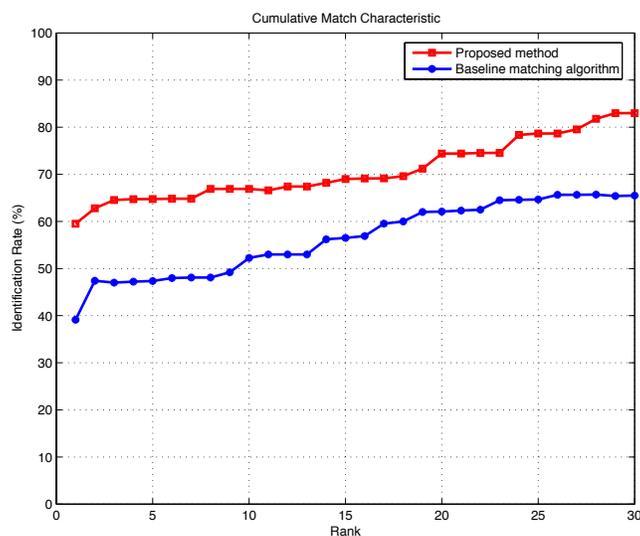
$$LR_{JR} = \frac{p(d_1, d_2, \dots, d_k|h^0)}{p(m_1, m_2, \dots, m_k)} \quad (2.55)$$

where m_1, m_2, \dots, m_k are the average between the corresponding minutiae. Rarity is defined to be the inverse of the joint distributions of the matched features in the denominator of LR_{JR} that measures how typical/unusual the set of features is. Rarity is computed using a generative model for all minutiae. LRs for the three pairs in Figure 2.23 are given in Table 2.6. An LR of 8.9×10^{-5} implies that the evidence seen is almost 13000 times more likely given the defense hypothesis than the prosecution hypothesis.

The methods were tested on the fingerprint databases of FVC2002 which contain four different databases (DB1, DB2, DB3 and DB4) [40]. Each contains 110 different fingers and 8 impressions of each finger yielding



(a)



(b)

Figure 2.24: Cumulative Match Characteristics (CMC) curves: (a) good, and (b) bad

Table 2.6: Joint probability and likelihood ratio for the three pairs of prints in Fig. 2.23

	Same with low distortion (Fig 2.23(a))	Same with high distortion (Fig 2.23(b))	Different (Fig 2.23(c))
$P(f)$	1.7×10^{-28}	$7.3e \times 10^{-65}$	1.9×10^{-184}
LR_S	9.5904	1.0103	8.9×10^{-5}
LR_J	31.2352	15.2362	2.3×10^{-7}
LR_{JR}	45.4822	14.2304	7.1×10^{-8}

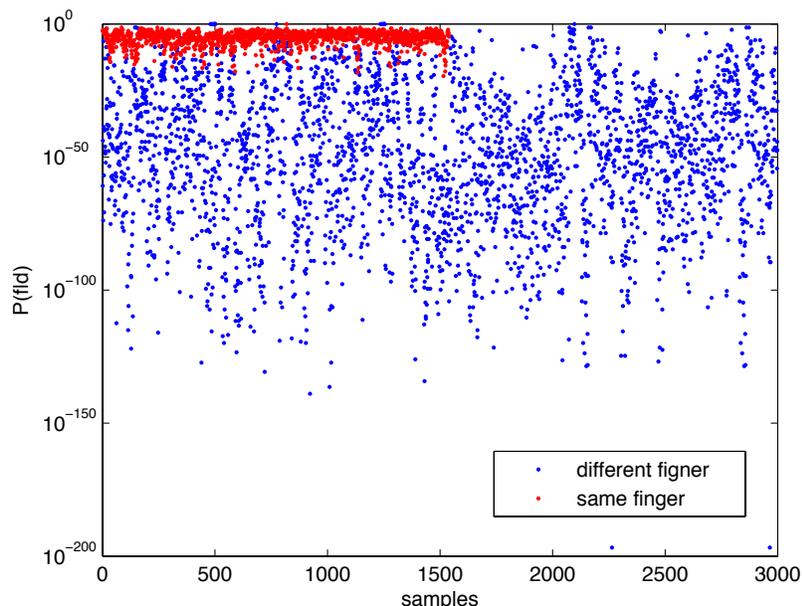


Figure 2.25: Distribution of joint probabilities $P(f_1, \dots, f_n)$ computed given test set of FVC2002 DB1

a total of 880 fingerprints. There are 3,080 $(110 \times (8 \times 7)/2)$ same finger pairs and we chose one print from each finger to make 5,995 $(110 * 109/2)$ different pairs. Half of the set was used as the training set and the remaining for testing. The probabilistic method is compared to Champod’s radial triangulation algorithm and two other fingerprint matching algorithms, Bozorth and coupled breadth first search (CBFS) (see Figure 2.14). The Bozorth algorithm constructs two intra-fingerprint minutiae pair table to capture relative position and orientation of a pair of minutiae and a inter-fingerprint tables and then search these tables for the longest path of linked compatible associations. CBFS converts a fingerprint represented by local k-minutiae models to an adjacency graph and then searches for minutiae correspondence between two fingerprints. For the Bozorth method, each intra-table construction involves $O(n^2)$ comparisons and the inter-table match has more than $O(n)$ comparisons. CBFS takes $O(n)$ to build the local models and $O(n^2)$ for coupled breadth-first search. The time complexity of above methods is no less than $O(n^3)$ when both fingerprints to be compared have exactly n minutiae. The time complexity of above methods is no less than $O(n^3)$ when both fingerprints to be compared have exactly n minutiae. The time complexity of the loopy belief propagation algorithm is $O(nd^2I)$, where n is the number of variables, d is the maximum size of neighborhoods of a variable and I is the number of iterations.

