Fingerprint Recognition: Models and Applications

By

Soweon Yoon

A Dissertation

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Computer Science – Doctor of Philosophy

2014

ABSTRACT

FINGERPRINT RECOGNITION: MODELS AND APPLICATIONS

By

Soweon Yoon

Fingerprint recognition has been successfully used in law enforcement and forensics to identify suspects and victims for over a century. Recent advances in automated fingerprint identification systems (AFIS), coupled with the growing need for reliable person recognition, have resulted in an increased deployment of AFIS in broad applications such as border control, employment background checks, secure facility access, and user authentication in laptops and mobile devices. Despite the success of fingerprint recognition technique in many large-scale and diverse person identification applications, several challenging issues in fingerprint recognition still need to be addressed. First, the persistence and uniqueness of fingerprints—the fundamental premises for fingerprint recognition—remain as presumptions rather than facts with solid scientific underpinnings. Although some studies have addressed the uniqueness of fingerprints, there has been no systematic study reported on the persistence of fingerprints. Given a large longitudinal fingerprint database, we have analyzed it with a multilevel statistical model to assess the impact of time interval between a genuine fingerprint pair on the corresponding match score and identification decision. Second, an appropriate mathematical model for fingerprint orientation field is necessary in addressing a number of challenging problems, including feature extraction (e.g., orientation field and minutiae) from noisy or partial fingerprints, detecting abnormality in fingerprint patterns, and representing fingerprints in a compact form. To this end, we have developed a global orientation field model in the form of ordinary differential equations which is constrained by the number of singular points in a fingerprint. The model represents the global orientation field with a small number of polynomial terms and outputs fingerprint-like orientation fields with the specific number of singular points after fitting an input pattern. We use

this model to check the *fingerprint-ness* of the input image and ensure the integrity of exemplar fingerprint databases. Third, given that automatic latent fingerprint matching is difficult due to poor quality of latent fingerprints found at crime scenes, we have developed a semi-automatic latent enhancement method. The proposed algorithm is based on robust orientation field estimation that is able to (i) reduce human intervention in feature markup and (ii) improve automatic feature extraction and matching accuracy of latents. Fourth, fingerprint obfuscation or alteration is a type of attack on AFIS that is of concern to law enforcement and border crossing agencies. We show that the fingerprint matching accuracy can greatly deteriorate when the altered fingerprints are presented to AFIS. To address this deficiency of AFIS, we have (i) analyzed the types of fingerprint obfuscation, (ii) developed a detection algorithm for altered fingerprints by measuring the abnormality in orientation field and minutiae distribution, and (iii) proposed a restoration and matching algorithm to possibly link an altered fingerprint to its pre-altered mate. By addressing these contemporary problems, this dissertation has advanced our understanding of fingerprint recognition and enabled development of robust and accurate fingerprint recognition algorithms.

ACKNOWLEDGMENTS

I thought that being a Ph.D. was to become an independent researcher. I do not dare say that I have matured into an independent researcher at this time as I finish my Ph.D. study. However, my six years at MSU gave me courage to step out into the wilderness and explore.

I am most grateful to my advisor, Dr. Anil K. Jain. He is not only a perfect advisor for me, but also a great mentor to me. He has been always supportive of my research, open to my gut feelings when determining research directions, and respected my way of conducting research with great patience. He opened up many opportunities for me during the course of study and let me experience the real world, instead of just sitting in the lab without knowing what is happening out there. I also appreciate his warmth and compassion as a person. He was there whenever I needed any help, sometimes even before I was aware that I needed help.

Jianjiang Feng, a great researcher and my model Ph.D., made my Ph.D. life complete. His flawless papers and brilliant ideas thrilled me. I can say, without hesitation, that the happiest moment during my Ph.D study was when I worked together with Dr. Jain and Jianjiang. I have really enjoyed doing research and learned how research should be done.

I was very lucky that I joined the PRIP lab along with four other rookies—Alessandra, Brendan, Serhat, and Kien. We went through formalities necessary to complete the study by checking on each other, traveled together to conferences and meetings, and shared joy and frustration we had. Alessandra, who is thoughtful and conscientious in nature, was not just a colleague in the lab, but a friend to me. Brendan's confidence and energy motivated me to endeavor to become more active and optimistic. I would like to thank Serhat for becoming a companion in the last tough year of Ph.D. study. I also thank: Jung-Eun, Unsang, Pavan, Abhishek, Radha, Koichiro, Sunpreet, Scott, Lacey, Charles, Kai, Hu, Di, Qijun, Heeseung,

iv

Table of Contents

LIST OF TABLES

List of Figures

xiii

Chapter 1

Introduction

Fingerprints have a long history of use as a means of reliably identifying individuals. Based on the persistence and uniqueness of fingerprints, fingerprint recognition systems have become one of the most popular biometric systems used in many applications, including law enforcement, border control, and forensics. In this chapter, we describe the fingerprint representation at three levels which are widely used in fingerprint matching, and explain two matching scenarios in fingerprint recognition: (i) exemplar fingerprint matching that compares fingerprints obtained in tenprint cards and (ii) latent fingerprint matching that searches latents found at crime scenes against exemplar fingerprint databases. We discuss challenging contemporary issues in fingerprint recognition, including latent matching and fingerprint obfuscation, and present the contributions of this dissertation that address these issues.

1.1 Friction Ridge Pattern for Human Identification

Friction ridge skin is located only on particular parts of human body: the palms of the hands and the soles of the feet. Although their biological function is presumably to increase frictional force between an object and the hand or the foot to grab things firmly or walk on the ground easily, another usage of the friction ridge skin was discovered thousands of years
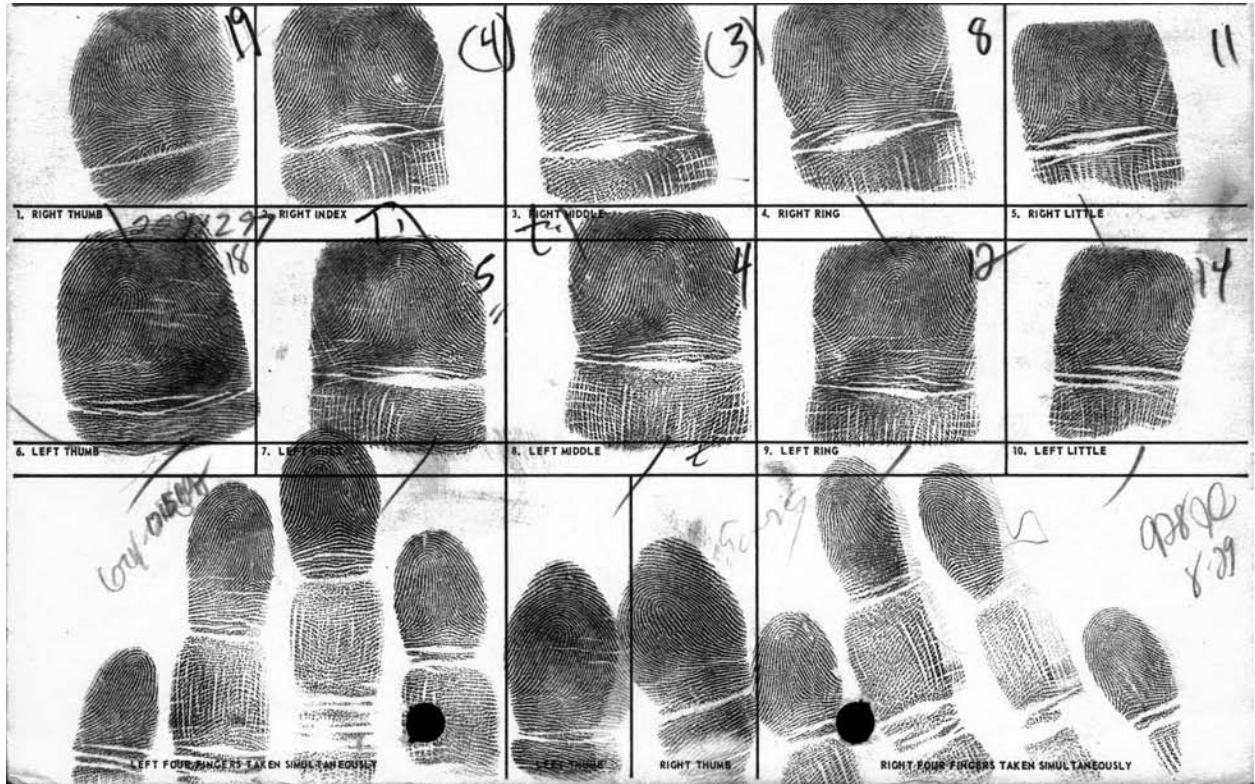
Figure 1.1: The FBI's tenprint card (from NIST SD29 [1]).

ago: human identification based on the patterns of the friction ridges. The historical records of human identification by friction ridge patterns can be found in Babylon (1955–1913 BC) and Ancient China (AD 600–700), where fingerprints were used to seal contracts and legal documents [69]. However, scientific study of fingerprints as a tool of human identification emerged only in the late 19th century [51, 60, 72].

The first reported case where fingerprints were officially accepted as an evidence to convict a suspect involved a murder case in Argentina in 1893; bloody fingerprints on the door made the mother confess to the murder of her two children [69]. After the Scotland Yard started to include fingerprints on the anthropometric identification cards which recorded measurements of various physical attributes of the criminals around 1900 [44], use of fingerprints spread rapidly worldwide to identify habitual criminals. In the United States, the Federal Bureau of Investigation (FBI) set up its fingerprint identification division in 1924, and their fingerprint

2

database initially contained about 810,000 tenprint cards [52]. Figure 1.1 shows an example of the FBI's tenprint card. With the increasing number of records in the database, it was not feasible to do manual matching of fingerprints. Research in the design and development of Automated Fingerprint Identification Systems (AFIS) was initiated by the FBI in 1970's. State and local law enforcement agencies took the lead from the FBI and started to operate their own AFIS; the system installed in Minneapolis-St. Paul, Minnesota, in 1978 was one of the first AFIS in the United States [95]. The FBI's Integrated AFIS (IAFIS) currently holds the tenprint records of over 75 million apprehended criminals and 39 million civilian government job applicants as of November 2013; the average response time for tenprint rapsheet requests is 1 minute and 9 seconds (97% of requests are completed within 15 seconds) [7].

1.1.1   Friction Ridge Pattern: Fingerprints and Palmprints

It is widely accepted that the formation of friction ridge skin is related to the skin layer, called *volar pads* [84]. The volar pads start to form at the 7th week of gestation, and the size and shape of the embryo's volar pads affect the fingerprint pattern; the area where volar pads are located tends to create loop or whorl patterns while the area without volar pads tends to have straight ridges. Primary ridges are developed after the volar pads disappear around the 10th week of gestation until the 19th week, and the ridge structures developed during this period do not change further and are established for life. Then, secondary ridges where sweat pores do not exist (also called incipient ridges in fingerprint representation) start to appear until the 24th week of fetal development. The volar pads also play a critical role in the formation of friction ridge skin on the palm; four pads are present in the interdigital regions (below the four digits from index finger to little finger) and one pad each is present in the hypothenar and thenar region [46].

3

(a)   (b)   (c)

Figure 1.2: A. E. H. Herschel's fingerprints [73] at (a) age 7, (b) age 17, and (c) age 40. The pairwise match scores (i.e., similarity between a pair of fingerprints) from a state-of-the-art fingerprint matcher for these three fingerprints are: (a) vs. (b) 6,217; (a) vs. (c) 5,032; (b) vs. (c) 5,997; the maximum impostor score of (a) against 10,000 fingerprints in NIST SD4 [2] is 3,325 and that of (b) is 2,935.

1.1.2 Premises of Fingerprints for Human Identification

Fingerprints are believed to be persistent during the lifetime of an individual. Figure 1.2 shows three fingerprint impressions captured by Herschel of his son's finger at different ages (age 7, 17, and 40) [73]. The pairwise match scores from a state-of-the-art fingerprint matcher indicate that these *genuine* pairs (i.e., fingerprint impressions from the same finger) have much higher similarity in fingerprint feature correspondences than the *impostor* pairs (i.e., fingerprints from different fingers). This supports the general belief that there is no significant change in fingerprint pattern of an individual over time.

Individual fingerprints are also known to be unique; even identical twins who share the same DNA information have different fingerprints. In [81], the experimental results with fingerprints collected from 94 pairs of identical twins showed that, even though the type of fingerprint pattern is highly correlated between identical twins (see Figure 1.3; fingerprints of the identical twins have the same pattern type, i.e., right-loop), the fingerprint recognition system is able to distinguish them as different individuals based on the ridge details.

4

| (a) | (b) | (c) |
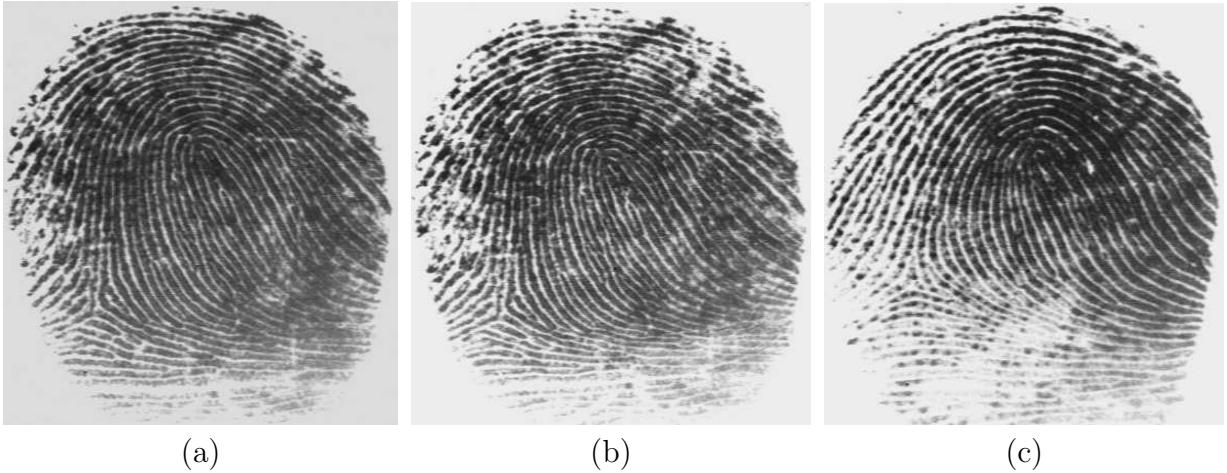
Figure 1.3: Fingerprints from the index fingers of identical twins [81]. (a) The first impression of the index finger of a twin, (b) the second impression of the same finger of the twin, and (c) an impression of the index finger of her sibling. The match score of (a) and (b), which is a genuine match, is 487 from the matcher used in [81] while the match score of (a) and (c), which is a twin-impostor match, is only 24.

1.1.3 Applications of Fingerprint Recognition

Fingerprint recognition is widely used in various applications ranging from law enforcement and international border control to personal laptop access. Almost all law enforcement agencies worldwide routinely collect fingerprints of apprehended criminals to track their criminal history [7]. To enhance border security in the United States, the US-VISIT program [8] acquires fingerprints of visa applicants to identify high profile criminals on a watch list and detect possible visa fraud. India's UIDAI project was initiated to issue a unique 12-digit identification number to each resident. Given the large population in India (approximately 1.2 billion), an identification number for an individual is associated with his biometric information (i.e., ten fingerprints and two irises) to ensure that each resident has only one identification number [9].

Fingerprint recognition systems are now pervasive in our daily life. Disney Parks, for example, captures fingerprints of visitors when they initially enter the park to link the ticket to the ticket holder's fingerprint. Fingerprint verification is performed whenever the same

ticket is presented for reuse to prevent fraudulent use of the ticket (e.g., sharing of a ticket by multiple individuals). Many automated teller machines (ATMs) in Brazil use fingerprint recognition as a replacement for personal identification numbers (PINs) [28]. Also, several laptop computer models are equipped with fingerprint sensors and authenticate users based on their fingerprints [29].

1.2 Fingerprint Representation

A fingerprint pattern consists of intervening ridges and valleys spaced almost equidistantly (about 9 pixels in 500ppi fingerprint images of adults). Fingerprints are typically described by features at three levels (see Figure 1.4): (i) Level-1 features: ridge flow and pattern type, (ii) Level-2 features: ridge endings and bifurcations, and (iii) Level-3 features: fine ridge details such as pores and incipient ridges [91].

1.2.1 Level-1 Features: Orientation Field and Pattern Type

1.2.1.1 Orientation Field

Orientation field[1] of a fingerprint, $\theta(x, y)$, describes the tangential direction of the ridges at a point (x, y), where $0 \leq \theta(x, y) < \pi$. Since the orientation has π-ambiguity (i.e., two different vectors, one directed at θ and the other directed at $(\theta + \pi)$, have the same orientation), the orientation field is often converted to a vector field by doubling the angles.

A 2-dimensional vector field can be represented by the first-order ordinary differential equations (ODE) as follows:

$$\dot{x} = f(x, y),$$
$$\dot{y} = g(x, y), \tag{1.1}$$

---

[1]Orientation field is also called *ridge flow*.
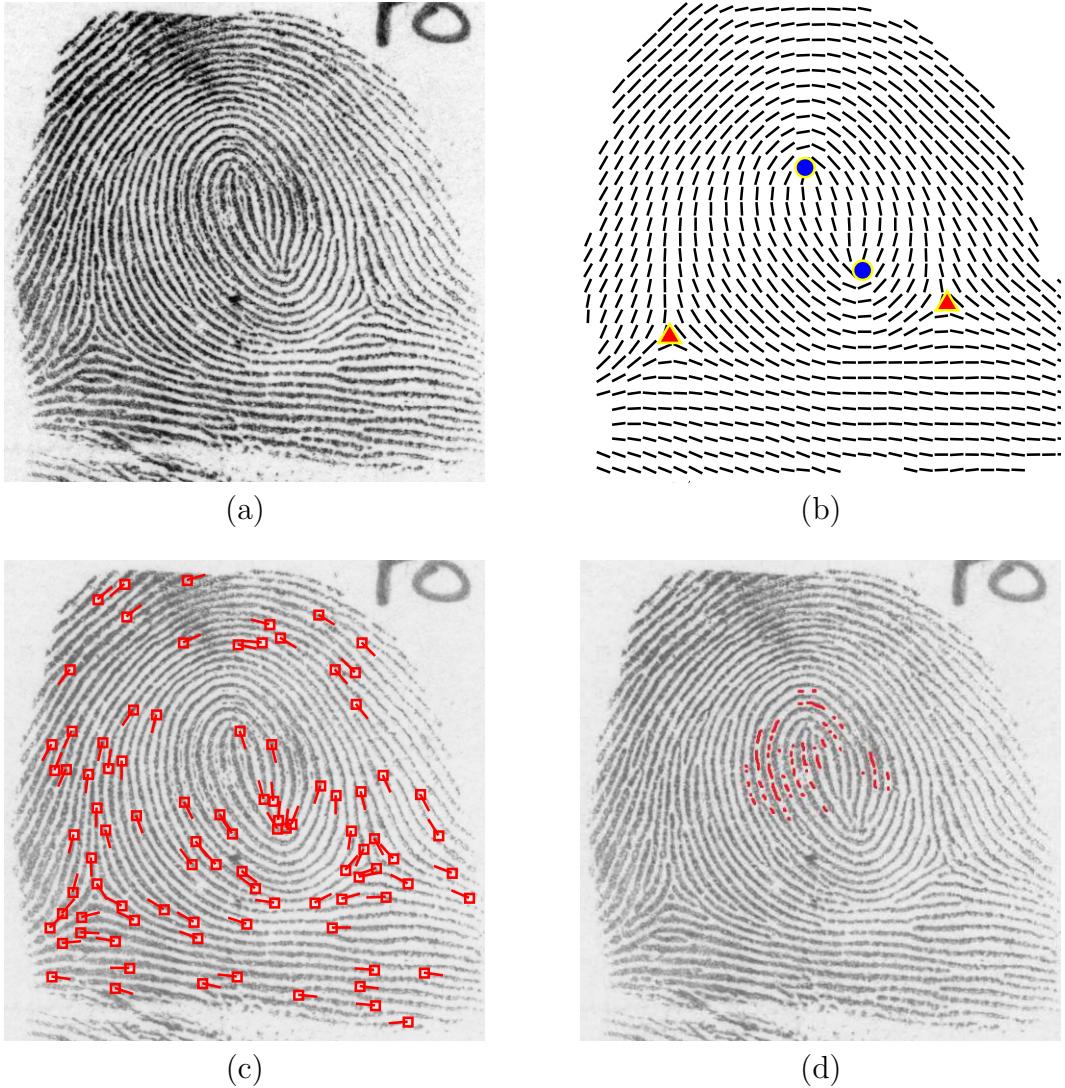
Figure 1.4: Fingerprint features at three different levels of detail. (a) Rolled fingerprint in NIST SD29 [1], (b) Level-1 features: orientation field and singular points (cores marked as circles and deltas marked as triangles), (c) Level-2 features: minutiae, and (d) Level-3 features: incipient ridges.

where $f(x, y)$ and $g(x, y)$ are real-valued functions. Once $f(x, y)$ and $g(x, y)$ are known, the orientation field can be uniquely determined as:

$$\theta(x, y) = \frac{1}{2} \tan^{-1} \left(\frac{\dot{y}}{\dot{x}} \right) = \frac{1}{2} \tan^{-1} \left(\frac{g(x, y)}{f(x, y)} \right). \tag{1.2}$$

The x-isocline in differential equations is the set of points where $\dot{x} = 0$; that is, the x-

isocline of the system in Equation (1.1) is the set of points where $f(x, y) = 0$ [75]. Similarly, the y-isocline of the system is the set of points where $\dot{y} = 0$, that is, $g(x, y) = 0$. A singular point[2] occurs at a point where both $\dot{x} = 0$ and $\dot{y} = 0$; that is, the x-isocline and the y-isocline intersect. In a fingerprint orientation field, the singular points correspond to *core* and *delta* (see Figures 1.5(a) and (c)).

A way of determining the characteristics of a singular point of an ODE system is by a *linearization* of the system. The local vector field can be approximated by a linear system [120]:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{b}, \tag{1.3}$$

where

$$\dot{\mathbf{x}} = \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad \text{and } \mathbf{b} = \begin{bmatrix} e \\ f \end{bmatrix}.$$

The matrix $\mathbf{A}$ is called the *characteristic matrix*. The stability of the singularity at the center of the local field is the same as that of the singularity at the origin of the linearized field when the real parts of the eigenvalues of $\mathbf{A}$ are non-zero (see *linearization theorem*) [75]. Singularities in a vector field such as a source and a sink with $|\mathbf{A}| > 0$, where $|\mathbf{A}|$ is the determinant of $\mathbf{A}$, correspond to a core in the fingerprint orientation field. On the other hand, a saddle in a vector field which has $|\mathbf{A}| < 0$ corresponds to a delta in the fingerprint orientation field. Figure 1.5 shows the orientation fields of cores and deltas at different orientations and the corresponding vector fields.

Another way of determining singular points is by using the Poincaré index [83], which is commonly used to detect singularities in a fingerprint. The Poincaré index of an orientation field along an arbitrary simple closed path γ is defined as:

$$I(\gamma) = \frac{1}{2\pi} \oint_{(x,y) \in \gamma} d2\theta(x, y), \tag{1.4}$$

---

[2]In a vector field, a singular point is also called an *equilibrium point* or a *critical point*.

8
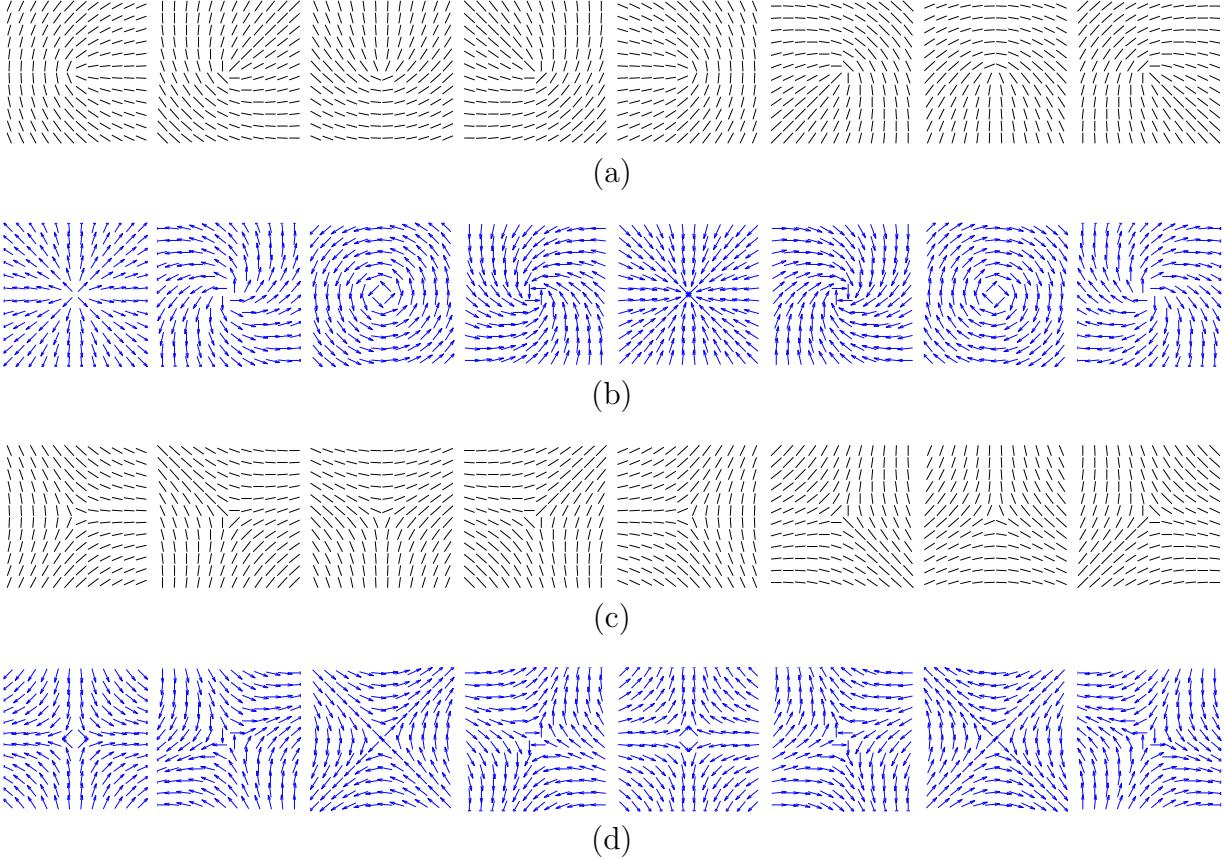
Figure 1.5: Orientation field and vector field around singular points. (a) Orientation fields of cores with different orientations, (b) vector fields of the corresponding cores, (c) orientation fields of deltas with different orientations, and (d) vector fields of the corresponding deltas. A core in the orientation field corresponds to a source or a sink with $|\mathbf{A}| > 0$ in the vector field while a delta in the orientation field corresponds to a saddle in the vector field which has $|\mathbf{A}| < 0$.

where $\theta(x, y)$ is the orientation field. To determine if there is a singularity at a point P, the simple closed path γ is generally set to the neighboring circular path surrounding the point P. Then, the Poincaré index always takes one of the following three integer values: $I(\gamma)$ is 1 if P is a core, -1 if P is a delta, and 0 if P is in a smooth orientation field region.
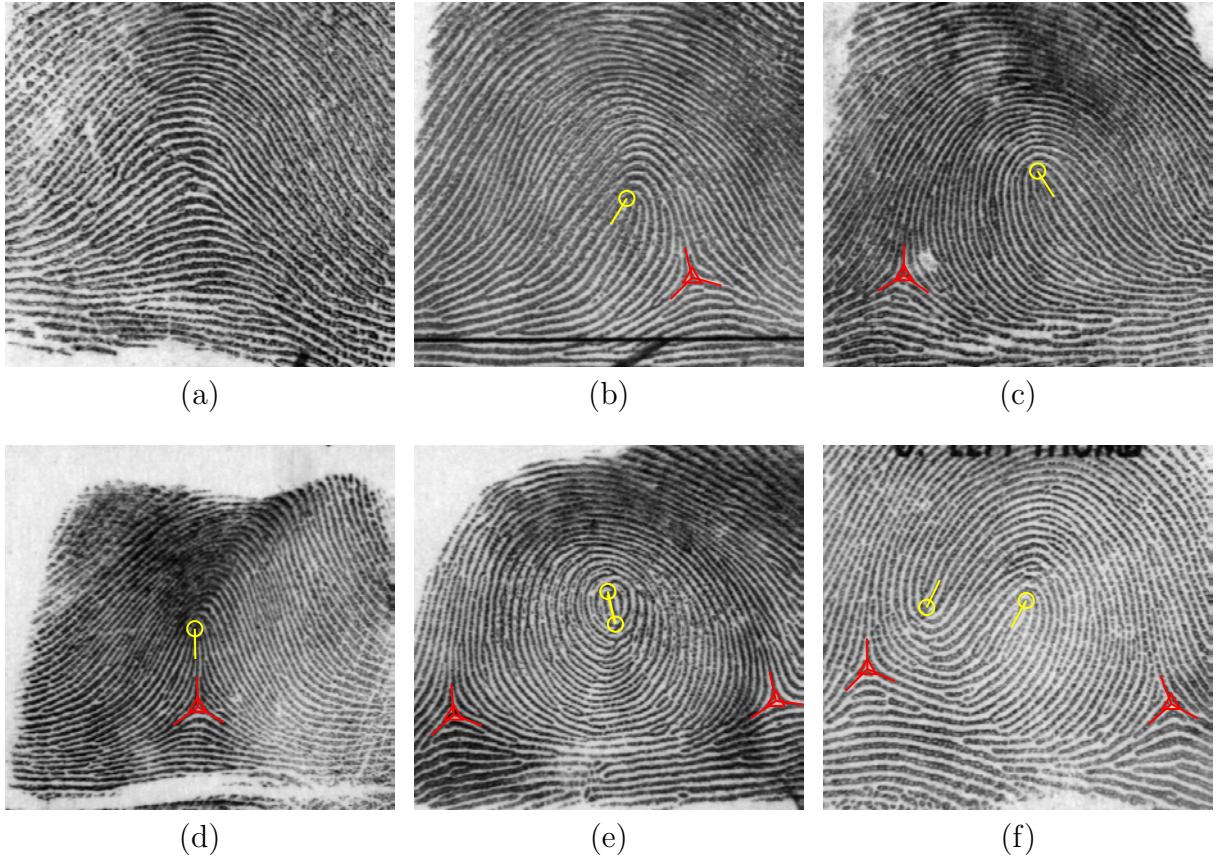
Figure 1.6: Fingerprint pattern types. (a) Arch, (b) left-loop, (c) right-loop, (d) tented-arch, (e) whorl, and (f) twin-loop. Core is marked as yellow circle, and delta is marked as red triangle.

1.2.1.2 Pattern Types

The Henry classification of fingerprint patterns [71] is based on the number and the spatial configuration of singular points in fingerprints, and almost all fingerprints[3] fall into one of the following three categories (see Figure 1.6): (i) no singularity (arch type), (ii) one core and one delta (left-loop, right-loop, and tented-arch type), and (iii) two cores and two deltas (whorl and twin-loop type). Note that a whorl can be viewed as consisting of two cores. The fingerprints in category (iii) are collectively called *double-loop* type.

---

[3]The statistics of fingerprint class or type distribution in a very large database containing more than 220 million fingerprints reported in [10] showed that almost all fingerprints belong to arch, loop or double-loop; only 0.2% of fingerprints are classified as accidental whorl or scar/mutilation, and 0.15% of fingerprints are determined as amputation/missing.

$$\text{(a)} \qquad\qquad\qquad\qquad \text{(b)}$$
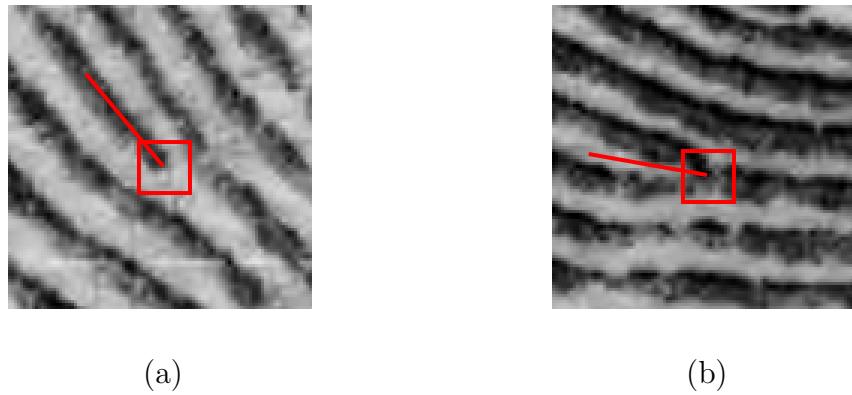
Figure 1.7: Minutia types in a fingerprint. (a) Ridge ending and (b) ridge bifurcation.

1.2.2 Level-2 Features: Minutiae

Minutia points are the most popular feature used in fingerprint matching. The two most common types of minutiae found in a fingerprint are ridge ending (Figure 1.7(a)) and ridge bifurcation (Figure 1.7(b)). The simplest way of representing a minutia is by using a 3-tuple vector: (x, y, θ), where (x, y) is the position of the minutia and θ is its direction. Most fingerprint matching algorithms are based on measuring the similarity in global configurations of two minutiae sets representing the two fingerprint images.

1.2.3 Level-3 Features: Pores, Dots, and Incipient Ridges

The fine ridge details in a fingerprint such as pores, dots, incipient ridges, etc. (see Figure 1.8) are defined as Level-3 features. They can be useful information to compare fingerprints when the Level-1 and Level-2 features present in the fingerprints are not sufficient (often the case with latent fingerprints found at crime scenes) to make a decision on a pair of fingerprints: the two fingerprint impressions (i) come from the same finger or (ii) come from two different fingers. The Level-3 features are easier to observe in high resolution images (at least 1000 ppi as opposed to the typical resolution of 500 ppi used to collect fingerprint images). Algorithms to extract Level-3 features are not only computationally expensive in

11
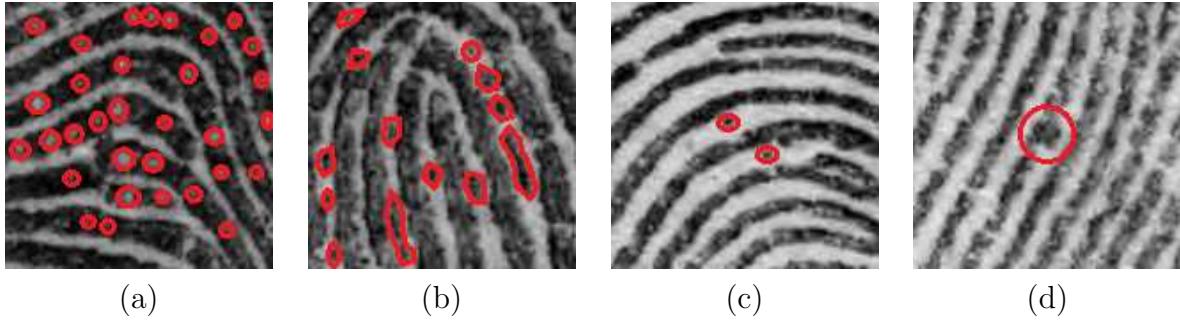
Figure 1.8: Level-3 features in a fingerprint. (a) Pores, (b) incipient ridges, (c) dots, and (d) ridge edge protrusion.

general, but are also quite sensitive to image noise.

1.3 Fingerprint Matching

Fingerprint matching scenarios generally fall into one of the following two categories based on the type of query fingerprint: (i) exemplar search and (ii) latent search. In exemplar search, rolled or plain impression of a subject's finger (Figures 1.9(a) and (b)) is searched against the exemplar fingerprint database (typically consisting of rolled and plain impressions) by using an AFIS. In law enforcement and border crossing applications, this generally involves all ten fingerprints (i.e., tenprint search) to ensure the individual's identity. In latent search, a fingerprint captured at a crime scene is fed into the AFIS along with the manually marked features; a latent examiner verifies the top N matching results (N varies according to the importance of a crime; N is typically around 100) from an AFIS to confirm the identify of the suspect.

1.3.1 Types of Fingerprints

1.3.1.1 Exemplar Fingerprints

Rolled and plain fingerprints, collectively called *exemplar* fingerprints, refer to the fingerprints obtained on a formatted tenprint card (see Figure 1.1; the first two rows contain rolled

12

(a)   (b)   (c)
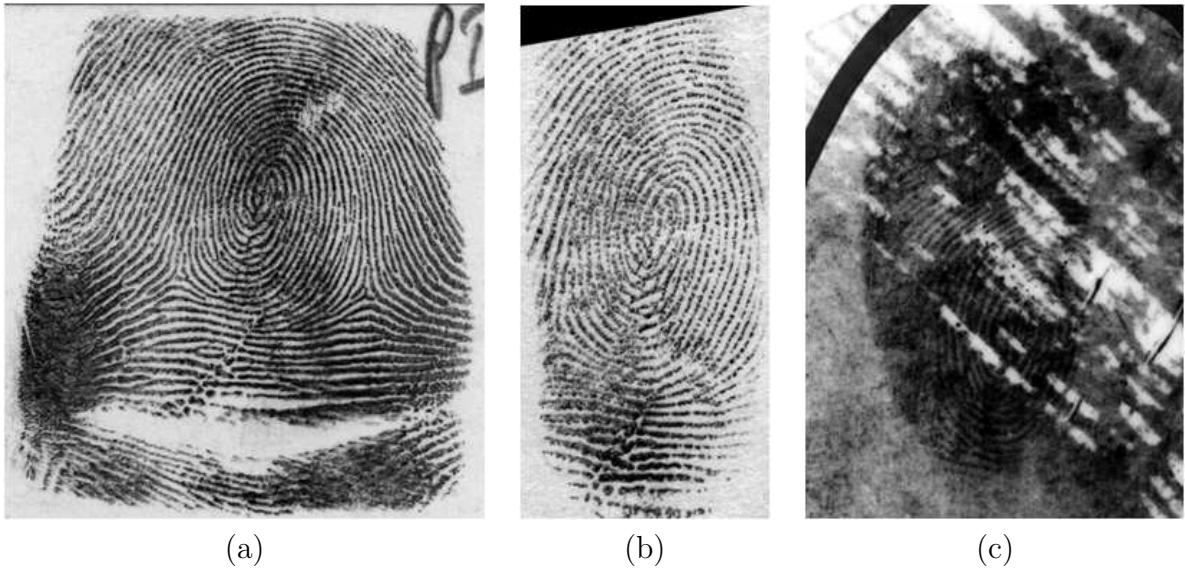
Figure 1.9: Three types of impressions of the same finger. (a) Rolled fingerprint, (b) plain fingerprint in NIST SD29 [1], and (c) latent fingerprint in NIST SD27 [3].

fingerprints of ten fingers and the last row contains plain fingerprints) in an attended mode (i.e., typically in the presence of a law enforcement officer). Rolled fingerprints are obtained by rolling a finger from one side to the other ("nail-to-nail") in order to capture all the ridge details of a finger (Figure 1.9(a)). Plain fingerprints, also called *flat* or *slap* fingerprints, are acquired by pressing fingertips onto a flat surface of either a paper for inking methods or a flatbed of a live-scan device (Figure 1.9(b)). To make sure that the indices of the fingerprints in rolled impressions are correct, plain or slap impressions are made by capturing four fingers of a hand (from index finger to little finger) together and then taking the thumb impression separately (called 4-4-2 acquisition). Rolled fingerprints contain a large number of minutiae (about 100), and a significant amount of skin distortion can be introduced during rolling of the finger. Plain fingerprints, on the other hand, capture relatively small finger area with a smaller number of minutiae (about 50), but have lower skin distortion. Both rolled and plain fingerprints are typically acquired under the supervision of a human operator to ensure that good quality impressions are collected. If the fingerprint image is determined to be of poor quality, it is recaptured.

The matching performance of state-of-the-art AFIS for exemplar search has already reached a highly satisfactory level in most fingerprint recognition applications. The Fingerprint Vendor Technology Evaluation (FpVTE) in 2003 [135] reported that the best performing commercial matcher achieved a 99.4% verification rate in searching plain fingerprints against an exemplar database with 10,000 plain fingerprints.

1.3.1.2 Latent Fingerprints

Latent fingerprints (see Figure 1.9(c)) refer to the fingerprints lifted from the surface of objects touched or handled by a person. Latents are an extremely important source of evidence in crime scene investigation to identify and convict suspects. Unlike rolled and plain fingerprints, latent fingerprints are often of poor quality: the latent fingerprints contain partial ridge patterns of a finger, incomplete or missing ridge structures, smudged or blurred ridges, mixture of ridge pattern and complex background noise or friction ridge structures from other fingers, and large nonlinear skin distortion due to pressure variations (see Figure 1.10).

Matching a latent fingerprint against an exemplar fingerprint database is of utmost importance in forensics and law enforcement to apprehend suspects. However, the latent fingerprint recognition is still a very challenging problem, particularly with respect to the following two issues: (i) eliminate or minimize human intervention which is currently needed to process latents and (ii) improve latent identification accuracy so that a large percentage of latents can be identified by verifying a very short list of candidates from the exemplar database (ideally the list contains only one candidate if there is a match). In contrast to the exemplar matching that is fully automated except for very poor quality query prints, current practice in latent matching involves a large degree of human intervention in marking features in the latents and comparing the latent to top N candidates retrieved by the AFIS from the database. Furthermore, the ELFT-EFS (Evaluation of Latent Fingerprint Technologies: Extended Feature Sets) in 2012 [78] showed that the state-of-the-art matching performance
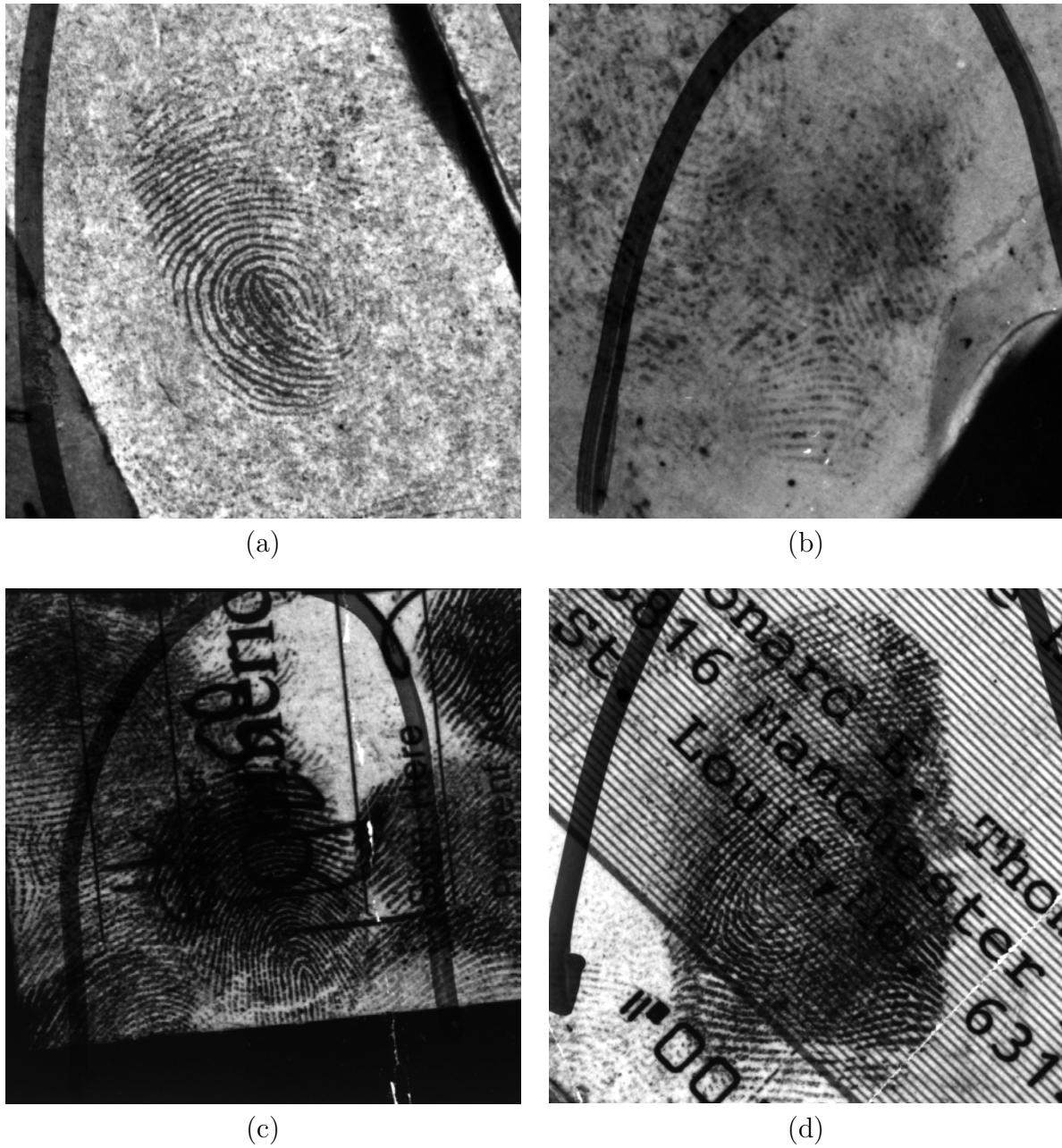
Figure 1.10: Challenging latent fingerprint images in NIST SD27 [3]. (a) Partial area, (b) unclear ridge structure, (c) overlapped with other fingerprints, and (d) complex background.

of latent fingerprints was 68.2% with a full feature set (i.e., latent image, minutiae, and extended features such as Level-3 features and fingerprint skeleton) in searching 418 latents against 100,000 tenprint exemplar database including both rolled and plain impressions.

1.3.2   Exemplar Fingerprint Matching

Exemplar fingerprint matching process consists of two phases: (i) enrollment and (ii) verification or identification [91]. In the enrollment phase, a user's fingerprint is acquired and the features extracted from the fingerprint images are stored as a template with the subject's ID and other demographic information. During the verification/identification phase, the fingerprint of a subject is used to determine if he was previously enrolled in the system.