Discriminative Power Evaluation

In order to evaluate the discriminative power for the method we carried out 1:1 fingerprint verification. The decision is based on whether the LR is larger than 1. For each pair of comparison, if the likelihood ratio is greater than 1, then the the two samples belong to the same person and if the ratio is less than 1, they belong to different persons. For LR_S , the similarity measures are computed from Bozorth matcher (LR-S-Bozorth), CBFS (LR-S-CBFS) and the probabilistic measure (LR-S-MRF). The two distributions of the numerator and the denominator of the likelihood ratio are estimated parametrically using EM algorithm. Figure 2.25 shows distribution of joint probabilities $P(f_1, \dots, f_n)$ for minutiae correspondence computed for the test set of FVC2002 DB1. For LR_J , we develop MRF for distances and compute overall joint probabilities based on our distortion model (LR-J-MRF). For LR_{JR} , The joint distribution in the denominator is represented by another MRF and the potentials of which are estimated using a rarity model for any minutiae (LR-JR-MRF).

The discriminatory power for these methods are compared. Table 2.7 shows the percentage distributions

Table 2.7: Percentage distributions of likelihood ratios (calculated for 1500 within-finger (same) comparisons and 3,000 between-finger (different) comparisons) for FVC2002 DB1. False positive (between-finger comparison giving a value of LR greater than 1) and false negative (within-finger comparison giving a value of LR smaller than 1) rates and overall error rates are given.

Likelihood Ratio	LR-S-Bozorth		LR-S-CBFS		LR-S-MRF		LR-J-MRF		LR-JR-MRF	
	different	same	different	same	different	same	different	same	different	same
$\leq 10^{-6}$	78.24	0	69.04	0	77.47	0	82.1	0	86.89	0
$10^{-6} - 10^{-5}$	5.32	0	16.24	0	8.63	0	6.11	0	5.21	0
$10^{-5} - 10^{-4}$	6.11	0.02	2.38	0	3.42	0	4.25	0	2.65	0
$10^{-4} - 10^{-3}$	2.35	0	2.97	0	2.21	0.01	2.33	0.04	1.25	0
$10^{-3} - 10^{-2}$	1.94	0.25	3.21	0.49	1.52	0.31	1.39	0.1	0.69	0.02
$10^{-2} - 10^{-1}$	2.28	0.38	2.45	0.15	2.38	0.33	1.05	0	0.33	0.02
$10^{-1} - 10^0$	1.35	0.25	1.49	0.02	1.8	0	0.51	0.03	0.65	0.04
$10^0 - 10^1$	0.67	0.02	1.24	0.14	1.64	0.3	0.49	0.01	0.92	0.39
$10^1 - 10^2$	1.28	1.24	0.89	2.21	0.23	0.12	0.27	0.05	0.36	0.55
$10^2 - 10^3$	0.41	10.2	0.04	0.78	0.34	2.35	0.81	4.21	0.31	0.52
$10^3 - 10^4$	0.03	15.21	0.04	4.8	0.13	0.56	0.17	6.37	0.24	1.46
$10^4 - 10^5$	0	46.85	0.01	56.24	0	62.19	0.21	52.52	0.27	23.22
$10^5 - 10^6$	0.01	25.57	0	35.21	0.02	33.55	0.3	32.11	0.23	66.55
$\leq 10^6$	0.01	0.01	0	0	0.01	0.18	0.01	4.56	0	7.23
False +/- Rate	2.41	0.9	2.22	0.8	2.37	0.75	2.26	0.17	2.33	0.08
Error Rate	3.31		3.02		3.12		2.43		2.41	

Table 2.8: Error rates for fingerprint identification on 4 datasets.

	LR-S-Bozorth	LR-S-CBFS	LR-S-MRF	LR-J-MRF	LR-JR-MRF
DB1	3.31%	3.02%	3.12%	2.43%	2.41%
DB2	3.56%	3.45%	3.32%	2.86%	2.80%
DB3	8.57%	9.43%	7.88%	6.75%	6.96%
DB4	4.74%	4.42%	4.21%	3.95%	3.58%
Overall	5.04%	5.07%	4.63%	3.99%	3.94%

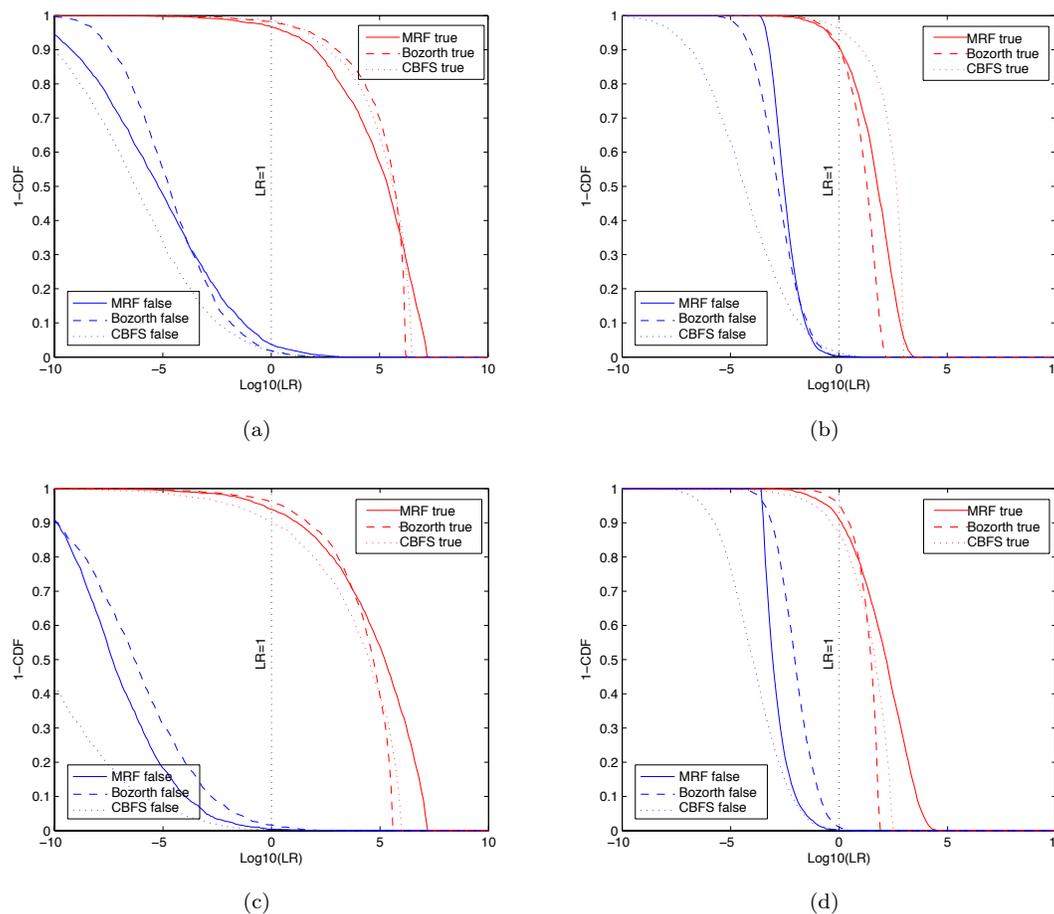


Figure 2.26: Tippet plots for LR of four different data sets.

of LRs calculated for 1500 within-finger (same) comparisons and 3000 between-finger (different) comparisons for FVC2002 DB1. Because of information lost in computing LR_S , identification accuracy for likelihood ratios computed using joint information of features are much better than similarity based method because it takes into account the dependencies among features. Especially, false negatives have been dramatically decreased. LR_{JR} performs slightly better than LR_J . It is also observed that for within-finger comparison, the values of likelihood ratio computed from LR_{JR} are much larger than those computed from LR_J and for between-finger comparisons, higher portion of LRs are smaller than 10^{-6} for LR_{JR} than LR_J . This is because the joint probability of a set of minutiae is usually smaller than the joint probability of a set of minutiae distance in the between-finger data. The variance of the former distribution is larger. The error rates for all four FVC2002 database are presented in Table 2.8. The results indicate our probabilistic measure based on MRF model outperforms deterministic measures such as Bozorth for all databases. CBFS works slightly better for DB1 and DB2. But for DB3, which is the most difficult and noisy among the four database in terms of image quality, our model have nearly 1.5% improvement than CBFS and 0.7% improvement than Bozorth. Both joint probability methods outperform the similarity-based methods. LR_J computed using rarity information is slightly better except for the noisy data DB3.

The result for evidence uncertainty is assessed using Tippet plots which are based on the distributions of LRs [76, 22]. The horizontal axis is graduated with increasing values of LRs while the vertical axis indicates the estimated probability that the result of the experiment exceeds a given value of LR. The Tippet plot includes two curves: the first one shows the evolution of the estimated LR when the hypothesis h^0 is verified and the second one shows the evolution of the estimated LR when the hypothesis h^1 is verified. The Tippet

Table 2.9: LR ranges and rate of misleading evidence in favor of prosecution (RMEP) and in favor of the defense (RMED) for four different data sets in FVC2002. The Tippett plot is shown in Fig.2.26

		Between Finger			Within Finger		
		LR_B min	LR_B max	RMEP ($LR_B > 1$)	LR_W min	LR_W max	RMED ($LR_W < 1$)
DB1	MRF	6.63×10^{-12}	1.72×10^3	1.65%	1.03×10^{-6}	1.50×10^7	2.82%
	Bozorth	2.18×10^{-18}	3.21×10^3	3.80%	$1.36e \times 10^{-9}$	1.58×10^6	1.81%
	CBFS	1.19×10^{-17}	1.06×10^4	1.50%	1.32×10^{-8}	4.21×10^6	2.20%
DB2	MRF	2.24×10^{-4}	0.87×10^1	0.35%	8.05×10^{-4}	3.34×10^3	4.90%
	Bozorth	2.99×10^{-7}	0.11×10^1	0.25%	5.81×10^{-5}	1.29×10^2	4.97%
	CBFS	2.20×10^{-11}	0.41×10^2	1.20%	2.58×10^{-6}	9.21×10^2	3.85%
DB3	MRF	3.17×10^{-14}	1.02×10^3	0.40%	4.88×10^{-9}	1.33×10^7	6.10%
	Bozorth	1.12×10^{-8}	0.31×10^2	1.25%	1.98×10^{-16}	0.31×10^2	6.55%
	CBFS	2.01×10^{-22}	0.67×10^2	0.20%	1.20×10^{-14}	8.93×10^5	9.15%
DB4	MRF	2.90×10^{-4}	0.75×10^1	0.15%	3.27×10^{-4}	3.44×10^4	5.20%
	Bozorth	1.94×10^{-5}	0.59×10^1	0.60%	0.61×10^{-2}	0.64×10^2	4.90%
	CBFS	3.87×10^{-9}	0.94×10^1	0.30%	3.90×10^{-6}	0.21×10^3	7.60%