1.3.2.1   Quality Assessment

Fingerprint image quality is an important factor in the matching accuracy; features extracted from poor quality fingerprint are likely to have many spurious or missing minutiae. Fingerprint image quality is influenced by intrinsic factors of the finger skin (i.e., skin condition such as dryness or salience of the ridges) and extrinsic factors (i.e., sensitivity of the fingerprint imaging sensor or positioning of the user's finger on the sensor). AFIS for exemplar search assess the fingerprint quality at the front end of the system, and ask users to provide another impression of their fingerprints if the fingerprints are of poor quality in the enrollment phase. In the recognition phase, the quality module rejects the poor quality fingerprints that are not adequate for matching in the verification/identification phase instead of making erroneous identification decisions.

Algorithms to assess fingerprint image quality mainly utilize the features to measure local properties (i.e., clarity of ridge structure in a grayscale image) and global properties (e.g., continuity of orientation field or energy concentration in the frequency domain over the entire fingerprint region) [31]. NIST Fingerprint Image Quality (NFIQ) is one of the de facto standards to determine fingerprint image quality, which gives one of five discrete quality levels ranging from 1 (the highest quality) to 5 (the lowest quality). Figure 1.11 shows fingerprint images with three different NFIQ values.
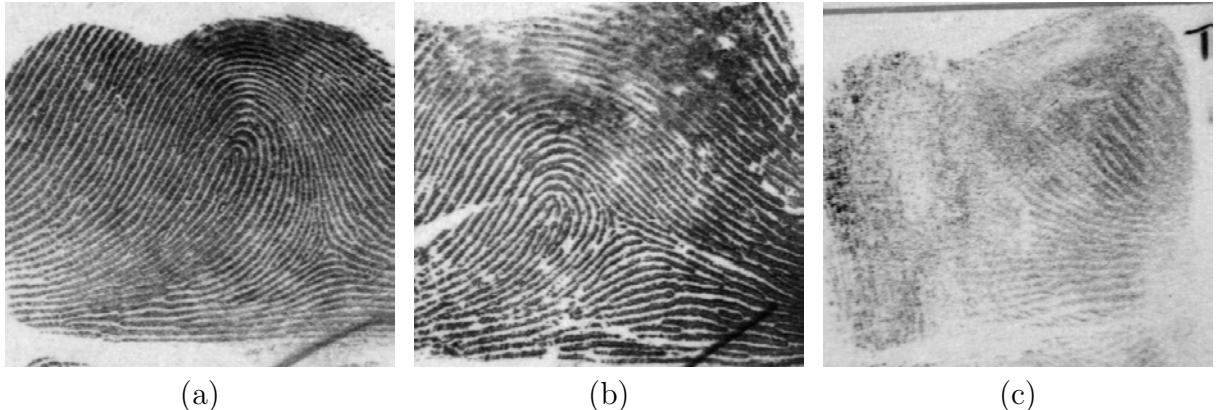
Figure 1.11: Fingerprints of various quality levels: (a) NFIQ of 1, (b) NFIQ of 3, and (c) NFIQ of 5.

1.3.2.2 Feature Extraction

During the enrollment phase, a template which contains a set of features such as orientation field, singular points, and minutiae from one or more fingerprint images of a finger is created and stored in the system along with the images. To compensate the intra-class variations in the fingerprint images (e.g., each image can capture a different part of the finger, or the quality of each impression can vary significantly due to the changes in acquisition condition), either a template integrating multiple fingerprint images is created [82], or or multiple prototype templates which represent the variations in a set of fingerprint images are selected [127].

1.3.2.3 Verification and Identification

An exemplar fingerprint matching is conducted as either (i) verification or (ii) identification. Verification is an one-to-one matching; when a user provides his fingerprint along with his identity, the system retrieves the template corresponding to the identity and determines if the two fingerprints match. On the other hand, in the identification scenario, a user provides his fingerprint without claiming his identity, and the system searches his fingerprint against all the fingerprints in the database (i.e., one-to-many matching) and determines if there is a

match or no match.

Two types of errors can be made in fingerprint recognition: (i) false acceptance and (ii) false rejection. False acceptance refers to the case where the input fingerprint is falsely matched to a fingerprint in the database which is not the mate of the input fingerprint. False rejection is the case where the input fingerprint is not matched to its true mate in the database. The performance of a fingerprint verification algorithm is usually measured by a Receiver Operating Characteristic (ROC) curve which indicates the true acceptance rate (TAR) with respect to the false acceptance rate (FAR).

A number of evaluations have been performed for fingerprint verification scenarios. The NIST FpVTE 2003 determined that the best commercial AFIS had a TAR of 99.4% at FAR of 0.01% in the medium-scale test with 10,000 plain fingerprints [135]. The best performing algorithm in Fingerprint Verification Competition (FVC) achieved an average equal error rate (EER)[4] of 2.07% on the FVC 2004 datasets [39]. For the evaluation of fingerprint identification performance, the FpVTE 2012 searching against the exemplar database with millions of fingerprints has been conducted (but the results have not yet been released) [98].

1.3.3   Latent Fingerprint Matching

1.3.3.1   Acquisition

Latent fingerprints are lifted from the surface of objects through a variety of means ranging from simply photographing the print to more complex dusting or chemical processing [87]. Unlike exemplar fingerprints which are obtained under a supervision of a well-trained operator, latent fingerprints are acquired from the residues of a fingermark such as sweat which can be easily destructible in natural environments over time. This explains why the quality of latent fingerprints is often poor and cannot be controlled.

---

[4]The equal error rate (EER) refers to the error rate where FAR = FRR.

1.3.3.2 ACE-V Methodology

In matching latents to exemplar prints, latent fingerprint examiners are expected to follow a methodology, called *Analysis, Comparison, Evaluation*, and *Verification* (ACE-V) [33]. In the analysis phase, an examiner evaluates the ridge information contained in latent images. If the latent is determined to have sufficient information for individualization or exclusion (called "of value" latent; see the definitions described in the evaluation phase), the features in the latent are manually marked by the examiner to search for their mates using an AFIS. In the comparison phase, the examiner compares the "of value" latent with the candidate mates retrieved from the exemplar database side-by-side and ascertains the similarity between the latent and mated exemplar print pairs using feature markup in the latent. In the evaluation phase, one of the following decisions is made about the latent in question: individualization, exclusion, or inconclusive[5]. Finally, in the verification phase, the decision made by the first examiner is confirmed by having a second examiner analyze the results independently.

1.3.3.3 Performance Evaluation

Since manual comparisons between a latent and a set of candidate mates retrieved from an exemplar database is involved in the comparison phase of the ACE-V methodology, the performance of latent matching algorithms is commonly measured by plotting a Cumulative Match Characteristic (CMC) curve. A CMC curve draws the cumulative rank-m identification rate (i.e., the percentage of the query latents whose mates are retrieved within rank m) with respect to the rank m.

---

[5]Individualization is the decision that a latent examiner makes on a pair consisting of a latent and an exemplar print indicating that the pair originates from the same finger based on a sufficient agreement between the two ridge patterns. Exclusion, on the other hand, is the decision where an examiner concludes that the pair did not originate from the same finger based on a sufficient disagreement between the two ridge patterns. An inconclusive decision is made when an examiner cannot make a decision of either individualization or exclusion due to insufficient ridge details or small corresponding area between latent and exemplar print [117].

1.4 Challenges in Fingerprint Recognition

Fingerprint recognition is regarded as one of the most reliable methods for human identification achieving a high level of matching accuracy [135] and throughput in the widespread deployment of AFIS in large-scale operational applications [7–9]. However, there exist some challenges in fingerprint recognition that need to be addressed: (i) demonstration of persistence and uniqueness of fingerprints to establish solid ground for fingerprint recognition, (ii) "lights-out" latent fingerprint identification, and (iii) ensuring that AFIS cannot be compromised by various types of attacks by adversary.

1.4.1 Persistence and Uniqueness of Fingerprints

The persistence and uniqueness of fingerprints—two fundamental properties of fingerprints enabling a high level of fingerprint recognition accuracy—have been presumed and accepted as facts without the support of any scientific basis. In situations where fingerprints are presented as forensic evidence in the courts of law, the linkage between a friction ridge evidence and the suspect is required to be supported by scientific methodology with a known error rate under the Daubert criteria [17].

Fingerprint uniqueness has been investigated based on statistical models of fingerprint features. These models provide either (i) a probability of random correspondence (PRC) [42, 105, 142] or (ii) a likelihood of a match in fingerprint comparison [101]. In [105], the probability that a template fingerprint with m minutiae has q matched minutiae with a query fingerprint with n minutiae was estimated under the restrictive assumptions that (i) fingerprint minutiae are uniformly distributed, (ii) the location and direction of a minutia are independent, and (iii) the correspondence of a minutiae pair is an independent event. A more realistic minutiae model was developed in [142] based on a mixture model that combines the distributions of minutiae location and direction; the PRC of two fingerprints was approximated by a Poisson distribution. While the studies in [105, 142] focused only on

minutiae models, a model incorporating fingerprint type (arch, loop, and whorl) and pores was also developed, which resulted in more realistic PRC values [42]. In a different approach to the uniqueness problem, the likelihood ratio of the comparison between a latent and an exemplar fingerprint was computed under the null hypothesis that the two impressions were made by the same finger. The likelihood ratio indicates the reliability of the comparison as forensic evidence [101]. That is, a small likelihood ratio value implies that the suspect identification based on the fingerprint pairing may be misleading. This model used the spatial configuration of minutiae, which is able to assess the latent-exemplar comparison with a small number of minutiae and account for skin distortion.

Fingerprint persistence, on the other hand, still remains as anecdotal evidence. To our knowledge, no large-scale studies on fingerprint persistence have been published. Perhaps, this is due to the difficulty in collecting a large-scale longitudinal fingerprint database. Galton [61] and Herschel [73] observed, from a set of genuine fingerprint pairs captured at two distinct time instances with a large interval of several decades, that the fingerprint ridge structure is stable over time. However, considering the current practice of fingerprint recognition that is based on AFIS, permanence of fingerprint ridge structure is not of the utmost interest. Instead, the temporal stability of fingerprint matching accuracy under variations in sensing modality, feature representation, and matching algorithms needs to be demonstrated to solidify the basis for fingerprint recognition.

1.4.2 "Lights-Out" Latent Identification

A "lights-out" fingerprint identification system refers to a system that requires only fingerprint images as input (query) and which returns a short list of exemplar prints as potential mates [79]. While virtually all AFIS for exemplar matching operate in "lights-out" identification mode, matching latent fingerprints in "lights-out" mode is still very challenging. In the 2012 ELFT-EFS [78], the best performing latent matcher achieved rank-1 identification rate of 63.4% in "lights-out" identification mode.

In the current practice of matching latent fingerprints, a large degree of human intervention is needed to mark up the features in the latents, including minutiae and extended features and verify the mate from the candidate list retrieved from a large exemplar database. Although the ACE-V methodology is widely accepted by the forensic community for latent print examination, the influence of human factors in the ACE-V procedure has raised concerns about their reliability and consistency. A noteworthy case is the erroneous identification of Brandon Mayfield as a suspect in the Madrid train bombing incident based on an incorrect match between Mayfield's exemplar fingerprint and the latent print captured at the bombing site [103, 104]. The National Research Council's report on limitations and recommendations of forensic science [100] pointed out two major shortcomings in the current forensic science discipline: (i) "lack of mandatory and enforceable standards" that can be globally referred to in crime labs and (ii) "unacceptable case backlogs in state and local crime labs which likely make it difficult for laboratories to provide strong evidence for prosecutions and avoid errors that could lead to imperfect justice". Along with the efforts to understand the human factors in latent fingerprint examination [99], standards and guidelines for latent examiners' practices have also been set up. As an example, the Science Working Group on Friction Ridge Analysis, Study and Technology (SWGFAST) published standards which define terminologies and establish the sufficiency level for decisions at each step of the ACE-V methodology to alleviate subjectivity involved in feature markups and decision makings among examiners [117].

Based on the guidelines in SWGFAST standard, latent examiners' practices have been evaluated from various aspects (e.g., reliability of decisions, degree of consensus and consistency of decisions) [74, 125, 126], mainly on two critical decisions that the examiners make in ACE-V methodology: (i) latent value determination in the analysis phase and (ii) latent individualization conclusion in the evaluation phase. This line of study indicates that a significant amount of subjectivity and variation exist in the practices of the ACE-V methodology.

The advantages of "lights-out" latent identification mode include: (i) alleviating subjectivity in latent print examination and (ii) increasing throughput of latent print matching. This will ease the burden on latent examiners, given their growing case workload. Hence, important research topics in latent matching include: (i) pre-processing of latents such as latent quality assessment, latent segmentation, and latent enhancement, and (ii) advanced matching algorithm based on minutiae and extended features. The FBI's Next Generation Identification (NGI) program [11] also specifies the improved latent processing services as one of the desirable tasks in advanced fingerprint identification technology.

1.4.3 Vulnerability of Automated Fingerprint Identification Systems

The widespread deployment of AFIS in law enforcement, international border crossing, and access control to secure facilities has prompted some individuals to engage in extreme measures for the purpose of circumventing the AFIS: (i) fingerprint spoofing and (ii) fingerprint obfuscation (or alteration). Fingerprint spoofing – the use of fake fingers made of glue, latex or silicone – is a well publicized method to adopt another person's identity. Figure 1.12 gives examples of fake fingerprints. Figure 1.12(a) shows a gummy fingerprint from a mold made by having a finger press down on the softened plastic and solidifying materials such as latex or silicone in the mold. Figure 1.12(b) shows a dummy finger made of silicone with detailed friction ridge pattern on the fingertip.

In order to detect attacks based on fake fingers, extensive research has been conducted in the biometrics literature to measure the liveness of a finger that have resulted in: (i) software solutions and (ii) hardware solutions. Software solutions utilize the conventional fingerprint sensing device, but adopt a series of image frames of the fingerprint to observe perspiration pattern [106] or skin distortion [32] of the finger. On the other hand, hardware solutions involve additional sensing devices to measure electrical resistance [118] or spectral characteristics [102] of the finger skin.
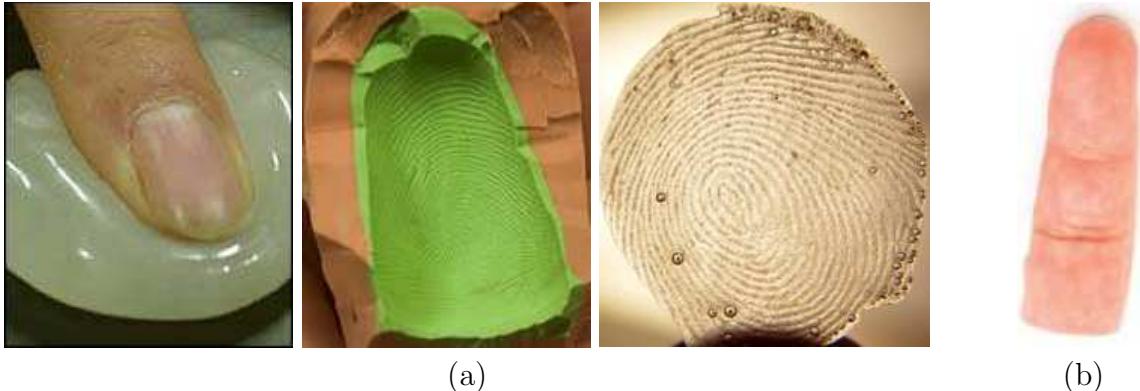
Figure 1.12: Fake fingerprints. (a) Gummy fingerprint: (Left) pressing a finger on the softened plastic [19, 94], (Center) the mold generated from a finger [21], (Right) a gummy fingerprint from the mold [4], and (b) dummy finger [5].

Unlike fake fingers, obfuscated or altered fingerprints[6] are real fingers whose ridge structure has been severely changed by abrading, cutting, burning, or performing plastic surgery on fingertips (see Figure 1.13). The purpose of fingerprint obfuscation is to conceal one's identity in order to evade AFIS. The problem of altered fingerprints has hitherto not been studied and there are no reported techniques to identify them. Although a number of algorithms and methods to measure the fingerprint image quality and to detect fake fingerprints have been proposed, none of them can be effectively used to detect altered fingerprints due to the following reasons. (i) Not all the altered fingerprints are of poor quality (see Figures 5.9 and 5.10) and (ii) the altered fingerprints are indeed from live fingers, so the fake fingerprint detectors which essentially measure the properties of the live fingers would not work.

To deal with the problem of attacks on AFIS by fingerprint obfuscation, the following research problems need to be investigated: (i) analyzing altered fingerprints, (ii) detecting altered fingerprints, and (iii) matching altered fingerprints to their pre-altered mates.

---

[6]Fingerprint *obfuscation* and fingerprint *alteration* will be used interchangeably in this dissertation.

Figure 1.13: Photographs of altered fingerprints. (a) Transplanted friction ridge skin from sole of the feet [6], (b) fingers that have been bitten [122], (c) fingers burnt by acid [67], and (d) stitched fingers [23].

1.5 Dissertation Contributions

The contributions of this dissertation are: (i) longitudinal analysis of genuine fingerprint match scores and identification decisions based on statistical model to demonstrate the persistence of fingerprint matching accuracy over time, (ii) global fingerprint orientation field modeling, (iii) latent fingerprint enhancement, and (iv) detection and matching of altered fingerprints.

- Objective of the longitudinal study of fingerprint recognition is to provide scientific basis for the stability of fingerprint matching accuracy over time. In order to tackle this problem, we first investigate how the time interval between two fingerprint impressions

obtained from the same finger at two different time instances accounts for the variation in genuine match scores. Based on a longitudinal fingerprint database provided by the Michigan State Police, we use a multilevel statistical model to infer the relationship between time interval and genuine match score. In order to assess the impact of time interval on genuine match scores, we compare it with other possible covariates including subject's age, gender, race, and fingerprint image quality. We find that (i) genuine match score tends to decrease as time interval between two fingerprints increases, (ii) although time interval and subject's age have significant influence on genuine match scores, fingerprint image quality is able to explain the variation in genuine match scores the best, (iii) subject's gender and race barely affect the fingerprint match scores, and (iv) true acceptance rate tends to be stable even if the time interval between two fingerprints in comparison increases.

- Global orientation field models are useful for estimating the orientation field in noisy fingerprint images. The models are also useful in representing the unique characteristics of fingerprint ridge flow patterns in a mathematical form. Global orientation field models reported in the literature integrate the properties of local field around singular points and arch-like global field. We develop a global orientation field model in the form of ordinary differential equations which represent x-derivative and y-derivative with a small number of polynomial terms and are constrained by the number of singular points[7]. The proposed model is used to determine if an arbitrary input image contains a fingerprint pattern. This property can be used to (i) check the validity (e.g., deviation from normal friction ridge pattern) of the input image to an AFIS, and (ii) investigate the integrity of fingerprint images in an exemplar database.

- The manual markup of features in latent fingerprints is not only a tedious job for latent examiners, but the markup features can also involve subjectivity (i.e., the feature

---

[7]Almost all the fingerprints belong to one of the following categories based on the number of singular points: (i) no singular point, (ii) one core and one delta, and (iii) two cores and two deltas.

sets marked by multiple examiners can be different). We propose a latent fingerprint enhancement algorithm which requires minimal markup (only the region of interest and singular points) to provide the enhanced latents to an AFIS that leads to more robust and consistent minutiae extraction and matching. The key component of the algorithm is the orientation field estimation in latents. For robust orientation field estimation, we use (i) a global orientation field model which decomposes a singular orientation field component with discontinuity around cores and deltas and a residual orientation field component which is continuous over the entire fingerprint region, and (ii) a randomized-RANSAC (Random Sample Consensus) algorithm to fit the model to latents.

- Fingerprint obfuscation is a serious attack that can compromise the performance and integrity of AFIS. However, unlike spoof fingerprints, fingerprint obfuscation has not received adequate attention. We analyze the patterns found in the altered fingerprints based on a large database of altered fingerprints provided by law enforcement agencies. We develop an algorithm to detect altered fingerprints based on the abnormality in orientation field and minutiae distribution that are observed in altered fingerprints. In order to improve the matching accuracy of altered fingerprints to their pre-altered mates, we propose an algorithm to restore the arrangement of local ridge patches, and evaluate the performance of AFIS when the altered region in an input fingerprint is restored or masked out.

Chapter 2

Longitudinal Study of Fingerprint Recognition

2.1   Introduction

Friction ridge skin on fingers and palms has been purportedly known to be a physical characteristic of an individual that does not change over time (i.e., persistence or permanence of friction ridge pattern) and can be used as a person's "seal" or "signature" (i.e., uniqueness or individuality of ridge pattern). Pioneering work of Herschel [73], Henry [71], Galton [61], and Faulds [51] established the way for recognizing humans based on their fingerprints. Starting with the first known case where the fingerprints found at the crime scene in Argentina in 1893 were officially accepted as an evidence to convict a suspect [69], friction ridge analysis has become one of the most crucial methods in crime scene investigations worldwide. The decision made in *Frye v. United States* in 1923 [16] is widely cited as the basis for the admissibility of forensic evidence, including friction ridge pattern; *Frye* standard states that a scientific principle or discovery which has gained a *general acceptance* in the relevant field is admissible in the courts.

In *Daubert v. Merrell Dow Pharmaceuticals, Inc.* in 1993 [17], however, the general

28

acceptance test of *Frye* was superseded by the Federal Rules of Evidence (FRE). The *Daubert* ruling established a guideline for admitting forensic evidence which consists of the following factors: (i) empirical testing, (ii) peer review and publication, (iii) the known or potential error rate, (iv) standards controlling the operation, and (v) the *Frye* standard of general acceptance. The *Daubert* standard provoked challenges to admissibility of friction ridge evidence in the courts. Although all of about 40 such challenges resulted in a decision that friction ridge analysis is acceptable as forensic evidence, the *Daubert* case highlighted a lack of scientific basis of persistence and uniqueness that support the human identification by friction ridge patterns and standards that can be universally referred to in friction ridge analysis.

The *Daubert* ruling led to the development of standards and guidelines for friction ridge analysis [117] and retraining of latent examiners [97], along with a body of research to demonstrate the foundations of fingerprint identification, i.e., uniqueness and persistence of friction ridge patterns. While a significant amount of effort has been made to study the uniqueness of fingerprints by (i) estimating the probability of a random correspondence (i.e., two randomly selected fingerprints will be sufficiently similar to be claimed as genuine mates) [42,105,142] or (ii) measuring the evidential value of latent fingerprint comparisons [101], no systematic study on persistence of fingerprints has been reported.

Early studies on persistence of fingerprints focused on demonstrating the invariance of ridge structure in the fingerprints with respect to time. As already shown in Figure 1.2, Herschel collected three fingerprints of his son when he was 7, 17, and 40 years old, and verified that all the ridge details in the three fingerprints did not change over time [73]. Galton collected 11 pairs of fingerprints from six different individuals at two different time instances [61]. The time interval between a pair of fingerprints in Galton's collection ranged from 11 years to 31 years. The six subjects in his study were selected from different age groups; the age at the second impression was as young as 15 years old and as old as 79 years old. Among the 389 minutiae pairs that were manually labeled by Galton, only a
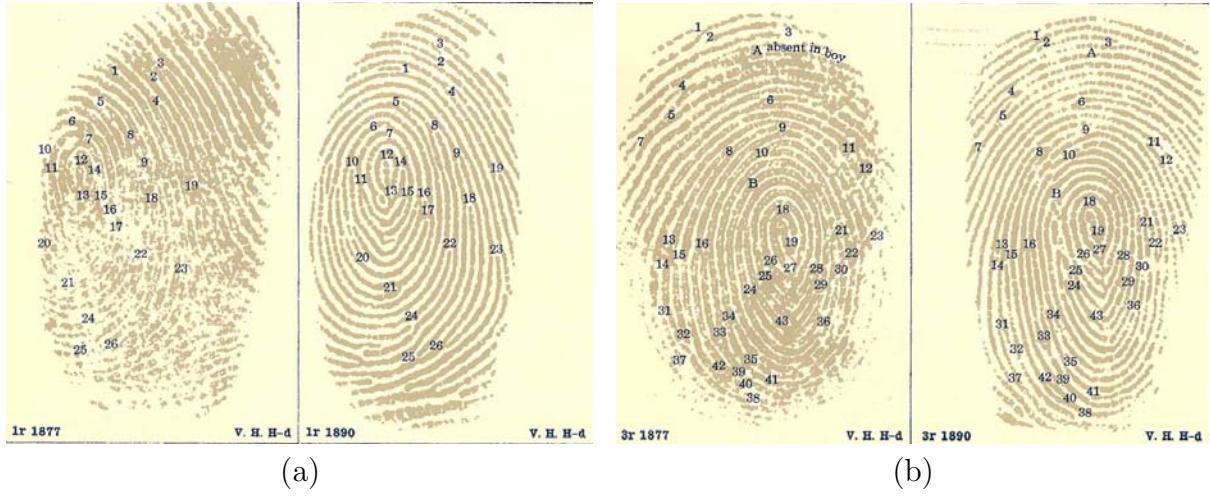
Figure 2.1: Fingerprint pairs with minutiae correspondences labeled by Galton in his study on fingerprint persistence [61]. (a) A pair of fingerprints with 13-year time interval showing perfect minutiae correspondences, and (b) a pair of fingerprints from another finger of the same subject with one minutia missing in the later age impression (denoted as 'A').

single minutia was missing in a fingerprint pair (see Figure 2.1). In addition to these case studies, the temporal invariance of fingerprint patterns has been explained by the anatomical structure of friction ridge skin; the ridge pattern formed in the inner (dermal) layer during gestation remains unchanged with the protection of the outer (epidermal) layer [46].

Fingerprint recognition technology has made great strides since the Federal Bureau of Investigation (FBI) initiated Automated Fingerprint Identification Systems (AFIS) in 1970's [95]. Since then, there has been a phenomenal increase in both the size of the fingerprint databases held by various law enforcement agencies as well as the speed at which they can be searched while maintaining high matching accuracy. Considering that the fingerprint recognition today essentially involves AFIS searching large-scale databases, the objective of research on fingerprint persistence should be to establish the temporal invariance of *fingerprint recognition accuracy*, rather than the permanency of the friction ridge pattern itself. The 2009 National Research Council (NRC) report [100] pointed out that

> *"Uniqueness and persistence are necessary conditions for friction ridge identifi-*
> *cation to be feasible, but those conditions do not imply that anyone can reliably*

discern whether or not two friction ridge impressions were made by the same person. Uniqueness does not guarantee that prints from two different people are always sufficiently different that they cannot be confused, or that two impressions made by the same finger will also be sufficiently similar to be discerned as coming from the same source."

This means that the representation and extraction of invariant and discriminative features for fingerprint identification may be deficient in a collection of fingerprint impressions captured under various conditions. Furthermore, fingerprint comparison algorithms may be imperfect to model the similarity among genuine fingerprints and the dissimilarity among impostor fingerprints and separate the two groups [80].

Fingerprint recognition exhibits two types of comparison errors: (i) false rejection: falsely determining two impressions from the same finger (a genuine fingerprint pair) as a non-match due to large *intra-class* variability, and (ii) false acceptance: falsely determining two different fingerprint impressions from two distinct fingers (an impostor fingerprint pair) as a match due to large *inter-class* similarity. The intra-class variability is observed due to changes in intrinsic skin condition (e.g., finger skin dryness, distortion), and changes in sensing method (inked or live-scan fingerprints) or sensing technology (known as the *interoperability* problem [12]). On the other hand, the inter-class similarity can be observed when the fingerprint impressions from two distinct fingers partially coincide. A notable case of false recognition by friction ridge analysis in forensic investigation is the erroneous identification of Brandon Mayfield as a suspect in the Madrid train bombing based on an incorrect match between Mayfield's exemplar fingerprint and the latent print captured at the bombing site [103].

In the biometrics literature, a phenomenon called *template aging* has been reported, which refers to the increase in the system error rate with respect to the time gap between the query and the template (or reference) [92]. A study comparing groups of fingerprint pairs with respect to time gap reported that the fingerprint comparisons with less than a 5-year

time gap show lower error rate than comparisons with a larger time gap [53]. Further, a scheme to mitigate this template aging effect has been proposed by updating the templates stored in the system with newly acquired data [127].

In order to determine the tendency of biometrics recognition accuracy with respect to time interval between acquisitions of a biometric trait, we need to (i) collect longitudinal biometrics data consisting of multiple acquisitions of a biometric trait from the sampled population over a reasonably long period of time, and (ii) conduct an appropriate analysis, considering the characteristics of the longitudinal data. If the longitudinal data is *balanced* and *time structured*[1], cross-sectional analysis can be readily applied to longitudinal data by grouping the data according to cohort (for example, short-term and long-term comparison groups) under the assumption of compound symmetry[2]. In reality, however, it is often not feasible to collect longitudinal biometrics data by following an identical measurement schedule over a target sample satisfying the compound symmetry. To handle the unbalanced and time-unstructured longitudinal data, several statistical models have been developed: linear mixed-effects model [85], multilevel statistical model [62], hierarchical linear models [112], and random coefficient regression [70]. Despite their different names and origins, these models are essentially identical.

A number of published studies on template aging in major biometrics modalities (e.g., [53] on fingerprint recognition, [56, 57] on iris recognition, and [86] on face recognition) made incorrect use of cross-sectional analysis on unbalanced and time-unstructured longitudinal data. On the other hand, the longitudinal study on iris recognition by the National Institute of Standards and Technology (NIST) [64] correctly used a nonlinear mixed-effects model to

---

[1]A longitudinal dataset is characterized by (i) the number of measurements per individual and (ii) the time schedule used to make the measurements [121]. *Balanced* dataset means that every subject has the same number of measurements. *Time-structured* dataset consists of the repeated measurements following an identical time schedule across individuals. The sequence of measurements for each individual can be spaced either regularly or irregularly.

[2]The compound symmetry requires (i) homoscedasticity of variance: the variance of the measurements at a time instance across all subjects is the same as that of the measurements at another time instance, and (ii) constant covariance: the correlation between the measurements at the first and second time instances, for example, is the same as that between the measurements at the first and third time instances, and so on.

show the relationship between genuine iris match scores and covariates such as time elapsed after enrollment and the difference in iris dilation. Despite the use of an appropriate modeling approach, the NIST study suffers from the following drawbacks: (i) the dataset used was truncated in the sense that the iris match scores from falsely rejected genuine comparisons were not included, and (ii) the validity of some of the assumptions in the mixed-effects model (i.e., normality of residuals and random effects) was not verified. The truncated data is problematic because the truncated portion of the data—erroneous identification decisions— is the very target of the analysis to determine the tendency of error rate with respect to time. Also, only if all the model assumptions are satisfied, the analysis and inferences from the data can be accepted.

We demonstrate whether the fingerprint matching accuracy changes over time, namely the time interval between two fingerprint impressions being compared. For this purpose, we obtained a longitudinal database of fingerprints collected from 15,597 subjects booked by the Michigan State Police over at least a 5-year time span. A multilevel statistical model [62, 121] is used to understand temporal tendency of fingerprint recognition performance in the longitudinal dataset which is unbalanced and time unstructured. More specifically, we address the following two questions to statistically demonstrate the persistence of fingerprints from the perspective of AFIS matching accuracy:

1. *What is the trend of genuine match scores over time?* It has been empirically noticed that match scores of genuine fingerprint pairs with a large time gap tend to decrease, possibly due to aging of finger skin and higher chance to have worn out or altered fingerprints. To examine the relationship of time interval between two fingerprints being compared with its genuine match score, a linear 2-level model with time interval as a covariate is fit to the genuine match scores generated by commercial matchers. Given the parameter estimates and confidence intervals, the null hypothesis "the slope of the linear model is zero" is tested.

2. *What is the trend of matching accuracy (or identification decisions) with respect to the*

time interval between fingerprints? The ultimate question of interest is whether or not the genuine match scores remain significantly higher than the decision threshold in the presence of time lapse between impressions being compared. This problem is formulated by making a binary identification decision on each genuine pair with a predetermined decision threshold and fitting a linear 2-level model to the binary decisions. The probability of true acceptance (i.e., a genuine fingerprint pair is correctly determined as genuine), for a predetermined threshold, with respect to time interval is estimated.

The rest of the chapter is organized as follows. Section 2.2 reviews the traditional and recent approaches for longitudinal data analysis, and introduce the multilevel statistical model used in this study. Section 2.3 explains the longitudinal fingerprint database analyzed in this study and gives the overview of longitudinal analysis of the dataset with multilevel model. In Section 2.4, the match scores of genuine fingerprint pairs are modeled by linear 2-level models with various covariates, including time interval, subject's age, and fingerprint image quality. The models with salient covariates are selected based on the evaluation of goodness-of-fit. After the underlying model assumptions are validated, the tendency of genuine match scores with respect to the selected covariates is investigated based on the parameter estimates and their confidence intervals. In Section 2.5, a linear 2-level model is fit to the binary decisions made on genuine fingerprint pairs by applying a decision threshold to determine whether the true acceptance rate is stable over time. Finally, Section 2.6 summarizes the study and suggests future research directions.

2.2 Longitudinal Data Analysis

Longitudinal data refers to repeated measurements on a collection of individuals sampled from a population over time. This is in contrast to cross-sectional data where a single measurement is made on each individual [62]. The approach for analyzing the longitudinal data
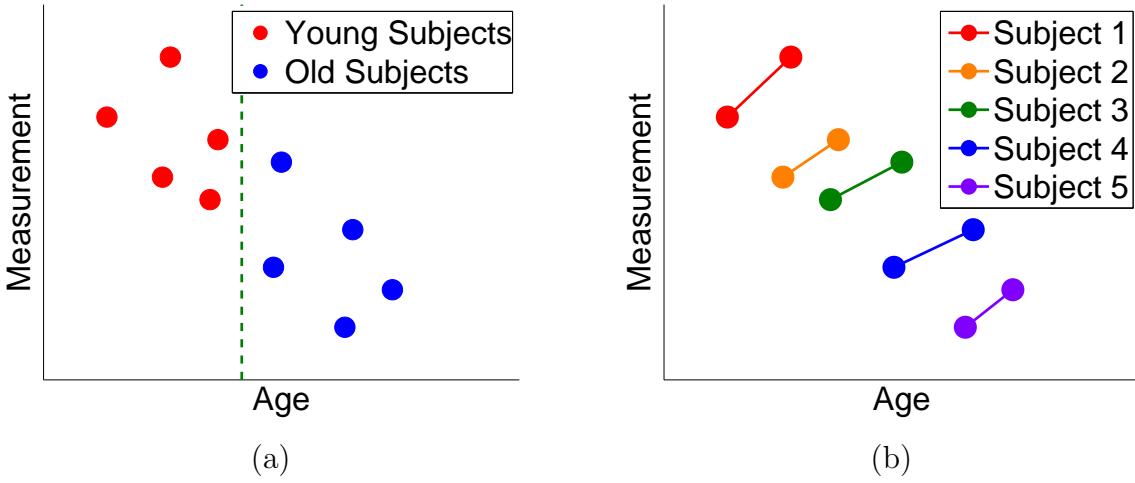
Figure 2.2: Cross-sectional analysis versus longitudinal analysis of balanced but time-unstructured longitudinal data (adapted from [49]). For this dataset (two measurements for each of 5 subjects), (a) the cross-sectional analysis that discards subject labels on data makes an inference that *the measurement values tend to decrease with respect to subject's age*, while (b) the longitudinal analysis interprets the data as *the measurement values tend to increase with respect to age*.

differs according to the data characteristics. Early models for longitudinal data analysis were derived from cross-sectional analysis, in particular *analysis of variance* (ANOVA) [113], assuming that (i) the longitudinal dataset is balanced and time structured (i.e., n repeated measurements for each one of the N individuals following an identical measurement schedule), and (ii) the dataset satisfies the compound symmetry. Figure 2.2 shows a hypothetical example that, if the dataset is unbalanced and/or time unstructured, the cross-sectional analysis makes incorrect inference against the actual longitudinal behavior.

In this section, we will review the models based on ANOVA for longitudinal data analysis, and then describe the details of multilevel statistical model which is used for our longitudinal study of fingerprint recognition.

2.2.1 ANOVA-based Approach

ANOVA is widely used to compare multiple groups, where each group consists of cross-sectional data. For example, consider an experiment to demonstrate the effectiveness of a

medicine for a disease. A simple experimental design would be to construct two groups of test subjects: one group is treated by the medicine, while the other group, a control group, is not treated by it. A popular null hypothesis tested in ANOVA is that the group means are equal. By rejecting the null hypothesis, we can claim that the medicine makes a significant difference in the treatment of the disease against the control group.

Although the traditional way of conducting ANOVA test is to use an F-statistic, a general linear model (GLM) is preferred to describe ANOVA due to the conceptual and practical advantages [116]. In the GLM framework, multiple concepts, such as regression, ANOVA, and analysis of covariance (ANCOVA), can be easily linked, and one can develop a unified methodology for statistical analysis under different circumstances. Also, the GLM makes it easy to test the underlying assumptions for ANOVA and ANCOVA.

In the GLM framework, ANOVA model can be described as follows:

$$y_{ij} = \mu + b_i + \varepsilon_{ij}, \text{ for } i = 1, \dots, N, \text{ and } j = 1, \dots, n, \tag{2.1}$$

where y_{ij} is the observed response (or measurement) for the j-th sample in group i, μ is the grand mean for the entire population, b_i is the random variable for group i, and the error ε_{ij} is normally distributed ($\varepsilon_{ij} \sim \mathcal{N}(0, \sigma_e^2)$).

This model can be viewed as either a fixed-effects model or a random-effects model according to the question of interest and the experimental design. If an investigator gives different treatments to each of N groups and wants to determine if there is any significant difference among treatments, for instance, the following null hypothesis can be tested: $b_1 = b_2 = \dots = b_N = 0$ (i.e., all N group means are the same). As the treatments applied to each group are *fixed*, this is called a fixed-effects ANOVA. On the other hand, one can view the treatment as a random variable sampled from a larger population of possible treatments. Then the objective of statistical analysis will be to determine to what extent the treatments account for the variance in the observed responses. This is called a random-effects ANOVA.

Under the assumptions, $b_i \sim \mathcal{N}(0, \sigma_b^2)$ and $y_{ij} \sim \mathcal{N}(\mu, \sigma_b^2 + \sigma_e^2)$, the primary goals of random-effects ANOVA include (i) testing the null hypothesis: $\sigma_b^2 = 0$ (i.e., the variance in treatments does not have any impact on the variance in the observed responses), or (ii) computing intra-class correlation coefficient (ICC) that is defined as:

$$ICC = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_e^2}. \tag{2.2}$$

In many situations, both fixed and random effects exist in the data, or it is difficult to distinguish between them. A univariate repeated-measures ANOVA (also called rANOVA or mixed-effects ANOVA) includes both fixed and random effects in the model [59]. In a longitudinal dataset, the j-th measurement of subject i, y_{ij}, is made along with p explanatory variables $(x_{ijk}, k = 1, \ldots, p$; also referred to as covariates or predictors) observed at the time of measurement. Then, a univariate repeated-measures ANOVA is described as:

$$y_{ij} = \beta_1 x_{ij1} + \beta_2 x_{ij2} + \cdots + \beta_p x_{ijp} + b_i + \varepsilon_{ij},$$
$$= \mathbf{x}_{ij}^T \boldsymbol{\beta} + b_i + \varepsilon_{ij}, \text{ for } i = 1, \ldots, N, \text{ and } j = 1, \ldots, n, \tag{2.3}$$

where $\mathbf{x}_{ij}$ is a design vector consisting of the p explanatory variables, $\boldsymbol{\beta}$ is a regression parameter vector, $b_i \sim \mathcal{N}(0, \sigma_b^2)$, and $\varepsilon_{ij} \sim \mathcal{N}(0, \sigma_e^2)$.

2.2.2 Multilevel Statistical Model

A multilevel model is primarily used to analyze hierarchically structured, nested, or clustered data [62]. The model, for example, will be appropriate if one is interested in the factors that affect the performance of students from multiple schools in a mathematics examination. Consider N schools with school i having n_i students. Let y_{ij} denote the test score of student j in school i. The explanatory variables of the observed responses can be either (i) time-variant predictor (e.g., student's age at the time of taking test) or (ii) time-invariant predictor

(e.g., student's gender).

Longitudinal data can be viewed as hierarchically structured data in the following sense. The first level (lower level) of the model consists of the repeated measurements taken from each individual, and the level-1 model is regressed to the data of the individual and accounts for the intra-subject changes. In the level-2 model (higher level), the population-averaged tendency and deviations of individuals are modeled to account for the inter-subject changes. The multilevel model is applicable for unbalanced and time-unstructured longitudinal data and does not require compound symmetry.

A simple 2-level model with a single covariate can be represented by:

$$\text{Level-1 Model: (Intra-subject change)} \quad y_{ij} = \varphi_{0i} + \varphi_{1i}x_{ij} + \varepsilon_{ij},$$

$$\varepsilon_{ij} \sim \mathcal{N}(0, \sigma_e^2), \tag{2.4}$$

$$\text{Level-2 Model: (Inter-subject change)} \quad \varphi_{0i} = \beta_{00} + b_{0i},$$

$$\varphi_{1i} = \beta_{10} + b_{1i},$$

$$\begin{bmatrix} b_{0i} \\ b_{1i} \end{bmatrix} = \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{10} & \sigma_1^2 \end{bmatrix} \right). \tag{2.5}$$

In the level-1 model, the variables and parameters are defined as follows:

- y_{ij} is the j-th measurement (observed response) for subject i at the time instance x_{ij}, for $i = 1, \ldots, N$, and $j = 1, \ldots, n_i$.

- φ_{0i} is the true parameter representing the intercept of the linear model for subject i.

- φ_{1i} is the true parameter representing the slope of the linear model for subject i.

- ε_{ij} is the error in the j-th measurement of subject i from the model fit. The error is assumed to be normally distributed with a zero mean and a variance of σ_ε^2.

In the level-2 model, the true parameters, φ_{0i} and φ_{1i}, can be modeled by a mixture of fixed and random effects.

- Parameters β_{00} and β_{10} are the fixed effects of the model, representing the grand means of intercept and slope across all N subjects in the data, respectively.

- Parameters b_{0i} and b_{1i} are the random effects of the model, representing the deviations of subject i's intercept and slope from β_{00} and β_{10}, respectively. The random effects are assumed to follow a Gaussian distribution.

Any covariates other than time instance can be introduced in the model. For instance, if we are interested in whether or not gender (time-invariant predictor) is a factor that explains differences in students' performance in math exam, the level-2 model can be updated in the following way:

$$\varphi_{0i} = \beta_{00} + \beta_{01}GENDER_i + b_{0i},$$
$$\varphi_{1i} = \beta_{10} + \beta_{11}GENDER_i + b_{1i}. \tag{2.6}$$

If the predictor of interest is time varying, it can be introduced in the level-1 model (for example, the area of residency of the student at the time of taking exam):

$$y_{ij} = \varphi_{0i} + \varphi_{1i}x_{ij} + \varphi_{2i}AREA_{ij} + \varepsilon_{ij}. \tag{2.7}$$

Note that, depending on the nature of the covariate, the model parameter associated with the covariate can have only fixed effects (i.e., $\varphi_{2i} = \beta_{20}$) or a mixture of fixed and random effects (i.e., $\varphi_{2i} = \beta_{20} + b_{2i}$).

The multilevel model is often represented in its composite form [121]. The composite form of the 2-level model in Equation (2.4) is:

$$y_{ij} = \varphi_{0i} + \varphi_{1i}x_{ij} + \varepsilon_{ij}$$
$$= [\beta_{00} + \beta_{10}x_{ij}] + [b_{0i} + b_{1i}x_{ij} + \varepsilon_{ij}]. \tag{2.8}$$

The first part of the composite model is called the *structural* component which contains only the fixed effects of the model. The second part of the composite model is called the *stochastic* component consisting of random effects [121]. The stochastic component shows that the compound symmetry assumption is not required in the model. For example, the term $b_{1i}x_{ij}$ indicates the heteroscedasticity of variance in the observed responses; the variance can vary with respect to time instance x_{ij} for the j-th measurement of subject i. In addition, the term b_{0i} allows the residuals to be autocorrelated; that is, the residuals within a subject can be linked across time. The composite form of the multilevel model shows that it is equivalent to *linear mixed-effects model.*

Equation (2.8) can be also represented in a matrix form:

$$
\begin{aligned}
y_{ij} &= (\beta_{00} + \beta_{10}x_{ij}) + (b_{0i} + b_{1i}x_{ij} + \varepsilon_{ij}) \\
&= \begin{bmatrix} 1 & x_{ij} \end{bmatrix} \begin{bmatrix} \beta_{00} \\ \beta_{10} \end{bmatrix} + \begin{bmatrix} 1 & x_{ij} \end{bmatrix} \begin{bmatrix} b_{0i} \\ b_{1i} \end{bmatrix} + \varepsilon_{ij} \\
&= \mathbf{x}_{ij}^{T}\boldsymbol{\beta} + \mathbf{z}_{ij}^{T}\mathbf{b}_{i} + \varepsilon_{ij}.
\end{aligned}
\tag{2.9}
$$

Hence,

$$
\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{b} + \boldsymbol{\varepsilon}, \tag{2.10}
$$

where $\mathbf{X}$ is a design matrix for fixed effects, $\mathbf{Z}$ is a design matrix for random effects, $\boldsymbol{\beta}$ is a fixed-effects coefficient vector, $\mathbf{b}$ is a random-effects coefficient vector, and $\boldsymbol{\varepsilon}$ is a residual vector. Note that, although $\mathbf{X}$ and $\mathbf{Z}$ are the same in Equation (2.9), they can be different in general when some of the covariates do not have random effects (thus, the components of the design vector in $\mathbf{Z}$ are a subset of those in $\mathbf{X}$).

2.2.2.1 Parameter Estimation Method

The maximum likelihood (ML) and generalized least-squares (GLS)[3] estimations are widely used to obtain parameters in the multilevel model [121]. The estimation based on ML requires that the residuals be normally distributed; in this case, GLS estimates are the same as ML estimates. The normality assumption is useful since, if the sample size is large, ML estimates are (i) asymptotically unbiased (converging to true population parameters), (ii) asymptotically normally distributed, and (iii) asymptotically efficient (smaller standard errors than those derived by other methods).

Under the normality assumption ($\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega})$), the estimate for parameter $\boldsymbol{\beta}$ in Equation (2.10) using GLS is given by [62]:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \boldsymbol{\Omega}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \boldsymbol{\Omega}^{-1} \mathbf{y}. \tag{2.11}$$

As the true error covariance $\boldsymbol{\Omega}$ is unknown, the GLS estimate is obtained by a two-stage approach:

1. Initialize $\boldsymbol{\beta}$, and estimate $\boldsymbol{\Omega}$ using the residuals from the model with the initial parameter value. The initial parameter value can be the OLS estimate of the composite model.