plots are constructed by resampling 2000 LRs computed for comparisons of fingerprints from the same source and 2,000 LRs for comparisons of fingerprints from different source. To compute LRs, same and different pairs of fingerprints are randomly selected from the test set for two hypothesis respectively. The first set of experiments were based on full fingerprint matching with all the minutiae points available for each. The second set of experiments were based on testing similarity measure at different number of minutiae for both prints.

Guidelines for human experts suggests that a minimum of 12 matched minutiae are required to make a decision [40]. However, a minutiae-based AFIS cannot make a decision using an absolute value alone as a human expert. Unlike human experts having the access to all the information that a fingerprint image has, such as ridge flows, singular points and scars, etc., the minutiae-based automatic systems only have the information from the minutiae representation of fingerprints.

The Tippett plot in Figure 2.26 shows of the distribution of computed LRs obtained under both hypotheses. The x-axis represents the $\log_{10}(LR)$. The inverse cumulative distribution of the LRs ($1 - cdf(LR)$) is given on the y-axis. The Tippett plot then gives inverse cumulative distribution for both hypothesis. There are two rates of misleading results defined as follows: RMED: rate of misleading evidence in favor of the defense. The percentage of all $LR \geq 1$ computed knowing h^0 is true. RMEP: rate of misleading evidence in favor of the prosecution. The percentage of all $LR \geq 1$ computed knowing h^1 is true. The results are summarized in Table 2.9. From the Tippett plots, we see a small portion of the both RMEP and RMED, that indicate a supportive strength of the evidence (accuracy of the system).

To compare the proposed probabilistic measure with Champod’s distance-based metric, different number of minutiae set were tested using fingerprint images from 400 individuals in the FVC2002 data set. Let n be the number of minutiae used. For within finger variability, each image X is compared against the other 7 images Y coming from the same finger. 1 minutia is randomly chosen from X and $n - 1$ nearest minutiae are selected. Then n corresponding minutiae from Y are used for comparison. There are total of 2,700 comparisons. For between finger variability, minutiae set from each image X is compared with the ones in another fingerprint Y coming from 99 different fingers. n minutiae are randomly chosen from X and Y respectively. n is tested in the following cases: 6,12,18,24. The results in 2.9 demonstrate great improvement.

The ability of MRF for model spatial correlated feature allows us to establish minutiae corresponding given to sets of fingerprint minutiae. We consider a mapping field that assigns each minutiae in one set to one from the other. A probability is associated with each mapping field by a pairwise MRF that encodes the node and edge information of corresponding minutiae. A Gaussian noise observation model that measures the dissimilarity/disparity between the two sets is incorporated into the MRF. Prior knowledge about the

Table 2.10: LR ranges and rate of misleading evidence in favor of prosecution (RMEP) and in favor of the defense (RMED) for different number minutiae.

		Between Finger			Within Finger		
		LR_B min	LR_B max	RMEP ($LR_B > 1$)	LR_W min	LR_W max	RMED ($LR_W < 1$)
n=6	MRF	3.45×10^{-14}	5.01×10^7	1.24%	7.22×10^{-7}	9.14×10^7	1.21%
	Champod	8.56×10^{-17}	4.14×10^8	1.95%	6.21×10^{-10}	6.48×10^9	2.24%
n=12	MRF	2.24×10^{-4}	0.87×10^1	0.85%	8.05×10^{-4}	3.34×10^3	0.82%
	Champod	2.99×10^{-7}	0.11×10^1	1.11%	5.81×10^{-5}	1.29×10^2	1.24%
n=18	MRF	1.62×10^{-12}	0.72×10^5	0.52%	9.18×10^{-8}	5.88×10^6	0.81%
	Champod	4.25×10^{-11}	2.12×10^5	1.86%	0.28×10^{-6}	0.31×10^7	0.72%
n=24	MRF	7.14×10^{-9}	6.23×10^2	0.45%	1.49×10^{-5}	5.32×10^8	0.74%
	Champod	2.44×10^{-13}	2.81×10^7	0.92%	4.15×10^{-9}	2.66×10^9	0.45%

mapping field, incorporated using a Bayesian framework, is used to determine the most probable assignment (MAP estimate) of the MRF. The posterior probability is inferred using message passing belief propagation. Most probable assignments of minutiae are determined after iteratively updating local node beliefs, which is an approximation of the marginal probability of every variable. Finally, the similarity of two fingerprints is measured as the overall belief of all variables in the MRF. Distributions of the joint probabilities of MRFs, given the alternative hypotheses, are used to calculate the LRs. Our experimental results indicate that we are able to perform minutiae matching using MRF framework and our model provides reasonable probabilistic measure for calculating a similarity between two fingerprints.

One important question about the method is the convergence of the loopy belief propagation. The algorithm works well for genuine pairs but fails to converge when corresponding unrelated fingerprints. The average convergence rate in our experiment is 92%.

2.3 Discussion of findings

The principal finding in this research are:

1. Probability distributions of friction ridge features can be modeled using probabilistic graphical models (PGMs). Such models are useful and necessary to handle the complexity of the data. Both directed (Bayesian networks) and undirected (Markov random fields) PGMs can be used.
2. A new solution to the fingerprint registration problem has been developed. The origin of the coordinate system for an input print, which is either its core point or a high curvature point, can be determined probabilistically using regression. Gaussian process regression, which provides a distribution as output, was found to perform better (using the MAP estimate) than a point estimate provided by a geometry-based method.
3. Rarity of configurations can be inferred from probability distributions. The probability of random correspondence of an input print in a database of a given size can be determined. Minutia configurations increase in rarity with number of minutia considered. When a large number of minutia (about a dozen or more) are included the rarity is quite high pointing to their uniqueness.
4. Likelihood ratios (LRs) can be readily converted into a probability of identification. These probabilities can be exceedingly small or large thereby providing support to a latent print examiner's opinion of individualization or exclusion. Since the true likelihood ratio based on all variables is difficult to compute, similarity (or distance) based methods have been suggested before. We have showed that combining rarity with probabilistic similarity is better than a distance only method.

5. Similarity of fingerprints can be expressed probabilistically. It was demonstrated using a Markov random field model of the features in two fingerprints.

2.4 Implications for policy and practice

Some implications of this research are: (i) development of practical tools useful to the latent print examiner, (ii) provide a response to criticisms in the legal community regarding uniqueness and individualization, and (iii) help quantify cognitive processes in ACE-V, particularly those relating to rarity of features and similarity of features. These are further expanded as follows:

1. The method for modeling probability distributions of features can be useful in developing software tools for training latent print examiners. One of these is to quantify the degree of uniqueness of features selected by the examiner to be used in the comparison. Rarity can also be regarded as supplementary information for an opinion. Rarity information can possibly help avoid errors, e.g., when the matching features are common ones then the degree of match should be discounted.
2. When the confidences, i.e., probability of identification, are extremely small/high then it justifies the argument of the latent print examiner that there is individualization/exclusion.
3. Characterization of rarity of features and similarity of prints is a step towards formalization of the cognitive process of the human examiner. Such a formalization can help improve human procedures and develop future automated procedures.

2.5 Implications for further research

1. The methods proposed are a first step towards complete modeling of probability distributions of friction ridge patterns. We only considered modeling level-two detail. They can be combined with other levels of detail and other features.
2. The distributions were determined using 4,000 NIST images. Larger and more representative distributions would make the models and predictions more accurate.
3. We have modeled the distribution of minutiae using probabilistic graphical models so that we can exploit independencies and make the task manageable. But inference with probabilistic graphical models itself become intractable. Thus approximate methods, e.g., sampling, may have to be used to determine the probability of evidence.
4. This work focused on latents to ten print comparison where latent fingerprint are compared with rolled fingerprints based only on minutiae (ridge ending and ridge bifurcation). Minutiae representation of fingerprint or minutiae-based approach is the most commonly used method, primarily because: (i) forensic examiners have successfully relied on minutiae to match fingerprints for more than a century, (ii) minutiae-based representation is storage efficient, and (iii) expert testimony about suspect identity based on mated minutiae is admissible in courts of law. Extension of comparison to other levels of detail is needed.

2.6 Dissemination

2.6.1 Publications

The following papers were published:

1. C. Su and S. N. Srihari, "Latent Fingerprint Core Point Prediction Based on Gaussian Processes," *Proceedings 20th International Conference on Pattern Recognition*, Istanbul, Turkey, Aug 23-26, 2010, pp. 1634-1637.
2. C. Su and S. N. Srihari. "Latent Fingerprint Rarity Analysis in Madrid Bombing Case," *Proceedings of International Workshop on Computational Forensics*, Tokyo, Japan, Nov 11-12, 2010.
3. C. Su and S. N. Srihari, "Evaluation of Rarity of Fingerprints in Forensics," *Proceedings of Neural Information Processing Systems (NIPS)*, Vancouver, Canada, December 6-9, 2010. NIPS is a top-rated computer science conference for probabilistic methods relating to images.
4. S. N. Srihari, "Computational Forensics," *IEEE Spectrum*, December 2010, has a brief summary on the topic of this research. See <http://spectrum.ieee.org/computing/software/beyond-csi-the-rise-of-computational-forensics>.
5. C. Su and S. N. Srihari, "Generative Models and Probability Evaluation for Forensic Evidence," in *Pattern Recognition, Machine Intelligence and Biometrics*, P. Wang (ed.), Springer, 2011.

2.6.2 Presentations

The following presentations were made:

1. An oral presentation at the *National Symposium on Indigent Defense* in Washington DC on February 18, 2010
2. A poster presentation at the *NIJ Conference* in Arlington VA on June 14, 2010.
3. A poster presentation at the *Neural Information Processing Systems Conference* in Vancouver, BC, Canada in December 2010.
4. Oral presentation at the *NIJ Impression and Pattern Evidence Symposium* in Tampa, FL on August 2, 2010. This presentation on the current state of the research was made to a well-attended segment of the impression evidence community:
5. Oral presentation at the *International Conference on Pattern Recognition (ICPR)* in Istanbul on August 24, 2010.

2.6.3 Students

1. Chang Su completed his doctoral dissertation titled *Machine Learning in Fingerprint Probability Evaluation* in July 2011.
2. Yu Liu will be defending her doctoral dissertation on *Fingerprint modeling using Markov Random Fields* in Spring 2013.

Chapter 3

References

The references in the Program Narrative are given below.