2. Refit the composite model using GLS with the estimated $\boldsymbol{\Omega}$, and then find $\boldsymbol{\beta}$.

The above two-stage scheme can be repeated to obtain a better fit by replacing the initial estimate of parameter value in step 1 with the updated parameter value in step 2. This is called an iterative GLS.

---

[3]Note that ordinary least-squares (OLS) estimation cannot be used to estimate the parameters in the multilevel model since the OLS assumes that the error terms are uncorrelated (recall that the multilevel model allows them to be autocorrelated). Although the estimates of fixed-effects parameters using OLS can be very close to the GLS results, the standard errors tend to be underestimated. This can result in incorrect inferences from data, or wrong conclusions made by the statistical hypothesis test.

2.3 Multilevel Model for Longitudinal Fingerprint Data

We introduce the longitudinal fingerprint data used in the study and give an overview of analyzing the dataset with multilevel models towards understanding the temporal tendency of genuine match scores and true acceptance decisions.

The following notations are consistently used throughout the paper.

- N: Total number of subjects in the longitudinal fingerprint database.

- n_i: Number of different tenprint acquisitions (cards or impressions) of subject i; $i = 1, \ldots, N$.

- $T_{i,j}$: Time stamp of the j-th tenprint impression of subject i; $i = 1, \ldots, N$, and $j = 1, \ldots, n_i$.

- The tenprint impressions of each subject are ordered according to the time sequence; a set of tenprints of subject i is labeled as follows: $\mathcal{F}_i = \{F_{i,1}, \cdots, F_{i,n_i}\}$, such that $T_{i,1} < \ldots < T_{i,n_i}$.

- $y_{i,jk}$: Genuine match score between the j-th and k-th fingerprint impressions of subject i.

- $\triangle T_{i,jk}$: Time interval between fingerprint impressions j and k; $\triangle T_{i,jk} = T_{i,k} - T_{i,j}$, for $j < k$.

- $AGE_{i,jk}$: Age of the subject i when his k-th tenprint impression was made, where $T_{i,j} < T_{i,k}$.

- $Q_{i,j}$: Quality value of fingerprint impression j of subject i. In this study, we use NIST Fingerprint Image Quality (NFIQ) measure [124], which assigns one of the five discrete values ranged from 1 (the highest quality) to 5 (the lowest quality), to define fingerprint image quality.

June 2001	July 2002	April 2003

September 2007	March 2008	October 2008
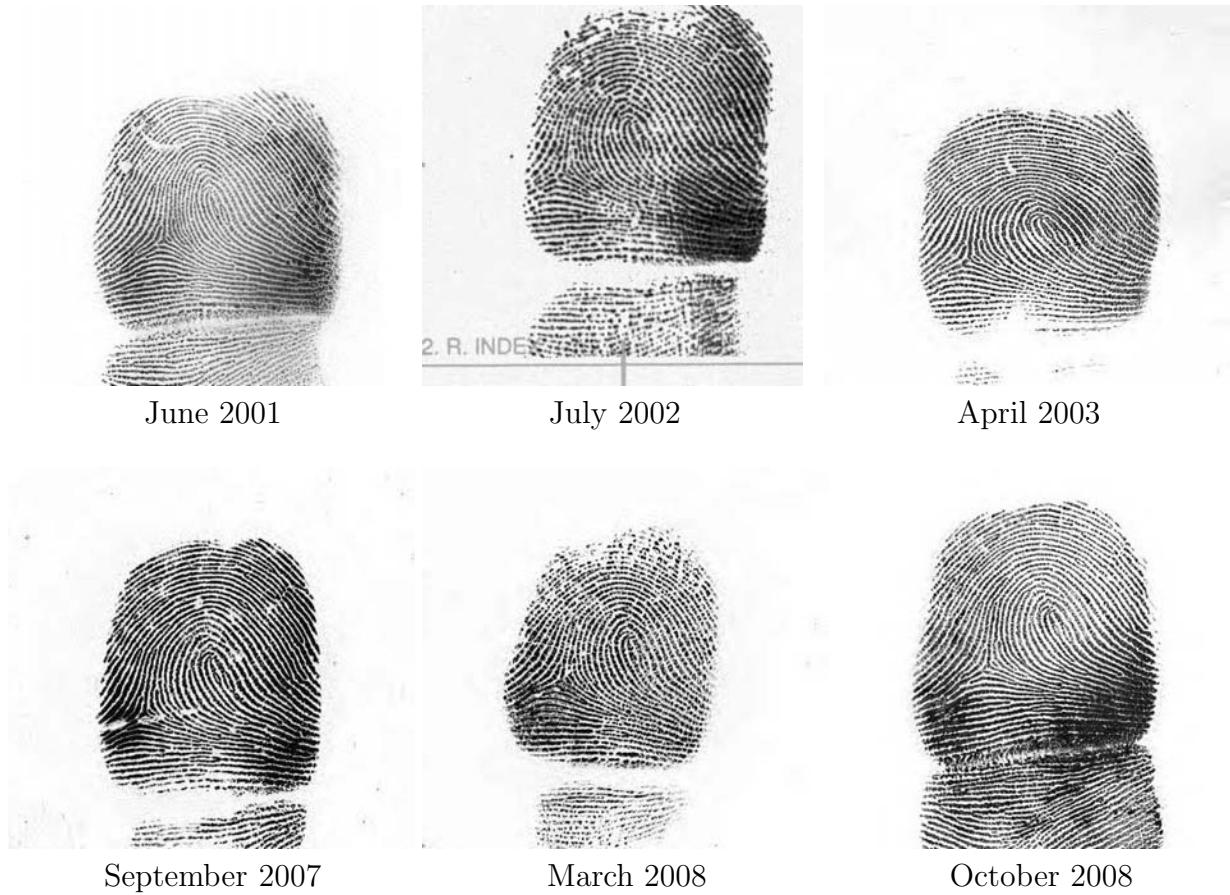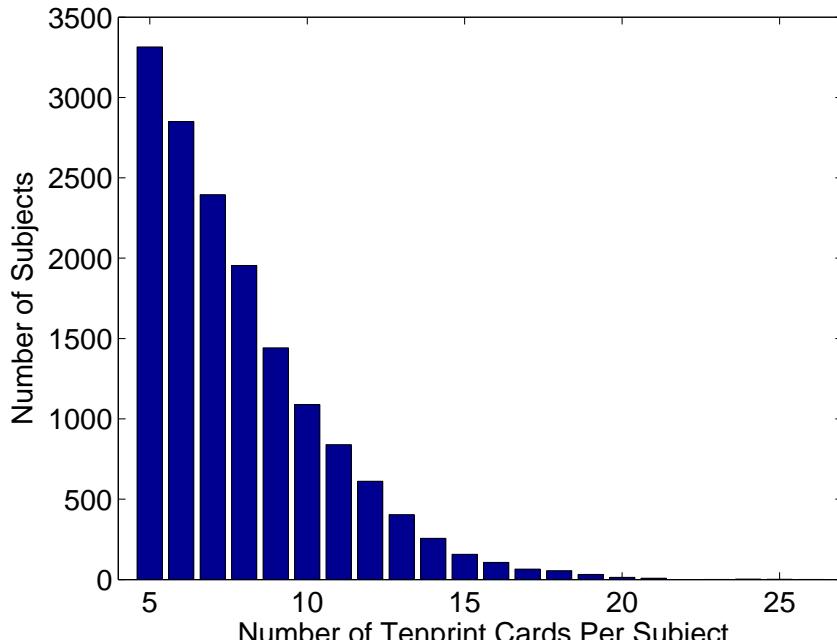
Figure 2.3: Six different impressions of the right index finger of a subject in the longitudinal fingerprint database.
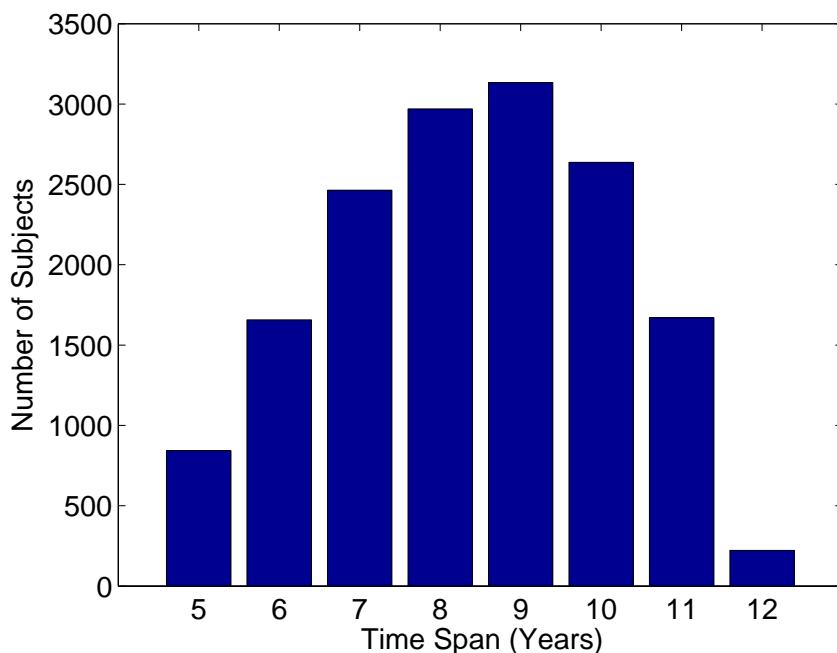
- $Q_{i,jk}$: The value corresponding to the lower of the j-th fingerprint and the k-the fingerprint qualities of subject i; according to the definition of NFIQ, $Q_{i,jk} = \max(Q_{i,j}, Q_{i,k})$.

- $bMALE_i$: A binary indicator for gender of subject i; 1 for male and 0 for female.

- $bWHITE_i$: A binary indicator for race of subject i; 1 for whites, and otherwise 0.

2.3.1   Longitudinal Fingerprint Database

A longitudinal database of fingerprint impressions was collected from the records of repeat offenders booked by the Michigan State Police. Figure 2.3 shows an example of six fingerprint impressions of the right index finger of a subject in the database acquired between June 2001 and October 2008.
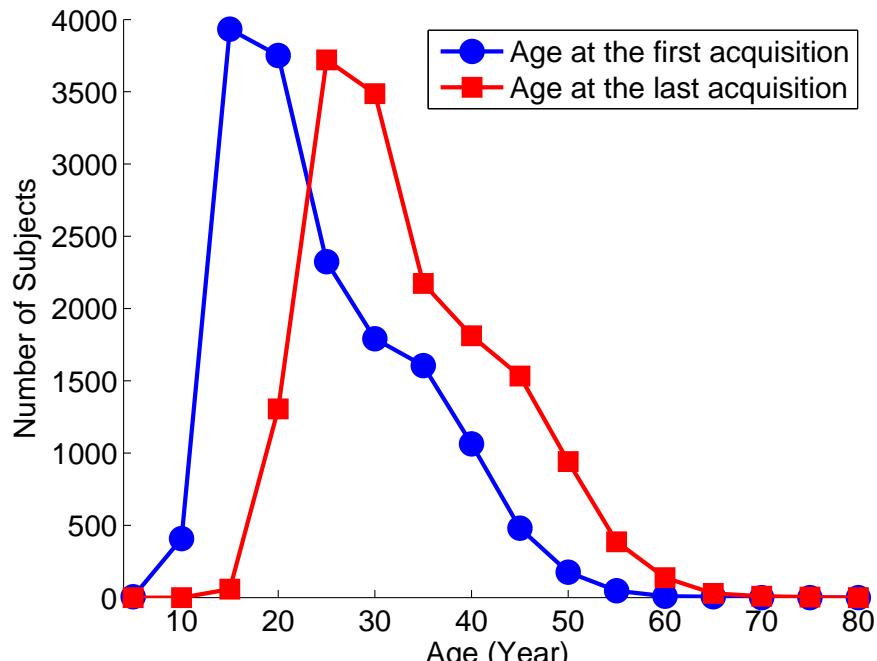
(a)
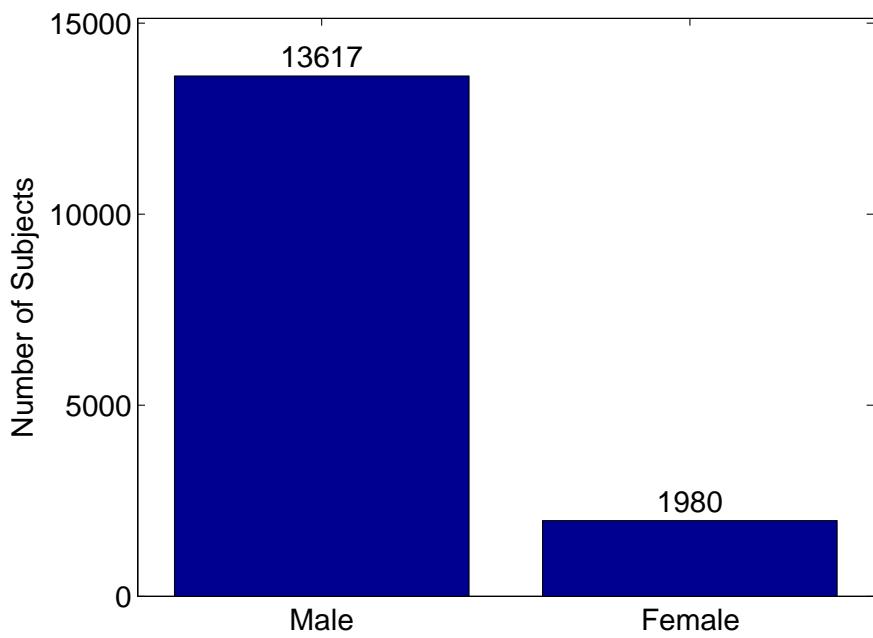
(b)

Figure 2.4: Statistics of the longitudinal fingerprint database used in this study. Histograms of (a) the number of tenprint cards per subject, (b) time span of data collection for a subject, (c) age at the first and last tenprint acquisitions, (d) gender, and (e) race.

(c)

(d)

Figure 2.4: (cont'd)

(e)
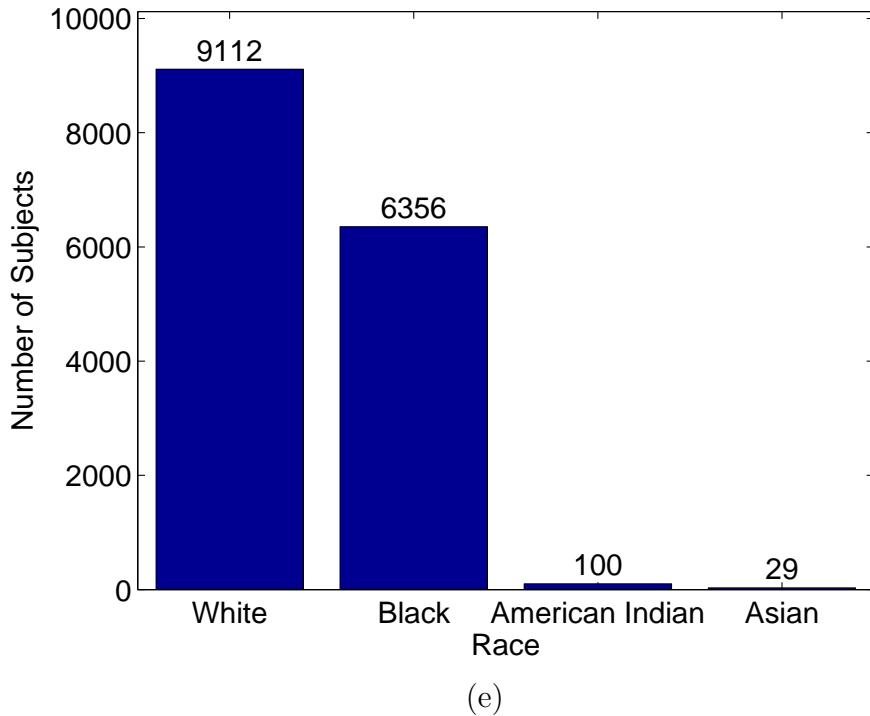
In constructing this database, we ensured that each subject had at least 5 tenprint impressions over a minimum of 5-year time span. A total of 15,597 subjects were identified whose fingerprint data met this requirement, and the fingerprints from these subjects were used for the longitudinal analysis. A summary of the fingerprint database is as follows:

- Each of the 15,597 subjects has at least 5 tenprint impressions. The average number of tenprint impressions per subject is 8, with the maximum of 26 cards for one of the subjects. The histogram of the number of tenprint impressions per subject is shown in Figure 2.4(a).

- The tenprint impressions of a subject have a minimum of 5-year time span (the time difference between the first and the last fingerprint acquisitions of a subject); that is, $\triangle T_{i,1n_i} \geq 5$ years for $i = 1, \ldots, N$. The average time span is 9 years, and the maximum time span in the database is 12 years. The histogram of time span for longitudinal

tenprint acquisitions for subjects is shown in Figure 2.4(b).

- Any two consecutive tenprint impressions of a subject are obtained with at least 2-month time gap; $(T_{i,j+1} - T_{i,j}) \geq 2$ months.

- Along with tenprint images, the following demographic information is also available for each subject:

  - Gender: Male or female

  - Race: White/Hispanic, Black, American Indian/Eskimo, or Asian/Pacific Islander

  - Age at the time of tenprint acquisition: The youngest subject at the time of the first impression is 8 years; the oldest subject at the time of the last impression is 78 years.

  The statistics of the demographics of subjects in the database are given in Figures 2.4(c)–(e).

Two commercial off-the-shelf (COTS) matchers (denoted as COTS-1 and COTS-2) are used to compute fingerprint match scores. For subject i with n_i fingerprint impressions, we conduct all pairwise comparisons; that is, $\binom{n_i}{2}$ genuine match scores are generated from each matcher. This is because law enforcement agencies store all the tenprint records for every booked subject and compare a query fingerprint to all the records in the database. The genuine match scores from each COTS matcher are normalized such that they have a mean of 0 and a standard deviation of 1; the normalized genuine match score $\tilde{y}_{i,jk}$ is obtained by:

$$\tilde{y}_{i,jk} = \frac{y_{i,jk} - \mu}{\sigma}, \tag{2.12}$$

where μ and σ are the mean and standard deviation of $\{y_{i,jk}\}$, respectively.

2.3.2 Overview

A fingerprint comparison essentially involves two fingerprint impressions to generate a single match score. The match score is then used to determine whether or not the two impressions are from the same finger; if the match score is greater than a predetermined threshold[4], a binary identification decision of genuine match is made. Since the impostor score distributions are considered to be stable under various acquisition conditions and time lapse, the decision threshold is typically fixed in advance of system deployment.

With multilevel modeling, this study targets at analyzing the following observed responses:

- Genuine match score of the j-th and k-th fingerprint impressions of subject i ($y_{i,jk}$)

- Binary identification decision of $y_{i,jk}$ by applying a decision threshold (Th) as follows.

$$y_{i,jk}^* = \begin{cases} 1 & \text{if } y_{i,jk} > Th \\ 0 & \text{otherwise.} \end{cases} \tag{2.13}$$

The covariates investigated in the study include:

- $\triangle T_{i,jk}$: Time interval between fingerprint impressions being compared

- $AGE_{i,jk}$: Subject's age at the latter fingerprint impression between two fingerprints being compared

- $Q_{i,jk}$: The lower quality value between two fingerprints being compared

- $bMALE_i$: Subject's gender

- $bWHITE_i$: Subject's race

---

[4]The decision threshold is determined by considering the desired performance of the fingerprint recognition system. Typically, the decision threshold is decided by looking at the impostor score distribution; for example, the threshold is set for a prespecified false acceptance rate, say 0.01%.

2.4 Longitudinal Analysis of Genuine Fingerprint Match Scores

We address the first question—what is the trend of genuine match scores over time?—by demonstrating the presence of a functional relationship between the time interval of two fingerprints being compared ($\triangle T_{i,jk}$) and the corresponding genuine match score that is commonly observed in the dataset. If the genuine match score is indeed affected by the time interval between two impressions, we want to understand how much of the variation in genuine match scores can be explained by time interval alone. To accomplish this, we quantify the impact of $\triangle T_{i,jk}$ on the change in genuine match scores using goodness-of-fit of the model, and compare it to the impacts of other covariates, namely $Q_{i,jk}$ (a time-independent covariate that is known to be closely related to genuine match score [65]), $AGE_{i,jk}$, $bMALE_i$, and $bWHITE_i$.

We analyze the data under two scenarios: (i) a single finger is used for recognition and (ii) all ten fingers are used for recognition. For (i), we carry out the analysis on the right index finger of the subjects that is typically chosen as the primary finger in the single-finger based recognition. For (ii), we fuse the genuine match scores of all ten fingers by a sum rule, and analyze the fused scores.

2.4.1 Modeling of Genuine Match Scores

The genuine match scores from each of the two COTS matchers are fit by linear 2-level models with various covariates (see Table 2.1):

- Model A (unconditional mean model) without any covariate

- Model B with a single covariate at level 1: Model B_T with $\triangle T_{i,jk}$, Model B_A with $AGE_{i,jk}$, and Model B_Q with $Q_{i,jk}$

Table 2.1: Models with different combinations of covariates.

Model	Level-1 Model	Level-2 Model
Model A	$y_{i,jk} = \varphi_{0i} + \varepsilon_{i,jk}$	$\varphi_{0i} = \beta_{00} + b_{0i}$
Model B_T	$y_{i,jk} = \varphi_{0i} + \varphi_{1i}\triangle T_{i,jk} + \varepsilon_{i,jk}$	$\varphi_{0i} = \beta_{00} + b_{0i}, \ \varphi_{1i} = \beta_{10} + b_{1i}$
Model B_A	$y_{i,jk} = \varphi_{0i} + \varphi_{1i}AGE_{i,jk} + \varepsilon_{i,jk}$	$\varphi_{0i} = \beta_{00} + b_{0i}, \ \varphi_{1i} = \beta_{10} + b_{1i}$
Model B_Q	$y_{i,jk} = \varphi_{0i} + \varphi_{1i}Q_{i,jk} + \varepsilon_{i,jk}$	$\varphi_{0i} = \beta_{00} + b_{0i}, \ \varphi_{1i} = \beta_{10} + b_{1i}$
Model C_G	$y_{i,jk} = \varphi_{0i} + \varphi_{1i}\triangle T_{i,jk} + \varepsilon_{i,jk}$	$\varphi_{0i} = \beta_{00} + \beta_{01}bMALE_i + b_{0i},$ $\varphi_{1i} = \beta_{10} + \beta_{11}bMALE_i + b_{1i}$
Model C_R	$y_{i,jk} = \varphi_{0i} + \varphi_{1i}\triangle T_{i,jk} + \varepsilon_{i,jk}$	$\varphi_{0i} = \beta_{00} + \beta_{01}bWHITE_i + b_{0i},$ $\varphi_{1i} = \beta_{10} + \beta_{11}bWHITE_i + b_{1i}$
Model D	$y_{i,jk} = \varphi_{0i} + \varphi_{1i}\triangle T_{i,jk} + \varphi_{2i}AGE_{i,jk}$ $+\varphi_{3i}Q_{i,jk} + \varepsilon_{i,jk}$	$\varphi_{0i} = \beta_{00} + b_{0i}, \ \varphi_{1i} = \beta_{10} + b_{1i},$ $\varphi_{2i} = \beta_{20} + b_{2i}, \ \varphi_{3i} = \beta_{30} + b_{3i}$

- Model C with an additional covariate at level 2 of Model B_T: Model C_G with $bMALE_i$, and Model C_R with $bWHITE_i$

- Model D with $\triangle T_{i,jk}$, $AGE_{i,jk}$, and $Q_{i,jk}$ as level-1 covariates

The true parameters of the linear models (φ_{ri}, $r = 0, 1, 2, 3$) for subject i have both fixed (β_{r0}) and random effects (b_{ri}).

We first assess the goodness-of-fit of each model, and compare various models to determine which of the covariates better explains the variation in genuine match scores. The models with covariates that are significantly influential in explaining genuine match scores are further verified to determine whether they satisfy the model assumptions. This is followed by parameter and confidential interval estimation to test the null hypothesis: slope of linear model is zero (i.e., $\beta_{10} = 0$ in Model B).

2.4.1.1 Assessment of Goodness-of-Fit

Goodness-of-fit of a model evaluates how well the model fits the data, and is also used to compare multiple models. The following three well-known criteria are typically used to measure the goodness-of-fit: (i) Deviance, (ii) Akaike Information Criterion (AIC), and (iii) Bayesian Information Criterion (BIC).

- Deviance (D): Deviance can be used to compare the goodness-of-fit of nested models. The nested property is easily determined by checking if one model becomes equivalent to the other by setting the coefficients for some of the covariates to zero. For example, Model A and Model B_T are nested; Model A and Model B_Q are nested; however, Model B_T and Model B_Q are not nested. The deviance is defined as:

$$D = -2\log(L), \tag{2.14}$$

  where L is the maximum value of the likelihood function for the model.

- AIC: AIC can be used for any model comparison tasks (models do not need to be nested). AIC is defined as:

$$AIC = 2k - 2\log(L), \tag{2.15}$$

  where k is the number of parameters in the model, and L is the maximum value of the likelihood function for the model.

- BIC: Under the assumption that the data distribution is in the exponential family, BIC is defined as:

$$BIC = k\log(n) - 2\log(L), \tag{2.16}$$

  where k is the number of parameters in the model, n is the number of data points, and L is the maximum value of the likelihood function for the model. BIC also can be used for comparisons of non-nested models.

AIC and BIC add a constant term to the deviance for the sake of comparing non-nested models. The smaller the deviance (AIC or BIC), the better the model fit.

Table 2.2 shows the goodness-of-fit measures of the models in Table 2.1. The model comparisons based on the goodness-of-fit lead to the following observations:

- A decrease in D is observed when each of Model B is compared to Model A. This

Table 2.2: Goodness-of-fit of the models shown in Table 2.1.
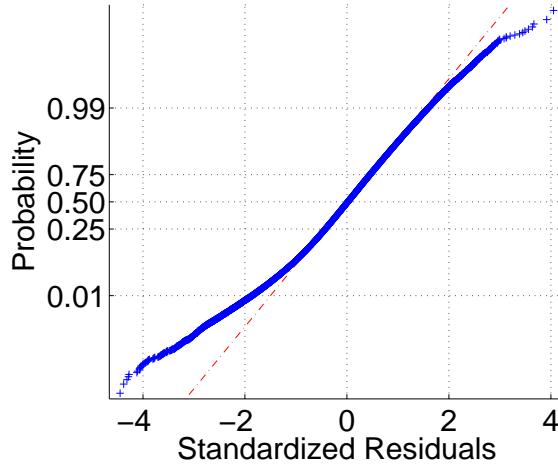
Model	COTS-1			COTS-2		
	Deviance	AIC	BIC	Deviance	AIC	BIC
Model A	1,114,948	1,114,954	1,114,988	1,142,532	1,142,538	1,142,571
Model B_T	1,099,980	1,099,992	1,100,058	1,115,191	1,115,203	1,115,269
Model B_A	1,100,979	1,100,991	1,101,057	1,120,911	1,120,923	1,120,990
Model B_Q	1,028,899	1,028,911	1,028,978	1,060,037	1,060,049	1,060,115
Model C_G	1,099,969	1,099,985	1,100,074	1,115,117	1,115,133	1,115,222
Model C_R	1,099,817	1,099,833	1,099,921	1,114,378	1,114,394	1,114,483
Model D	1,003,908	1,003,938	1,004,105	1,019,412	1,019,442	1,019,608

means that each individual covariate used in Model B ($\triangle T_{i,jk}$, $AGE_{i,jk}$, and $Q_{i,jk}$) can explain some of the variation in genuine match scores.
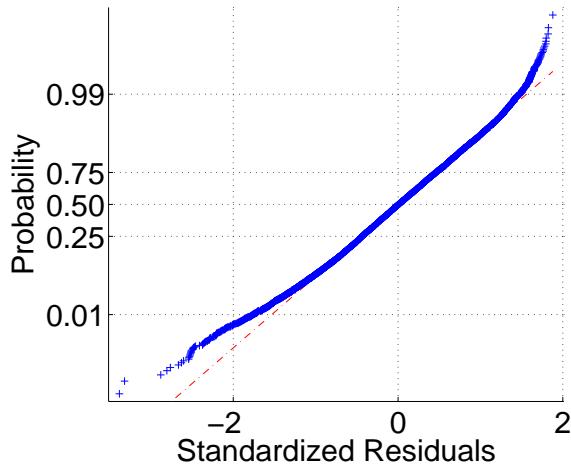
- Model B_Q provides a better fit to the data than Model B_T and Model B_A. This implies that fingerprint image quality is the best covariate to explain the variation in genuine match scores among the three covariates used in Model B.

- Gender and race are not important covariates to explain changes in genuine match scores since the deviance barely decreases from Model B_T to Model C_G or Model C_R.

- Model D has the best fit to the genuine match scores, compared to the other models, as it has the smallest goodness-of-fit value. In other words, including all three covariates ($\triangle T_{i,jk}$, $AGE_{i,jk}$, and $Q_{i,jk}$) as covariates in the multilevel model better explains the trend in genuine match scores compared to including a single covariate.

2.4.1.2 Validation of Model Assumptions
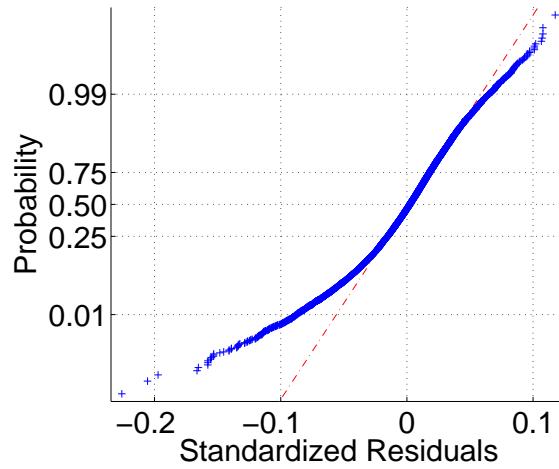
The multilevel model assumes that the residuals ($\varepsilon_{i,jk}$) and random effects (b_{ri}) follow normal distributions. The inference made based on the model fitting is valid only if the underlying assumptions of the multilevel model are satisfied. The normal probability plot is a way to visually verify the normality of the data. If the normal probability plot is linear, one can ascertain the data is from a normal distribution. Figures 2.5 and 2.6 show the normal

52
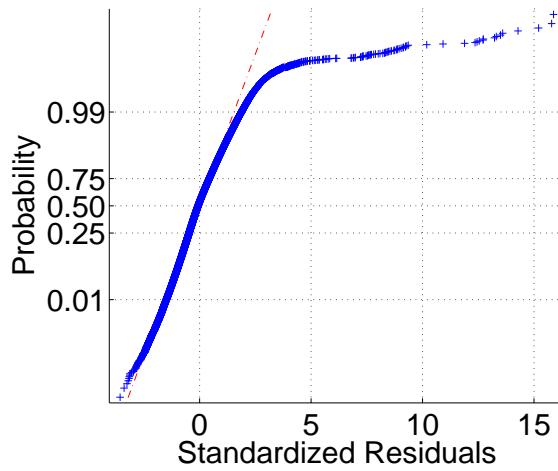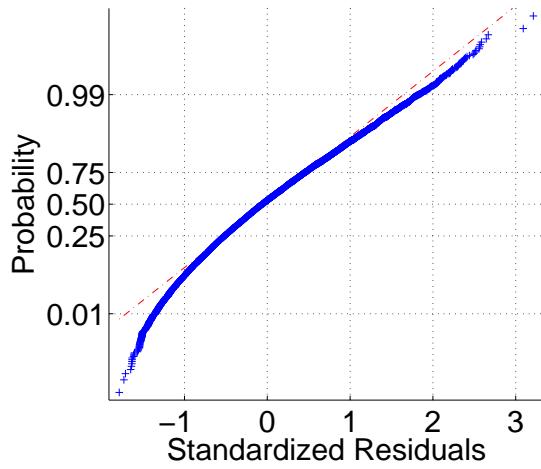
Figure 2.5: Normal probability plots of (a) residuals at level 1 ($\varepsilon_{i,jk}$), (b) random-effects for intercept at level 2 (b_{0i}), and (c) random-effects for slope at level 2 (b_{1i}) of Model B_T fitting the match scores obtained from COTS-1

probability plots of $\varepsilon_{i,jk}$, b_{0i}, and b_{1i} when Model B_T is fit to the genuine match scores obtained from COTS-1 and COTS-2 matchers, respectively.
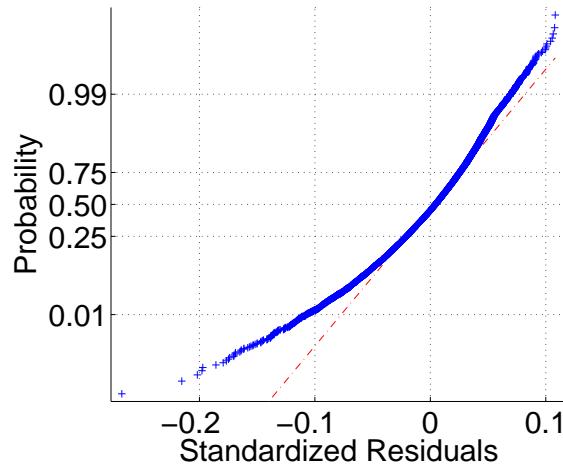
While the residuals generally follow normal distributions, significant departures from normality are observed at the tails for the scores output by both the matchers. A possible cause of the non-normality at tails is that the scores from COTS fingerprint matchers are typically censored, i.e., very low and high match scores are trimmed so that the output scores

Figure 2.6: Normal probability plots of (a) residuals at level 1 ($\varepsilon_{i,jk}$), (b) random-effects for intercept at level 2 (b_{0i}), and (c) random-effects for slope at level 2 (b_{1i}) of Model B_T fitting the match scores obtained from COTS-2.

are in a finite range. To resolve the issue of model assumption violation, a bootstrapping is conducted to obtain reliable parameter estimates and establish empirical confidence intervals for hypothesis test.
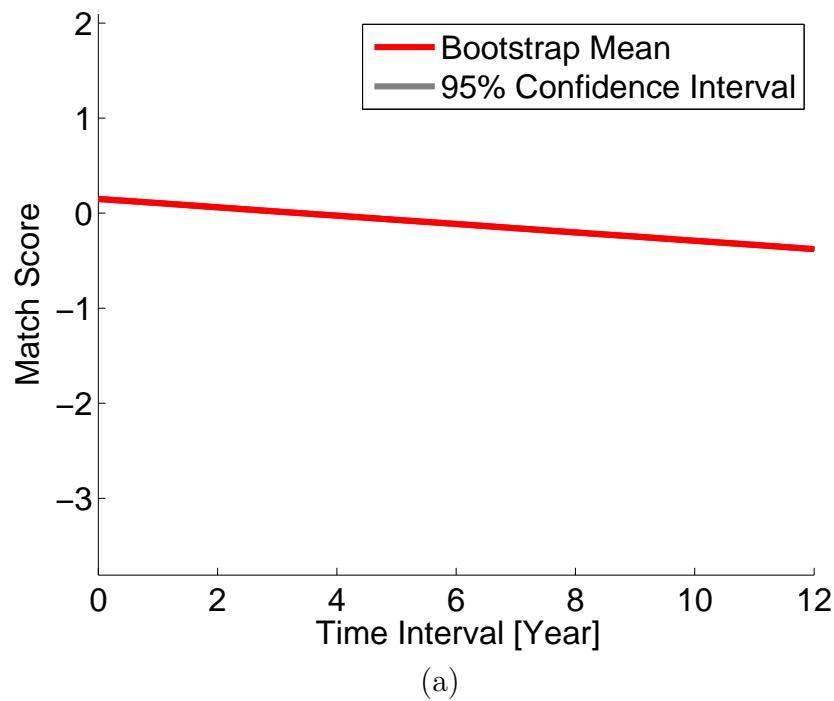
2.4.1.3 Parameter Estimation and Hypothesis Test

When the model assumptions are violated, the parameter estimates for fixed and random effects tend to be reliable while the standard errors (consequently, the confidence intervals) tend to be underestimated [90]. In this case, bootstrapping is a useful way to estimate parameters and confidence intervals [88]. We perform a fully nonparametric bootstrap [62] by repeating parameter estimation for bootstrap samples and then deriving the mean of parameter estimates and percentile confidence intervals. To preserve the hierarchy in the longitudinal data, a cluster bootstrap is used: for each bootstrap sample, N subjects with replacement are resampled at level 1 and all the level-2 data belonging to those subjects are included in the sample. We generate 1,000 bootstrap samples to establish 95% confidence intervals.
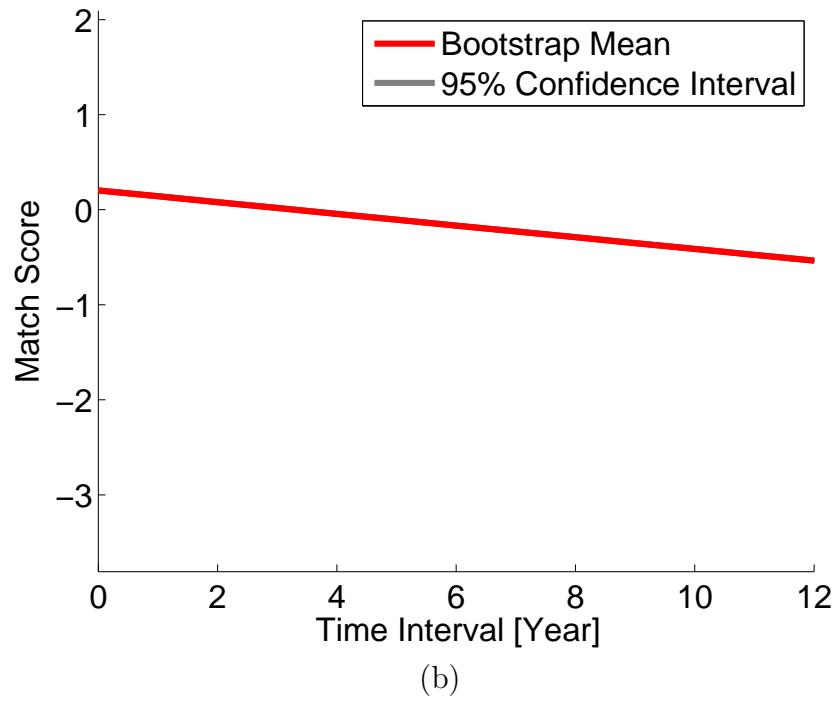
Considering that all the three covariates ($\triangle T_{i,jk}$, $AGE_{i,jk}$, and $Q_{i,jk}$) were shown to have impact on genuine match scores, we estimate the parameters and their 95% confidence intervals in Models B_T, B_A, and B_Q. The population-mean trends of genuine match scores from two COTS matchers and 95% confidence intervals with respect to $\triangle T_{i,jk}$ (Figure 2.7), $AGE_{i,jk}$ (Figure 2.8), and $Q_{i,jk}$ (Figure 2.9) are illustrated. Tables 2.3–2.5 summarize the parameter estimates and 95% confidence intervals for fixed effects in Models B_T, B_A, and B_Q when a single finger is used for recognition.

The null hypothesis we want to test is: $\beta_{10} = 0$ in Model B (i.e., the slope of linear model is zero). This null hypothesis is rejected for all three models, Models B_T, B_A, and B_Q, at the significance level of 0.05 since the 95% confidence interval for β_{10} does not contain zero. Hence, the genuine match scores from both the matchers have negative correlation with the above three covariates.

Model D incorporates all the three influential covariates into the model. Table 2.6 shows the parameter estimates and 95% confidence intervals in Model D with two COTS matchers. The fixed-effects parameter estimates for $\triangle T_{i,jk}$ (β_{10}), $AGE_{i,jk}$ (β_{20}), and $Q_{i,jk}$ (β_{30}) remain as negative in Model D. The estimated covariance matrix indicate the relationship (i) between

Figure 2.7: Population-mean trends of genuine match scores and 95% confidence intervals with respect to $\triangle T_{i,jk}$. Match scores are obtained from (a) COTS-1 and (b) COTS-2 matchers.
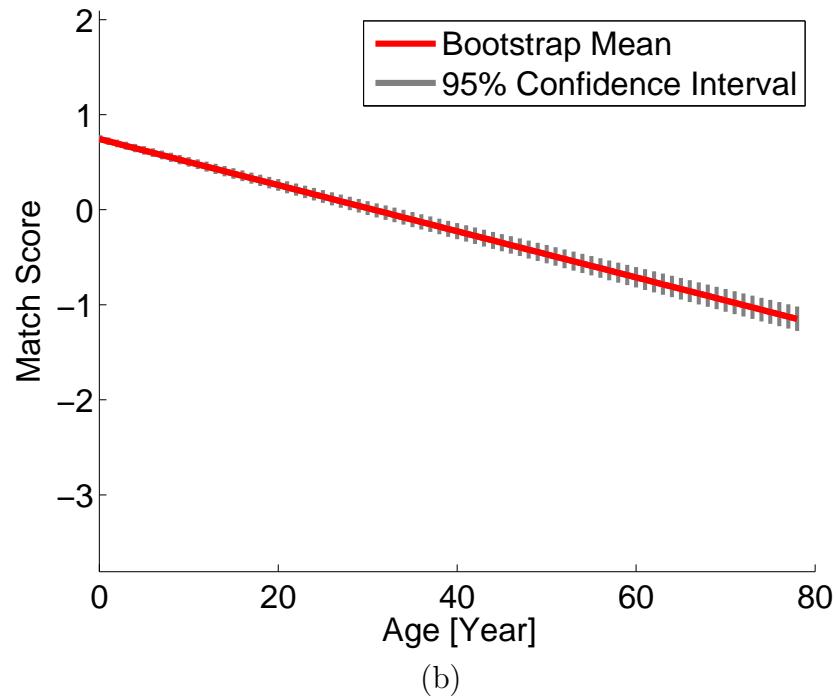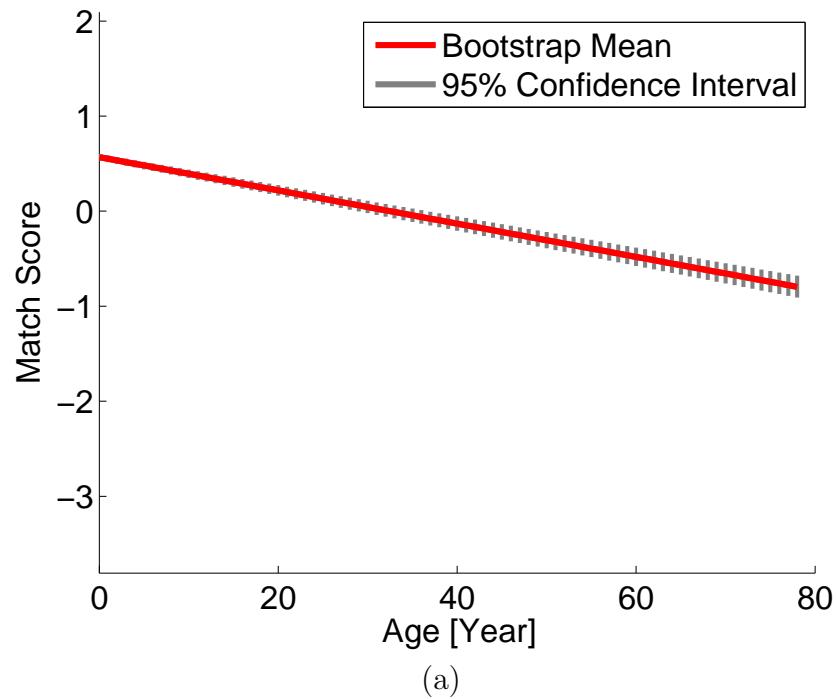
Figure 2.8: Population-mean trends of genuine match scores and 95% confidence intervals with respect to $AGE_{i,jk}$. Match scores are obtained from (a) COTS-1 and (b) COTS-2 matchers.
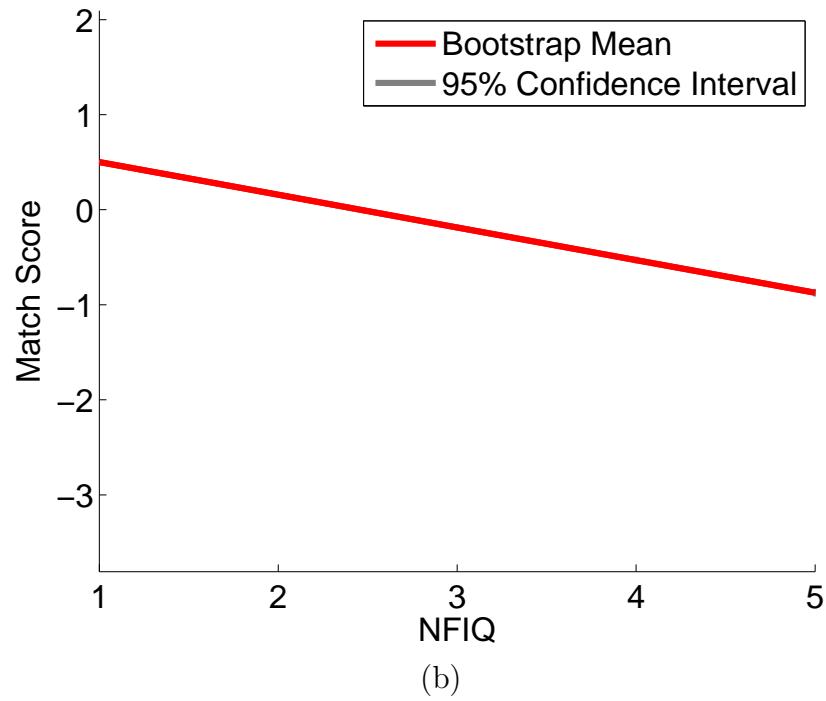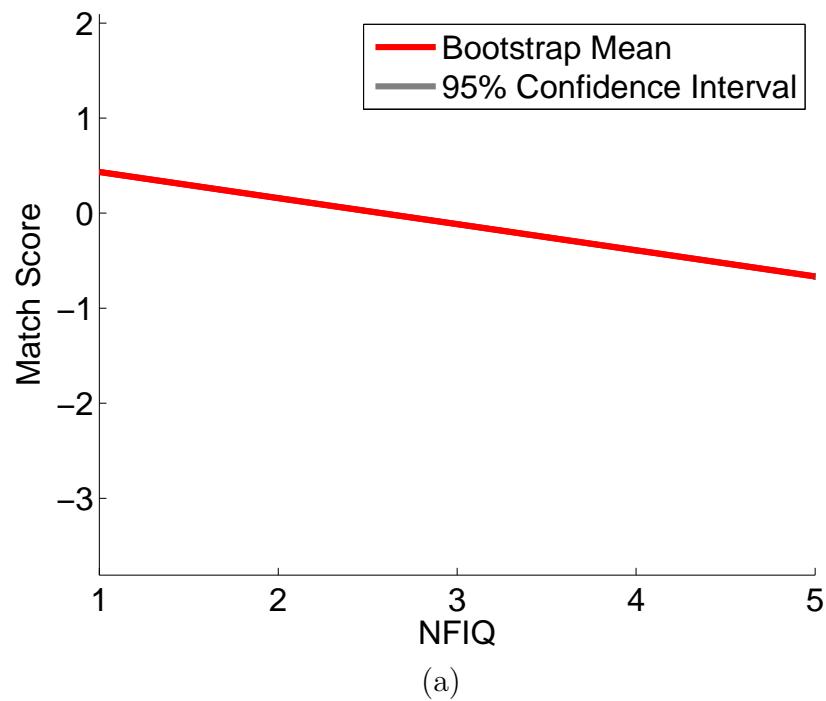
Figure 2.9: Population-mean trends of genuine match scores and 95% confidence intervals with respect to $Q_{i,jk}$. Match scores are obtained from (a) COTS-1 and (b) COTS-2 matchers.