Bibliography

- [1] AITKEN, C., AND TARONI, F. *Statistics and the Evaluation of Evidence for Forensic Scientists*. Wiley, 2004.
- [2] AITKEN, C. G. G., AND LUCY, D. Evaluation of trace evidence in the form of multivariate data. *Journal of the Royal Statistical Society* 53, 1 (2004), 109–122.
- [3] ASHBAUGH, D. *Quantitative-Qualitative Friction Ridge Analysis: An Introduction to Basic and Advanced Ridgeology*. CRC Press, 1999.
- [4] ASHBY, F. G., AND ENNIS, D. M. Similarity measures. *Scholarpedia* 2, 12 (2007), 4116.
- [5] ASHBY, F. G., AND PERRIN, N. A. Toward a unified theory of similarity and recognition. *Psychological Review* 95, 1 (1988), 124–150.
- [6] ASHBY, F. G. E. Multidimensional models of perception and cognition. *Hillsdale, NJ: Erlbaum* (1992).
- [7] BAZEN, A., AND GEREZ, S. Systematic methods for the computation of the directional fields and singular points of fingerprints. *IEEE Trans. Pattern Anal. Mach. Intell.* 24, 7 (2002), 905–919.
- [8] BISHOP, C. *Pattern Recognition and Machine Learning*. Springer, New York, 2006.
- [9] BOLCK, A., WEYERMANN, C., DUJOURDY, L., ESSEIVA, P., AND BERG, J. Different likelihood ratio approaches to evaluate the strength of evidence of MDMA tablet comparisons. *Forensic Science International* 191, 1 (2009), 42–51.
- [10] BOYKOV, Y., VEKSLER, O., AND ZABIH, R. Markov random fields with efficient approximations. *IEEE Conference on Computer Vision and Pattern Recognition* (1998), 648–655.
- [11] CAO, G., SUN, Q., ZHANG, M., AND MEI, Y. Detection of core points in fingerprint images based on edge map. In *Electronic Computer Technology, 2009 International Conference on* (Feb. 2009), pp. 126–129.
- [12] CHAMPOD, C., AND MARGOT, P. Computer assisted analysis of minutiae occurrences on fingerprints. In *Proceeding of Int. Symp. Fingerprint Detection and Identification* (August 1996).
- [13] CHARLES WILSON. *Fingerprint Vendor Technology Evaluation 2003: Summary of Results and Analysis Report (FoPTE)*. U.S. Department of Commerce, National Institute of Standards and Technology, 2004.
- [14] CHEN, J., AND MOON, Y. A statistical study on the fingerprint minutiae distribution. In *ICASSP 2006 Proceedings*. (2006), vol. 2, pp. II–II.
- [15] CHEN, J., AND MOON, Y. The statistical modelling of fingerprint minutiae distribution with implications for fingerprint individuality studies. In *Proc. Computer Vision and Pattern Recognition* (2008).
- [16] CHEN, Y., AND JAIN, A. Beyond minutiae: A fingerprint individuality model with pattern, ridge and pore features. In *ICB '09 Proceedings* (Berlin, Heidelberg, 2009), Springer-Verlag, pp. 523–533.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [17] CHIKKERUR, S., CARTWRIGHT, A. N., AND GOVINDARAJU, V. K-plet and cbfs: A graph based fingerprint representation and matching algorithm. In *International Conference on Biometrics* (2006), pp. 309–315.
- [18] D’AGOSTINO, R., AND STEPHENS, M. *Goodness-of-fit Techniques*. CRC Press, 1986.
- [19] DUDA, R., HART, P. E., AND STORK, D. G. *Pattern Classification, Second Edition*. Wiley-Interscience, 2001.
- [20] DVORNYCHENKO, V. N., AND GARRIS, M. D. Summary of nist latent fingerprint testing workshop, nistir 7377, 2006.
- [21] EGLI, N., CHAMPOD, C., AND MARGOT, P. Evidence evaluation in fingerprint comparison and automated fingerprint identification systems—modeling within finger variability. *Forensic Science International 167*, 2-3 (2007), 189–95.
- [22] EVETT, I. W., AND BUCKELTON, J. S. Statistical analysis of str data. *Advances in Forensic Haemogenetics 6* (1996), 79–86.
- [23] EVETT, I. W., LAMBERT, J. A., AND BUCKLETON, J. S. A Bayesian approach to interpreting footwear marks in forensic casework. *Science and Justice 38*, 4 (1998), 241 – 247.
- [24] EXPERT WORKING GROUP ON HUMAN FACTORS IN LATENT PRINT ANALYSIS. *Latent Print Examination and Human Factors: Improving the Practice through a Systems Approach*. U.S. Department of Commerce, National Institute of Standards and Technology, 2012.
- [25] FENG, J. Combining minutiae descriptors for fingerprint matching. *Pattern Recognition 41*, 1 (January 2008), 342–352.
- [26] FIGUEIREDO, M. A. T., AND JAIN, A. K. Unsupervised learning of finite mixture models. *IEEE Trans. Pattern Anal. Mach. Intell. 24* (2002), 381–396.
- [27] GALTON, F. *Finger Prints*. McMillan, London, 1892.
- [28] GARRIS, M., AND MCCABE, R. NIST special database 27: Fingerprint minutiae from latent and matching tenprint images. <http://www.nist.gov/srd/nistsd27.htm>, June 2000.
- [29] GERMAIN, R. S., CALIFANO, A., AND COLVILLE, S. Fingerprint matching using transformation parameter clustering. *IEEE Computational Science and Engineering* (1997), 42–49.