Table 2.3: Parameter estimates and 95% confidence intervals in Model B_T with bootstrapping when a single finger is used for recognition.

	Parameters	COTS-1	COTS-2
Fixed Effects	β_{00}	0.1496 (0.1406; 0.1590)	0.2032 (0.1939; 0.2127)
	β_{10}	-0.0440 (-0.0450; -0.0430)	-0.0616 (-0.0625; -0.0606)
Variance Components	σ_ε^2	0.7057	0.7185
	σ_0^2	0.5298	0.5744
	σ_1^2	0.0034	0.0039
	σ_{01}	-0.0134	-0.0277

Table 2.4: Parameter estimates and 95% confidence intervals in Model B_A with bootstrapping when a single finger is used for recognition.

	Parameters	COTS-1	COTS-2
Fixed Effects	β_{00}	0.5682 (0.5335; 0.6020)	0.7447 (0.7072; 0.7843)
	β_{10}	-0.0175 (-0.0185; -0.0164)	-0.0243 (-0.0254; -0.0231)
Variance Components	σ_ε^2	0.6998	0.7120
	σ_0^2	5.6003	7.5575
	σ_1^2	0.0050	0.0066
	σ_{01}	-0.1574	-0.2136

Table 2.5: Parameter estimates and 95% confidence intervals in Model B_Q with bootstrapping when a single finger is used for recognition.

	Parameters	COTS-1	COTS-2
Fixed Effects	β_{00}	0.7087 (0.6954; 0.7221)	0.8456 (0.8316; 0.8595)
	β_{10}	-0.2750 (-0.2798; -0.2702)	-0.3439 (-0.3489; -0.3385)
Variance Components	σ_ε^2	0.6489	0.6738
	σ_0^2	0.9096	0.9105
	σ_1^2	0.1163	0.1027
	σ_{01}	-0.2543	-0.2473

$\triangle T_{i,jk}$ and $Q_{i,jk}$ (σ_{13}), and (ii) between $AGE_{i,jk}$ and $Q_{i,jk}$ (σ_{23}). Although the estimated values for σ_{13} and σ_{23} are negative, the correlations are not strong. σ_{13} in Model D with COTS-2 matcher cannot be said it is significantly different from 0 since the null hypothesis $\sigma_{13} = 0$ is not rejected at the 0.05 significance level. Also, when we get correlation matrix

Table 2.6: Parameter estimates and 95% confidence intervals in Model D with bootstrapping when a single finger is used for recognition.

	Parameters	COTS-1	COTS-2
	β_{00}	0.9137	1.0353
		(0.8765; 0.9471)	(0.9947; 1.0750)
	β_{10}	-0.0368	-0.0533
		$(-0.0378; -0.0358)$	$(-0.0543; -0.0522)$
Fixed Effects	β_{20}	-0.0030	-0.0024
		$(-0.0042; -0.0019)$	$(-0.0036; -0.0011)$
	β_{30}	-0.2509	-0.3064
		$(-0.2558; -0.2463)$	$(-0.3112; -0.3015)$
	σ_ε^2	0.6033	0.6127
	σ_0^2	6.6328	7.8996
	σ_1^2	0.0041	0.0040
	σ_{01}	0.0944	0.0800
	σ_2^2	0.0068	0.0082
Variance Components	σ_{02}	-0.1941	-0.2335
	σ_{12}	-0.0036	-0.0033
	σ_3^2	0.1165	0.1039
	σ_{03}	-0.2092	-0.1466
	σ_{13}	-0.0007	-0.0004^*
	σ_{23}	-0.0013	-0.0030

* The hypothesis test indicates that the null hypothesis *the parameter is zero* is not rejected at the significance level of 0.05.

from the covariance matrix, the correlation coefficients for σ_{13} and σ_{23} are very close to 0. For COTS-1 matcher, the correlation coefficients for σ_{13} and σ_{23} are -0.0324 and -0.0464, respectively; for COTS-2 matcher, they are -0.0174 and -0.1035.

When all ten fingerprints of a subject are utilized for identification, the match scores from the ten fingers can be fused by a sum rule; the fusion score $y_{i,jk}^\dagger$ is obtained by:

$$y_{i,jk}^\dagger = \sum_{m=1}^{10} y_{i,jk}^{(m)}, \tag{2.17}$$

where $y_{i,jk}^{(m)}$ is the genuine match score of finger m of subject i.

Models B_T and B_A were also fit to the genuine match scores from all ten fingers fused by a sum rule. Figures 2.10 and 2.11 show the population-mean trends and 95% confidence

Table 2.7: Parameter estimates and 95% confidence intervals in Model B_T with bootstrapping when the match scores from ten fingers are fused by a sum rule.

	Parameters	COTS-1	COTS-2
Fixed Effects	β_{00}	0.1896 (0.1800; 0.1995)	0.2258 (0.2159; 0.2360)
	β_{10}	−0.0603 (−0.0612; −0.0594)	−0.0726 (−0.0736; −0.0717)
Variance Components	σ_ε^2	0.6562	0.6615
	σ_0^2	0.5986	0.6599
	σ_1^2	0.0037	0.0040
	σ_{01}	−0.0144	−0.0304

Table 2.8: Parameter estimates and 95% confidence intervals in Model B_A with bootstrapping when the match scores from ten fingers are fused by a sum rule.
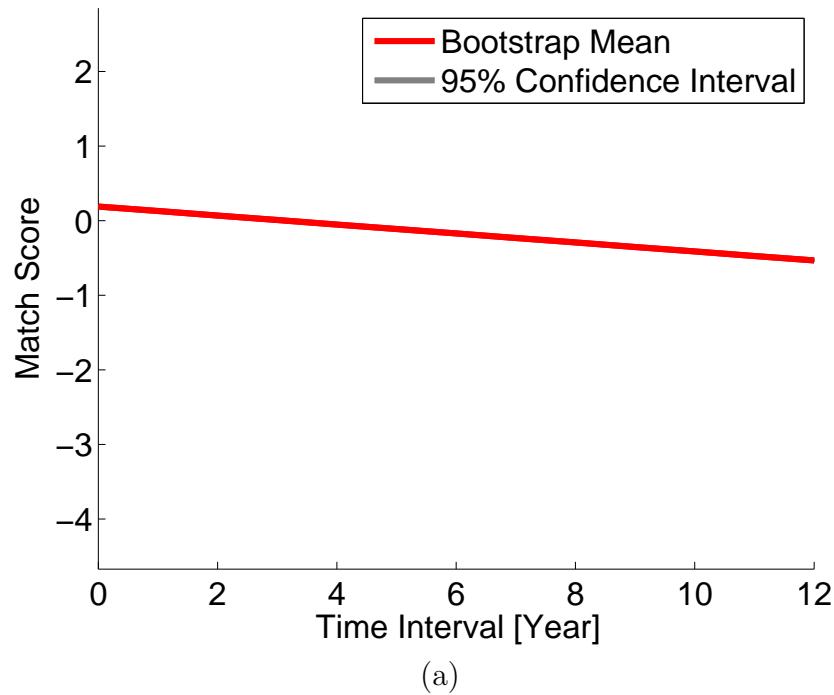
	Parameters	COTS-1	COTS-2
Fixed Effects	β_{00}	0.5867 (0.5231; 0.6841)	0.7537 (0.7056; 0.8996)
	β_{10}	−0.0185 (−0.0223; −0.0163)	−0.0249 (−0.0295; −0.0235)
Variance Components	σ_ε^2	0.6651	0.6602
	σ_0^2	5.0636*	8.6870
	σ_1^2	0.0046	0.0074
	σ_{01}	−0.1409*	−0.2423

* The hypothesis test indicates that the null hypothesis *the parameter is zero* is not rejected at the significance level of 0.05.
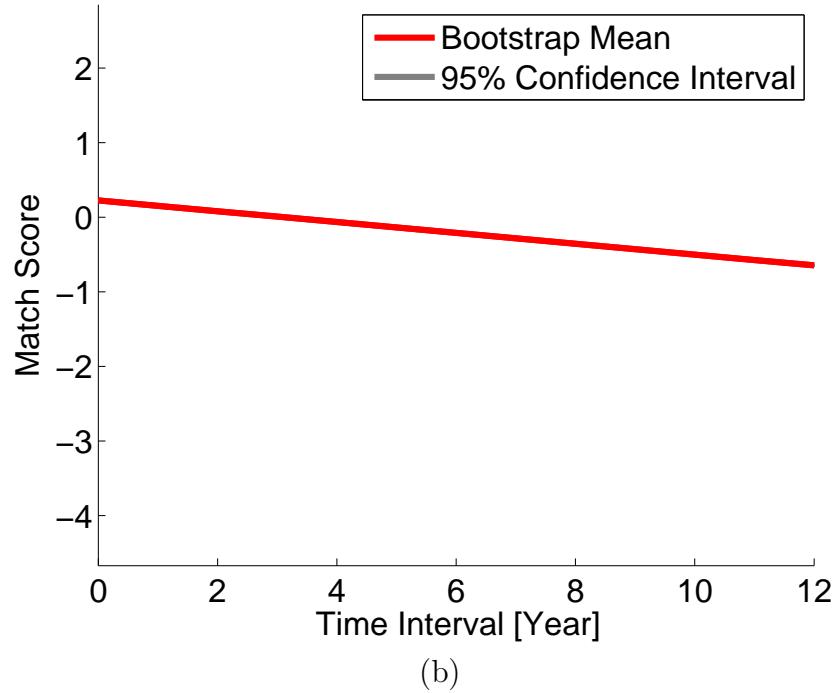
intervals, and Tables 2.7 and 2.8 show the parameter estimates and confidence intervals. We observe a negative relationship between genuine match scores and the two covariates ($\triangle T_{i,jk}$ and $AGE_{i,jk}$). The null hypothesis ($\beta_{10} = 0$ in Model B) is rejected for both of the models. Note that Models B_Q and D are not used in ten-finger fusion scenario since the covariate corresponding to $Q_{i,jk}$ is difficult to define in this case.

2.4.1.4 Genuine Match Score Trends of Individual Subjects

The random effects (b_{ri}) at level 2 in the multilevel model represent the deviation of subject i from the population-mean trend (β_{ri}). The parameter estimates of ($\varphi_{0i}, \varphi_{1i}$) for each subject in Model B_T are shown in Figure 2.12 in addition to the population-mean trend (β_{00}, β_{11}).
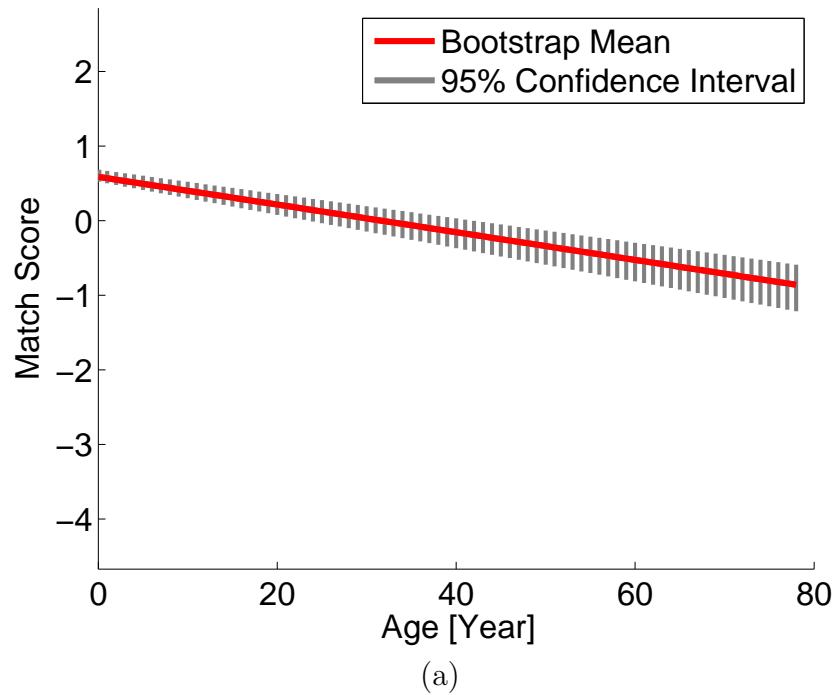
(a)
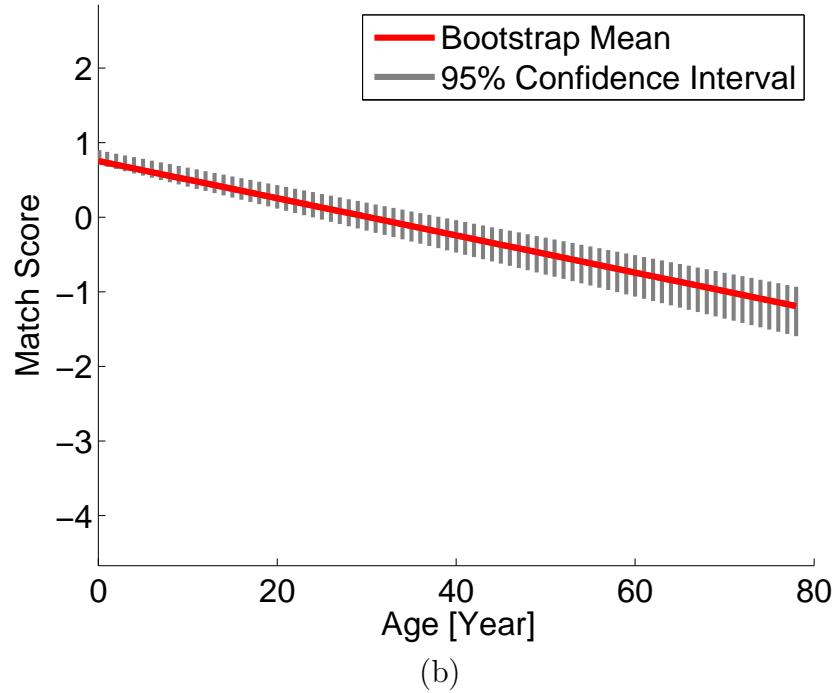
(b)

Figure 2.10: Population-mean trends of genuine match scores and 95% confidence intervals with respect to $\triangle T_{i,jk}$, when the scores from ten fingers are fused. Match scores are obtained from (a) COTS-1 and (b) COTS-2 matchers.

Figure 2.11: Population-mean trends of genuine match scores and 95% confidence intervals with respect to $AGE_{i,jk}$, when the scores from ten fingers are fused. Match scores are obtained from (a) COTS-1 and (b) COTS-2 matchers.
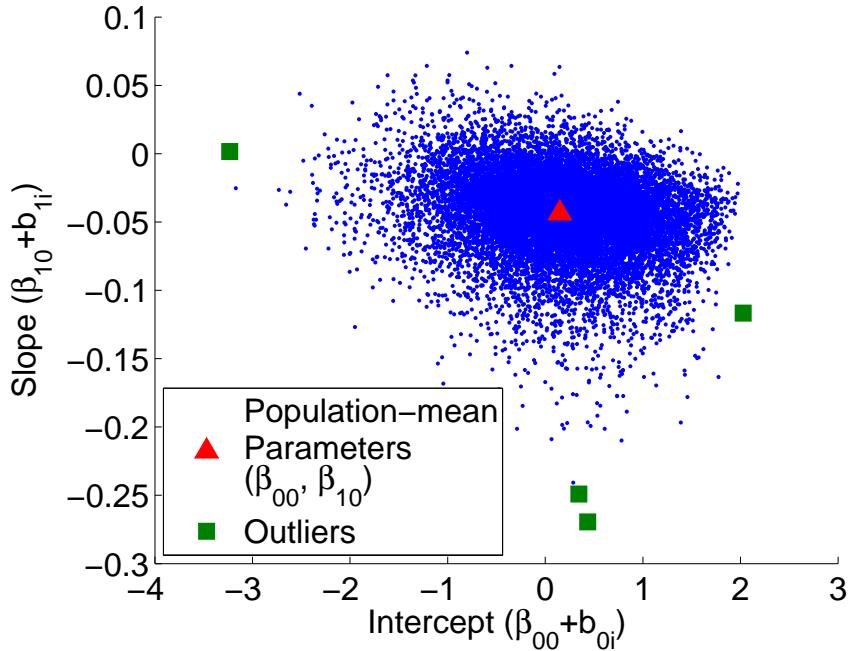
Figure 2.12: Parameter estimates of Model B_T with genuine match scores provided by COTS-1 matcher. The estimates for the population-mean parameters (β_{00}, β_{10}) and the parameters for each subject $(\varphi_{0i}, \varphi_{1i})$ are represented as a triangle and dots, respectively. The parameters associated with four outlier subjects are marked as squares.

The parameter estimates associated with several outlying subjects who markedly deviate from the population-mean trend are also indicated. Figures 2.13–2.16 show the individual trends of the outlying subjects and their fingerprint impressions.
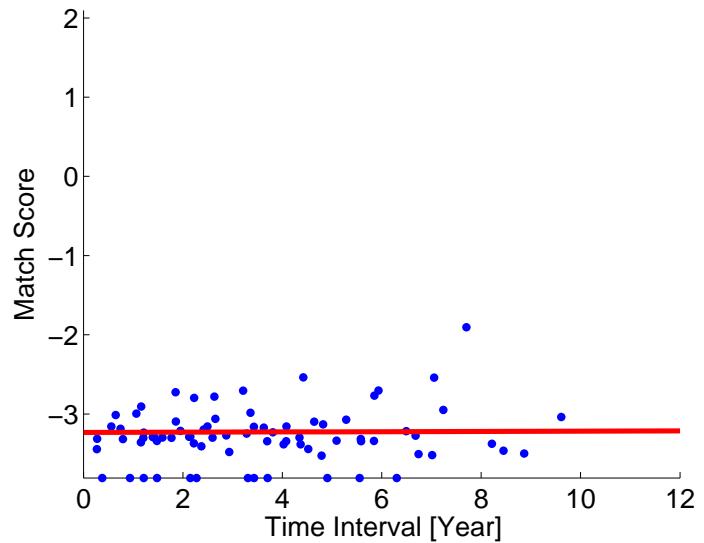
- Outlier case 1 (Figure 2.13): The estimated intercept of this subject is very small. The subject consistently gives low genuine match scores since his fingerprints have been severely scarred, resulting in poor fingerprint quality (the NFIQ values of the fingerprints are $\{4, 4, 5, 5, 4, 4\}$). This subject can be called a "goat" in the Doddington's biometric zoo nomenclature [50] which refers to the subjects that can easily lead to false rejections.

- Outlier case 2 (Figure 2.14): The intercept of the fitted model for this subject is rather large while the slope is negative. This subject consistently gives high genuine match scores because his fingerprint impressions are of good quality (the NFIQ values

64

are $\{3, 1, 2, 1, 1, 3\}$). This subject can be viewed as a "sheep" in the Doddington's biometric zoo who tends to be successfully verified.

- Outlier case 3 (Figure 2.15): This subject shows a very sharp decrease in genuine match scores as a function of time interval. In Figure 2.15(a), the genuine match scores involving the first fingerprint impression are very low. This fingerprint impression is indeed an impostor fingerprint (see Figure 2.15(b)) since it is of tented arch type while the actual pattern of this finger is a right loop. This shows that the operational fingerprint data can be mislabeled.

- Outlier case 4 (Figure 2.16): This subject also has a steep slope. It turns out that the fingerprint impressions of this subject were collected during his adolescence (starting at the age of 11 until the age of 21). It is known that an increase in the size of finger during the adolescent period can cause false rejections in fingerprint matching when a juvenile's fingerprint impression is compared with his adult fingerprint impression [63].

2.5 Longitudinal Analysis of Fingerprint Matching Accuracy

The ultimate question of interest with regard to fingerprint persistence is as follows: what is the trend of matching accuracy with respect to the time interval between fingerprints in the comparison? This question is addressed by analyzing the binary identification decisions made on genuine fingerprint pairs using the linear 2-level model. We introduce the multilevel model for binary responses, and report the bootstrap results of parameter estimates and confidence intervals.

(a) $(\varphi_{0i}, \varphi_{1i}) = (-3.2326, 0.0016)$

Age 25   Age 28   Age 29

Age 31   Age 32   Age 33

(b)
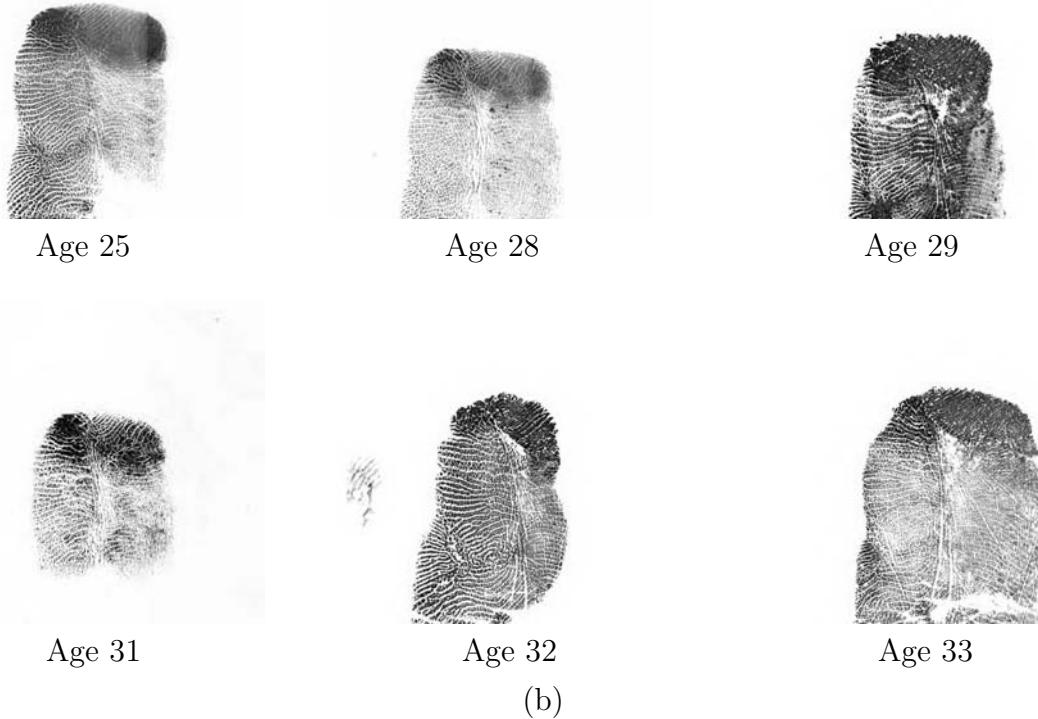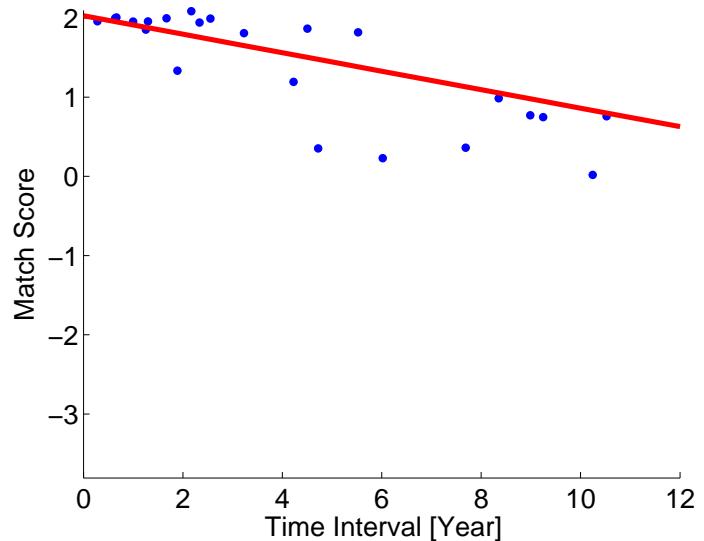
Figure 2.13: A subject whose intercept in Model B_T is very small due to the severe alteration of the fingerprint pattern (outlier case 1). (a) The observed responses and fitting result of the subject, and (b) fingerprint impressions of the subject at different ages.

(a) $(\varphi_{0i}, \varphi_{1i}) = (2.0281, -0.1167)$

Age 34   Age 40   Age 42
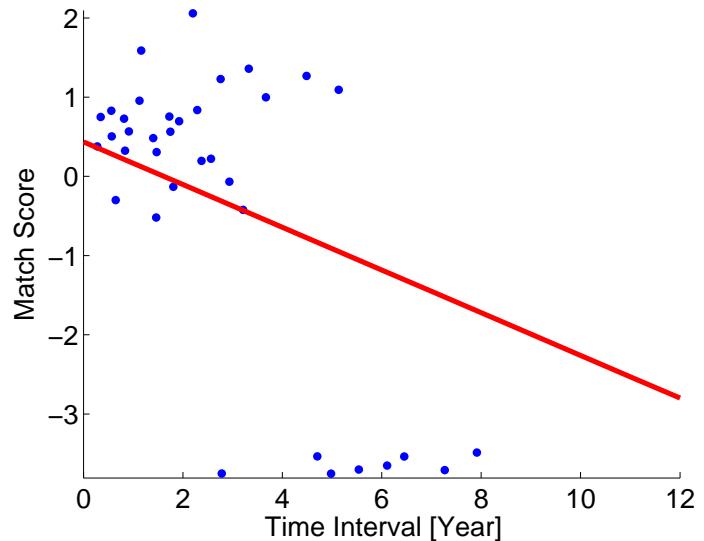
Age 43   Age 44   Age 45

(b)

Figure 2.14: A subject with high quality ridge pattern resulting in the large intercept in Model B_T (outlier case 2). (a) The observed responses and fitting result of the subject, and (b) fingerprint impressions of the subject at different ages.

(a) $(\varphi_{0i}, \varphi_{1i}) = (0.4353, -0.2695)$

Age 18   Age 21   Age 22

Age 23   Age 24   Age 26

(b)
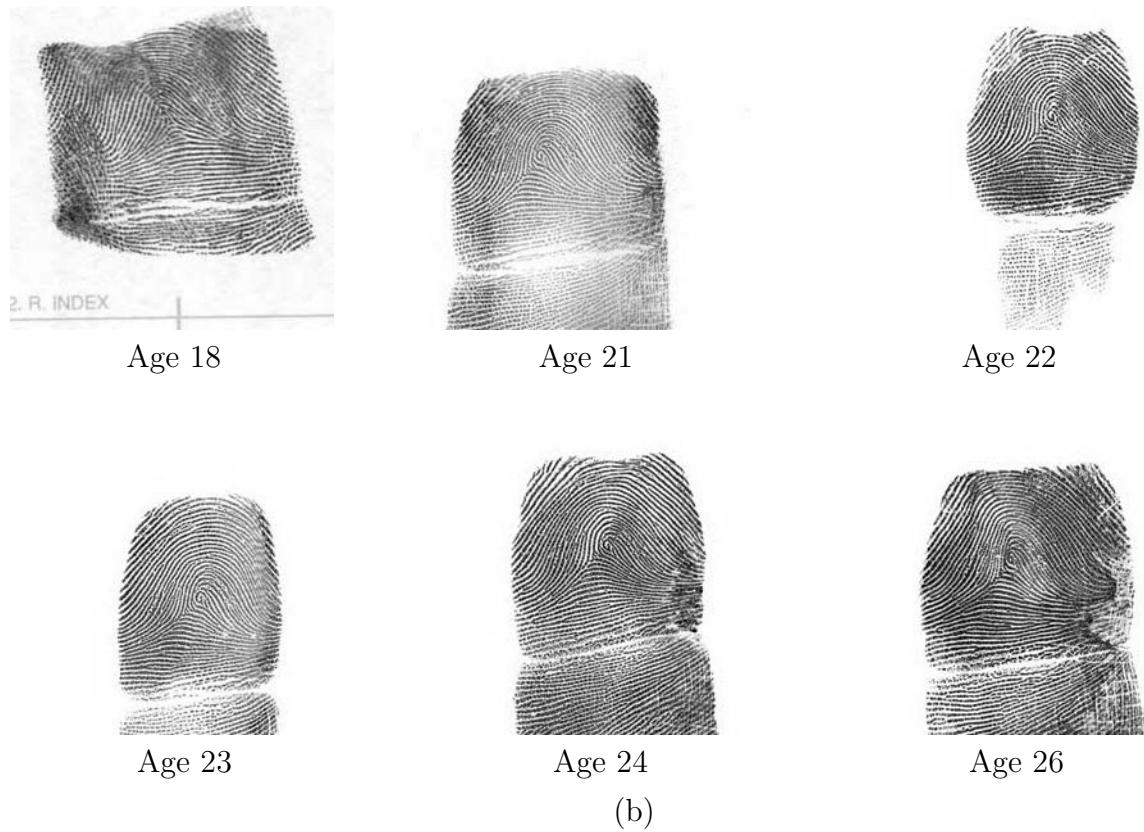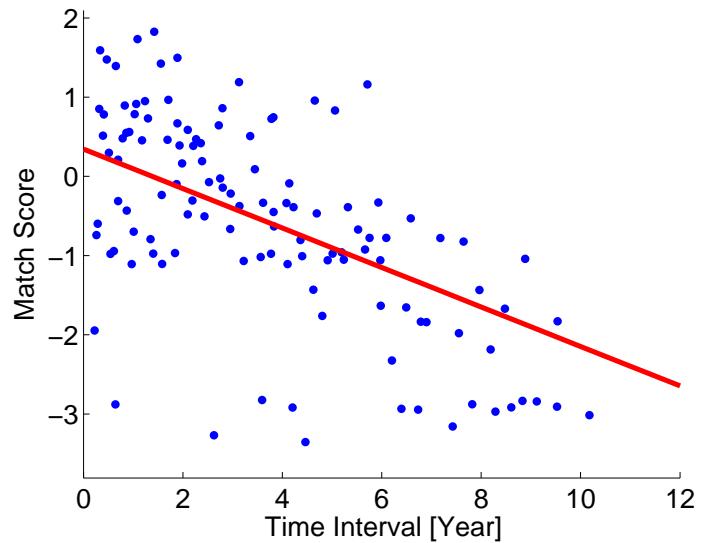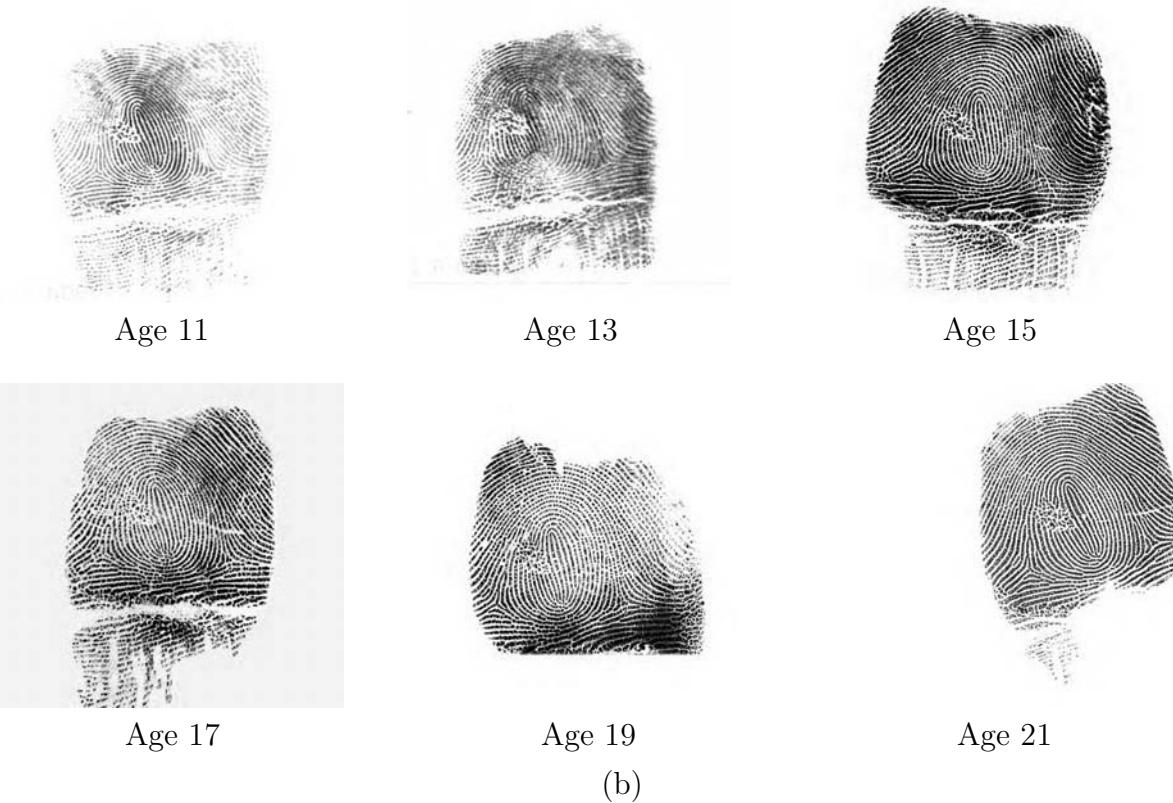
Figure 2.15: A subject with steep negative slope (outlier case 3) resulting from a mislabeled fingerprint (the first impression). (a) The observed responses and fitting result of the subject, and (b) fingerprint impressions of the subject at different ages.

(a) $(\varphi_{0i}, \varphi_{1i}) = (0.3443, -0.2491)$

Age 11   Age 13   Age 15

Age 17   Age 19   Age 21

(b)

Figure 2.16: A subject with steep negative slope due to fingerprint impressions made during his adolescence (outlier case 4). (a) The observed responses and fitting result of the subject, and (b) fingerprint impressions of the subject at different ages.

2.5.1 Multilevel Model for Binary Responses

A binary decision for each genuine fingerprint pair $(y_{i,jk}^*)$ in the database was made by applying a predetermined decision threshold as shown in Equation (2.13). We set the decision threshold corresponding to false acceptance rate of 0.01%, assuming that the impostor distribution is stable over time.

The multilevel model cannot be directly fit to the binary responses since the model assumptions (normality of residuals and random effects) will be obviously violated. Thus, a binary identification decision $y_{i,jk}^*$ is modeled by a Bernoulli trial with the probability of true acceptance $\pi_{i,jk}$, and the expected $\pi_{i,jk}$ is modeled using a logit link function.

$$\text{Level-1 Model: } g(\pi_{i,jk}) = \varphi_{0i} + \varphi_{1i}\triangle T_{i,jk} + \varepsilon_{i,jk},$$

$$y_{i,jk}^* \sim Bin(1, \pi_{i,jk}),$$
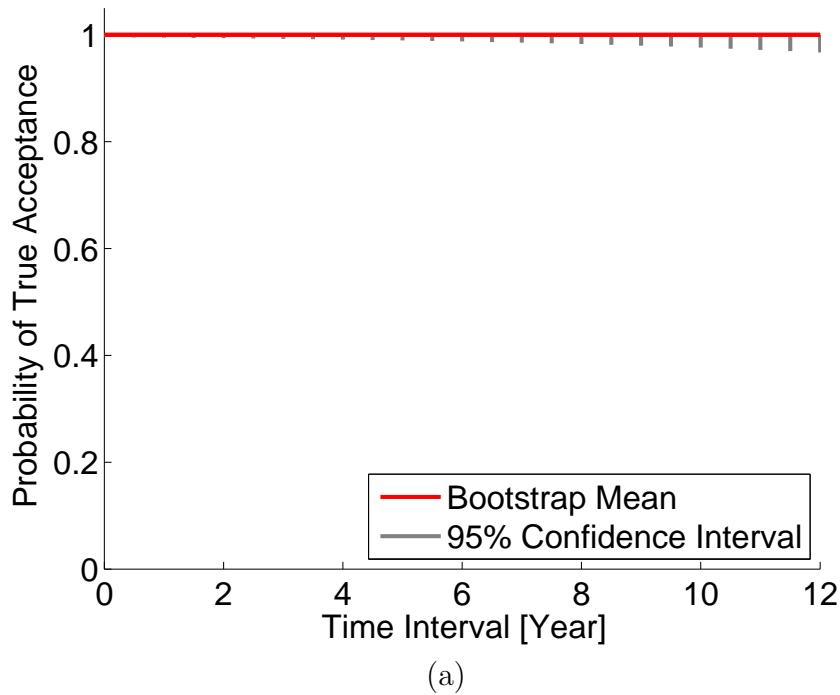
$$\text{Level-2 Model: } \varphi_{0i} = \beta_{00} + b_{0i},$$

$$\varphi_{1i} = \beta_{10} + b_{1i}, \tag{2.18}$$
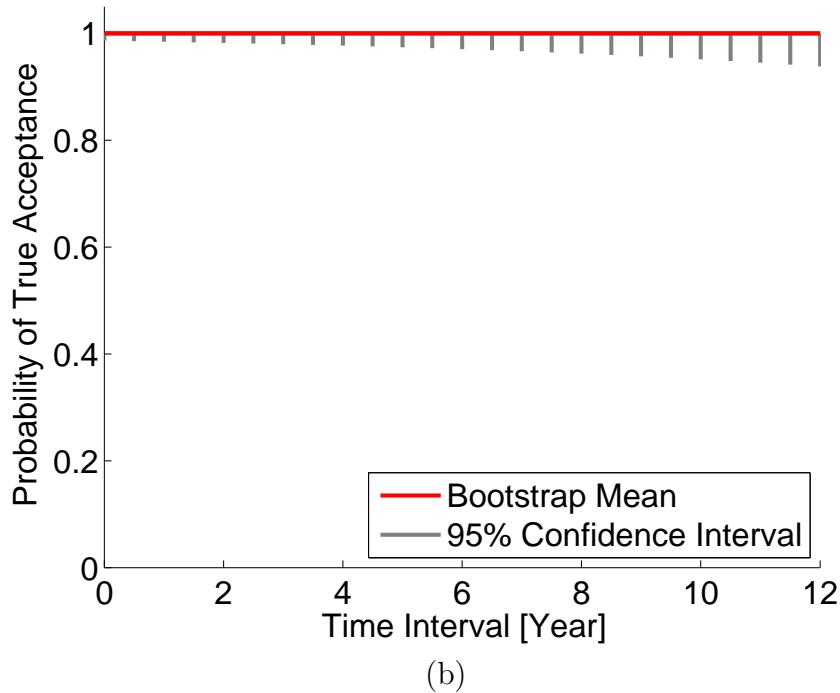
where $g(\cdot)$ is a link function (for binary responses, $g(\cdot)$ is a logit function; $g(\pi_{i,jk}) = \log\left(\frac{\pi_{i,jk}}{1-\pi_{i,jk}}\right)$).

2.5.2 Trend of True Acceptance Rate over Time

The parameter estimates and 95% confidence intervals of the model in Equation (2.18) are obtained by bootstrap. The population-mean trends of $\pi_{i,jk}$ are shown in Figure 2.17 when a single finger is used for recognition, and in Figure 2.18 when the match scores from all ten fingers are fused by a sum rule. Both results indicate that the probability of true acceptance for genuine fingerprint pairs tends to remain very close at 1 even though the time interval between two fingerprints in comparison increases up to 12 years (recall that the longitudinal fingerprint data in this study has a maximum time span of 12 years). In other words, the fingerprint pairs can be correctly determined as genuine regardless of the time interval
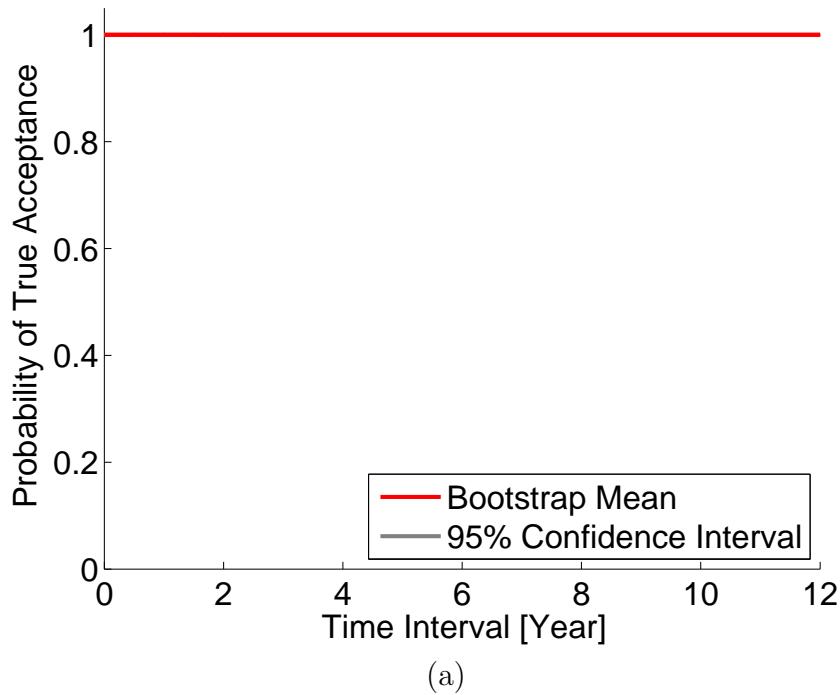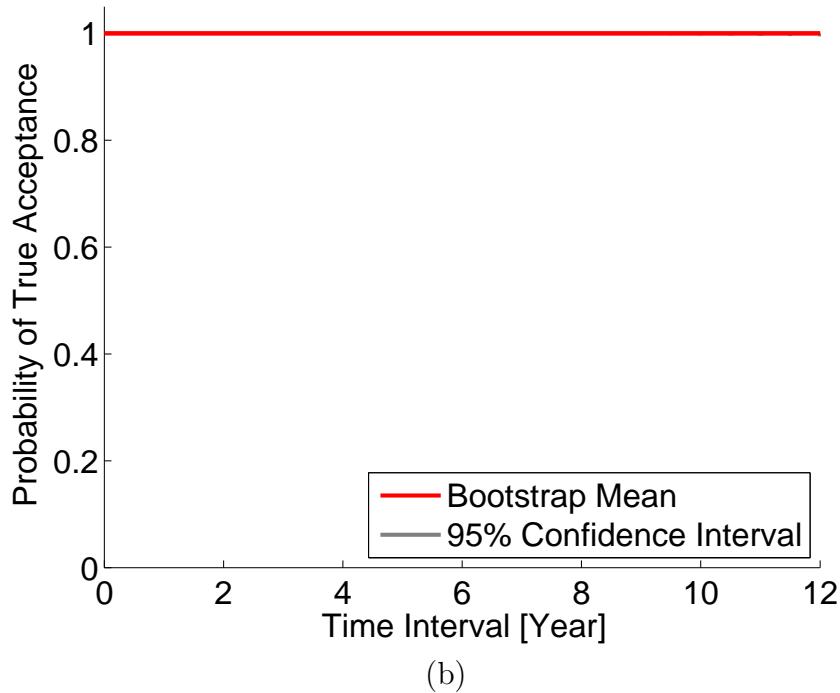
Figure 2.17: Probability of true acceptance with respect to time interval when match scores from a single (right index) finger are used. (a) COTS-1 and (b) COTS-2.

Figure 2.18: Probability of true acceptance with respect to time interval when match scores from all ten fingers are fused by the sum rule. (a) COTS-1 and (b) COTS-2.

(up to 12 years here) between the fingerprint impressions with high confidence. The 95% confidence intervals become negligibly small when all ten fingerprints from a subject are used in matching, compared to using a single finger.

2.6 Summary

Since ancient times, fingerprints have been recognized as persistent and unique to an individual. Early scientific studies on fingerprint recognition in the late 19th century claimed that there is no significant change in the friction ridge structure over time by examining small sets of genuine fingerprint pairs captured with large time interval. Although fingerprint recognition is now used to distinguish an extremely large number of individuals (for example, India's UIDAI [9] involving more than 1 billion people), the persistence of fingerprints can only be claimed based on anecdotal evidence.

To understand the temporal behavior of fingerprint matching accuracy, multiple fingerprint records of 15,597 subjects booked by the Michigan State Police over a duration of 5–12 years were collected. The genuine match scores obtained by two COTS fingerprint matchers have been analyzed by linear multilevel statistical models with various covariates, including time interval between two fingerprints being compared. Our longitudinal study of fingerprint recognition have led to the following conclusions:

- A formal test of hypothesis indicates that the genuine match scores tend to decrease as the time interval between the two fingerprints being compared increases.

- The genuine match scores also tend to decrease as the subject's age increases or when the fingerprint image quality decreases.

- The goodness-of-fit measures of multilevel models with different covariates show that:

  - Time interval, subject's age, and fingerprint image quality best explain the variation in genuine match scores, while subject's gender and race are not significant

covariates.

- – Among the three covariates (time interval, subject's age, and fingerprint image quality), fingerprint image quality is the most influential covariate.

- There exist subjects in our database that do not conform to the population trend as determined by model fit. These outlying subjects illustrate (i) the presence of individuals in the database used in the study that follow the nomenclature in the Doddington's biometric zoo, (ii) the degradation in genuine match scores when a juvenile fingerprint is compared to his adult fingerprint, and (iii) labeling errors in the operational fingerprint database.

- Despite the downward trend in genuine match scores over time, the probability of true acceptance, at a predetermined decision threshold, remains close to 1 even as the time interval between fingerprints in comparison increases (up to 12 years which is the maximum time span in our database). In other words, the fingerprint matching accuracy tends to be stable over time under the assumption that the impostor match scores are not affected by the time interval between fingerprints being compared.

- The inference made from fingerprint recognition results with single finger (right index finger) applies to the inference from ten-finger fusion results.

- The results from two different COTS matchers used in the study coincide.

As future work, we plan to pursue the longitudinal study of fingerprint recognition in the following ways:

- The match scores and identification decisions of impostor fingerprint pairs will be analyzed by the multilevel model to determine whether the false acceptance rate is stable over time.

- The interactions among covariates (time interval, subject's age, and fingerprint image quality) will be incorporated in the model.

74

- Given that we make all pairwise comparisons of the fingerprint impressions from each subject, the correlation among the genuine match scores of a subject needs to be reflected in the model.

- Nonliear multilevel models can be investigated and compared to the linear models.

Chapter 3

Fingerprint Orientation Field Modeling
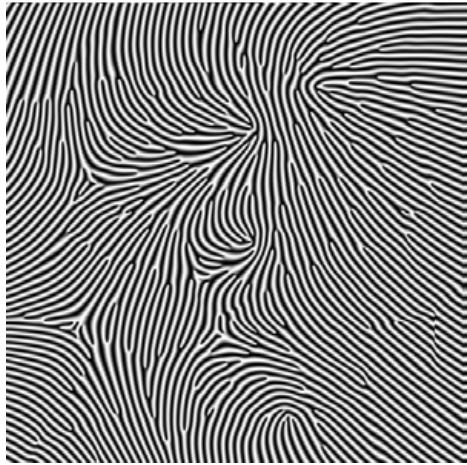
3.1 Introduction

Fingerprint patterns can be characterized by (i) ridge frequency, (ii) orientation field, and (iii) minutiae distribution. Fingerprint ridges are spaced almost equidistantly; ridge period is typically around 10 pixels in 500 ppi fingerprint images. Fingerprint orientation field has distinct characteristics which can differentiate fingerprints from any other flow pattern: it has a specific number of singular points, the configuration of singular points follows a certain spatial distribution [40], and its global field is shaped like an arch. Minutiae distribution in a fingerprint has a tendency to be dense around singular points since the ridge flow converges or diverges around singular points while maintaining the ridge frequency [42].

To assess the quality of a fingerprint pattern, features based on local properties (i.e., clarity of ridge structure in a grayscale image) and global properties (e.g., continuity of orientation field or energy concentration in the frequency domain over the entire fingerprint region) have been utilized [31]. However, lack of global constraints that can specifically distinguish a valid fingerprint pattern from any type of non-fingerprint images limits of the

ability of existing fingerprint quality measures to detect abnormalities in the images shown in Figure 3.1. Figure 3.1(a) is a synthesized texture image based on the iterative contextual filtering approach [38]; while it has a valid fingerprint ridge frequency, its orientation field does not follow that of a fingerprint pattern. Figure 3.1(b) shows a fingerprint image whose ridge structure in the central part of the finger has been transplanted from some other friction ridge skin. Although the quality of local ridge structure in Figure 3.1(b) is good, the orientation field shows discontinuity at the boundaries of the altered region. Figure 3.1(c) is a synthesized fingerprint image using the SFinGe (Synthetic Fingerprint Generator) model [38]; while the ridge frequency and the orientation field of this fingerprint image appear to be realistic, a large portion of the fingerprint area has ridges that flow parallel to each other and hence contain no minutiae. Figure 3.1(d) is an example of a typical latent fingerprint which is found at crime scenes. Although the quality of the latent print in this image is fairly good, the background line structure with similar frequency value as the fingerprint pattern makes it difficult to localize the fingerprint pattern in the image and extract features (e.g., orientation field and minutiae).

Fingerprint modeling has been developed to capture the unique characteristics of fingerprint patterns in orientation field and minutia. In the literature, fingerprint modeling has been approached in terms of (i) stochastic modeling or (ii) deterministic modeling. In stochastic modeling, properties of singularity distribution [40] and minutiae distribution [42], for example, are modeled as parameterized probabilistic functions (e.g., Gaussian mixture model) based on a large number of fingerprints. In orientation field modeling, learning schemes based on a large set of fingerprints have been proposed; for instance, local orientation patches are learned as dictionaries to estimate orientation field in noisy or missing regions of latent fingerprints [55]. In other approaches, the parameter space for the coefficients of the Fourier basis function is reduced by a set of fingerprints [130]. The stochastic modeling inherently requires a large number of *real* fingerprints in order to establish the correct model. However, the size of the fingerprint database available for research is typically

(a) NFIQ = 1, (Model does not fit)

(b) NFIQ = 3, $F = 0.74$

(c) NFIQ = 2, $F = 0.83$

(d) NFIQ = 2, (Model does not fit)

Figure 3.1: Images containing fingerprint-like pattern. (a) Synthesized image obtained by iterative contextual filtering [38], (b) altered fingerprint where the central part of the fingerprint has been transplanted from a different friction ridge skin, (c) synthesized fingerprint by SFinGe [38], and (d) latent fingerprint from NIST SD27 [3]. The NFIQ value [124] and the *fingerprint-ness* score (F) of the orientation field are shown. NFIQ value of 1 indicates the highest quality and NFIQ of 5 is the lowest quality. The *fingerprint-ness* score ranges from 0 (the lowest) to 1 (the highest); if the proposed model cannot find a feasible solution satisfying the constraints, there is no valid output from the model.

extremely small compared to the population size.

Deterministic models represent fingerprints in a mathematical form, and are often designed to provide fingerprint-like textured pattern as an output of model fitting. Deterministic modeling does not require fingerprint images for learning; instead, a set of functions

incorporating the unique characteristics of fingerprint patterns is designed. Generally, the more constrained the model is to the fingerprint pattern, the less flexible it is to represent the actual pattern in a given fingerprint.

Since the most salient characteristic of a fingerprint pattern that distinguishes it from any other textured patterns is its orientation field, we propose a global orientation field model for fingerprints that is based on constrained ordinary differential equations (ODE) [140]. First, the rational polynomial model [119] which gives a nice topology of fingerprint ridge flow with singular points is converted to the ODE form whose phase portrait is the same as the orientation field from the rational polynomial function. It can be shown that the x-derivative and y-derivative of the ODE system for double-loop fingerprints (the most complicated case) are represented as the 4th order polynomial and the 3rd order polynomial, respectively, in geometric forms. Second, a generalized ODE model is proposed, which consists of much simpler polynomials and a smaller number of independent coefficients (i.e., 16 parameters in the model and 1 parameter for rotation) compared to the polynomial model in [66] and the constrained nonlinear phase portrait model in [89][1]. Third, the constraints on the number of singular points (either 0, 2, or 4) based on the Hermite's theorem are applied to reduce the parameter space during optimization. Finally, the degree of similarity of a given image to a fingerprint pattern, called *fingerprint-ness*, in terms of orientation field is measured by computing the difference between the orientation field extracted from the given fingerprint image and the orientation field fit by the proposed model.

This chapter is organized as follows. We survey on fingerprint orientation field models that have been proposed in the literature in Section 3.2. Section 3.3 describes a proposed global orientation field model which incorporates the distinctive characteristics of singular point in fingerprints. Section 3.4 summarizes this chapter and indicates how the orientation

---

[1]The numbers of coefficients in the polynomial model and the constrained nonlinear phase portrait model are $(n+1)(n+2)$ and $(n+1)(n+2) - 3N_s$, where n is the order of the polynomial and N_s is the number of singular points detected in the image (see Table 3.1). In [66], $n = 4$ was used, which involves 30 coefficients; in [89], $n = 8$ was used, which involves $90 - 3N_s$ coefficients. Note that N_s can be at most 4 in case of double-loop type of fingerprints.

field models are used in this dissertation.

3.2   Fingerprint Orientation Field Models

Fingerprint orientation field can be represented in two ways: (i) a set of local orientations, $\Theta = \{\theta(x,y)|(x,y) \in \mathcal{R}\}$, where $\mathcal{R}$ is the region of interest (ROI) of the fingerprint image, or (ii) a 2-dimensional function describing the global orientation field, $\theta(x,y) = f(x,y)$, where $f(x,y)$ is a real-valued function. Local orientations are extracted by finding the dominant ridge direction in each local block. For good quality fingerprints, the local orientation field extraction is sufficient to process the fingerprint images. On the other hand, global orientation field models are very useful to represent fingerprint ridge structure in noisy fingerprint images or fingerprint images with missing parts. As the orientation field representation in global models is much more compact (e.g., a coefficient vector) than the collection of local orientations at each location, some researchers have used the global orientation field representation for fingerprint classification and indexing [130]. In addition, if the global model is specifically designed to represent fingerprints, it can be used to synthesize realistic fingerprint images [38], simulate fingerprint pattern formation [84], and determine abnormality in the given image as a fingerprint pattern (e.g., distinguish non-fingerprint patterns [140] or detect altered fingerprints [139]).

3.2.1   Local Orientation Field Extraction

Ridge orientation in a local region is perpendicular to the gradient of the ridges. The gradient vector of a fingerprint image $I(x,y)$ is given by [34]:

$$\begin{bmatrix} G_x(x,y) \\ G_y(x,y) \end{bmatrix} = \begin{bmatrix} \frac{\partial I(x,y)}{\partial x} \\ \frac{\partial I(x,y)}{\partial y}) \end{bmatrix}. \tag{3.1}$$

The fingerprint orientation field is usually defined in local blocks instead of at each pixel

to make it robust to noise. The orientation in a block W can be computed by averaging the doubled gradients [34]:

$$\theta = \frac{1}{2} \tan^{-1} \left(\frac{G_{xx} - G_{yy}}{2G_{xy}} \right),$$ (3.2)

where

$$G_{xx} = \sum_{(x,y) \in W} G_x^2(x, y),$$

$$G_{yy} = \sum_{(x,y) \in W} G_y^2(x, y), \text{ and}$$

$$G_{xy} = \sum_{(x,y) \in W} G_x(x, y) G_y(x, y).$$

As the gradient-based orientation field estimation extracts a dominant orientation based on the maximum magnitude of the averaged gradients in a local block, it is liable to extract the orientations of structured noise rather than the fingerprint ridges if the local block contains strong creases in the fingerprint or lines in the background. To overcome this limitation, the orientation field extraction algorithm based on short-time Fourier transform (STFT) [43] obtains multiple orientations characterized by both orientation and frequency in a block. The Fourier transform of the fingerprint image in a local block can be represented by $F(r, \theta)$. A set of dominant orientations with mated frequency corresponds to the local maxima of the amplitude of $F(r, \theta)$, that is $|F(r, \theta)|$. The averaged dominant orientation in a local block also can be found by a probabilistic approximation. Considering the probability density function $p(r, \theta)$ in Equation (3.3), the marginal density functions for frequency r and orientation θ are defined in Equation (3.4). The expectation values for $p(r)$ and $p(\theta)$ are the average frequency and orientation in the block.

$$p(r, \theta) = \frac{|F(r, \theta)|^2}{\int_r \int_\theta |F(r, \theta)|^2},$$ (3.3)

Table 3.1: Global fingerprint orientation field models. Note that n is the order of the basis function or polynomial and N_s is the number of singular points detected in the image. If the model requires singular points as input, the number of singular points is not included in the number of parameters to be estimated.

Model	# Parameters	Singular Points	Fingerprint-specific Features
Approximation Methods			
Piecewise linear functions [128]	$10N_s$	Estimated	Not utilized
Polynomials [66]	$(n+1)(n+2)$	Estimated	Not utilized
Fourier series [130]	$2(2n+1)^2$	Estimated	Not utilized
Legendre polynomials [111]	$(n+1)(n+2)$	Estimated	Not utilized
Deterministic Mathematical Models			
Rational polynomial function [119]	1	Required	Not utilized
Constrained nonlinear model [89]	$(n+1)(n+2)-3N_s$	Estimated	Constraints on coefficients from singular points
Quadratic differentials [77]	5	Required	Arch-shaped global field
Peripheral models [129]	2 for each model[a]	Estimated	Arch-shaped global field
Constrained nonlinear model [140]	17	Not required	Number of singular points

[a] The number of coefficients here is only for the peripheral models, and does not include parameters in the Fourier expansion [130].

$$p(r) = \int_\theta p(r, \theta)d\theta,$$

$$p(\theta) = \int_r p(r, \theta)dr. \tag{3.4}$$

3.2.2 Global Orientation Field Models

Global orientation field models for fingerprints have been developed primarily based on two approaches: (i) approximation method and (ii) deterministic mathematical model (see Table 3.1).

3.2.2.1 Approximation Method

Approximation methods represent the global orientation field with a set of piecewise linear functions [128], a set of polynomials [66], or a set of basis functions such as Fourier series [130] or Legendre polynomials [111]. In Equation (1.1), x-derivative and y-derivative of ridge flow are represented by functions $f(x, y)$ and $g(x, y)$, respectively. The basis functions used in the approximation methods are as follows:

- Polynomial [66]:

$$f(x, y) = \sum_{k=0}^{n} \sum_{l=0}^{k} a_{kl} x^{k-l} y^{l}, \tag{3.5}$$

- Fourier series [130]:

$$f(x, y) = \sum_{k=0}^{K-1} \sum_{l=0}^{L-1} a_{kl} e^{i2\pi\left(\frac{k}{K}x + \frac{l}{L}y\right)}, \tag{3.6}$$

- Legendre polynomial [111]:

$$\begin{aligned} f(x, y) &= \sum_{k=0}^{n} \sum_{l=0}^{k} a_{kl} \phi_{kl}(x, y) \\ &= \sum_{k=0}^{n} \sum_{l=0}^{k} a_{kl} \phi_{k-l}(x) \phi_{l}(y), \end{aligned} \tag{3.7}$$

where

$$\phi_k(x) = \frac{1}{2^k k!} \frac{d^k}{dx^k} \left[(x^2 - 1)^k \right],$$

a_{kl} is a coefficient for each term in the basis function, n is the order of the basis function or polynomial, K and L are the width and height of a block. Note that $g(x, y)$ for y-derivative in each method is of the same form as $f(x, y)$, but with different coefficients.

The approximation method with piecewise linear functions proposed in [128] is a variation of the deterministic mathematical model with rational polynomial function [119]. From the orientation field modeled by the rational polynomial function shown in Equation (3.12)[2], a

---

[2]Equation (3.12) will appear in the next section
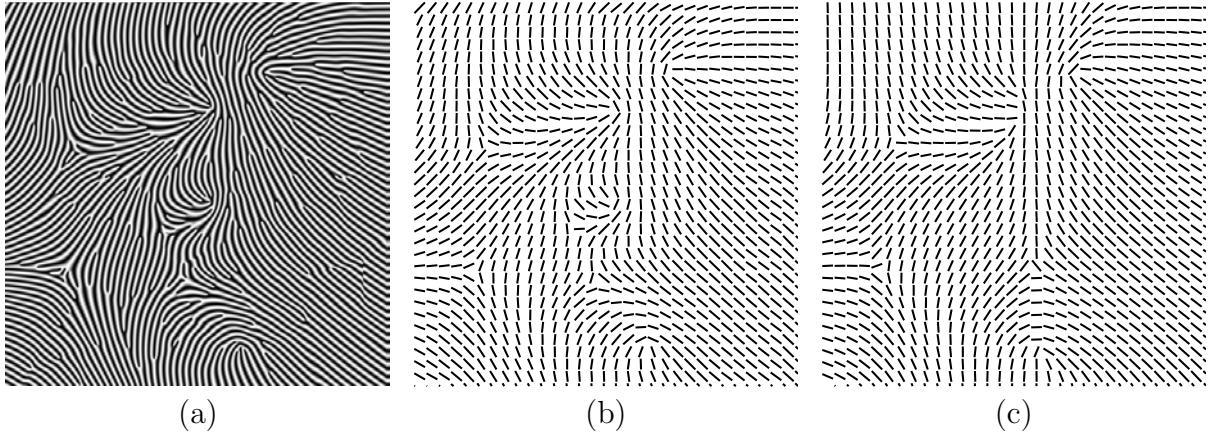
Figure 3.2: A synthesized image with fingerprint-like ridge structure and its orientation field estimation. (a) Synthesized image obtained by iterative contextual filtering [38], (b) orientation field extracted from the image in (a) using the gradient-based method [34], and (c) orientation estimated by an approximation method, the polynomial model [66].

set of piecewise linear functions, $\{f_k\}$ and $\{g_l\}$, is introduced to reflect the actual ridge flow in the model:

$$\theta(z) = \theta_\infty + \frac{1}{2}\left[\sum_{k=1}^{K} f_k(\arg(z - q_k)) - \sum_{l=1}^{L} g_l(\arg(z - p_l))\right]. \qquad (3.8)$$

The parameter space for the coefficients of the basis functions is often reduced by training the coefficient vectors on a large number of fingerprints. However, since the coefficient vectors are not invariant to rotation and translation, the location and orientation of the fingerprint pattern have to be determined first before projecting to the basis functions. In addition, when high order basis functions are used to represent fingerprints with high curvature (e.g., fingerprints with singular points), these models can fit well to virtually any flow pattern. Figure 3.2(c) shows the orientation field estimated by the polynomial model [66] with the 4th order polynomial for x-derivative and the 3rd order polynomial for y-derivative of an image with fingerprint-like ridge structure shown in Figure 3.2(a). The coefficients are found by the following objective function:

$$\min \sum_{(x,y) \in \mathcal{R}} \sin^2[\theta(x,y) - \hat{\theta}(x,y)], \tag{3.9}$$

where $\theta(x,y)$ is the orientation field extracted from the image by the gradient-based method, $\hat{\theta}(x,y)$ is the orientation field approximated by the polynomial model, and $\mathcal{R}$ is the ROI where the ridge structure is present. Since the polynomial model is not constrained for fingerprint patterns, it also fits well to the non-fingerprint patterns and results in an orientation field which is not likely to be found in fingerprints as an output of the model fitting.

3.2.2.2 Deterministic Mathematical Models

Deterministic mathematical models, on the other hand, incorporate the unique characteristics of fingerprint orientation field such as local field around singular points [77, 89, 119] and global field shaped like an arch [77, 129]. Sherlock *et al.* [119] proposed a rational polynomial function which gives topological behavior of the global fingerprint ridge flow induced by singular points. The rational polynomial function is defined as:

$$Q(z) = \frac{(z - q_1)(z - q_2) \cdots (z - q_K)}{(z - p_1)(z - p_2) \cdots (z - p_L)}, \tag{3.10}$$

where $z = x + iy$, $\{q_k\}_{1 \leq k \leq K}$ and $\{p_l\}_{1 \leq l \leq L}$ are zeros and poles, respectively, in the complex domain. Given this, the orientation field from $Q(z)$ is obtained by:

$$\theta(z) = \frac{1}{2} \arg \left(e^{i2\theta_\infty} \cdot Q(z) \right), \tag{3.11}$$

where θ_∞ is the orientation of the finger. Since the Poincaré indices are $+1$ at zeros and -1 at poles, cores and deltas in fingerprints correspond to zeros and poles in Equation (3.10), respectively. Equation (3.11) can be equivalently represented by:

$$\theta(z) = \theta_\infty + \frac{1}{2} \left[\sum_{k=1}^{K} \arg(z - q_k) - \sum_{l=1}^{L} \arg(z - p_l) \right]. \tag{3.12}$$

85

As the model in [119] requires only the location and type of the singularities, it does not reflect the actual ridge flow of the fingerprint. Also, in Sherlock *et al.*'s approach [119], there is no model for arch type of fingerprints, which does not contain any singular points.

Li *et al.* [89] proposed a constrained nonlinear phase portrait model which represents the x-derivative and y-derivative of the doubled orientation field with nth order polynomials as shown in Equation (3.5). Then, a set of constraints on the coefficients of the polynomials is derived from the local orientation field behavior around singular points. That is, the local orientation field with a singular point at (x_o, y_o) in the Cartesian coordinates can be represented by a polynomial system in the following way:

$$\dot{x} = \sum_{k=0}^{n} \sum_{l=0}^{k} a_{kl}(x - x_o)^{k-l}(y - y_o)^l = a_{00}^* + a_{10}^* x + a_{11}^* y + \tilde{f}(x, y),$$

$$\dot{y} = \sum_{k=0}^{n} \sum_{l=0}^{k} b_{kl}(x - x_o)^{k-l}(y - y_o)^l = b_{00}^* + b_{10}^* x + b_{11}^* y + \tilde{g}(x, y), \qquad (3.13)$$

where $\tilde{f}(x, y)$ and $\tilde{g}(x, y)$ are the functions containing the second order or higher polynomial terms. Since the linearization of the system preserves the local behavior around the singular point, the following approximation can be established:

$$a_{00}^* = 0, \ a_{10}^* = a, \ a_{11}^* = b,$$

$$b_{00}^* = 0, \ b_{10}^* = c, \ b_{11}^* = d, \qquad (3.14)$$

where a, b, c, and d are the elements of the characteristic matrix $\mathbf{A}$ in Equation (1.3).

In this model, the global orientation field is reconstructed by combining the orientation field computed from the image with the trained orientation field from a set of fingerprint images according to the coherence[3] of the local block; for example, if a local block has low coherence due to existence of singularity or noise, the orientation field from training data

---

[3]Coherence is a measure of consistency of orientations in a local region; coherence of 1 occurs when all the orientations point in the same direction while coherence of 0 occurs when the orientations point in all different directions so that the sum in the vector field cancels out.

has more weight than the orientation field extracted from the image for the block.

Huckemann *et al.* [77] proposed a model based on quadratic differentials that describes local fields around singular points (core, delta, and whorl), and global field simulating arch-shaped field. The local field around a singular point can be represented by $Q_{local}(z)dz^2 > 0$, where $z = x + iy$ and $dz^2 = \dot{z}^2$ which is the quadratic differential; the rational function $Q_{local}(z)$ for each singularity is as follows:

$$Q_{local}(z) = \begin{cases} z & \text{for a delta} \\ \frac{1}{z} & \text{for a core} \\ -\frac{1}{z^2} & \text{for a whorl.} \end{cases}$$

In general,

$$Q_{local}(z) = \frac{(z - Rq_1)(z - Rq_2)}{(z - Rp_1)(z - Rp_2)}, \tag{3.15}$$

where Rp_j and Rq_j, for $j = 1, 2$, are the locations of cores and deltas, respectively. If the fingerprint is a loop type (i.e., one core and one delta), the model has $p_2 = q_2$; in case of a whorl, $p_1 = p_2$.

The global field which resembles arch-like flow is modeled by locating virtual poles at $\pm R$:

$$Q_{global}^{(k)}(z) = \frac{1}{(z^2 - R^2)^{2k}}, \tag{3.16}$$

where k is the number of poles at each $+R$ and $-R$ that determines the curvature of the global field. The orientation field integrating local and global field is given by:

$$Q(z) = Q_{local}(z)\bar{Q}_{local}(z)Q_{global}^{(k)}(z), \tag{3.17}$$

where $\bar{Q}_{local}(z)$ is the complex conjugate of $Q_{local}(z)$ to ensure that the quadratic differentials are positive.

The orientation field can be represented by the solution curve of the quadratic differen-

tials, $Q(z)dz^2 > 0$:

$$\theta(z) = \left\{ \left(z, \frac{dz^2}{|dz|^2} \right) : Q(z) \neq 0 \text{ or } \infty \right\}. \tag{3.18}$$

The coordinates of the fingerprint (translation, rotation, scale, and aspect ratio of the axes) are estimated during the model fitting, while the singular points are required as input to determine the local orientation field.

Wang *et al.* [129] proposed two peripheral models: a modified cosine function used in [37] and a fluid model which describes the 2-dimensional flow of a uniform stream toward a cylinder.

Most deterministic mathematical models require information about singular points (location and type). The rational polynomial model [119] and the quadratic differential model [77] assume that actual location and type of the singular points are given. The constrained non-linear phase portrait model [89] estimates the singular points by computing the Poincaré index [34] of the orientation field obtained from image gradient. However, the singular point information is not typically available as a prior knowledge in practice and is extracted automatically from the image. Accurate detection of singular points is still a challenging problem, especially in noisy fingerprints. If the information about singular points is not correct or otherwise not available, the deterministic mathematical models give poor estimates of the orientation field.

3.3 Global Orientation Field Modeling

We propose a global orientation field model for fingerprints that is based on constrained ordinary differential equations (ODE) [140]. First, the rational polynomial model [119] which gives a nice topology of fingerprint ridge flow with singular points is converted to the ODE form whose phase portrait is the same as the orientation field from the rational polynomial function. It can be shown that the x-derivative and y-derivative of the ODE system for double-loop fingerprints (the most complicated case) are represented as the 4th

order polynomial and the 3rd order polynomial, respectively, in geometric forms. Second, a generalized ODE model is proposed, which consists of much simpler polynomials and a smaller number of independent coefficients (i.e., 16 parameters in the model and 1 parameter for rotation) compared to the polynomial model in [66] and the constrained nonlinear phase portrait model in [89][4]. Third, the constraints on the number of singular points (either 0, 2, or 4) based on the Hermite's theorem are applied to reduce the parameter space during optimization. Finally, the degree of similarity of a given image to a fingerprint pattern, called *fingerprint-ness*, in terms of orientation field is measured by computing the difference between the orientation field extracted from the given fingerprint image and the orientation field fit by the proposed model.

3.3.1 From Rational Polynomial Function to Ordinary Differential Equations

Sherlock *et al.* [119] proposed a fingerprint orientation field model with rational polynomial function which is given in Equation (3.10). The rational polynomial function can be converted to differential equations whose phase portrait is equivalent to the vector field from the rational polynomial function. For a loop type of fingerprint with a core at $z_c = x_c + iy_c$ and a delta at $z_d = x_d + iy_d$, the rational polynomial function is given by:

$$Q_{Loop}(z) = \frac{z - z_c}{z - z_d} = \frac{a_c e^{i\phi_c}}{a_d e^{i\phi_d}} = \frac{a_c}{a_d} e^{i(\phi_c - \phi_d)}, \tag{3.19}$$

where $a_c = [(x - x_c)^2 + (y - y_c)^2]^{1/2}$, $a_d = [(x - x_d)^2 + (y - y_d)^2]^{1/2}$, $\phi_c = \tan^{-1}\left(\frac{y - y_c}{x - x_c}\right)$, and $\phi_d = \tan^{-1}\left(\frac{y - y_d}{x - x_d}\right)$.

For simplicity, denote $u = \frac{u_y}{u_x} = \frac{y - y_c}{x - x_c}$ and $v = \frac{v_y}{v_x} = \frac{y - y_d}{x - x_d}$. Then, the phase part in

---

[4]The numbers of coefficients in the polynomial model and the constrained nonlinear phase portrait model are $(n+1)(n+2)$ and $(n+1)(n+2) - 3N_s$, where n is the order of the polynomial and N_s is the number of singular points detected in the image (see Table 3.1). In [66], $n = 4$ was used, which involves 30 coefficients; in [89], $n = 8$ was used, which involves $90 - 3N_s$ coefficients. Note that N_s can be at most 4 in case of double-loop type of fingerprints.

Equation (3.19) can be written as:

$$\phi_c - \phi_d = \tan^{-1} u - \tan^{-1} v$$

$$= \tan^{-1}\left(\frac{u-v}{1+uv}\right)$$

$$= \tan^{-1}\left(\frac{u_y v_x - u_x v_y}{u_x v_x + u_y v_y}\right). \tag{3.20}$$

Considering that the set of points where $\phi_c - \phi_d = \frac{(2k+1)\pi}{2}$, k is an integer, corresponds to the x-isocline, the x-isocline of the vector field for a loop is: $u_x v_x + u_y v_y = 0$, which is equal to:

$$(x - x_o)^2 + (y - y_o)^2 - r_o^2 = 0, \tag{3.21}$$

where

$$x_o = \frac{x_c + x_d}{2}, \ y_o = \frac{y_c + y_d}{2}, \ \text{and} \ r_o^2 = \frac{(x_c - x_d)^2 + (y_c - y_d)^2}{4}.$$

The x-isocline for the loop type of fingerprint corresponds to a circle.

Similarly, the set of points where $\phi_c - \phi_d = k\pi$, k is an integer, corresponds to y-isocline where $u_y v_x - u_x v_y = 0$ in Equation (3.20). Thus, the y-isocline of the vector field for a loop is:

$$ax + by + c = 0, \tag{3.22}$$

where $a = -y_c + y_d$, $b = x_c - x_d$, and $c = y_c x_d - x_c y_d$. The y-isocline for the loop type of fingerprint is a line. Figure 3.3 illustrates the orientation field of a loop model obtained from the rational polynomial function and its vector field with x-isocline and y-isocline.

For double-loop type fingerprints, the rational polynomial function with two cores at $z_{cj} = x_{cj} + iy_{cj}$ and two deltas at $z_{dj} = x_{dj} + iy_{dj}$ for $j = 1, 2$ is as follows:

$$Q_{DoubleLoop}(z) = \frac{(z - z_{c1})(z - z_{c2})}{(z - z_{d1})(z - z_{d2})}. \tag{3.23}$$

The x-isocline and y-isocline for double-loop fingerprints also can be represented in geo-

90

(a)

(b)
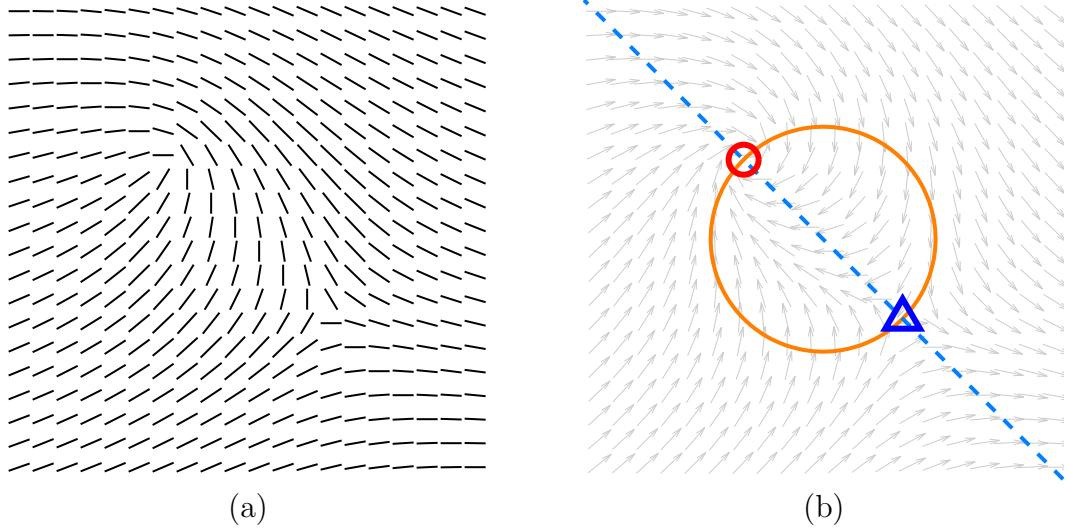
Figure 3.3: Loop model from the rational polynomial function. (a) Orientation field and (b) vector field with x-isocline (orange solid line) and y-isocline (blue dotted line) with singular points (a core marked as a circle and a delta marked as a triangle).

(a)

(b)
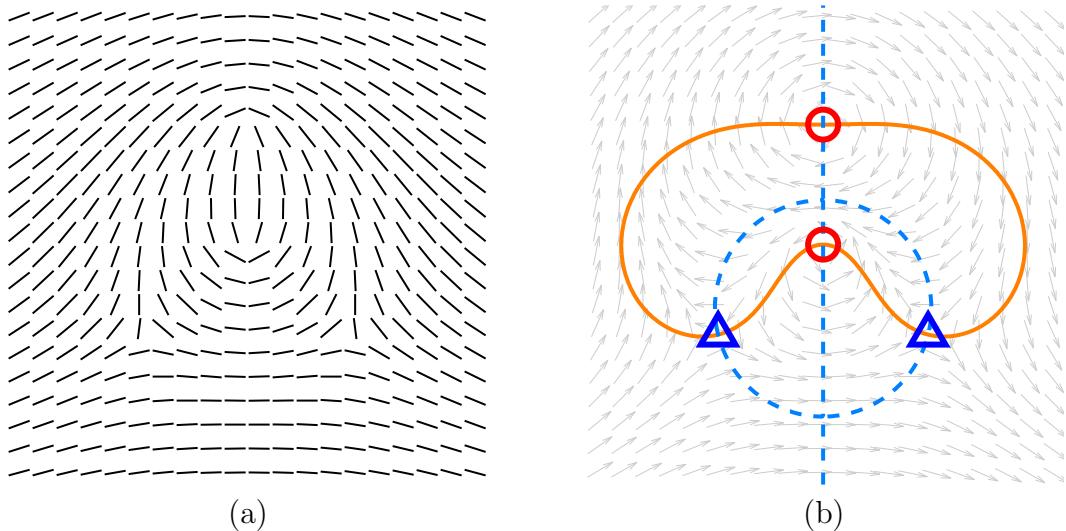
Figure 3.4: Double-loop model from the rational polynomial function. (a) Orientation field and (b) vector field with x-isocline (orange solid line) and y-isocline (blue dotted line) with singular points (cores marked as circle and deltas marked as triangle).

metric form. Denote $u_j = \frac{u_{jy}}{u_{jx}} = \frac{y - y_{cj}}{x - x_{cj}}$ and $v_j = \frac{v_{jy}}{v_{jx}} = \frac{y - y_{dj}}{x - x_{dj}}$, $j = 1, 2$, for cores and deltas, respectively. Then, the x-isocline corresponds to the curve of $(u_{1x}u_{2x} - u_{1y}u_{2y})(v_{1x}v_{2x} -$

91

$v_{1y}v_{2y}) + (u_{1y}u_{2x} + u_{1x}u_{2y})(v_{1y}v_{2x} + v_{1x}v_{2y}) = 0$, which is equal to:

$$(x^2 + y^2)\{(x - x_o)^2 + (y - y_o)^2 - r_o^2\} + (Ax^2 + Bxy - Ay^2 + Cx + Dy + E) = 0, \quad (3.24)$$

$$x_o = \frac{\alpha_1 + \alpha_2}{2}, \ y_o = \frac{\beta_1 + \beta_2}{2}, \ r_o^2 = \frac{(\alpha_1 + \alpha_2)^2 + (\beta_1 + \beta_2)^2}{4} - (\alpha_1\alpha_2 + \beta_1\beta_2),$$

$$A = \hat{\gamma}_1 + \hat{\gamma}_2, \ B = 2(\gamma_1 + \gamma_2),$$

$$C = -\alpha_1\hat{\gamma}_2 - \alpha_2\hat{\gamma}_1 - \beta_1\gamma_2 - \beta_2\gamma_1, \ D = -\alpha_1\gamma_2 - \alpha_2\gamma_1 + \beta_1\hat{\gamma}_2 + \beta_2\hat{\gamma}_1, \ E = \gamma_1\gamma_2 + \hat{\gamma}_1\hat{\gamma}_2,$$

where

$$\alpha_1 = x_{c1} + x_{c2}, \ \alpha_2 = x_{d1} + x_{d2}, \ \beta_1 = y_{c1} + y_{c2}, \ \beta_2 = y_{d1} + y_{d2},$$

$$\gamma_1 = x_{c1}y_{c2} + x_{c2}y_{c1}, \ \gamma_2 = x_{d1}y_{d2} + x_{d2}y_{d1}, \ \hat{\gamma}_1 = x_{c1}x_{c2} - y_{c1}y_{c2}, \ \text{and} \ \hat{\gamma}_2 = x_{d1}x_{d2} - y_{d1}y_{d2}.$$

The y-isocline for double-loop occurs when $(u_{1y}u_{2x} + u_{1x}u_{2y})(v_{1x}v_{2x} - v_{1y}v_{2y}) - (u_{1x}u_{2x} - u_{1y}u_{2y})(v_{1y}v_{2x} + v_{1x}v_{2y}) = 0$, which is equal to:

$$(x^2 + y^2)(ax + by + c) + (Ax^2 + Bxy - Ay^2 + Cx + Dy + E) = 0, \quad (3.25)$$

$$a = -\beta_1 + \beta_2, \ b = \alpha_1 + \alpha_2, \ c = -\alpha_1\beta_2 + \alpha_2\beta_1,$$

$$A = \gamma_1 - \gamma_2, \ B = -2(\hat{\gamma}_1 - \hat{\gamma}_2), \ C = \alpha_1\gamma_2 - \alpha_2\gamma_1 - \beta_1\hat{\gamma}_2 + \beta_2\hat{\gamma}_1,$$

$$D = -\alpha_1\hat{\gamma}_2 + \alpha_2\hat{\gamma}_1 - \beta_1\gamma_2 + \beta_2\gamma_1, \ E = \gamma_1\hat{\gamma}_2 - \gamma_2\hat{\gamma}_1.$$

The x-isocline and y-isocline for double-loop fingerprints also consist of geometric forms: line, circle, and hyperbola. Figure 3.4 shows the double-loop model from the rational polynomial function with x-isocline and y-isocline in the vector field.

92

3.3.2 Generalized Orientation Field Model

Based on the rational polynomial function converted into the ODE form, a generalized fingerprint orientation field model is derived from the double-loop model, which is the most complicated case. The generalized model is shown in Equation (3.26): $f(x, y)$ is the 4th order polynomial with a reduced number of polynomial terms and the coefficients of the 3rd order polynomial terms are correlated; $g(x, y)$ is the 3rd order polynomial with full polynomial terms, but the coefficients of the 3rd order polynomial terms are correlated. Thus, there are 8 coefficients each for $f(x, y)$ and $g(x, y)$.

$$\dot{x} = f(x, y)$$
$$= x^4 + 2x^2 y^2 + y^4 + ax^3 + bx^2 y + axy^2 + by^3 + Ax^2 + Bxy + Cy^2 + Dx + Ey + F,$$
$$\dot{y} = g(x, y)$$
$$= a'x^3 + b'x^2 y + a'xy^2 + b'y^3 + A'x^2 + B'xy + C'y^2 + D'x + E'y + F'. \tag{3.26}$$

3.3.2.1 Constraints on Singularity

As a global constraint which characterizes a fingerprint pattern, the number of singular points in the modeled orientation field needs to be restricted to 0, 2, or 4. Considering that the singular points occur when x-isocline and y-isocline intersect, we want to find a constraint on the number of real roots of the following polynomial system:

$$\begin{cases} f(x, y) = 0 \\ g(x, y) = 0. \end{cases}$$

According to Hermite's theorem, the constraint which specifies the number of real roots of a polynomial system can be found as follows [109]:

1. Find the Gröbner bases [30] of the polynomial system, G.

2. Construct a set of monomials that are not multiples of the leading terms of the polynomials in G, $M = \{m_1, m_2, \ldots, m_P\}$.

3. Compute normal forms[5] of each monomial in M multiplied by x and y; that is, for each monomial m_p in M, the normal forms of xm_p and ym_p are computed.

4. Construct coefficient matrices, $\mathbf{P}_x$ and $\mathbf{P}_y$, where the (i, j)-th entry of $\mathbf{P}_x$ is the coefficient of the term m_i in the normal form of xm_j (similarly, the (i, j)-th entry of $\mathbf{P}_y$ is the coefficient of the term m_i in the normal form of ym_j).

5. Construct a trace matrix $\mathbf{T}_1$ which is defined as follows:

$$\mathbf{T}_1 = \begin{bmatrix} Tr(m_1 m_1) & \cdots & Tr(m_1 m_P) \\ \vdots & & \vdots \\ Tr(m_P m_1) & \cdots & Tr(m_P m_P) \end{bmatrix}, \tag{3.27}$$

where $Tr(m_i m_j)$ is the trace of a matrix $\mathbf{P}_x^{k_i+k_j} \mathbf{P}_y^{l_i+l_j}$ when $m_i = x^{k_i} y^{l_i}$ and $m_j = x^{k_j} y^{l_j}$.

6. **Hermite's Theorem** The signature of $\mathbf{T}_1$ is equal to the number of real roots of the polynomial system [108]; the signature of a real symmetric matrix is the difference between the numbers of positive and negative eigenvalues.

3.3.2.2 Parameter Estimation

A foreground mask, $\mathcal{R}$, consists of the local blocks with sufficient dynamic range in image intensity [132]. The orientation field in $\mathcal{R}$ is computed by image gradients [34]. To ensure that the modeled orientation field accurately reflects the high curvature ridge structure, the absolute value of $|\mathbf{A}|$, where $\mathbf{A}$ is the characteristic matrix, in Equation (1.3) from linearization[6] is used in parameter estimation. Figure 3.5 shows the map of $|\mathbf{A}|$ for a given

---

[5]Given the Gröbner bases G of a polynomial system, the normal form of any polynomial is a polynomial which does not contain any term that is divisible by the leading terms of the polynomials in G after a finite number of reductions.

[6]Recall that the characteristic matrix in the linearized differential equations retains the local properties of the original differential equations.

(a)   (b)   (c)
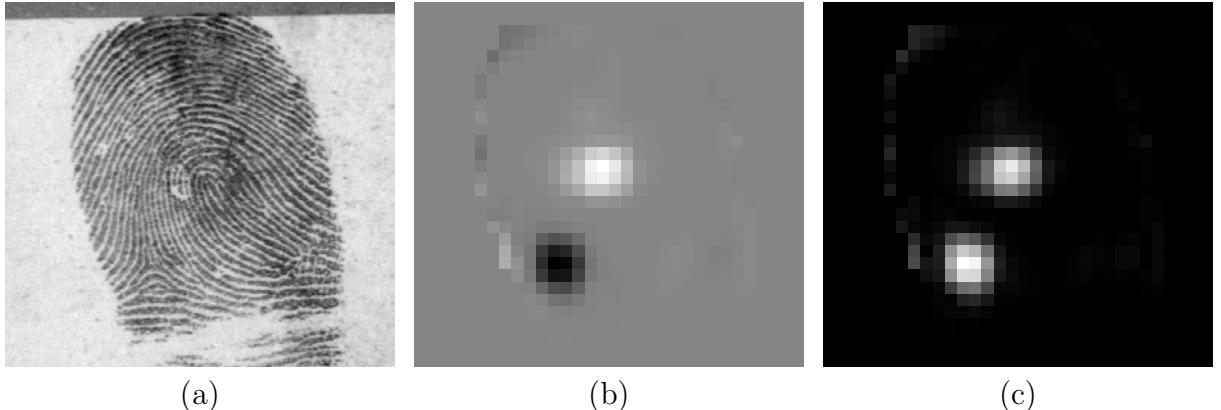
Figure 3.5: Weight matrix $c(x, y)$ used in Equation (3.28). (a) Input image, (b) the map of $|\mathbf{A}|$ where $\mathbf{A}$ is the characteristic matrix in Equation (1.3), and (c) the absolute value of (b).

image and the map of the absolute value of $|\mathbf{A}|$.

The objective function for parameter estimation is:

$$\min \sum_{(x,y)\in\mathcal{R}} \sin^2[\theta(x,y) - \hat{\theta}(x,y)] \cdot [\omega + (1-\omega)c(x,y)]$$

$$\text{subject to the signature of } \mathbf{T}_1 \text{ as } 0, 2, \text{ or } 4, \tag{3.28}$$

where $\theta(x, y)$ is the orientation field extracted from the image, $\hat{\theta}(x, y)$ is the modeled orientation field, ω is a constant ($\omega = 0.2$ is used), $c(x, y)$ is the map of the absolute value of $|\mathbf{A}|$, and $\mathbf{T}_1$ is the trace matrix in Equation (3.27).

While the origin of the coordinates is set to the center of the image (translation parameters are not needed), the rotation parameter needs to be estimated. The modeled orientation field with rotation angle r is obtained as follows:

$$\hat{\theta}(\mathbf{x}) = \frac{1}{2} \tan^{-1} \left(\frac{\mathbf{R}^2 g(\mathbf{R}^{-1}\mathbf{x})}{\mathbf{R}^2 f(\mathbf{R}^{-1}\mathbf{x})} \right), \tag{3.29}$$

95

where

$$\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix} , \text{ and } \mathbf{R} = \begin{bmatrix} \cos r & \sin r \\ -\sin r & \cos r \end{bmatrix}. \tag{3.30}$$

In total, 17 parameters, including a rotation parameter, are estimated for the model fitting. For the initial parameter estimation, the objective function in Equation (3.28) is used without applying the constraints. Then the parameters are estimated by applying one of the three constraints for the number of intersections (0, 2, and 4) at each time. Only the estimated parameters which do not violate the constraint are considered as valid. Figure 3.6 shows the orientation field estimated by the proposed model for various types of fingerprints.

3.3.3 Defining Fingerprint Pattern: *Fingerprint-ness*

One of the most interesting applications of the global fingerprint orientation field model is to determine if a given image (e.g., an input to AFIS or an image in an exemplar database) contains a *valid* fingerprint pattern. If a valid fingerprint image given as input is of good quality, a global orientation field model specifically designed to represent fingerprint patterns is expected to fit well to the orientation field of the input image. If the input image is a poor quality fingerprint which does not provide reliable orientation field as input to the model or the input is not a fingerprint image, the orientation field obtained from the model is expected to show a large difference from the orientation field of the image. Figure 3.7 shows an example of a good quality fingerprint image and its orientation field difference map (i.e., difference of image orientation field and modeled orientation field) and an example of a non-fingerprint image and its orientation field difference map.