- [30] HENRY, E. *Classification and Uses of Fingerprints*. London: Routledge, 1900.
- [31] JAIN, A., PRABHAKAR, S., AND HONG, L. A multichannel approach to fingerprint classification. *IEEE Trans. Pattern Anal. Mach. Intell. 21*, 4 (1999), 348–359.
- [32] JAIN, A., PRABHAKAR, S., HONG, L., AND PANKANTI, S. Filterbank-based fingerprint matching. *IEEE Transactions on Image Processing 9* (2000), 846–859.
- [33] JAIN, A. K., AND FENG, J. Latent fingerprint matching. *IEEE Trans. Pattern Anal. Mach. Intelligence 33(1)* (2011), 88–100.
- [34] KAWAGOE, M., AND TOJO, A. Fingerprint pattern classification. *Pattern Recognition 17*, 3 (1984), 295–303.
- [35] KOLLER, D., AND FRIEDMAN, N. *Probabilistic Graphical Models*. MIT Press, 2009.
- [36] KRUMHANSI, C. L. Concerning the applicability of geometric models to similarity data: The interrelationship between similarity and spatial density. *Psychological Review 85* (1978), 445–463.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [37] LI, S. Z. *Markov random field modeling in image analysis*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2001.
- [38] LINDLEY, D. A problem in forensic science. *Biometrika* 64(2) (1977), 207–213.
- [39] LIU, M., JIANG, X., AND KOT, A. Fingerprint reference-point detection. *EURASIP J. Appl. Signal Process. 2005* (2005), 498–509.
- [40] MALTONI, D., MAIO, D., JAIN, A. K., AND PRABHAKAR, S. *Handbook of Fingerprint Recognition*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2003.
- [41] MARDIA, K. V., AND JUPP, P. E. *Directional Statistics*. Wiley, 2000.
- [42] MARR, D. *Vision: A Computational Investigation Into the Human Representation and Processing of Visual Information*. MIT Press, 1982.
- [43] NANDAKUMAR, K., CHEN, Y., DASS, S. C., AND JAIN, A. K. Likelihood ratio-based biometric score fusion. *IEEE Trans. Pattern Anal. Mach. Intell.* 30, 2 (2008), 342–347.
- [44] NATIONAL ACADEMY OF SCIENCES. *Strengthening the Forensic Sciences in the United States: A Path Forward*. National Academies Press, 2009.
- [45] NEUMANN, C., CHAMPOD, C., PUCH-SOLIS, R., EGLI, N., ANTHONIOZ, A., AND BROMAGE-GRIFFITHS, A. Computation of likelihood ratios in fingerprint identification for configurations of any number of minutiae. *Journal of Forensic Sciences* 52 (2007), 54–64.
- [46] NEUMANN, C., CHAMPOD, C., PUCH-SOLIS, R., MEUWLY, D., EGLI, N., AND ANTHONIOZ, A. Computation of likelihood ratios in fingerprint identification for configurations of three minutiae. *Journal of Forensic Sciences* 51 (2006), 1255.
- [47] N.K.RATHA, R.M.BOLLE, V.D.PANDIT, AND V.VAISH. Robust fingerprint authentication using local structural similarity. *IEEE Workshop Fifth Applications of Computer Vision* (2000), 29–34.
- [48] PANKANTI, S., PRABHAKAR, S., AND JAIN, A. On the individuality of fingerprints. *IEEE Trans. Pattern Anal. Mach. Intell.* 24, 8 (2002), 1010–1025.
- [49] PAULINO, A. A., FENG, J., AND JAIN, A. K. Latent fingerprint matching using descriptor-based Hough transform. In *Proceedings International Joint Conference on Biometrics, Washington, DC* (2011).
- [50] PAULSEN, R. R., AND HILGER, K. B. Shape modelling using Markov random field restoration of point correspondences. *Information Processing in Medical Imaging, LNCS 2732* (2003), 1–12.
- [51] PEARSON, K. *The Life and Letters of Francis Galton*, vol. IIIA. University Press, Cambridge, 1930.
- [52] PHILLIPS, D. A fingerprint orientation model based on 2d Fourier expansion (fomfe) and its application to singular-point detection and fingerprint indexing. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29, 4 (2007), 573–585.
- [53] RASMUSSEN, C., AND WILLIAMS, C. *Gaussian Processes for Machine Learning*. MIT Press, 2006.
- [54] SAKS, M. J., AND KOEHLER, J. J. The individualization fallacy in forensic science evidence. *Vanderbilt Law Review* 61 (2008), 199–219.
- [55] SCHOLKOPF, B. The kernel trick for distances. *Advances in Neural Information Processing Systems* 13 (2001).
- [56] SCIENTIFIC WORKING GROUP ON FRICTION RIDGE ANALYSIS, AND TECHNOLOGY. Friction Ridge Examination Methodology for Latent Print Examiners, Version 1.01, 2002.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [57] SCIENTIFIC WORKING GROUP ON FRICTION RIDGE ANALYSIS, AND TECHNOLOGY. Standards for Conclusions, Version 1.0, September 2003.
- [58] SCOLVE, S. The occurrence of fingerprint characteristics as a two dimensional process. *Journal of the American Statistical Association* 367, 74 (1979), 588–595.