As a measure of similarity of a given image to fingerprint pattern (i.e., *fingerprint-ness*) in terms of orientation field, the gray-level co-occurrence matrix (GLCM) [68] is computed from the orientation difference map. A 5-dimensional feature vector consists of contrast, energy, and homogeneity of the GLCM, and the mean and standard deviation of the orientation difference map in the foreground region $\mathcal{R}$. Prior to applying the proposed model to the
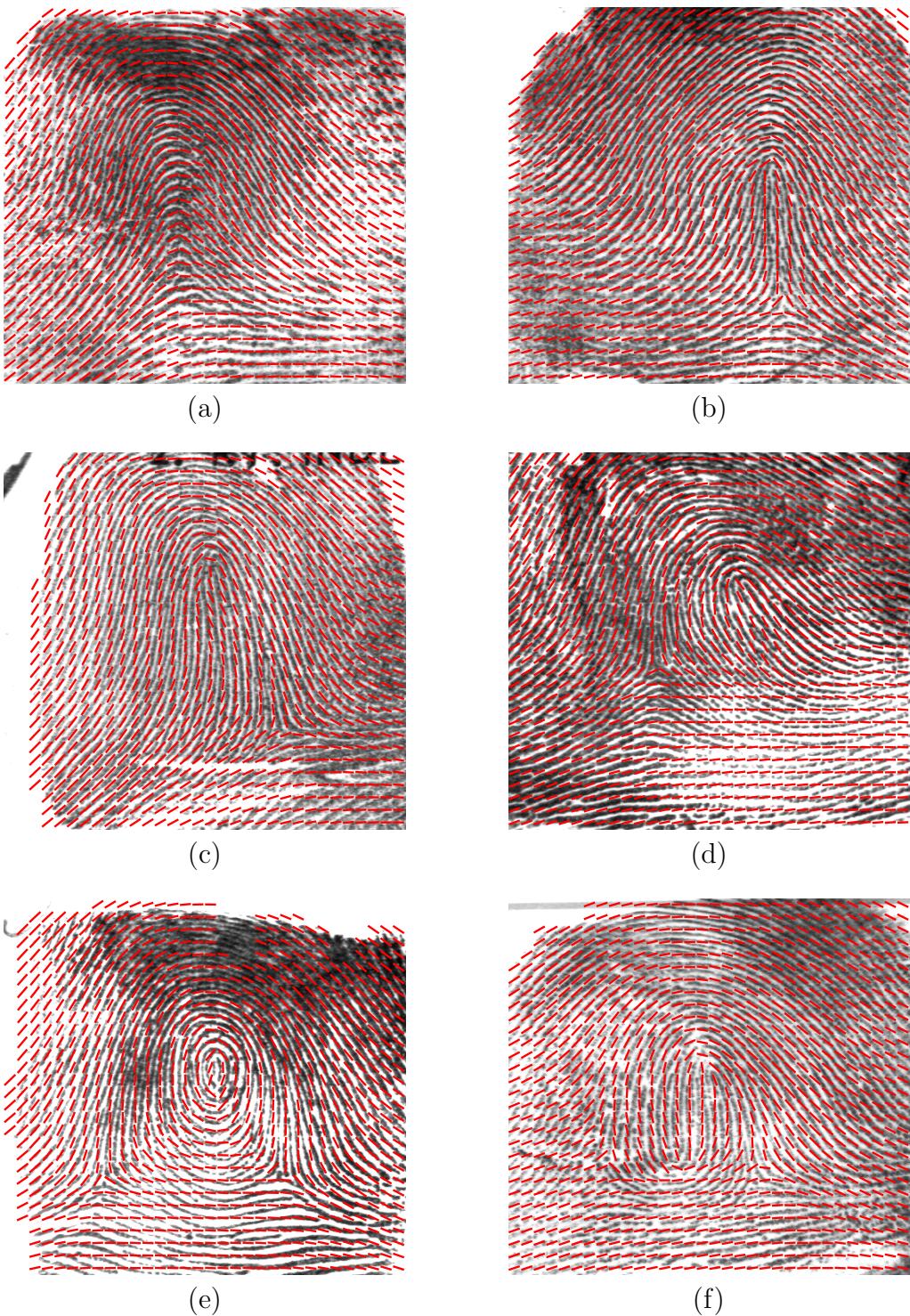
Figure 3.6: Orientation field fit by the proposed model for (a) arch, (b) tented-arch, (c) left-loop, (d) right-loop, (e) whorl, and (f) twin-loop type fingerprint.
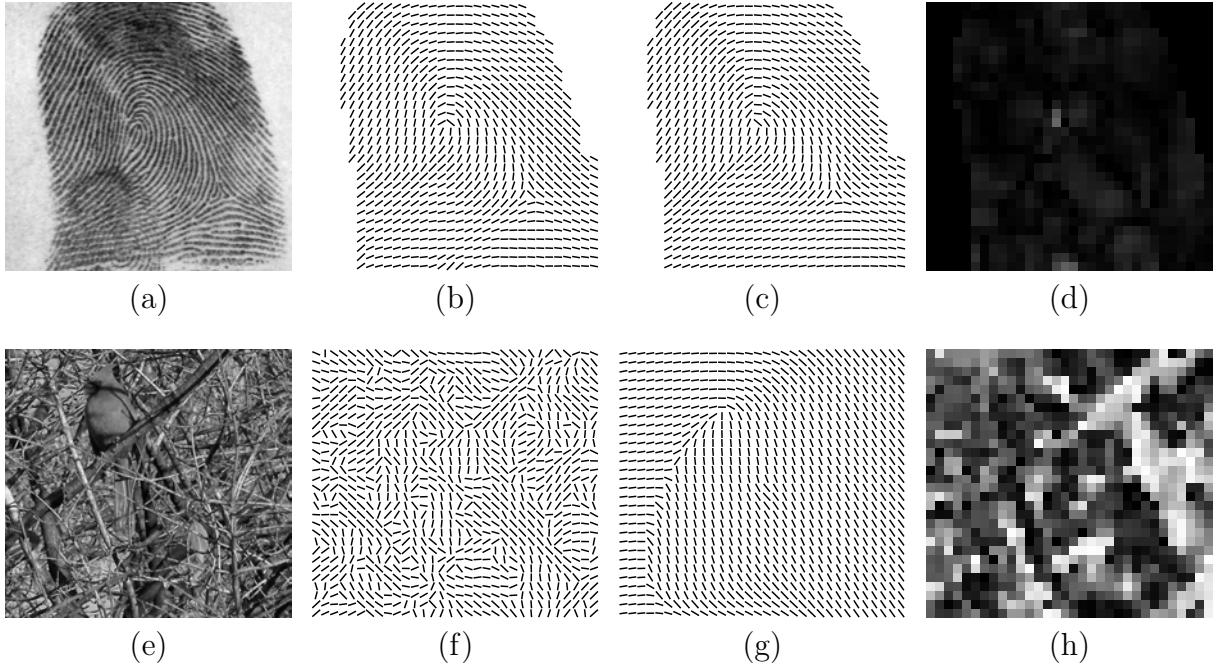
Figure 3.7: Orientation field difference map. (a)–(d) Input fingerprint image, orientation field from the image, orientation field from the model, and the orientation field difference map of the fingerprint image; (e)–(h) input non-fingerprint image, orientation field from the image, orientation field from the model, and the orientation field difference map of the non-fingerprint image.

orientation field of a given image, the image is rejected if $\mathcal{R}$ takes less than 30% of the entire image area. If the model does not find any feasible solution satisfying any of the three constraints on the number of singular points, the image is also rejected. For example, non-fingerprint images such as Figure 3.1(a) cannot be fit by the model satisfying the required number of singular points. Even for fingerprint images, if heavy structured noise is present in the image (e.g., the latent fingerprint image shown in Figure 3.1(d)), the model cannot find a good fit to the orientation field of the image.

3.3.4 Experimental Results

The proposed fingerprint orientation field model is evaluated as a detector of a valid fingerprint pattern in a given image. As a fingerprint database, 2,000 images of the first impressions
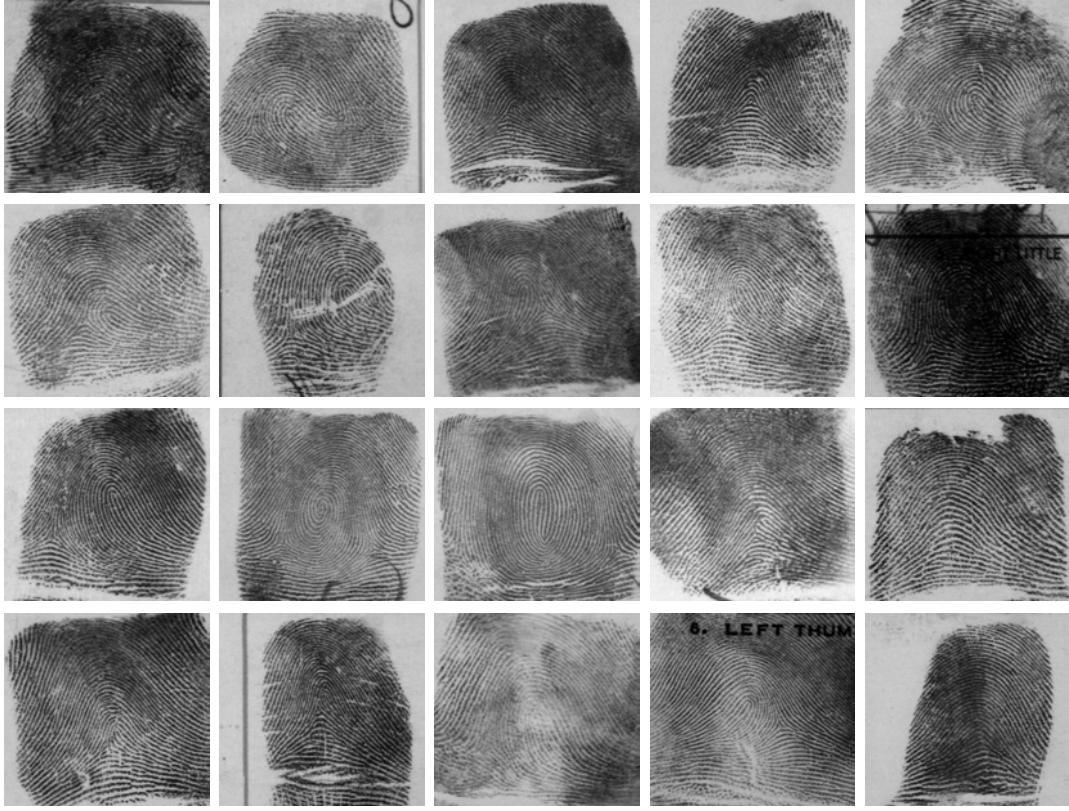
Figure 3.8: Fingerprint images in NIST SD4 [2].

in NIST SD4 [2] are used. As a non-fingerprint database, 2,000 images are drawn from the ImageNet dataset [48]; 100 object classes in the ImageNet dataset are randomly selected with 20 images per class. Figures 3.8 and 3.9 show the images used in the experiments.

The fingerprint images are set as a positive class while the non-fingerprint images are set as a negative class for a 2-class classification experiment. The 5-dimensional feature vector extracted from the orientation field difference map is used to determine if a given image is a valid fingerprint or not. The SVM with radial basis function [41] is used as a classifier and the performance is evaluated by 10-fold cross validation. Figure 3.10 shows the ROC curve with the average performance of the cross validation; the minimum and maximum performances are also indicated. Figures 3.11 and 3.12 show the images that are successfully classified based on the orientation field of the images.

Figure 3.13 shows false positive cases. If the model fit is wrong (Figure 3.13(a) is a double-
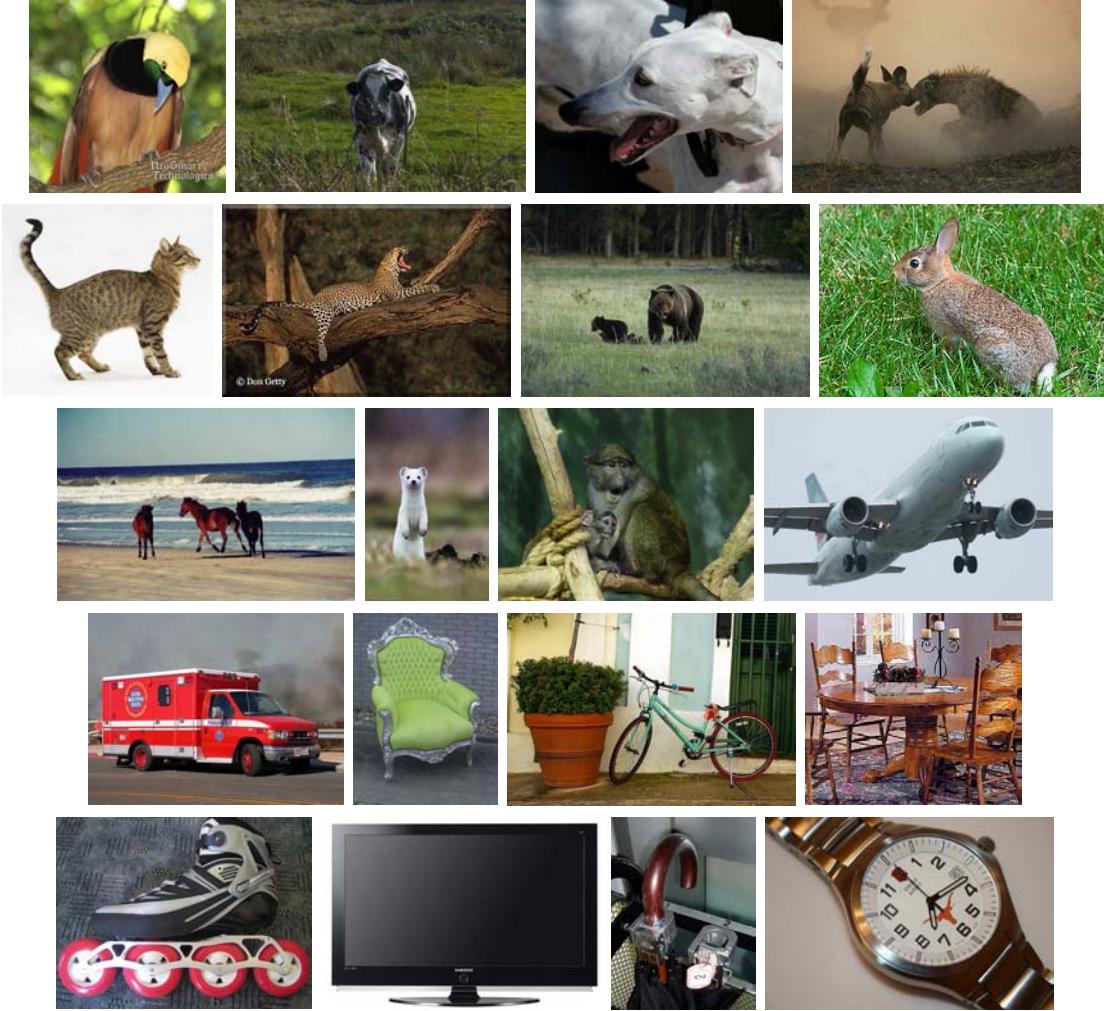
Figure 3.9: Non-fingerprint images in the ImageNet dataset [48].

loop fingerprint, but the model fits to a loop fingerprint), a high response in the orientation field difference map is observed. On the other hand, if the orientation field extracted from the image is noisy due to the poor quality of the fingerprint ridges (e.g., Figure 3.13(b)), the orientation field difference can also be high. However, the modeled orientation field for the partially noisy image in Figure 3.13(b) is still able to detect a fingerprint-like flow pattern. Figure 3.14 shows false negative examples where the images contain large areas of textured pattern with a smooth orientation field.
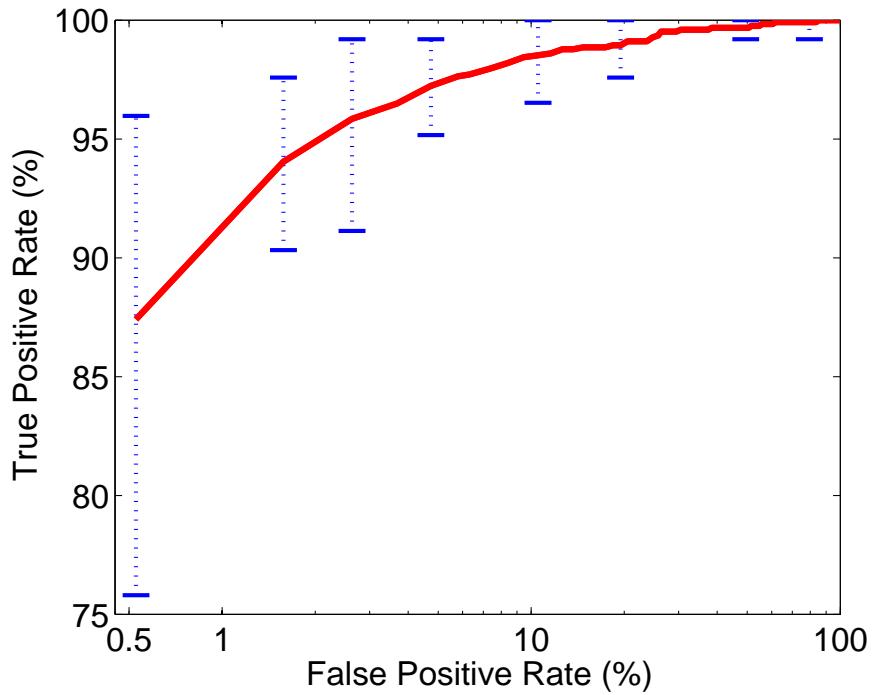
Figure 3.10: ROC curve. The average performance (red line) in the 10-fold cross validation is shown; the minimum and maximum performances (blue bars) over the 10 folds are also indicated.

3.4 Summary

Global fingerprint orientation field model is ideal for capturing unique characteristics of fingerprint patterns. It will assist in distinguishing fingerprints from any other texture pattern, representing the fingerprint orientation field in a compact form, and estimating the orientation field from noisy or partially missing fingerprints. There are two approaches to model the fingerprint orientation field: (i) approximation methods and (ii) deterministic mathematical models. In approximation methods, the global orientation field is represented by a set of basis functions, and the coefficient vectors are sometimes used as an index of the fingerprint. Deterministic mathematical models, on the other hand, are designed to specifically represent fingerprint ridge flow and incorporate unique characteristics of fingerprint orientation field into the model.

The global orientation field models published in the literature are highly dependent on
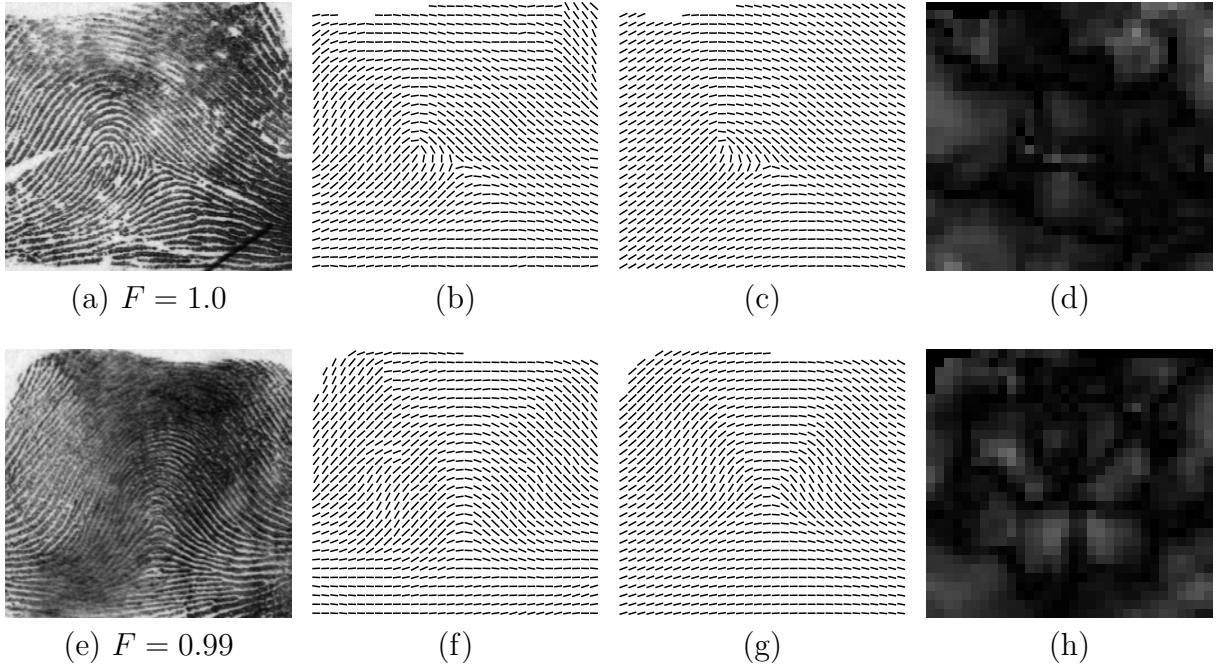
Figure 3.11: True positive examples. (a) and (e) Valid and good quality fingerprint images, (b) and (f) orientation field from the image, (c) and (g) modeled orientation field, and (d) and (h) orientation field difference map.

singular point information; if incorrect information about singularities is given or if the singular points are missing, the models are likely to fail to fit to the image. We propose an ODE representation of the global orientation field derived from the rational polynomial model and a set of constraints for the number of singular points in fingerprint is applied to estimate coefficients of the model. The proposed model is used to detect abnormality in the orientation field of a given image as a fingerprint and then classify it as a fingerprint or non-fingerprint pattern.

Global orientation field models also have been used in this dissertation for (i) latent fingerprint enhancement (for robust orientation field estimation in latents) in Chapter 4 and (ii) altered fingerprint detection (for abnormality measure in orientation field) in Chapter 5.

Future work on global orientation field modeling and fingerprint-ness measure includes:

1. Improve the global orientation field model by incorporating fingerprint-specific features

(a) $F = 0$ (b)   (c)   (d)

(e) $F = 0$ (f)   (g)   (h)
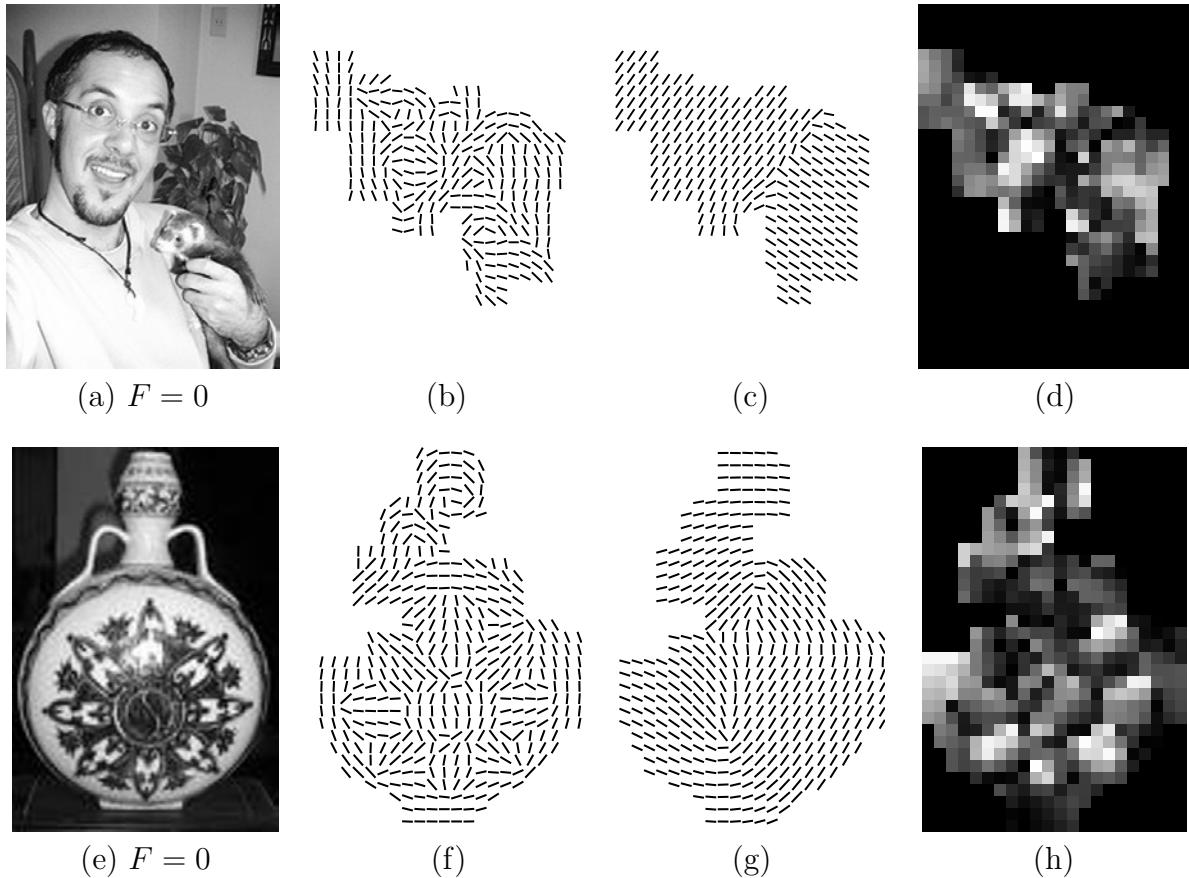
Figure 3.12: True negative examples. (a) and (e) Non-fingerprint image, (b) and (f) orientation field from the image, (c) and (g) modeled orientation field, and (d) and (h) orientation field difference map.

such as the configuration of singular points and arch-shaped global field.

2. Include features from ridge frequency and minutiae distribution into the definition of fingerprint-ness.

(a) $F = 0.43$ (b)   (c)   (d)

(e) $F = 0.59$ (f)   (g)   (h)
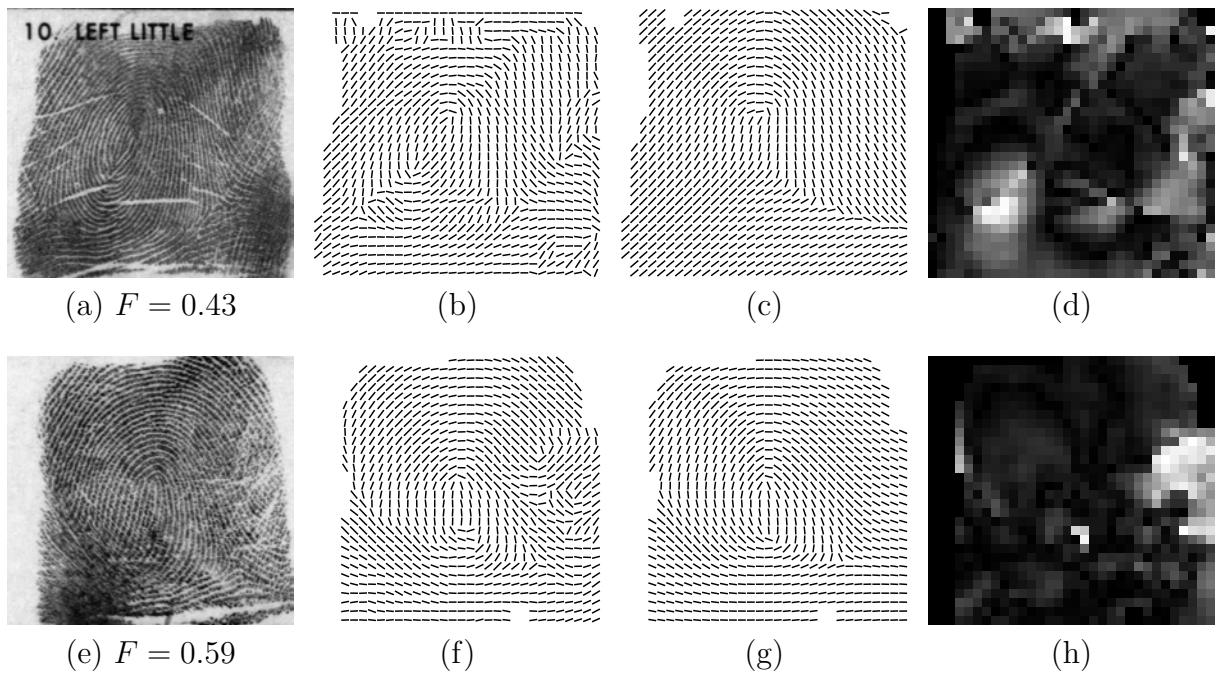
Figure 3.13: False positive examples. (a) and (e) Fingerprint images, (b) and (f) orientation field from the image, (c) and (g) modeled orientation field, and (d) and (h) orientation field difference map. The model estimates the fingerprint image in (a) as a loop fingerprint although it is indeed a double-loop fingerprint. The orientation difference map for the fingerprint image in (e) shows high response in the noisy region; the estimated orientation field from the model shows fingerprint-like behavior even in the noisy region.

(a) $F = 0.85$ (b)   (c)   (d)

(e) $F = 0.83$ (f)   (g)   (h)
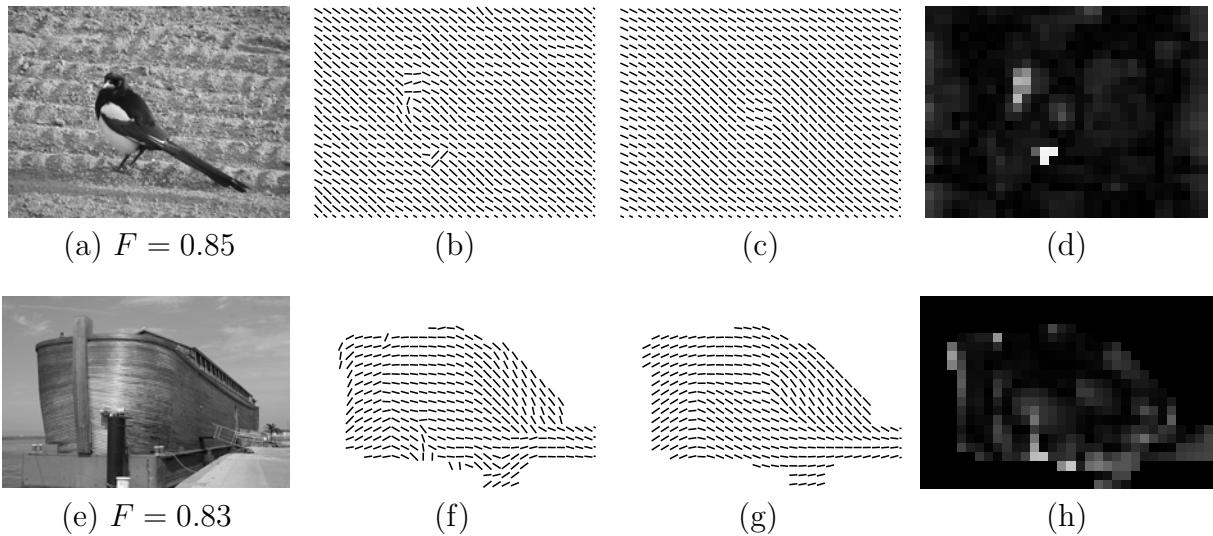
Figure 3.14: False negative examples. (a) and (e) Non-fingerprint image, (b) and (f) orientation field from the image, (c) and (g) modeled orientation field, and (d) and (h) orientation field difference map. Both images have textured background which provides a smooth orientation field.

Chapter 4

Latent Fingerprint Enhancement

4.1 Introduction

The success of AFIS in forensics and law enforcement applications is based on the availability of large legacy databases of rolled and plain prints of the ten fingers of all apprehended criminals. The size of these background databases is huge and is constantly increasing; for example, as of November 2012, the FBI's IAFIS fingerprint database contained the tenprints of about 73 million criminals. While the "lights-out" mode of operation for exemplar matching is a reality, a "lights-out" identification mode for latent search is still a research topic of high priority. One of the major goals of the FBI's NGI program [11] is to develop a latent matching system that will accept a latent image as input and return a short list of exemplar candidates from the database with no human intervention [79]. In a recent NIST report highlighting the results of ELFT-EFS [78], the best performing latent matcher achieved rank-1 identification rate of 63.4% in searching 418 latents against a 100,000 tenprint exemplar database when only latent images are given as input. As a comparison, the best performing matcher in verifying plain fingerprints against 10,000 plain fingerprints showed a true acceptance rate of 99.4% at a false acceptance rate of 0.01% in the 2003 FpVTE [135].

The primary focus in developing automated latent identification algorithms is to reduce

the human (latent examiners) intervention to as little as possible while preserving the matching accuracy that can be achieved with full manual markup. Automatic fingerprint feature extractors designed for rolled/plain prints do not give acceptable results on latents since many true minutiae are likely to be missed due to low ridge clarity and many spurious minutiae are likely to be extracted due to background noise.

We propose a semi-automatic latent fingerprint enhancement algorithm [137, 138] which requires minimal markup, namely ROI and singular points. Note that, in the proposed approach, a latent examiner is not required to mark minutiae or ridges, which demands significant effort. The objectives of the latent fingerprint image enhancement are to: (i) improve automatic feature extraction and matching performance of latent fingerprints, and (ii) provide visually enhanced images to latent examiners to enhance their ability to mark fingerprint features.

Estimating the orientation field of a latent fingerprint is a crucial stage in the latent enhancement algorithm. The significance of the proposed orientation field estimation algorithm in latents are: (i) extracting all dominant ridge orientations in a block, called *orientation elements*, considering the presence of background noise in the image, (ii) adopting randomized-RANSAC (R-RANSAC) algorithm [93], which is powerful in fitting a model to the dataset containing outliers, and (iii) using a global orientation field model to find the best hypotheses generated from the R-RANSAC algorithm. Once the orientation field is estimated, the quality of fingerprint ridges can be improved by Gabor filtering which enhances the local ridge clarity along the ridge orientation and suppresses noise in other directions.

The proposed latent fingerprint enhancement algorithm consists of the following four steps:

1. Manual markup of ROI and singular points.

2. Orientation element computation using the short-time Fourier transform (STFT) [43].

3. Orientation field estimation using R-RANSAC.

4. Fingerprint enhancement using Gabor filters [76].

4.2   Manual Markups of Latents

We assume that the following information is manually marked by latent examiners: (i) fingerprint region in the latent image, or the ROI, and (ii) singular points. ROI is a closed region that is bounded at the outermost trim of the latent. Figure 4.1 shows the boundary of ROI in the latent images. Only the features extracted in the ROI are regarded as valid.

Singularities observed in almost all the fingerprints fall into one of the following categories: (i) no singularity (i.e., arch type of fingerprints), (ii) one core and one delta (i.e., loop and tented-arch type), and (iii) two cores and two deltas (i.e., whorl and twin-loop type). Note that not all singular points may be observed in a given fingerprint image; for example, plain impression tends to capture only the central area of a finger. Based on this observation, latent examiners are expected to mark singular points using the following convention:

- If the latent does not contain any singularity, no singular points are marked (see Figure 4.1(a)). Even though the latent fingerprint may in fact come from a finger with singularity, the latent will be treated as plain arch, the simplest fingerprint pattern.

- If the latent contains singular points, cores and deltas should be marked in pairs. For example, the latent fingerprint in Figure 4.1(b) has one core and one delta in the ROI. These singular points are called *real* singular points since they are present in the ROI. On the other hand, when the numbers of core and delta are not equal, the missing singular points are marked with the examiner's best guess. These are called *virtual* singular points since they are not observable in the ROI. For example, the latent in Figure 4.1(c) contains only two real cores, and the two virtual deltas are also marked with the examiner's best guess.
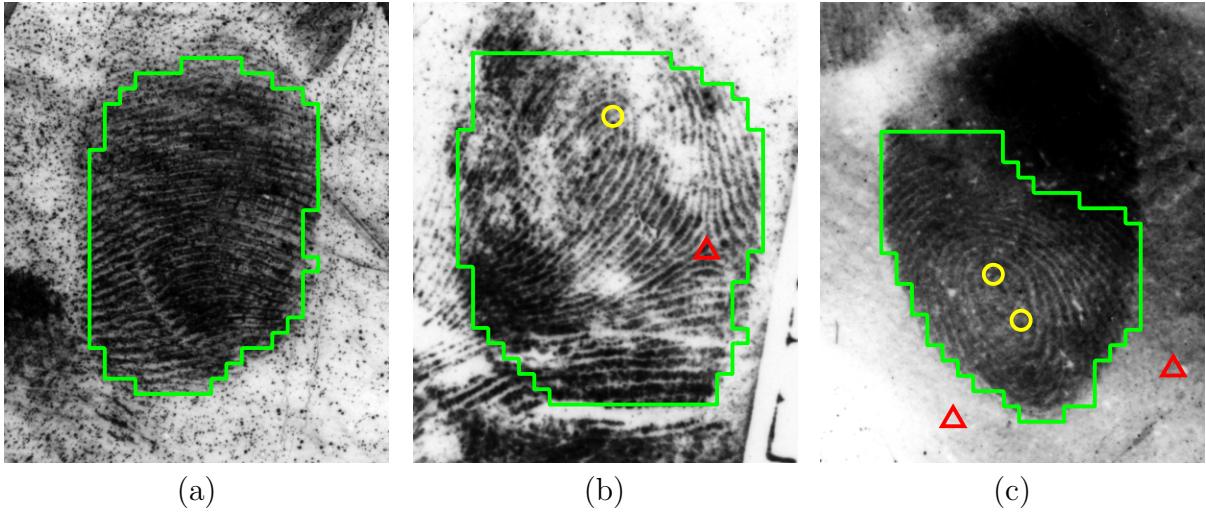
Figure 4.1: Manual markup of ROI and singular points. (a) No singularities presented in the image, (b) two real singular points, and (c) two real cores and two virtual deltas. The ROI is outlined by green boundary.

4.3 Orientation Field Estimation

In this stage, we use a three-level approach to estimate the orientation field of a latent: (i) the orientation elements are computed in each block in the ROI, (ii) the orientation elements in a neighborhood are merged into an orientation group whose elements are compatible with each other, and (iii) a robust estimate of the global orientation field is obtained by a set of orientation groups. The following sections describe the orientation field model and the details of the orientation field estimation algorithm.

4.3.1 Orientation Element Computation

Fingerprint ridges in a local image block generally contain only one dominant orientation if the fingerprint image is of good quality. However, in latent fingerprints, multiple orientations may appear in a block due to the structured background noise such as lines (see Figure 4.2(a)). A popular approach to compute multiple dominant orientations (orientation elements) in a block is based on the STFT [43], which detects the local maxima in the mag-

(a)   (b)
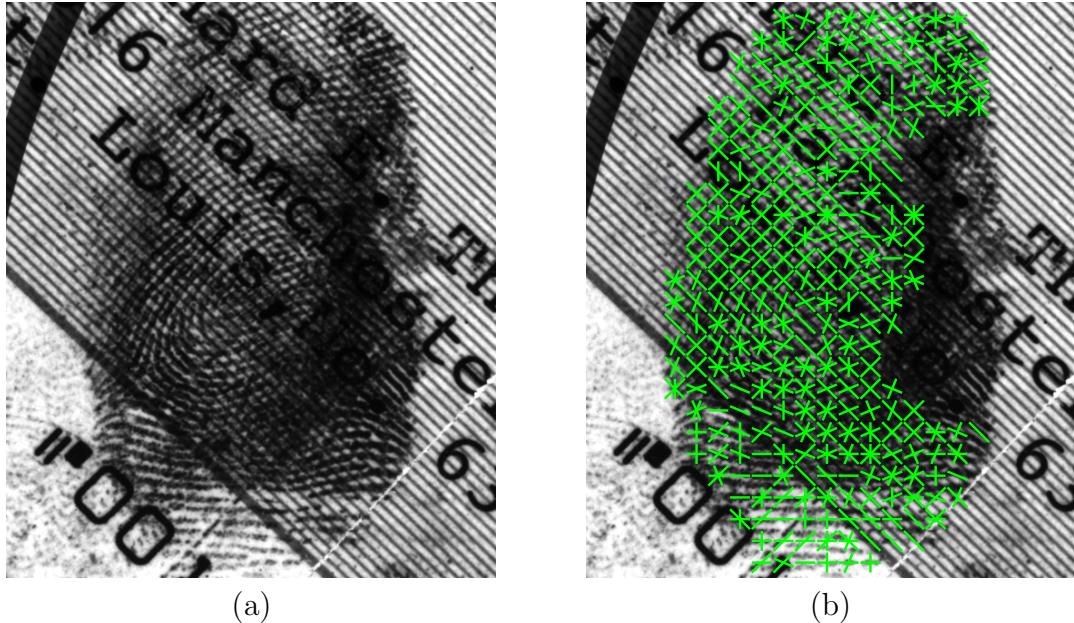
Figure 4.2: Orientation elements extracted from a latent using STFT. (a) Latent (NIST SD27, G027) and (b) orientation elements.

nitude spectrum of the local image. As the local maximum corresponding to the fingerprint ridges can be easily confused with the local maxima corresponding to the structured noise, we determine the true ridge orientation in the next stage. Figure 4.2(b) shows the extracted orientation elements in each block in the ROI of the latent image in Figure 4.2(a).

For each local block of 16×16 pixels in the ROI, the orientation elements are computed as follows:

1. Compute the STFT of a local image block (see Figures 4.3(a) and (b)).

2. Multiply the magnitude spectrum of the image with a set of directional filters (Figure 4.3(c)) in Fourier domain; each directional filter is constructed by multiplying a binary mask (1 for $\theta \in \left[\frac{\pi}{L} \left(i - \frac{1}{2} \right), \frac{\pi}{L} \left(i + \frac{1}{2} \right) \right)$, $i = 0, 1, \ldots, L - 1$; 0, otherwise) and a band-pass filter whose frequency range includes possible fingerprint ridge frequencies. $L = 8$ is selected.

3. Sum the energy from each filtered response (Figure 4.3(d)).
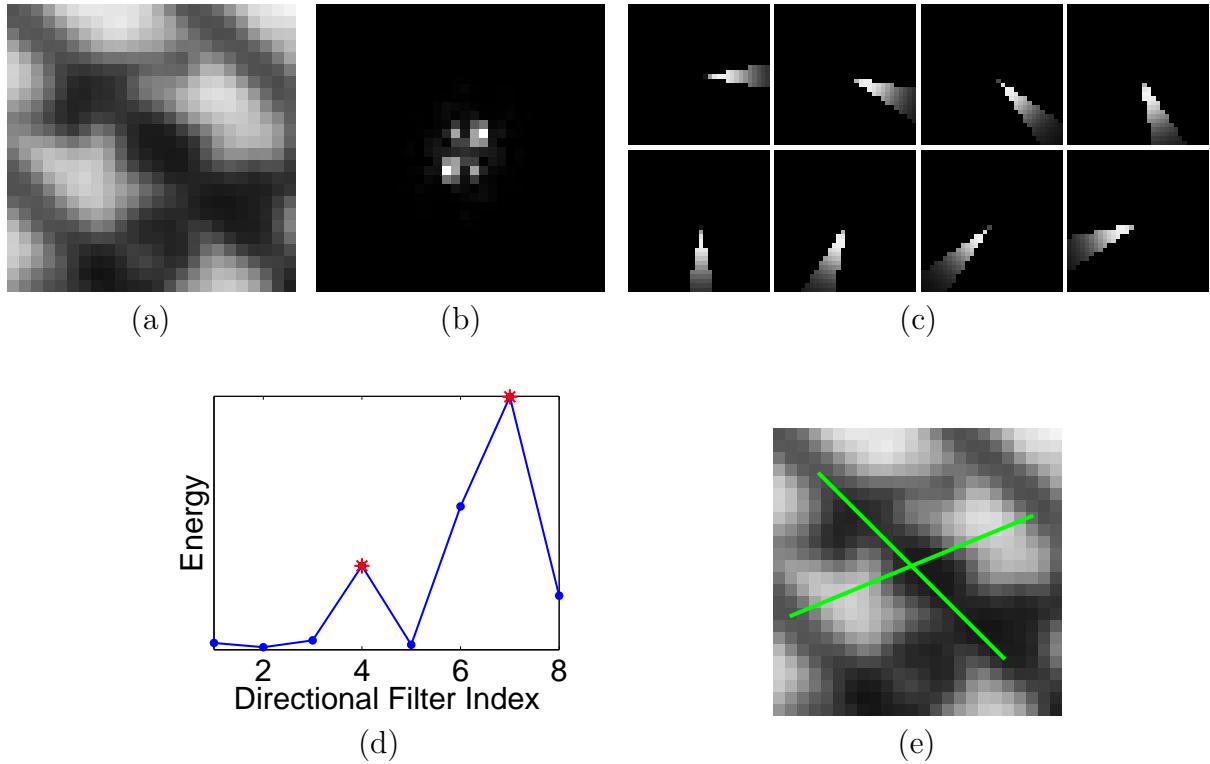
Figure 4.3: Orientation element computation. (a) A local image block, (b) magnitude spectrum of (a), (c) a set of directional filters, (d) energy of filtered responses by each directional filter, and (e) two orientation elements in this local block that correspond to the two local maxima in (d).

4. Find peaks (local maxima) in the one-dimensional energy plot. Orientation corresponding to the peak from the i-th directional filter is $\left[\frac{\pi}{L}i + \frac{\pi}{2}\right] \mod \pi$.

Figure 4.3(e) shows two orientation elements extracted from the image block in Figure 4.3(a).

4.3.2 Orientation Field Model

Orientation field, $\theta(x, y)$, of a fingerprint can be decomposed into two components:

$$\theta(x, y) = [\theta_s(x, y) + \theta_r(x, y)] \mod \pi, \tag{4.1}$$

111

where $\theta_s(x, y)$ represents the singular orientation field and $\theta_r(x, y)$ represents the residual orientation field.

Singular orientation field, $\theta_s(x, y)$, describes the abstract ridge flow determined by only the singular points where the fingerprint orientation field changes abruptly. The singular orientation field is obtained based on the rational polynomial function [119]:

$$\theta_s(z) = \left[\frac{1}{2} \arg \left(e^{i2\theta_\infty} \frac{(z - q_1)(z - q_2) \cdots (z - q_K)}{(z - p_1)(z - p_2) \cdots (z - p_L)} \right) \right] \bmod \pi, \tag{4.2}$$

where $z = x + iy$, $\{q_k\}$ and $\{p_l\}$ are the positions of cores and deltas, and θ_∞ is the orientation of the finger.

Residual orientation field, $\theta_r(x, y)$, represents natural changes in ridge flow of a fingerprint that are not influenced by the singularities; it is continuous everywhere. A set of polynomials can represent the x-derivative and y-derivative of the doubled residual orientation field as follows [66]:

$$\dot{x}_r = f_n(x, y) = \sum_{k=0}^{n} \sum_{l=0}^{k} a_{kl} x^{k-l} y^l, \tag{4.3}$$

$$\dot{y}_r = g_n(x, y) = \sum_{k=0}^{n} \sum_{l=0}^{k} b_{kl} x^{k-l} y^l, \tag{4.4}$$

where $\{a_{kl}\}$ and $\{b_{kl}\}$ are the polynomial coefficients for $f_n(x, y)$ and $g_n(x, y)$, respectively. The order of the polynomials is set as 4 (i.e., $n = 4$).

Figure 4.4 shows an example where the orientation field model is applied to a fingerprint.

4.3.3 Orientation Element Grouping

Neighboring orientation elements which are compatible with each other are grouped in order to facilitate the subsequent global orientation field estimation. Based on that the residual orientation field is continuous, the compatibility of the orientation elements is defined in the residual orientation field, and a unique label is assigned to each group. Let $\theta_r^{(i)}(x, y)$ be the i-th residual orientation element in a block at (x, y), and $\mathcal{L}^{(i)}(x, y)$ be the label assigned to

112
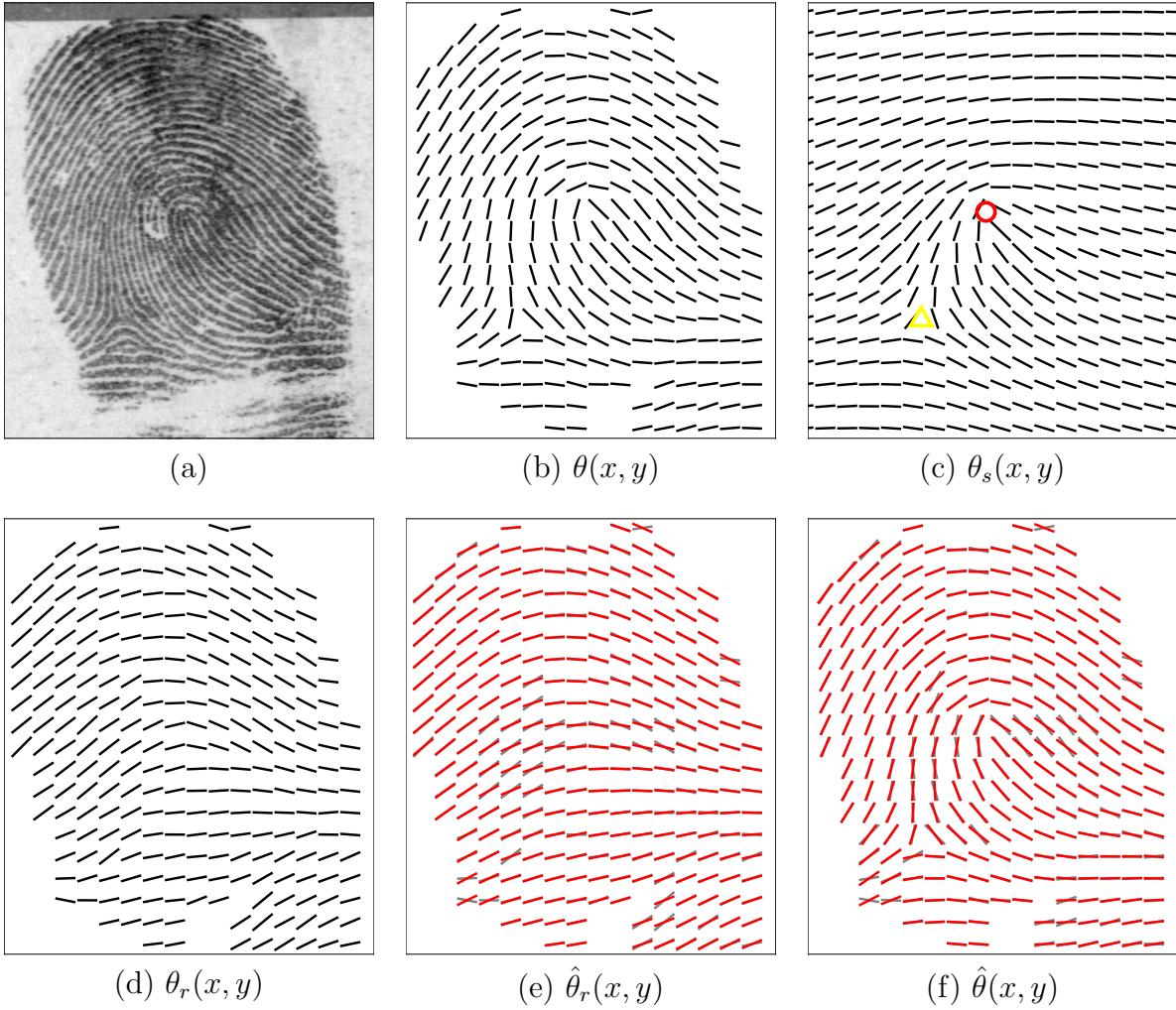
Figure 4.4: Orientation field model combining singular orientation field and residual orientation field. (a) Fingerprint (NIST SD4, F0004), (b) orientation field extracted from (a), (c) singular component of the orientation field, (d) residual component of the orientation field, (e) modeled residual orientation field using the 4th order polynomial model, and (f) reconstructed orientation field, $\hat{\theta}(x,y) = \theta_s(x,y) + \hat{\theta}_r(x,y)$.

this orientation element. Note that the number of orientation elements in a block varies from 1 to L. Initially, $\mathcal{L}^{(i)}(x,y) = 0$ for all i in a block at (x,y). Orientation element grouping starts from the elements which are the only dominant orientation in a block; they become seed points. For a seed point $\theta_r^{(i)}(x,y)$, an orientation element in a neighboring block $\theta_r^{(j)}(x',y')$ is deemed as a candidate to belong to the same group as this seed point if the following three conditions are satisfied:
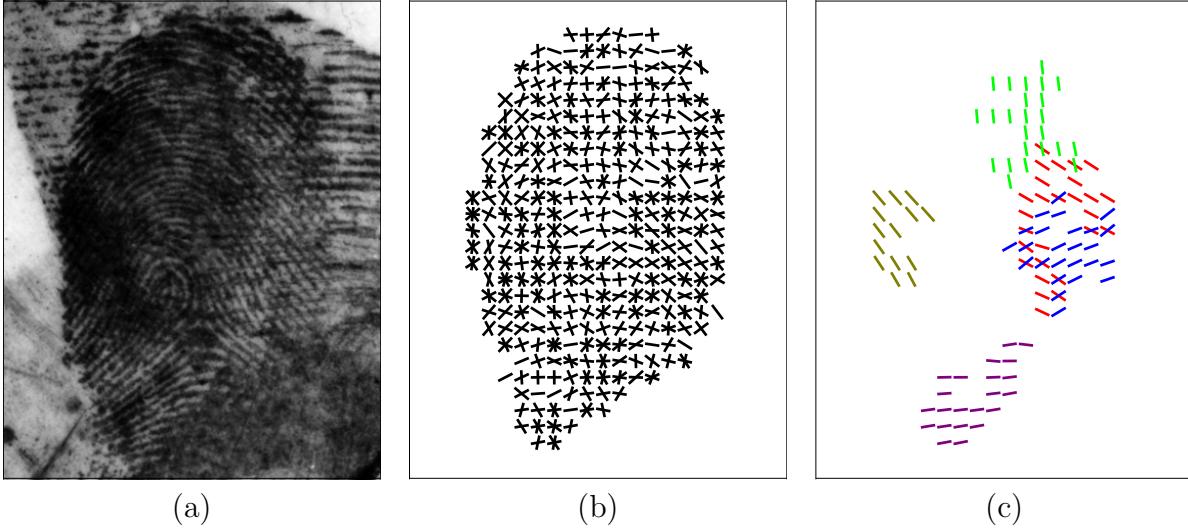
Figure 4.5: Orientation element grouping. (a) Fingerprint (NIST SD27, G017), (b) residual orientation elements, and (c) five orientation element groups (each shown in a different color).

(i) $\mathcal{L}^{(j)}(x', y') = 0$: no label is yet assigned to this neighboring orientation element;

(ii) The distance between two blocks is within 5 blocks; and

(iii) $|\theta_r^{(i)}(x, y) - \theta_r^{(j)}(x', y')| < \pi/2L$: the orientation difference is less than a threshold.

The candidates which are 4-connected to the seed point (i.e., a 4-connected path can be found among the set of candidate blocks) are assigned the same label as the seed point. Once all the seed points are grouped, the orientation elements which are the only one unlabeled element in a block become new seed points. This procedure is repeated until every orientation element has been assigned a label. Figure 4.5 shows five of the orientation element groups in a latent fingerprint.

4.3.4 Hypotheses Generation

Given a set of orientation element groups as the input, hypotheses for the residual orientation field are built based on the R-RANSAC algorithm [93]. RANSAC algorithms produce a number of hypotheses during iterations, and an iteration generally consists of three basic

steps: (i) select a set of initial data points randomly, (ii) build a hypothesis, and (iii) evaluate the hypothesis. A set of data points that are consistent with a given hypothesis is called the *consensus set*. The differences between R-RANSAC [93] and the basic RANSAC [58] are: (i) the number of initial data points selected in R-RANSAC is more than the minimum number of points required to build a hypothesis, and (ii) the hypothesis evaluation against all data points is conducted only if all initial data points are consistent with the hypothesis.

Let G be the set of all orientation elements and N be the total number of the orientation groups. An orientation group is randomly selected from G based on the following probability density function (pdf) which is constructed from the area of each group. The probability for random selection, p_i, of the i-th group with area A_i is defined as:

$$p_i = \frac{A_i}{\sum_{j=1}^{N} A_j}. \tag{4.5}$$

Orientation elements in the selected group are added to a set S; the size of S increases monotonically during an iteration. Initially, non-overlapping orientation groups are selected randomly according to the pdf of the size of groups until the size of S is greater than the minimum number of data points (m) to build a hypothesis; m is 15 for the 4th order polynomial.

Once the size of S is greater than m, a hypothesis is built using all orientation elements in the current S. The hypothesized model refers to the coefficients of the polynomials, $\{a_{kl}\}$ and $\{b_{kl}\}$, in Equations (4.3) and (4.4), which are obtained using the least-squares estimation by minimizing

$$\sum_{\theta_r^{(i)}(x,y)\in S} |f_n(x,y) - \cos 2\theta_r^{(i)}(x,y)|^2, \tag{4.6}$$

$$\sum_{\theta_r^{(i)}(x,y)\in S} |g_n(x,y) - \sin 2\theta_r^{(i)}(x,y)|^2. \tag{4.7}$$

If all orientation elements in S are within some error tolerance of the hypothesis (which is set to π/L), the consensus set S^* is determined by testing the hypothesis against all orientation elements in G and this hypothesis is recorded if the size of the S^* is the largest in

115

this iteration. The current iteration continues by adding a new non-overlapping orientation group to S. Figure 4.6 shows the procedure of hypothesis generation for one iteration.

Suppose not all orientation elements in S are consistent with the hypothesis. Then, one of the following two actions is taken: (i) if the inconsistent group is the most recently added one, remove this group from S, add a new group to S, and continue the current iteration; (ii) otherwise, terminate the iteration and evaluate the best hypothesis at this iteration. Action (ii) is also taken when no more non-overlapping orientation groups can be added to S.

The best hypothesis at each iteration is verified by checking if the hypothesized residual orientation field includes any singularity in the ROI. Singularities are found using the Poincaré index [83]. The hypothesis is accepted if there is no singularity present in the ROI. Figure 4.7 shows an accepted hypothesis, and the reconstructed orientation field combining the hypothesized residual orientation field with singular orientation field represents the true ridge flow well. If the hypothesized residual orientation field contains any singularity, it is rejected (see Figure 4.8).

The algorithm terminates when the number of iterations exceeds (i) a predetermined maximum number of trials, or (ii) the minimum number of trials k satisfying [58]

$$k \geq \frac{\log(1 - p_f)}{\log(1 - \varepsilon^m)}, \tag{4.8}$$

where p_f is the desired probability of having at least one S that all the orientation groups in S are from the target latent fingerprint and ε is the true fraction of orientation elements from the latent in G. Since ε is typically unknown, it is updated with the size of the best consensus set during the iterations.

Once the top-10 best candidate hypotheses for the orientation field are found, the latent fingerprint is enhanced using Gabor filters [76], which is described in the following section. The orientations of the Gabor filters are tuned to each orientation field hypothesis and their frequencies are set to a fixed value (1/8 ridges per pixel). All ten enhanced latents are input

Figure 4.6: One iteration of hypothesis generation procedure for the latent in Figure 4.5(a). An orientation group is added to S at each step from (1) to (6) (newly added groups are marked as red and existing groups in S are marked green in the left-hand side image of each pair); a hypothesis is built at every update in S and checked to see if all groups in S are consistent with the hypothesis (red blocks in the right-hand side image in each pair indicate consistency in S with the hypothesis). The final hypothesis M is obtained at the end of the iteration when no more groups can be added to S. Then, M in (6) is evaluated by checking the existence of singularity (see Figure 4.7).

(a)   (b)
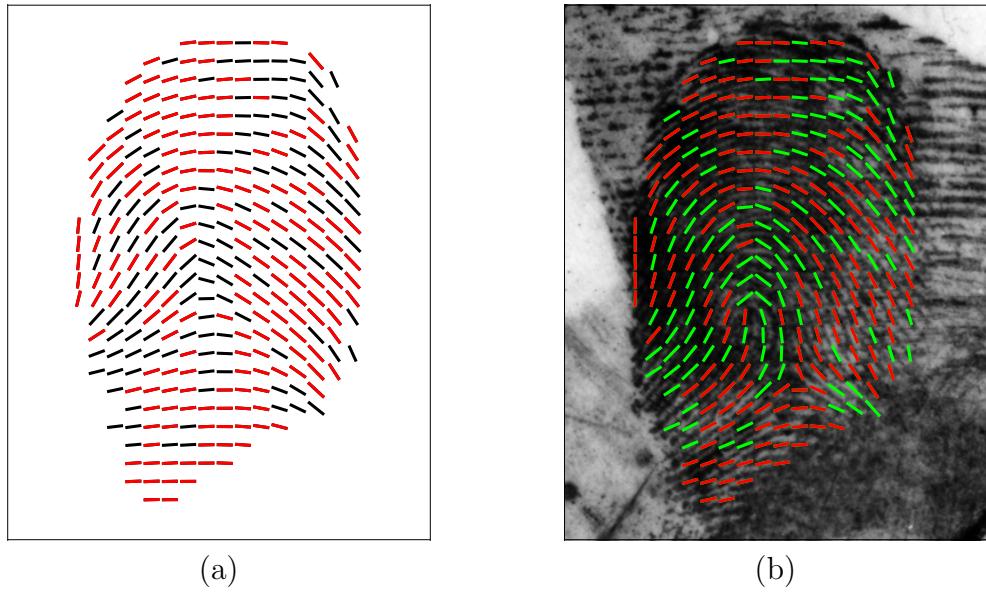
Figure 4.7: Accepted hypothesis. (a) Accepted hypothesized residual orientation field and (b) reconstructed orientation field using (a). Red blocks in (a) and (b) indicate the consensus set.

(a)   (b)
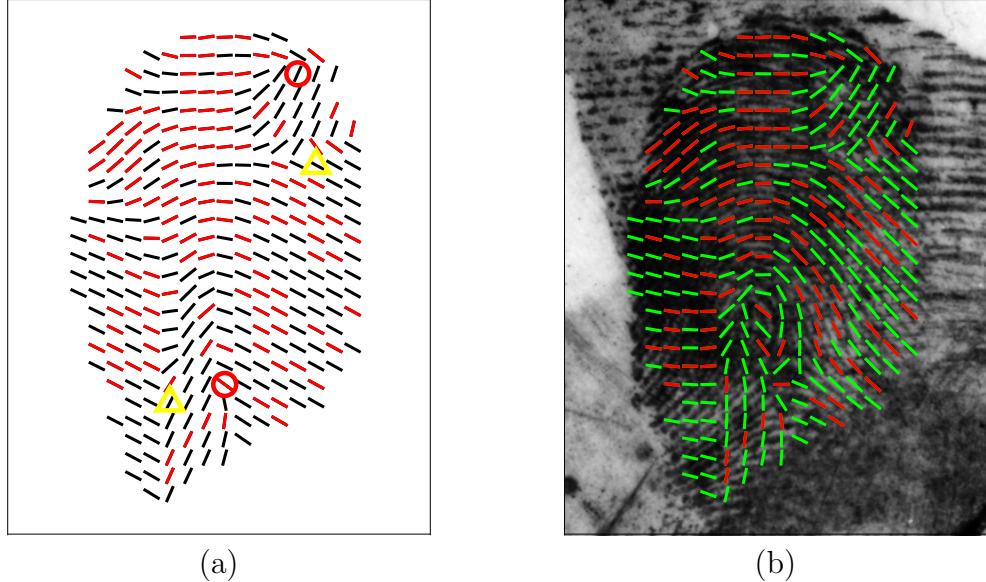
Figure 4.8: Rejected hypothesis. (a) Rejected hypothesized residual orientation field with singularities (red circle for core and yellow triangle for delta) and (b) reconstructed orientation field using (a).

to the fingerprint matcher. Then, the ten match scores output by the matcher are fused by the max rule [115].

4.4 Fingerprint Enhancement

Fingerprint ridges in a local region can be enhanced by applying a set of Gabor filters [76]. Gabor filter is widely used to select local structure with specific orientation and frequency, and the Gabor function is given by:

$$h(x, y; \theta, f) = \exp\left(-\frac{1}{2}\left(\frac{x_\theta^2}{\delta_x^2} + \frac{y_\theta^2}{\delta_y^2}\right)\right)\cos(2\pi f x_\theta), \qquad (4.9)$$

$$\begin{bmatrix} x_\theta \\ y_\theta \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix},$$

where θ is the orientation of the Gabor filter, f is the frequency of the filter, and δ_x and δ_y are standard deviations which determine the envelope of the Gabor filter.

4.5 Experimental Results

4.5.1 Database

The experiments are conducted on the public domain latent fingerprint database, NIST SD27 [3], which contains 258 latent fingerprints and their corresponding rolled fingerprints. Each latent image in this database was assigned one of three (subjective) quality levels - *good*, *bad*, and *ugly* - by latent fingerprint examiners. The numbers of latents belonging to the "good", "bad", and "ugly" category are 88, 85, and 85, respectively. To make the latent matching problem more realistic and challenging, the background database was enlarged to 27,258 fingerprints by including 27,000 rolled prints in NIST SD14 [10].

<u>Algorithm.</u> **Hypotheses Generation for Orientation Field Estimation**

m: Minimum number of blocks to build a hypothesis.
G: The set of orientation elements in all groups.
S: A set of orientation elements in selected groups (initially, S is empty).
p_i for $i = 1, \ldots, N$: A pdf for random selection of orientation groups.
N: A total number of orientation groups.

I. Hypothesis

1. Initialize S: Non-overlapping groups are selected randomly based on the given p such that the total number of orientation elements in S is greater than m.

2. Build a hypothesis M: $\{a_{kl}\}$ and $\{b_{kl}\}$ are estimated by the least-squares estimation using S.

3. If all orientation elements in S are consistent with the M, go to step 4. If the inconsistent orientation elements are only from the most recently added group, discard this group from S and go to step 5. Otherwise, go to step 6.

4. Find the consensus set S^*: S^* is determined by a subset of G which are within some error tolerance of the regularized residual orientation field from M. If $|S^*| > |S_{best}|$, update S_{best} and M_{best} with S and M. Go to step 5.

5. Add a group to S which is selected randomly in G and that does not overlap with any orientation elements in S and go back to step 2. If there are no more groups to be added, go to step 6.

6. Accept/reject the hypothesis: Check if the regularized residual orientation field from the hypothesis M_{best} contains any singularity [34]. If singularity exists in the ROI, reject the hypothesis. Otherwise, proceed to evaluation stage.

II. Evaluation

The accepted hypotheses are ranked in a descending order according to the size of the consensus set S^*_{best}.
Go back to I.

III. Output
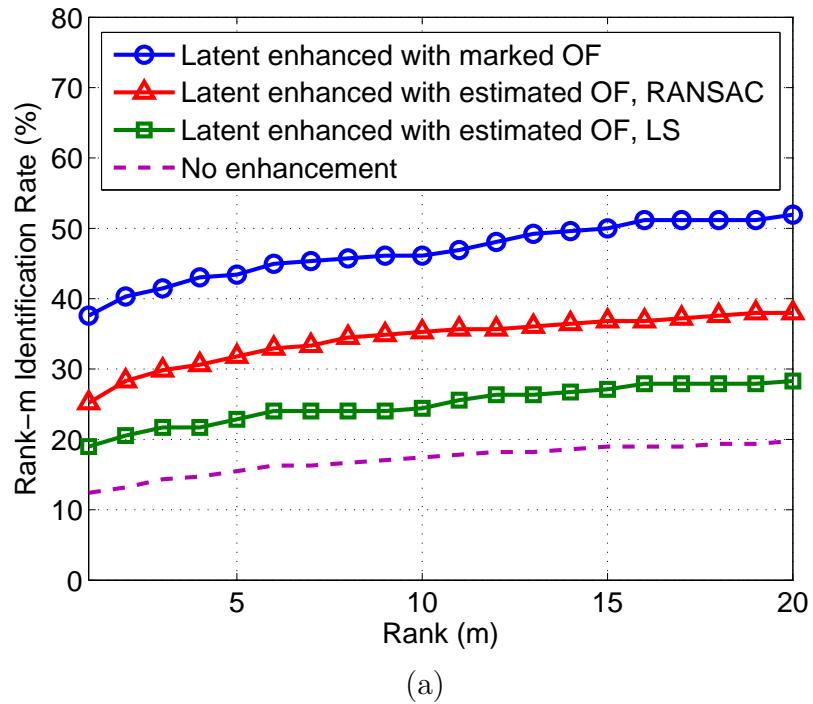
Top-10 best hypotheses are retrieved from the accepted hypotheses.

4.5.2 Performance Evaluation
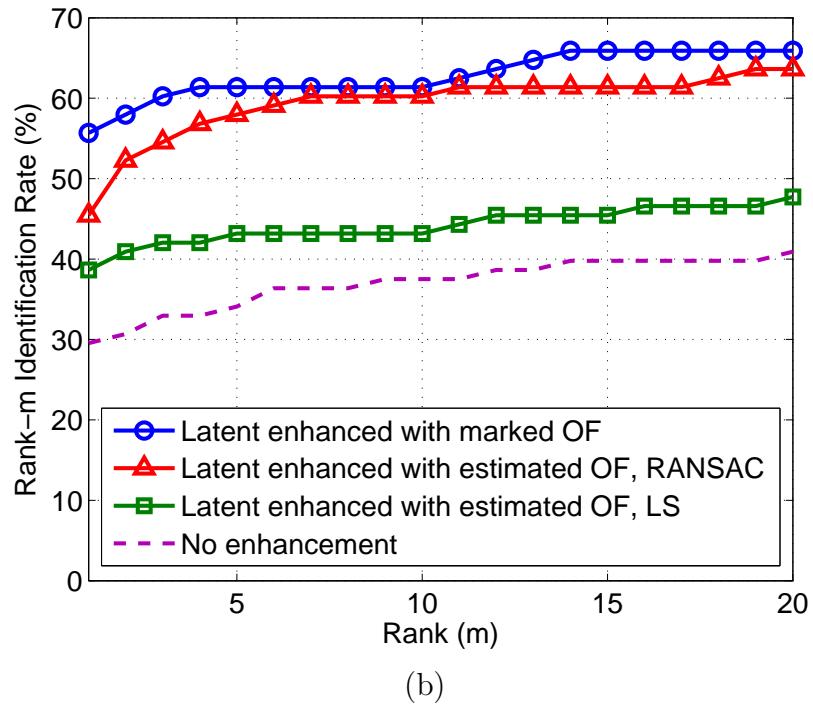
The accuracy of the proposed latent fingerprint enhancement algorithm is evaluated by measuring the latent matching performance before and after enhancement using a commercial matcher, Neurotechnology VeriFinger SDK 4.2 [13]. The CMC curves of four types of the input to the fingerprint matcher are shown in Figure 4.9. These inputs are: (i) enhanced latent by manually marked orientation field; (ii) enhanced latent by the estimated orientation field using the R-RANSAC algorithm [138]; (iii) enhanced latent by the estimated orientation field using the least-squares method [137]; and (iv) latent image with no enhancement.

Note that the performance of using a manually marked orientation field provides the upper bound. As shown in Figure 4.9(a), the automatic matching performance is significantly improved when the enhanced images are used as input to the matcher. Furthermore, the proposed algorithm based on R-RANSAC performs better than the least-squares estimation. The CMC curves are also shown separately according to the quality of the latents. For good quality latents, the proposed algorithm achieves performance close to the upper bound (manually marked orientation field; see Figure 4.9(b)). For bad and ugly quality latents, while the proposed algorithm performs much better compared to the least-squares estimation method or the no enhancement case (see Figures 4.9(c) and (d)), there is a significant gap in performance between the proposed algorithm and the upper bound. This shows that the automatic orientation field estimation in poor quality latents is still a challenging problem.

Figure 4.10 shows three successful examples in each quality category where the proposed enhancement algorithm improved the match score of the latent with the mated rolled fingerprint. In these three cases, the mated rolled print can be retrieved at a high rank from the exemplar database containing 27,258 images. Figure 4.11 shows a failure case where the failure in orientation field estimation leads to a lower genuine match score after enhancement compared to no enhancement.

121
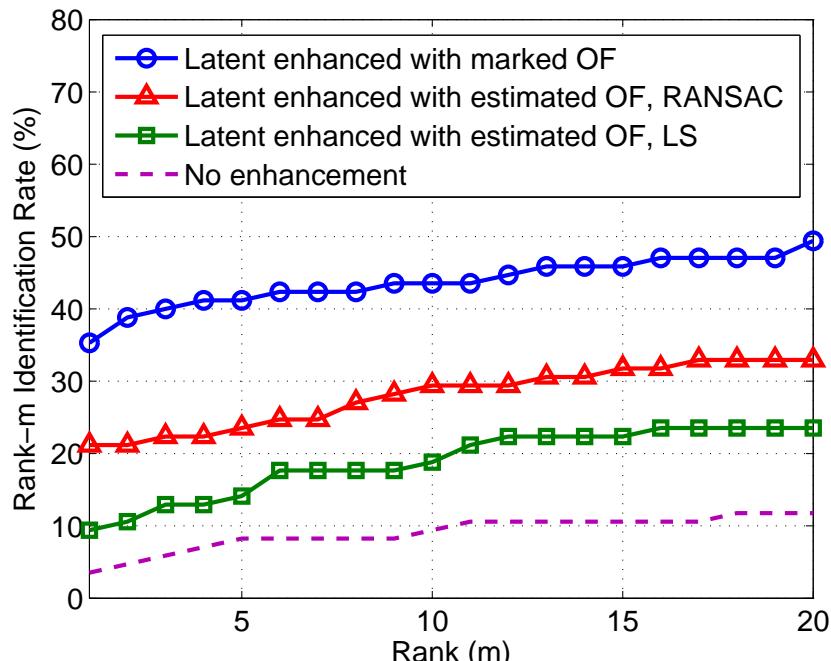
(a)
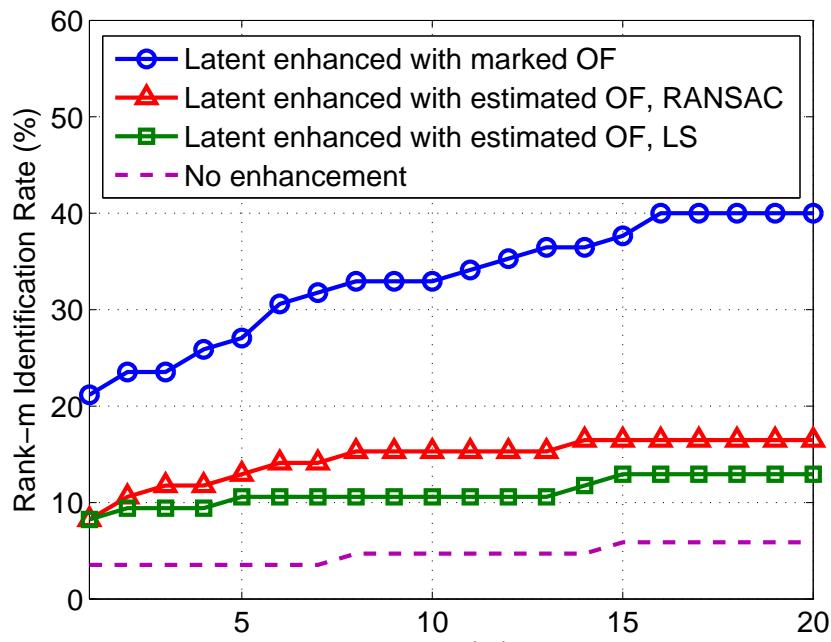
(b)

Figure 4.9: CMC curves showing the performance improvement due to latent enhancement. (a) All latents in NIST SD27, (b) good quality latents, (c) bad quality latents, and (d) ugly quality latents..
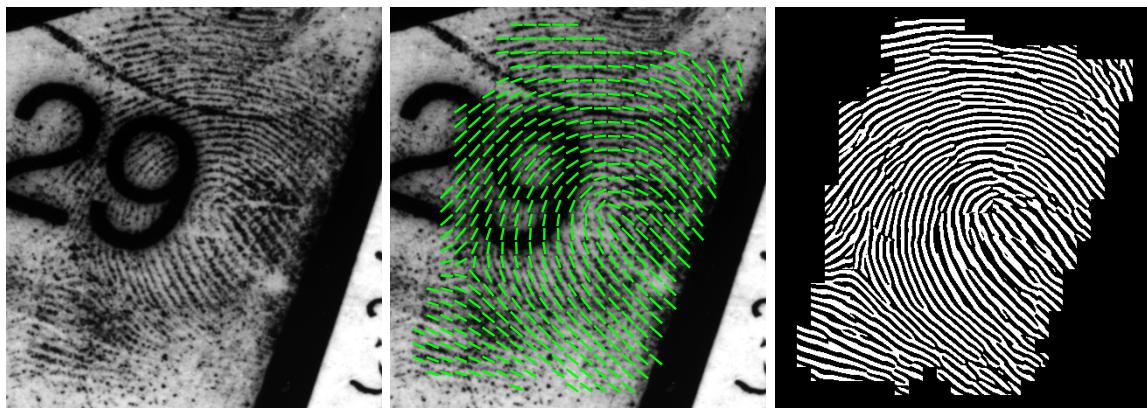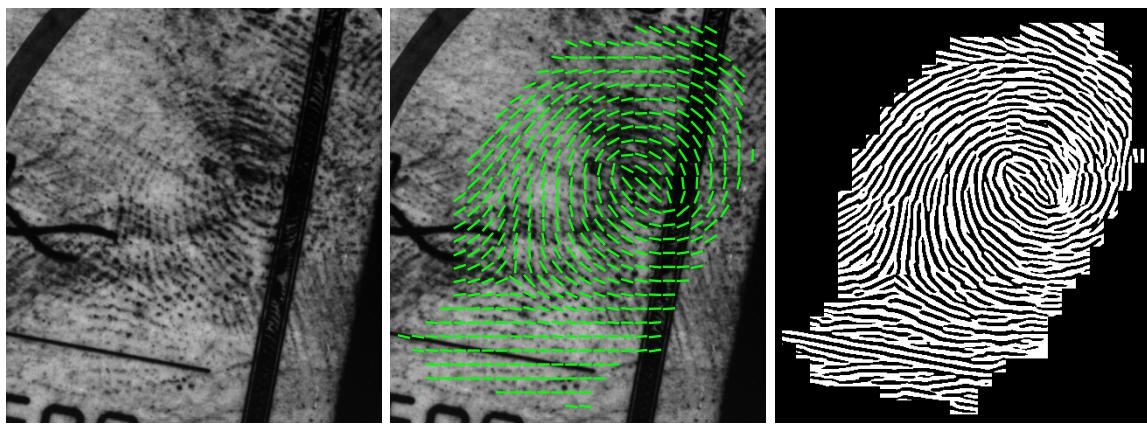
(c)
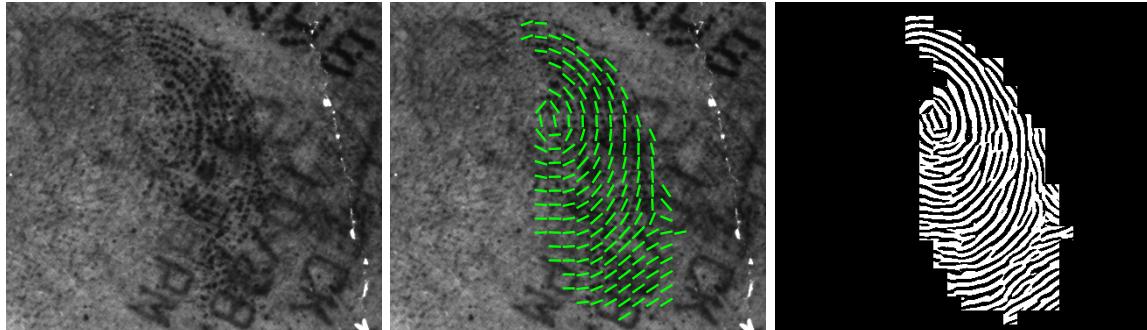
(d)

(a) Good quality latent (G019): {4,5277}→{54,1}

(b) Bad quality latent (B142): {3,1253}→{34,1}

(c) Ugly quality latent (U282): {0,27258}→{24,1}

Figure 4.10: Successful examples where the proposed latent enhancement algorithm improves the latent matching performance. Left: latent image, Center: estimated orientation field, and Right: binarized enhanced image. Match score of the latent with the mated rolled print and its retrieval rank from a database of 27,258 rolled prints are shown as: {score, rank} before enhancement → {score, rank} after enhancement.
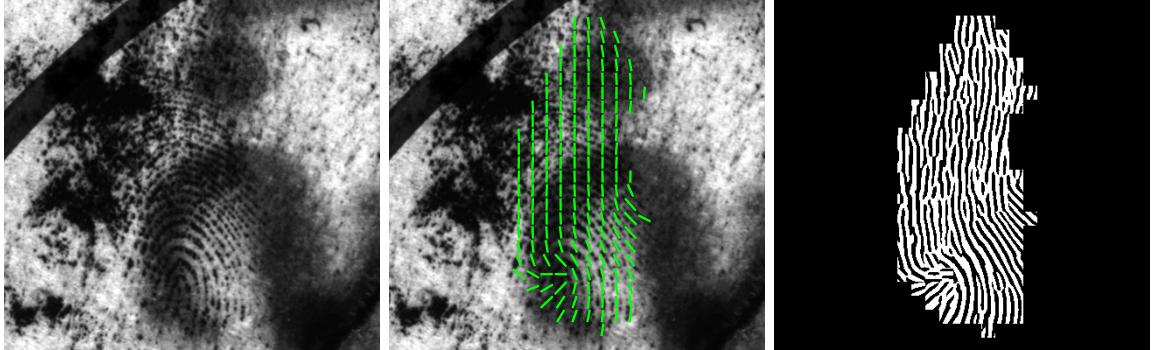
Figure 4.11: Failure case (G029). Due to the failure in orientation field estimation, the latent-to-rolled match score drops from 23 (before enhancement) to 3 (after enhancement), and the retrieval rank for the true mate increases from 12 to 15,801.

4.6 Summary

Latent fingerprints found at crime scenes provide crucial evidence to law enforcement agencies. The latents are typically searched against a large fingerprint database which is the collection of rolled/plain fingerprints of previously apprehended criminals. Due to the generally poor quality of latents, latent examiners perform manual feature markup and visual comparison between the latent and the candidate fingerprints from the database. A "lights-out" mode for latent identification is desired to reduce the burden on latent examiners and to introduce a level of consistency in fingerprint matching, particularly in searching ever-growing exemplar fingerprint databases.

We have proposed a latent fingerprint enhancement algorithm that requires minimal markup (i.e., ROI and singular points) to improve the automatic latent matching accuracy. The orientation field of the latents is estimated by the R-RANSAC algorithm, which is effectively used to find a correct orientation field in the noisy and partial latent fingerprints. The estimated orientation field is used to enhance ridge structures by Gabor filtering. The proposed algorithm significantly improves the matching performance of a commercial matcher when the enhanced latents are fed into the matcher.

We propose to extend our work on latent enhancement as follows:

- Improve the performance of the orientation field estimation algorithm for bad and ugly quality latents in NIST SD27.

- Reduce the human markup even further; ideally, the input of the algorithm should only be the latent image.

- Assess the quality of the latents automatically. The reliability of the features extracted from the latents can be adjusted according to the quality in order to automatically determine the degree of human intervention for feature markup and improve the overall matching accuracy.

Chapter 5

Fingerprint Obfuscation

5.1 Introduction

Fingerprint obfuscation refers to the deliberate alteration of the fingerprint pattern [45] using techniques varying from abrading, cutting, and burning fingers to performing plastic surgery (see Figure 1.13). The use of altered fingerprints to mask one's identity constitutes a serious "attack" against a border control biometric system since it defeats the very purpose for which the system was deployed in the first place, i.e., to identify individuals in a watch list.

The problem involving altered fingerprints falls under a broader category of attacks known as *biometric obfuscation*. Obfuscation can be defined as a purposeful attempt by an individual to mask his identity, so that it cannot be recognized by a biometric system, by altering the biometric trait prior to its acquisition by the system. Examples include mutilating the ridges of one's fingerprint by using abrasive material, perturbing the texture of the iris by wearing theatrical lenses, or altering facial attributes such as nose and lips via surgical procedures. In this study, we will concern ourselves with the problem of fingerprint obfuscation for the following reasons: (i) fingerprint-based biometric systems are much more widespread for large-scale identification than any other biometric modality; (ii) it is relatively easy to alter one's fingerprints using chemicals and abrasives compared to, say, one's iris or face where a

Figure 5.1: Inked impressions before and after fingerprint alteration (a) of Gus Winkler [45] (pattern type is altered from twin-loop to left-loop), and (b) of Jose Izquierdo [134] (altered by switching two parts of a 'Z' shaped cut on the fingertip).

more elaborate surgical procedure may be necessary; and (iii) mutilated fingerprints are being routinely encountered by law enforcement and immigration officials in several countries, thereby underscoring the urgency of finding a solution to this problem.

Fingerprint alteration has a long history. As early as 1933, Gus Winkler, a murderer and bank robber, was found to have altered the fingerprints of his left hand except for the thumb by slashing and tearing the flesh of the fingers [45]. Further, the pattern type of one of his fingers was altered from twin-loop to left-loop (see Figure 5.1(a)).

In more recent cases, a man using the name Alexander Guzman, arrested by Florida officials in 1995 for possessing a false passport, was found to have obfuscated fingerprints (see Figure 5.1(b)). After a two-week search based on manually reconstructing the damaged fingerprints and searching the FBI database containing 71 million records, the reconstructed fingerprints of Alexander Guzman were linked to the fingerprints of Jose Izquierdo who was an absconding drug criminal [134]. In this particular case, the mutilation process consisted of three steps: (i) making a 'Z' shaped cut on the fingertip, (ii) lifting and switching two triangular skin patches, and (iii) stitching them together. In September 2005, a drug dealer named Marc George was apprehended because his limping gait as a result of surgery involving transplantation of friction ridge skin taken from the sole of his feet caught the attention of border officials (see Figure 1.13(a)) [14].

Table 5.1: High Profile Cases of Fingerprint Alteration

Case	Year	Alteration Type	Description
		Criminal Cases	
Gus Winkler [45]	1933	Imitation	Pattern type was changed from twin-loop to left-loop (Figure 5.1(a)).
John Dillinger [14]	1934	Obliteration	Fingerprints were mutilated by applying acid.
Robert J. Philipps [14]	1941	Obliteration	Skin from the chest was transplanted to the fingertips.
Jose Izquierdo [134]	1997	Distortion	A fingerprint with strange pattern was formed by 'Z' cut (Figure 5.1(b)).
Marc George [14]	2005	Imitation	Friction ridge skin from sole of the feet was implanted to the fingertips (Figure 1.13(a)).
A man arrested for vehicle theft [122]	2007	Obliteration	Fingers were bitten (Figure 1.13(b)).
Mateo Cruz-Cruz [67]	2007	Obliteration	Fingerprints were blackened as a result of applying acid (Figure 1.13(c)).
Gerald Perez [107]	2008	Obliteration	Fingertips with thick stitches (Figure 1.13(d)).
		Non-criminal Cases	
A woman crossing U.S. border [67]	2007	Obliteration	A surgery was performed on fingertips to generate strange fingerprint pattern.
Asylum seekers to EU [20, 24]	2008	Obliteration	Fingertips were abraded and burned.
A woman attempting to evade the Japanese border control system [25]	2009	Imitation	Friction ridge skins from thumbs and index fingers were swapped between left and right hands.
Three people charged in Boston with conspiring to mutilate fingerprints [27]	2010	Obliteration	A physician, a broker, and a patient were involved in a scheme to mutilate or surgically remove the fingerprints to conceal illegal aliens from detection.

It is not just criminals who have been found to alter their fingerprints. In December 2009, a woman successfully evaded the Japanese immigration AFIS by surgically swapping fingerprints of her left and right hands [25]. Although she was originally arrested for faking a marriage license, scars on her hands made the border police suspicious.

Fingerprint alteration has even been performed at a much larger scale involving a group of individuals. It has been reported that hundreds of asylum seekers had cut, abraded, and burned their fingertips to prevent identification [20, 24] by EURODAC [22], a European Union-wide fingerprint system for identifying asylum seekers. Table 5.1 lists reported cases
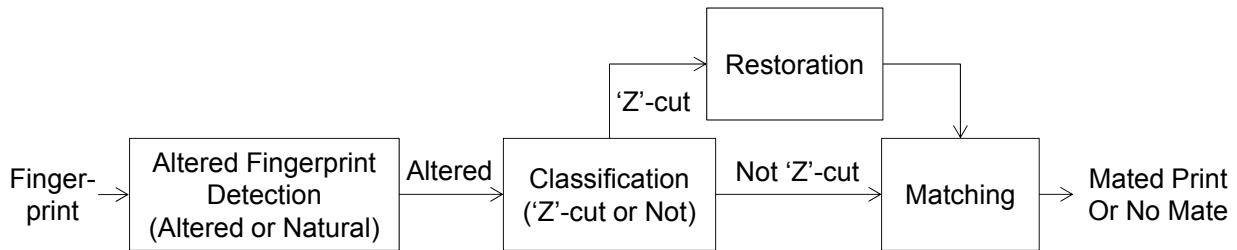
Figure 5.2: Flow chart for detecting and matching altered fingerprints.

of fingerprint alteration.

Although the number of publicly disclosed cases of altered fingerprints is not very large, it is extremely difficult to estimate the actual number of individuals who have successfully evaded identification by fingerprint systems as a result of fingerprint alteration. Almost all the persons that have been identified so far to have altered their fingerprints were not detected by AFIS, but by some other means such as improper documents [25] or by a product of the surgery (e.g., limping gait) [14].

With the growing use of AFIS in law enforcement and border control, it is expected that there will be more instances where altered fingerprints will be encountered by the authorities. Therefore, it is necessary that AFIS should have the capability to determine the true identity of individuals with altered fingerprints. To tackle the problem of fingerprint alteration, new algorithms for detecting and matching altered fingerprints are urgently needed. An overview of the procedure to deal with altered fingerprints is presented in Figure 5.2. Fingerprints submitted to AFIS need to pass through an altered fingerprint detector at the front end of the AFIS. Abnormality in orientation field and minutiae distribution in altered fingerprints provides indications of possible alterations [54, 139]. Once the altered fingerprints are detected, the suspect is sent to a secondary inspection in order to verify (i) that the unusual fingerprint pattern is indeed due to alteration, and (ii) whether these altered fingerprints could still be matched to the suspect's pre-altered fingerprints if present in the databases [141].
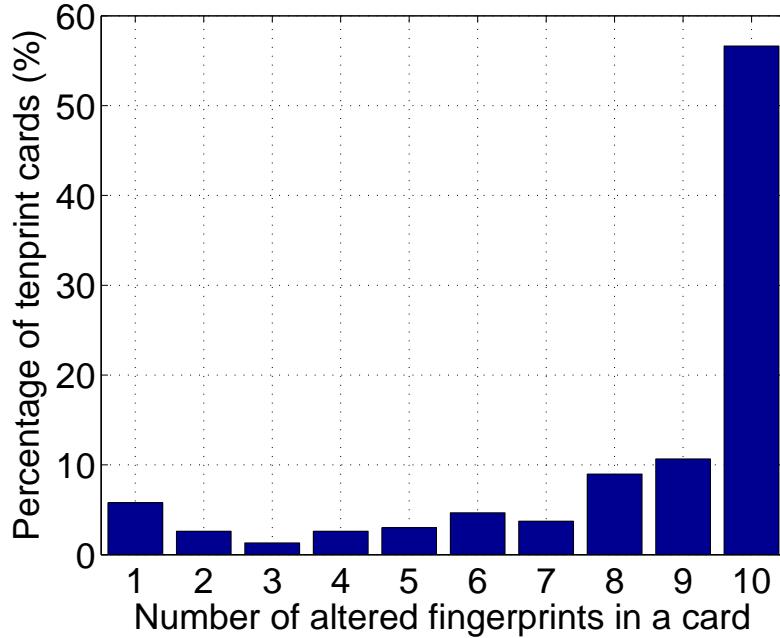
Figure 5.3: Distribution of the number of altered fingerprints in tenprint cards in our database.

This chapter is organized as follows. In Section 5.2, the impact of the fingerprint obfuscation on the matching performance is investigated and three different categories of altered fingerprints and their potential counter-measures are described. The proposed approach for detecting altered fingerprints is presented in Section 5.3. Section 5.4 describes an approach to improve the performance of matching altered fingerprints to their pre-altered mates by masking out the altered region and restoring 'Z'-cut cases. Finally, Section 5.5 proposes future directions for research on this topic.

5.2 Analysis of Altered Fingerprints

Based on a database of altered fingerprints made available to us by law enforcement agencies, we (i) determine the impact of fingerprint alteration on the matching performance, (ii) categorize altered fingerprints into three types[1]: *obliteration*, *distortion*, and *imitation*, and

---

[1]While the categorization is exclusive, there is ambiguity in some altered fingerprints.

(iii) assess the utility of a well-known fingerprint quality measure for altered fingerprint detection.

5.2.1   Altered Fingerprint Database
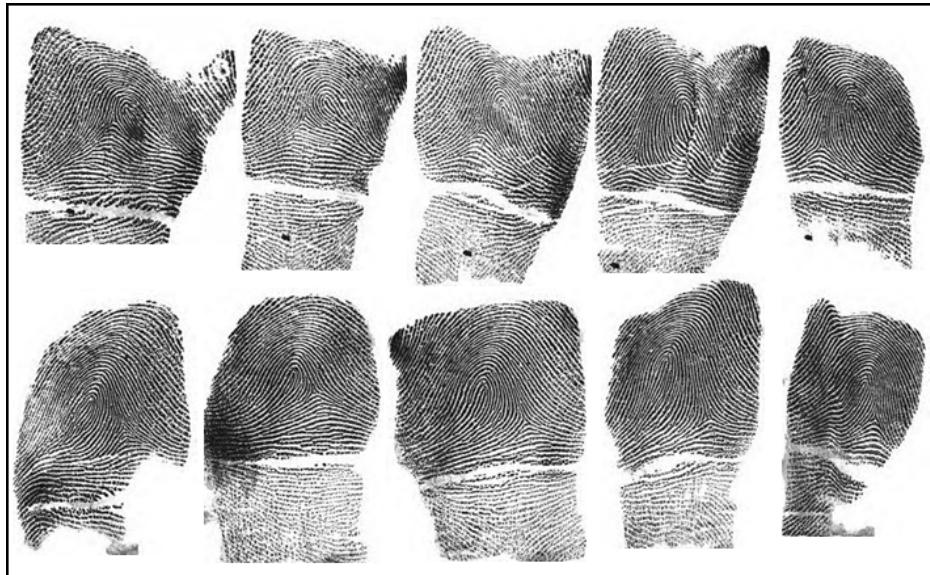
The database contains 4,433 altered fingerprints from 535 tenprint cards of 270 subjects. We summarize some of the characteristics of this database below.