- [59] SHEN, Y., NG, A., AND SEEGER, M. Fast Gaussian process regression using kd-trees. In *Advances in Neural Information Processing Systems 18*, Y. Weiss, B. Schölkopf, and J. Platt, Eds. MIT Press, Cambridge, MA, 2006, pp. 1225–1232.
- [60] SNELSON, E., AND GHARAMANI, Z. Sparse Gaussian processes using pseudo-inputs. In *Advances in Neural Information Processing Systems 18*, Y. Weiss, B. Schölkopf, and J. Platt, Eds. MIT Press, Cambridge, MA, 2006, pp. 1257–1264.
- [61] SPITZER, F. Markov random fields and gibbs ensembles. *The American Mathematical Monthly* 78, 2 (February 1971), 142–154.
- [62] SRIHARI, S., AND SRINIVASAN, H. Comparison of ROC and Likelihood Decision Methods in Automatic Fingerprint Verification. *International J. Pattern Recognition and Artificial Intelligence* 22, 1 (2008), 535–553.
- [63] SRIHARI, S., SRINIVASAN, H., AND FANG, G. Discriminability of fingerprints of twins. *Journal of Forensic Identification* 58 (2008), 109–127.
- [64] SRIHARI, S. N. Beyond CSI: The Rise of Computational Forensics. *IEEE Spectrum* 47 (2010), 38–43.
- [65] STONEY, D. Distribution of epidermal ridge minutiae. *American Journal of Physical Anthropology* 77 (1988), 367–376.
- [66] STONEY, D., AND THORNTON, J. A method for the description of minutia pairs in epidermal ridge patterns. *Journal of Forensic Sciences* 31 (1986), 1217.
- [67] STONEY, D. A. Measurement of fingerprint individuality. In *Advances in Fingerprint Technology*, H. Lee and R. Gaensslen, Eds. CRC Press, 2001.
- [68] SU, C., AND SRIHARI, S. Generative models for fingerprint individuality using ridge models. In *Proc. Int. Conf. Pattern Recognition (ICPR)* (2008), IEEE Computer Society Press, pp. 1–4.
- [69] SU, C., AND SRIHARI, S. Probability of random correspondence for fingerprints. In *Proceedings, International Workshop on Computational Forensics (IWCF)* (Berlin, Heidelberg, 2009), Springer-Verlag, pp. 55–66.
- [70] SU, C., AND SRIHARI, S. Evaluation of rarity of fingerprints in forensics. In *Advances in Neural Information Processing Systems 23*, J. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R. Zemel, and A. Culotta, Eds. NIPS, 2010, pp. 1207–1215.
- [71] SU, C., AND SRIHARI, S. N. Latent fingerprint core point prediction based on Gaussian processes. In *Proceedings 20th International Conference on Pattern Recognition, Istanbul, Turkey* (2010), pp. 1634–1637.
- [72] SU, C., AND SRIHARI, S. N. Generative models and probability evaluation for forensic evidence. In *Pattern Recognition, Machine Intelligence and Biometrics*, P. Wang, Ed. Springer, 2011.
- [73] TANG, Y., AND SRIHARI, S. N. Learning Bayesian networks for likelihood ratio computation. In *Proceedings International Workshop on Computational Forensics, Tsukuba, Japan* (2012), Springer.
- [74] TANG, Y., SRIHARI, S. N., AND SRINIVASAN, H. Handwriting individualization using distance and rarity. In *Proceedings Document Recognition and Retrieval, San Francisco, CA* (2012), SPIE.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [75] TARONI, F., AITKEN, C. G., GARBOLINO, P., AND BIEDERMANN, A. *Bayesian networks and probabilistic inference in forensic science*. Wiley, 2006.
- [76] TIPPETT, C., EMERSON, V., FEREDAY, M., LAWTON, F., RICHARDSON, A., JONES, L., AND LAMPERT, S. The evidential value of the comparison of paint flakes from sources other than vehicles. *Journal of Forensic Sciences* (1968), 61 – 65.
- [77] TVERSKY, A., AND GATI, I. Similarity, separability and the triangle inequality. *Psychological Review* 89 (1982), 123–154.
- [78] U.S. DEPARTMENT OF JUSTICE, OFFICE OF THE INSPECTOR GENERAL. A Review of the FBI’s handling of the Brandon Mayfield case (Unclassified and Redacted). http://www.justice.gov/oig/special/s0601/PDF_list.htm, March 2006.
- [79] WANG, X., LI, J., AND NIU, Y. Definition and extraction of stable points from fingerprint images. *Pattern Recognition* 40, 6 (2007), 1804–1815.
- [80] WATSON, C., GARRIS, M., TABASSI, E., WILSON, C., MCCABE, R., AND JANET, S. *User’s Guide to NIST Fingerprint Image Software 2 (NFIS2)*. NIST, 2004.
- [81] WATSON, C., AND WILSON, C. NIST special database 4: 8-bit gray images of fingerprint image groups. <http://www.nist.gov/srd/nistsd4.htm>, March 1992.
- [82] X.D.JIANG, AND W.Y.YAU. Fingerprint minutiae matching based on the local and global structures. *IEEE 15th ICPR* (2000), 1042–1045.
- [83] XIAO, P., BARNES, N., CAETANO, T., AND LIEBY, P. Mrf and gaussian curvature based shape representation for shape matching. *CVPR* (2007).
- [84] ZHU, Y., DASS, S., AND JAIN, A. Statistical models for assessing the individuality of fingerprints. *IEEE Transactions on Information Forensics and Security* 2, 3-1 (2007), 391–401.