- Not all the ten fingers in a tenprint card may have been altered. The distribution of the number of altered fingers in a tenprint card is shown in Figure 5.3; in 57% of the tenprint cards, all ten fingerprints were altered; 85% of the tenprint cards have more than five of the ten fingerprints altered.

- The number of tenprint cards for a subject varies from 1 to 16; a total of 87 subjects out of the 270 subjects have multiple tenprint cards due to multiple arrests.

- For subjects with multiple tenprint cards, there exist 1,335 pairs of pre-altered (natural) and post-altered fingerprints. Figure 5.4 shows an example of pre-altered and post-altered tenprint cards of a subject.
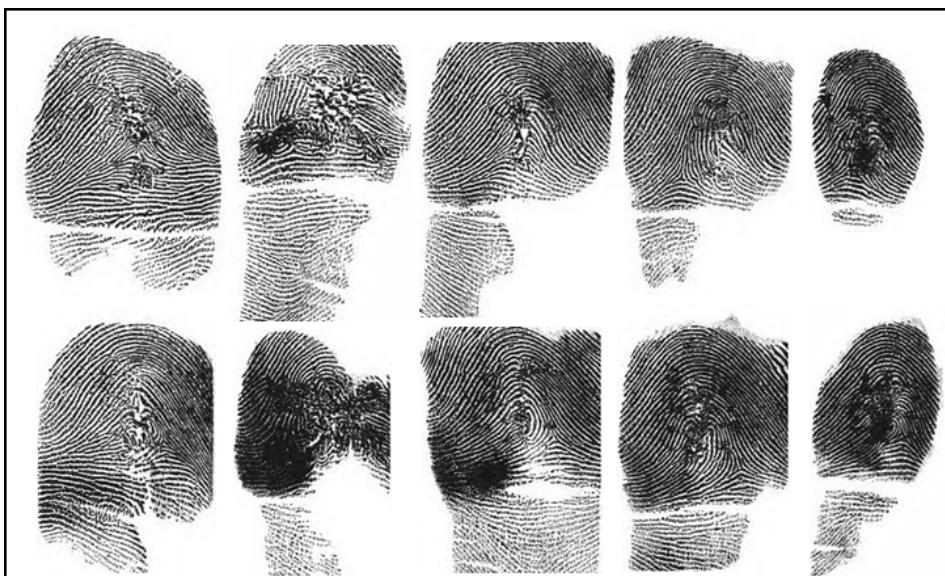
5.2.2   Vulnerability of Fingerprint Identification Systems

Fingerprint alteration is a serious threat to AFIS, since it revokes one of the fundamental premise that fingerprint is persistent during one's lifetime. To understand the vulnerability of AFIS to fingerprint alteration, we used a commercial matcher, VeriFinger SDK 4.2 [13], to match 1,335 altered fingerprints to their mated pre-altered fingerprints. To establish a baseline, NIST SD4 database [2], which consists of 2,000 fingers with two impressions per finger, was used to obtain genuine and impostor match score distributions using the matcher[2].

---

[2]Note that while the analysis is based on a specific fingerprint matcher, the results of fingerprint alteration are likely to affect all commercial matchers in a similar manner. VeriFinger, like other state-of-the-art fingerprint matchers, utilizes ridge pattern for matching, which is what the culprits are trying to change through fingerprint alteration.

(a)

(b)

Figure 5.4: Mated pre/post altered tenprint cards from a subject. (a) Pre-altered fingerprints and (b) post-altered fingerprints.

Figure 5.5 shows the score distributions for pre/post altered fingerprint pair matches according to type of alteration and genuine and impostor matches in NIST SD4. The key observations based on these matching results are:

1. The match score distributions of pre/post altered fingerprint pairs for all alteration
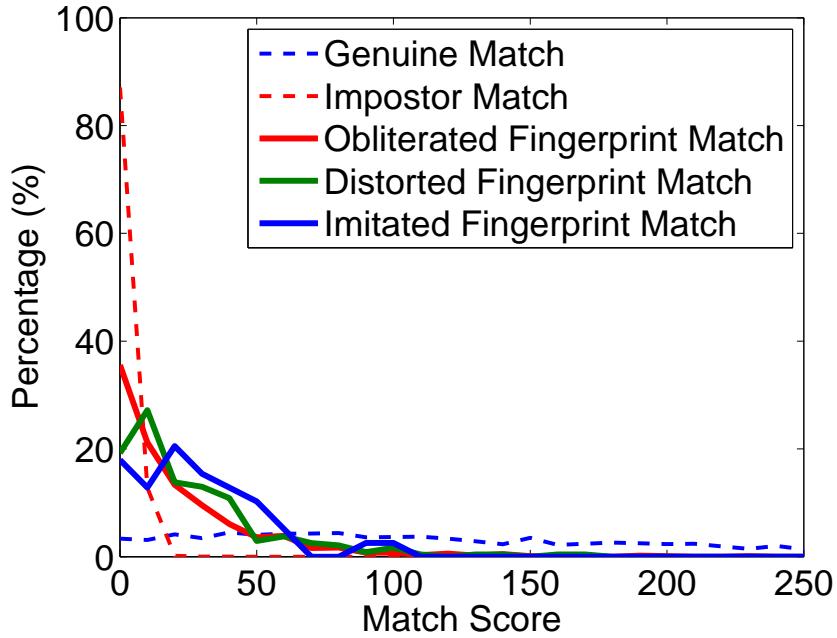
Figure 5.5: Match score distributions of pre/post alteration pairs according to type and genuine and impostor pairs in NIST SD4.

types follow the impostor score distribution.

2. Heavy tails in pre/post altered match score distributions indicate that fingerprint alteration, as observed in our database, is not always successful in evading AFIS.

3. At a threshold of 41, which corresponds to 0% false acceptance rate on NIST SD4, 83% of the pre/post altered fingerprint pairs have genuine match scores below the threshold. This means that an AFIS is unable to link most of the altered fingerprints to their true mates.

Figure 5.6 shows examples where altering a fingerprint leads to failure in matching it to its true mate. The process of fingerprint mutilation destroys the ridge structure itself so that minutiae extraction is not possible in the altered region (Figure 5.6(a)). A severe ridge distortion such as ridge structure transformation (Figure 5.6(b)) or ridge deformation due to scars alters the spatial distribution of the minutiae.

(a)   (b)
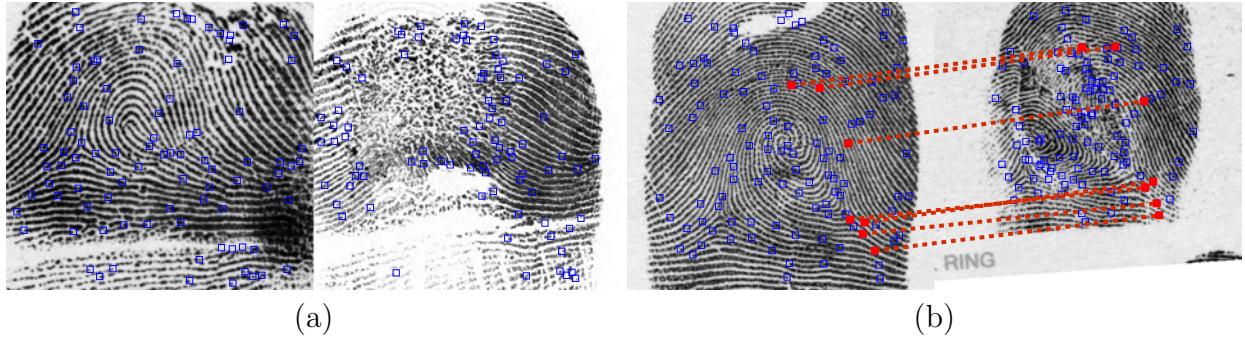
Figure 5.6: Examples where fingerprint alteration severely degrades the matching score with the pre-altered mates. (a) Mutilation over a large area and (b) ridge transformation. These altered fingerprints have a match score of 0 with their true mates. All squares indicate minutiae extracted from the image and squares filled with red color represent matched minutiae between the pre/post altered fingerprints. The left-hand side image in each pair shows the pre-altered fingerprint, and the right-hand side image shows the post-altered fingerprint.

(a) Match score = 233   (b) Match score = 173
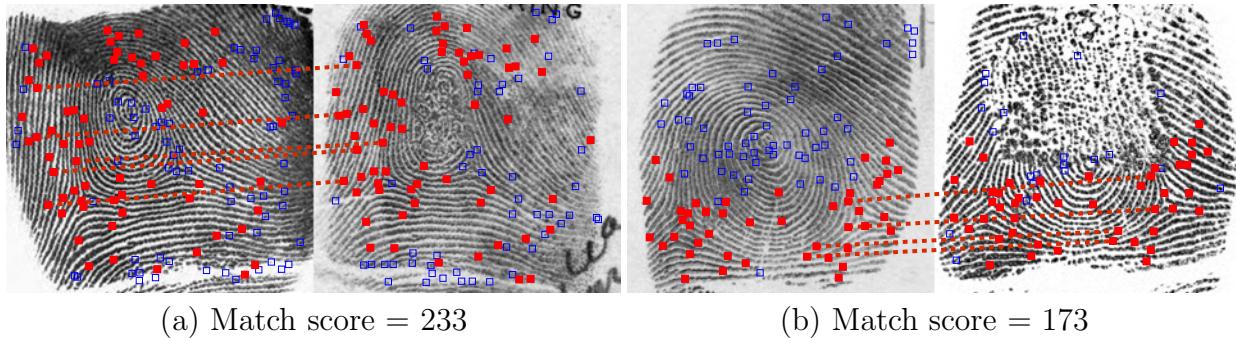
Figure 5.7: Examples where the pre/post altered fingerprint mates are correctly matched despite fingerprint alteration. (a) Alteration with a small damaged area in the fingerprint and no ridge distortion and (b) sufficient number of minutiae in the unaltered area even with severe fingerprint alteration. Only a subset of the corresponding minutiae are connected with dotted lines for visibility.

There is no guarantee that fingerprint alteration will always be successful in evading an AFIS (see Figure 5.7). As long as there are a sufficient number of minutiae that can be extracted in the unaltered area, pre/post fingerprint mates can be successfully matched by AFIS.

Table 5.2: Exclusive Categorization of the Altered Fingerprints into Three Types

Type	Obliteration		Distortion		Imitation
	Scar	Mutilation	Z-cut	Transplantation	
Number of Images	1,457	2,480	297	148	51

5.2.3 Types of Altered Fingerprints

We classify altered fingerprints into three categories (i.e., obliteration, distortion, and imitation) based on the change in ridge pattern due to alteration. This categorization will assist us in three ways: (i) getting a better understanding of the nature of alterations that can be encountered, (ii) detecting altered fingerprints by modeling types of alteration, and (iii) developing methods for altered fingerprint restoration and matching.

Table 5.2 shows the exclusive categorization of 4,433 altered fingerprints in our database. Note that this classification is not based on the method of alteration which is not known to us. Our classification is subjective and is based on our examination of the ridge patterns in a large number of altered fingerprint images in the database.

5.2.3.1 Obliteration

Friction ridge patterns on fingertips can be obliterated by abrading [36], cutting [45], burning [18, 20, 24], applying strong chemicals (Figure 1.13(c)), and transplanting smooth skin [14]. Further, factors such as skin disease (such as leprosy [131]) and side effects of a cancer drug [136] can also obliterate fingerprints. Friction ridge structure is barely visible within the obliterated region. According to Table 5.2, obliteration appears to be the most popular form of fingerprint alteration. This may be because obliteration which completely destroys ridge structures is much simpler to perform than distortion/imitation which requires a surgical procedure. Furthermore, detecting distorted or imitated fingerprints is much more difficult for human examiners than obliterated fingerprints.

Obliterated fingerprints can evade fingerprint quality control software, depending on the

Figure 5.8: Fingerprint obliteration. Examples of (a) scar and (b) mutilation.

area of the damage. If the affected finger area is small, the existing fingerprint quality assessment softwares may fail to detect it as an altered fingerprint (the fingerprint in Figure 5.8(a) has an acceptable NFIQ value of 3), but AFIS is likely to successfully match the damaged fingerprint to the original mated fingerprint (Figure 5.7(a)). But, if the altered area is sufficiently large, fingerprint quality control software can easily detect the damage. For example, the obliterated fingerprint in Figure 5.8(b) has the lowest NFIQ value of 5.

To identify individuals with severely obliterated fingerprints, it may be necessary to treat these fingerprints as latent images, perform an AFIS search using manually marked features, and adopt an appropriate fusion scheme for tenprint search [96]. In rare cases, even if the finger surface is completely damaged, the dermal papillary surface, which contains the same pattern as the epidermal pattern, may be used for identification [110].

5.2.3.2 Distortion

Friction ridge patterns on fingertips can be turned into unnatural ridge patterns as a result of obfuscation [23, 26, 134]. This can happen, for example by removing portions of skin from a fingertip and either grafting them back in different positions (Figure 5.9(a)) or replacing them with friction ridge skin on the palm or sole (Figure 5.9(b)). Distorted fingerprints have unusual ridge patterns which are not found in natural fingerprints. These abnormalities include abnormal spatial distribution of singular points or abrupt changes in orientation

<div align="center">(a)   (b)</div>

Figure 5.9: Fingerprint distortion. Examples of (a) transplantation within a finger by 'Z' cut and (b) transplantation from other friction ridge skin, e.g., from palm.

field along the scars. Note that an orientation field discontinuity in natural fingerprints is usually observed only at singular points.

Distorted fingerprints can also successfully pass the fingerprint quality test since their local ridge structures remain similar to natural fingerprints; it is their global ridge pattern that is abnormal. For instance, a distorted fingerprint as a result of swapping skin patches within the same finger (e.g. Figure 5.9(a)) retains the same ridge property (e.g. ridge frequency and width) over the entire fingerprint area. Figure 5.9(a) is indeed assigned the highest quality level, NFIQ of 1. Similarly, the altered fingerprint in Figure 5.9(b) is assigned the second highest quality level, NFIQ = 2.

Fingerprints altered by 'Z' cut are of special interest since they retain their original ridge structure, enabling a reconstruction of the original fingerprint pattern before alteration. Therefore, it is imperative to upgrade current fingerprint quality control software to detect the distorted fingerprints. Once detected, the following operations may be performed to assist the AFIS: (i) identify unaltered regions of the fingerprint and manually mark the features (i.e., the minutiae) in these regions and (ii) reconstruct the original fingerprint as in the 'Z' cut case [134].

$$(a) \qquad\qquad\qquad\qquad (b)$$

Figure 5.10: Fingerprint imitation. Left: pre-altered fingerprint, and Right: its post-altered fingerprint mate. (a) Removal of a portion of skin and (b) exquisite transplantation from other friction ridge skin on the body.

5.2.3.3 Imitation

Friction ridge patterns on fingertips can still preserve fingerprint-like pattern after an elaborate procedure of fingerprint alteration: (i) a portion of skin is removed and the remaining skin is pulled and stitched together (Figure 5.10(a)), (ii) friction ridge skin from other parts of the body is used to fill the removed part of the fingertip to reconcile with the remaining ridge structure (Figure 5.10(b)), or (iii) transplantation of the entire fingertip. As reported in [25], simply swapping the skin on fingertips between the left and right hands successfully evaded an AFIS.

Imitated fingerprints can not only successfully pass the fingerprint quality assessment software, they can also confound human examiners. Figure 5.10 shows pre-altered and post-altered fingerprint mates. The altered fingerprint in Figure 5.10(a) has a very smooth orientation field over the entire fingerprint area (which looks like an arch type fingerprint) and the only evidence of possible alteration is a thin scar. This fingerprint has the highest NFIQ value of 1. However, its pre-altered mate is indeed of right-loop type, and the match score between this pair of fingerprints is only 19, much below the threshold (a score of 41) on match score corresponding to 0% FAR of the matcher. The altered fingerprint in Figure 5.10(b) was generated by an exquisite surgical procedure to have very natural ridge flow even along the surgical scars. This fingerprint also has the highest NFIQ value of 1 with a

139
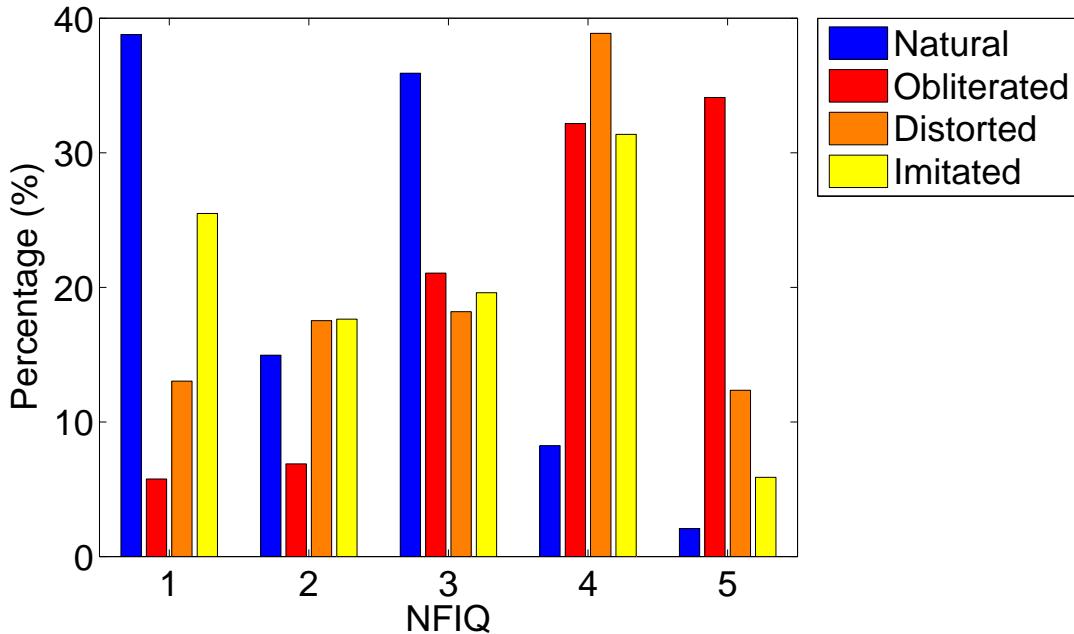
Figure 5.11: Histograms of NFIQ values for 27,000 natural fingerprints in NIST SD14 and 4,433 altered fingerprints in our altered fingerprint database according to the type of alteration. Recall that NFIQ = 1 indicates the highest quality.

match score between pre/post altered fingerprint pair of only 28.

To match altered fingerprints in Figure 5.10, matching algorithms that are robust to distortion and inconsistency in the fingerprint flow pattern need to be developed. In the case where fingerprints from different fingers are swapped, fingerprint matching without using finger position information (i.e., left thumb is allowed to match to right index finger) may help in determining the true identity at the expense of significantly higher matching time.

5.2.4 Effectiveness of Fingerprint Quality Assessment Algorithm

To determine the effectiveness of the commonly used fingerprint quality control softwares in detecting altered fingerprints, quality levels of altered fingerprints and natural fingerprints were estimated using the NFIQ software [124], which is the de facto standard of fingerprint quality. To construct a natural fingerprint database, we used the 27,000 fingerprints in NIST

SD14 [10]. From the histograms of NFIQ values for altered and natural fingerprints shown in Figure 5.11, we can observe that:

1. A significant portion of altered fingerprints have the lowest quality level of 5 while only a small percentage of natural fingerprints have this lowest quality level. In particular, the obliterated fingerprints have the largest portion at the NFIQ level of 5. By contrast, the distorted and imitated fingerprints have relatively small portion at the level of 5.

2. A large number of altered fingerprints have good quality; about 7% of altered fingerprints have the highest quality level of 1 in total, and a significant portion of distorted and imitated fingerprints has the highest quality level.

3. If the NFIQ value of 5 is used as a criterion for detecting altered fingerprints, it will lead to a true positive (an altered fingerprint is correctly classified as an altered fingerprint) rate of 31.6% at a false positive (a natural fingerprint is misclassified as an altered fingerprint) rate of 2.1%.

5.3   Detection of Altered Fingerprints

Developing an automatic solution to detect altered fingerprints is the first step in countering the threats due to fingerprint alteration. Fingerprint quality assessment routines used in most fingerprint identification systems, such as the open source NFIQ software [124], may be useful in detecting altered fingerprints if the corresponding images are indeed of poor quality. But, as mentioned earlier, not all altered fingerprint images have poor quality (see Figures 5.9 and 5.10). Since existing fingerprint quality assessment algorithms [31] are designed to examine if an image contains sufficient information (say, minutiae) for matching, they have limited capability in determining if an image is a natural fingerprint or an altered fingerprint. For example, while the synthesized ridge pattern in Figure 3.2(a) is not likely to appear on fingertips, it is declared to be of the best quality according to the NFIQ measure (i.e., NFIQ of 1).

Given that the altered fingerprints are likely to be encountered in large-scale national identification or border control systems, the automatic detector must satisfy the following three requirements:

1. Given the large throughput requirement of these systems, the algorithm must be extremely fast. In other words, it should not increase the computational burden of the matcher by any significant amount. State-of-the-art AFIS can process fingerprints at the rate of about 1 million matches per second. This implies that the feature extraction and decision rule used to automatically detect altered fingerprints must be efficient.

2. In operational scenarios, the number of individuals with altered fingerprints that will be encountered by an AFIS will be very small. Keeping this in mind, the altered fingerprint detection algorithm should operate at a very low false positive rate, say 1% or lower. Since subjects that are suspected to have altered fingerprints will need to go through a secondary inspection stage, it is desirable not to have many individuals to be sent for secondary inspection.

3. The altered fingerprint detector should be easily integrated into any AFIS.

In the previous section, we showed that the NFIQ algorithm is not suitable for detecting altered fingerprints, especially the distortion and imitation types. In fact, the distorted and imitated fingerprints are very hard to detect for any fingerprint image quality assessment algorithm that is based on local image quality alone. In this section, we consider the problem of automatic detection of alterations based on properties of the ridge orientation field and minutiae distribution [139]. The flow chart of the proposed alteration detector is given in Figure 5.12.

5.3.1 Abnormality in Orientation Field

Good quality fingerprints have a smooth orientation field except near the singular points. Based on this fact, many orientation field models have been developed by combining the
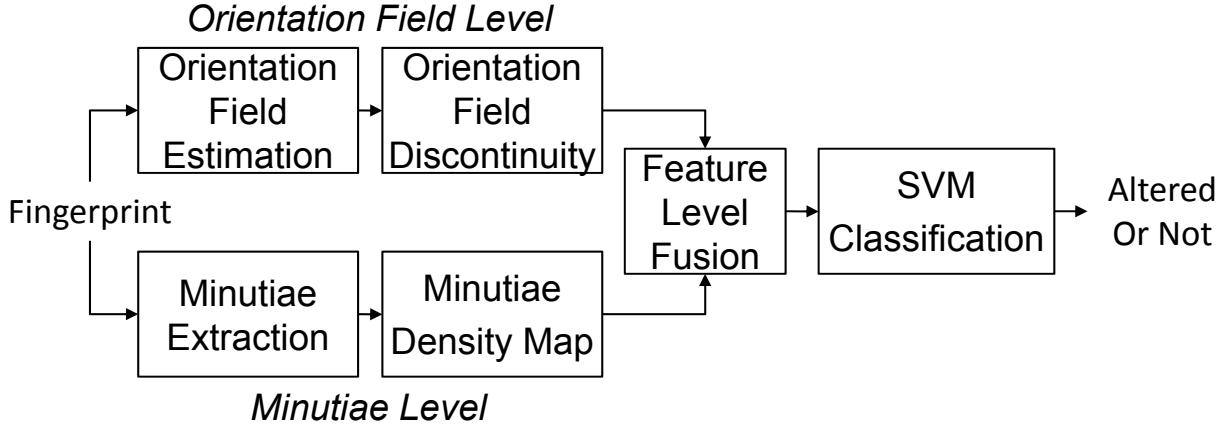
Figure 5.12: Flow chart of the altered fingerprint detection algorithm.

global orientation field model for the continuous flow field of the fingerprint with the local orientation field model around the singular points [66, 77, 129]. The global orientation field model represents either arch type fingerprints, which do not have any singularity, or the overall ridge orientation field except singularity in fingerprints. If the global orientation field model alone is used for orientation field approximation, the difference between the observed orientation field and the model will ideally be non-zero only around the singular points. On the other hand, for obfuscated fingerprints, the model fitting error is observed in the altered region as well. Thus, we use the difference between the observed orientation field extracted from the fingerprint image and the orientation field approximated by the model as a feature vector for classifying a fingerprint as natural fingerprint or altered one. The main steps of the proposed algorithm to detect altered fingerprints at orientation field level are described below:

1. Normalization: An input fingerprint image is normalized to 512×480 pixels by cropping a rectangular region of the fingerprint, which is located at the center of the fingerprint and aligned along the longitudinal direction of the finger, using the NIST Biometric Image Software (NBIS) [132]. This step ensures that the features extracted in the subsequent steps are invariant with respect to translation and rotation of the finger.

2. Orientation field extraction: The orientation field of the fingerprint, $\theta(x, y)$, is com-

143

puted using the gradient-based method [34]. The initial orientation field is smoothed by a 16×16 averaging filter, followed by averaging the orientations in 8×8 pixel blocks. Foreground mask is obtained by measuring the dynamic range of gray values of the fingerprint image in local blocks and morphological process for filling holes and removing isolated blocks is performed.

3. Orientation field approximation: The orientation field $\theta(x, y)$ is approximated by a polynomial model [66] to obtain $\hat{\theta}(x, y)$.

4. Feature extraction: The error map, $\varepsilon(x, y)$, is computed as the absolute difference between $\theta(x, y)$ and $\hat{\theta}(x, y)$ and used to construct the feature vector.

More details of steps 3 and 4 are given below.

5.3.1.1   Orientation Field Approximation

To represent the global orientation field, a set of polynomials is used, which is not only computationally efficient, but also provides a good approximation in orientation field modeling. Let $\theta(x, y)$ denote the orientation field. Then, the x-derivative and y-derivative of the doubled orientation at (x, y) can be represented by polynomials of order n:

$$\dot{x} = f_n(x, y) = \sum_{k=0}^{n} \sum_{l=0}^{k} a_{kl} x^{k-l} y^l, \tag{5.1}$$

$$\dot{y} = g_n(x, y) = \sum_{k=0}^{n} \sum_{l=0}^{k} b_{kl} x^{k-l} y^l, \tag{5.2}$$

where $\{a_{kl}\}$ and $\{b_{kl}\}$ are the polynomial coefficients for $f_n(x, y)$ and $g_n(x, y)$, respectively.

As the order of the polynomials increases, the model becomes more flexible in representing abrupt changes in the orientation field. When the order of the polynomial model is too low, the orientation field approximated by the model is quite different from the true orientation field. However, the order of the polynomial model does not need to be very high; polynomial

models with 6 or higher order do not make significant difference in the fitting results. Thus, we select the order of the polynomial model as 6 (i.e., $n = 6$).

Using the orientation field $\theta(x, y)$ obtained in step 2 (orientation field extraction), the polynomial coefficients a_{kl} and b_{kl} can be estimated by the least-squares method. For simplicity, we represent Equations (5.1) and (5.2) in matrix form:

$$f_n(x, y) = \mathbf{x}^T \mathbf{a}, \quad g_n(x, y) = \mathbf{x}^T \mathbf{b}, \tag{5.3}$$

where $\mathbf{x} = [1\ x\ y\ x^2\ xy\ y^2\ \cdots\ x^n\ \cdots\ y^n]^T$, and $\mathbf{a}$ and $\mathbf{b}$ are the corresponding coefficient vectors. The problem of estimating $\mathbf{a}$ and $\mathbf{b}$ can be formulated as:

$$\hat{\mathbf{a}} = \arg\min_{\mathbf{a}} \|\mathbf{f}_n - \mathbf{X}\mathbf{a}\|^2, \quad \hat{\mathbf{b}} = \arg\min_{\mathbf{b}} \|\mathbf{g}_n - \mathbf{X}\mathbf{b}\|^2, \tag{5.4}$$

where

$$\mathbf{f}_n = \begin{bmatrix} f_n(x_1, y_1) \\ f_n(x_2, y_2) \\ \vdots \\ f_n(x_N, y_N) \end{bmatrix}, \mathbf{g}_n = \begin{bmatrix} g_n(x_1, y_1) \\ g_n(x_2, y_2) \\ \vdots \\ g_n(x_N, y_N) \end{bmatrix}, \text{ and } \mathbf{X} = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_N^T \end{bmatrix}$$

from N observations of $\theta(x, y)$, where $(x, y) \in \mathcal{R}$ and $\mathcal{R} = \{(x, y): (x, y)$ in foreground$\}$.
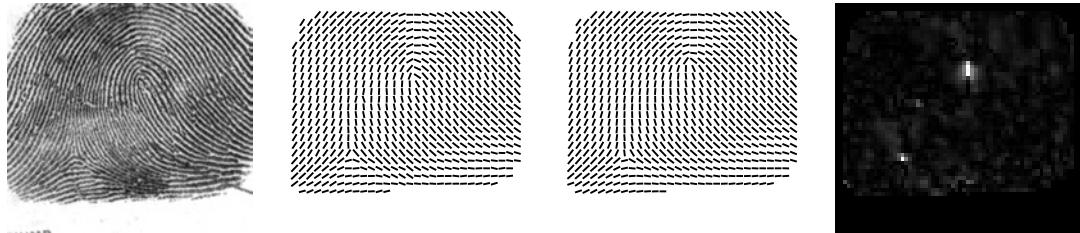
Finally, the orientation field approximated by the polynomial model, $\hat{\theta}(x, y)$, is obtained by:

$$\hat{\theta}(x, y) = \frac{1}{2} \tan^{-1} \left(\frac{\hat{g}_n(x, y)}{\hat{f}_n(x, y)} \right), \tag{5.5}$$
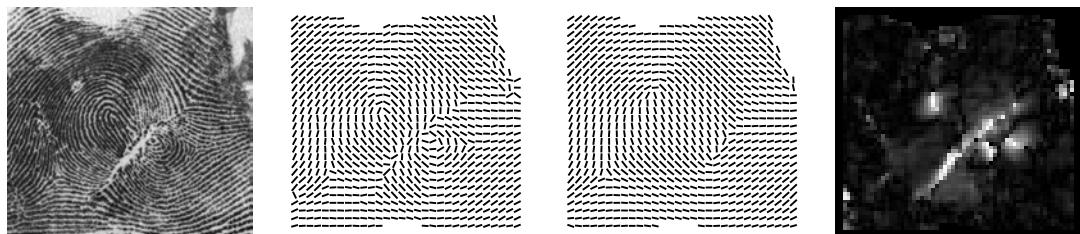
where $\hat{f}_n(x, y) = \mathbf{x}^T \hat{\mathbf{a}}$ and $\hat{g}_n(x, y) = \mathbf{x}^T \hat{\mathbf{b}}$.

5.3.1.2 Feature Extraction
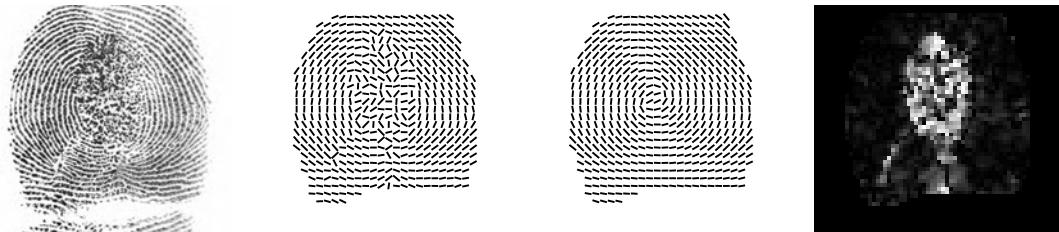
While the low order polynomial model can adequately represent smooth (global) changes in the orientation field, it cannot accurately model the abrupt changes in the orientation field
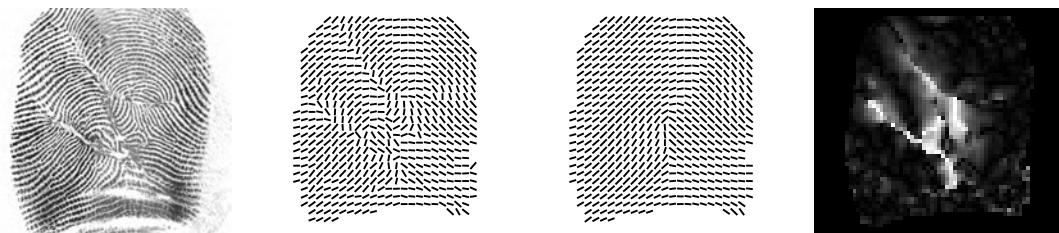
(a) Natural fingerprint (NIST SD14, F0000001)
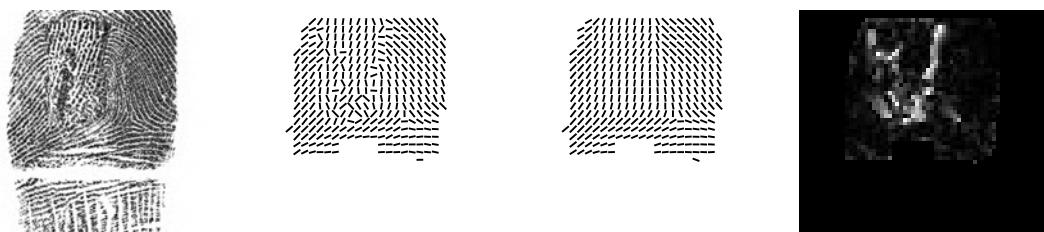
(b) Scarred fingerprint

(c) Mutilated fingerprint

(d) Distorted fingerprint by 'Z' cut

(e) Distorted fingerprint by transplantation from other friction ridge skin

Figure 5.13: Orientation field discontinuity. Column 1: fingerprint image; Column 2: orientation field extracted from the image, $\theta(x, y)$; Column 3: orientation field approximated by the polynomial model, $\hat{\theta}(x, y)$; and Column 4: error map, $\varepsilon(x, y)$.

in local areas, e.g., around cores and deltas in natural fingerprints. One of the observed characteristics of altered fingerprints is that their ridge flow can be discontinuous in non-singular regions as well, such as severely scarred areas (Figure 5.8(a)), mutilated areas (Figure 5.8(b)), and distorted ridge areas (Figures 5.9(a) and (b)). The difference between the observed orientation field and the modeled orientation field indicates the locations and the amount of abrupt changes in the ridge flow.

We define the error map $\varepsilon(x, y)$ as:

$$\varepsilon(x, y) = \min\left(|\theta(x, y) - \hat{\theta}(x, y)|, \pi - |\theta(x, y) - \hat{\theta}(x, y)|\right)/(\pi/2). \tag{5.6}$$

Figure 5.13 shows the error maps of a natural fingerprint and four different altered finger-prints. The size of the error map is 60×60 blocks after removing two columns from each side of the error map.

The feature vector extracted from the error map consists of histograms of local spatial regions [47]. The error map is divided into 3×3 cells, where each cell is of size 20×20 blocks. A histogram of the error map in each cell is computed in 21 bins in the range $[0, 1]$, and the histograms from all the 9 cells result in a 189-dimensional feature vector.

5.3.2 Abnormality in Minutiae Distribution

A minutia in the fingerprint indicates ridge characteristics such as ridge ending or ridge bi-furcation. Almost all fingerprint recognition systems use minutiae for matching. In addition to the abnormality observed in the orientation field, we also noted that minutiae distribution of altered fingerprints often differs from that of natural fingerprints.
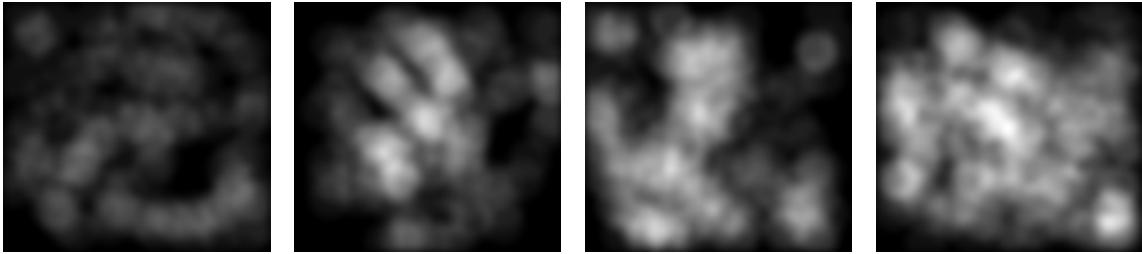
Based on the minutiae extracted from a fingerprint by the open source minutiae extrac-tor in NBIS, a minutiae density map is constructed by using the Parzen-window method with uniform kernel function. Let $\mathbf{S_m}$ be the set of minutiae of the fingerprint, i.e., $\mathbf{S_m} = \{\mathbf{x}|\mathbf{x} = (x, y) \text{ is the position of minutia}\}$. Then, the minutiae density map from $\mathbf{S_m}$

(a) Fingerprint image

(b) Minutiae extracted from the image

(c) Minutiae density map

Figure 5.14: Minutiae density map. Column 1: Natural fingerprint (NIST SD14, F0001826); Column 2: distorted fingerprint with dense minutiae along scars; Column 3: obliterated fingerprint with dense minutiae distribution in the altered region; Column 4: obliterated fingerprint with dense minutiae distribution over the entire altered region due to ridge-like pattern formed by alteration. Note that the minutiae density maps have been scaled to the same grayscale range.

is computed as follows:

1. Initial estimation: Initial minutiae density map, $M_d(\mathbf{x})$, is obtained by

$$M_d(\mathbf{x}) = \sum_{\mathbf{x_o} \in \mathbf{S_m}} K_r(\mathbf{x} - \mathbf{x_o}), \qquad (5.7)$$

where $K_r(\mathbf{x} - \mathbf{x_o})$ is a uniform kernel function centered at $\mathbf{x_o}$ with radius r (r is set to 40 pixels).

2. Low-pass filtering: $M_d(x, y)$ is smoothed by a Gaussian filter of size 30×30 pixels with a standard deviation of 10 pixels.

3. Normalization: $M_d(x, y)$ is transformed to lie in the interval $[0, 1]$ by

$$M_d(x, y) = \begin{cases} M_d(x, y)/T & \text{if } M_d(x, y) \leq T \\ 1 & \text{otherwise,} \end{cases} \tag{5.8}$$

where T is a predetermined threshold.

Figure 5.14 shows the minutiae density maps of a natural fingerprint and three altered fingerprints. In the natural fingerprint, minutiae are well spread and distributed almost uniformly. In the altered fingerprints, on the other hand, the distributions of minutiae are quite different: (i) many spurious minutiae are extracted along scars and in the obliterated region due to ridge discontinuity, and (ii) an excessive number of minutiae appear when a new ridge-like pattern is formed after alteration. These examples demonstrate that minutiae distribution can be useful for detecting altered fingerprints.

A feature vector from the minutiae density map is also constructed by local histograms in 3×3 cells. Then, the feature vectors from the orientation field discontinuity map and the minutiae density map are combined by concatenating local histograms in each cell, and fed into a support vector machine (SVM) for classification.

5.3.3 Experimental Results

The proposed algorithm was evaluated at two levels: finger level (one finger) and subject level (all ten fingers). At the finger level, we evaluate the performance of distinguishing between natural and altered fingerprints. At the subject level, we evaluate the performance of distinguishing between subjects with natural fingerprints and those with altered finger-prints. Since most AFIS used in law enforcement, national identification, and border control
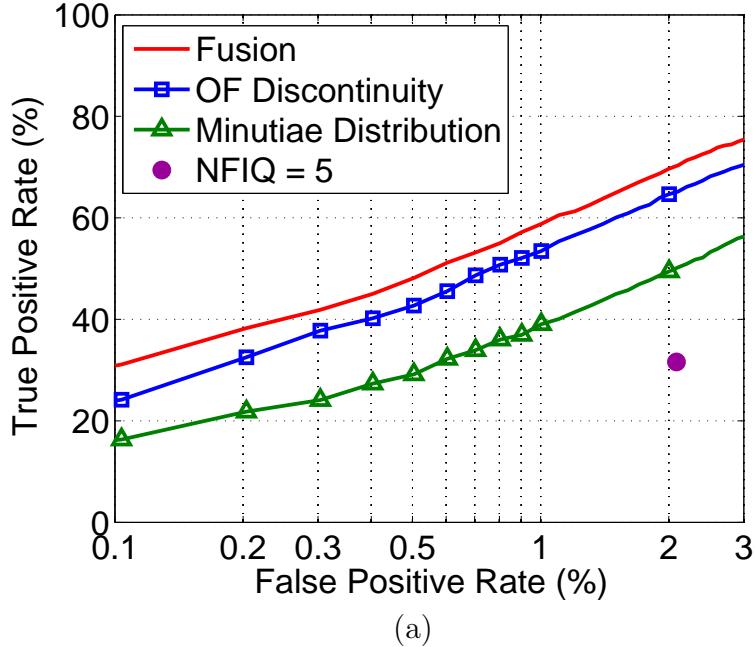
applications process all ten of a person's fingerprints, the subject-level performance utilizes this domain-specific information for detecting individuals with altered fingerprints.

5.3.3.1 Finger-Level Evaluation
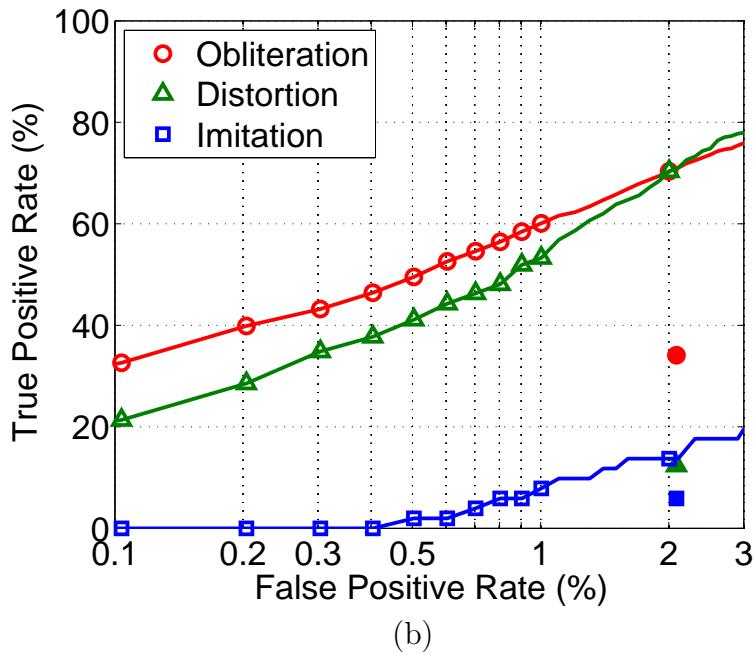
The altered fingerprint database available to us contains 4,433 fingerprints from 535 tenprint cards. For the non-altered fingerprint database, we use 27,000 fingerprints from the 2,700 tenprint cards in the NIST SD14 [10]. This database contains two impressions for each finger, called file and search; the file impression is used in our experiments.

LIBSVM [41] with radial basis kernel function is used for classification with 10-fold cross-validation. The scores output by LIBSVM are linearly scaled to the range $[0, 1]$. The normalized score is used as a measure of the deviation from natural fingerprint pattern, and denoted as F. When F of an input image is smaller than a predetermined threshold value, the system raises an alarm to indicate that the image is a possible altered fingerprint. If this image is indeed an altered fingerprint, it is deemed to be a true positive; otherwise, it is deemed to be a false positive. Similarly, true negative indicates that a natural fingerprint is correctly classified as natural and false negative indicates that an altered fingerprint is not detected as altered.

The ROC curves of the proposed approach and NFIQ software for detecting altered fingerprints are given in Figure 5.15. At the false positive rate of 2.1%, where natural fingerprints in NIST SD14 with the NFIQ value of 5 are determined as altered fingerprints, the proposed algorithm attains a 70.2% true positive rate while the true positive rate of the NFIQ is only 31.6%. Figure 5.15(a) shows the ROC curves for the three approaches for detecting altered fingerprints (orientation field discontinuity, minutiae distribution, and their feature-level fusion) and the NFIQ algorithm. Figure 5.15(b) shows the ROC curves of the proposed fusion algorithm and the NFIQ algorithm according to alteration type. Both obliterated and distorted fingerprints can be detected by the proposed algorithm at similar accuracy while NFIQ can only identify obliterated cases. On the other hand, imitated

Figure 5.15: ROC curves of the proposed algorithm and NFIQ criterion in detecting altered fingerprints. (a) The ROC curves of the three approaches in the proposed algorithm and the NFIQ algorithm and (b) the ROC curves of the proposed fusion algorithm and the NFIQ algorithm for each type of altered fingerprints. The ROC curve of NFIQ criterion is shown as a set of points (only one point is visible in the range of false positive rate plotted here) because its output can only take one of the five quality levels.

Table 5.3: NFIQ Distribution for False Positive Cases at the False Positive Rate of 1% by the Proposed Algorithm

NFIQ Value	1	2	3	4	5
Number of Images	2	1	39	145	83

(a) F=0.56, NFIQ=1   (b) F=0.42, NFIQ=4   (c) F=0.36, NFIQ=1   (d) F=0.48, NFIQ=1

Figure 5.16: True positive detection cases by (a) orientation field discontinuity, (b) minutiae distribution, (c) and (d) fusion of both approaches.

fingerprints are challenging for both algorithms.

At the false positive rate of 1% (which means 270 fingerprints among the 27,000 in NIST SD14 would be misclassified as altered fingerprints), the threshold value for F is 0.60. Figure 5.16 shows examples of successfully detected alterations using the proposed algorithm even though the NFIQ measure assigns acceptable quality level to these images.

Not all the altered fingerprints can be detected by the proposed algorithm. If the altered area is too small (Figure 5.17(a)), the evidence of alteration is difficult to detect. In the imitation case, the ridge structure is very natural even at the boundary of the altered region; the orientation field is continuous and there is insignificant abnormality in minutiae density along scars (Figure 5.17(b)).

The main reasons for false positive cases are: (i) poor image quality, leading to incorrect fingerprint feature extraction (see Figure 5.18(a)) and (ii) ground truth error; some of the fingerprints in NIST SD14 may possibly have been altered (see Figures 5.18(b), (c), and (d))! Table 5.3 shows the NFIQ distribution of the false positive examples found by the proposed algorithm at the false positive rate of 1%. Most of the false positive images have NFIQ of 4 or 5. Note that it is acceptable to raise alarms on poor quality fingerprints since (i) poor

(a) F = 0.92, NFIQ = 1   (b) F = 0.86, NFIQ = 1
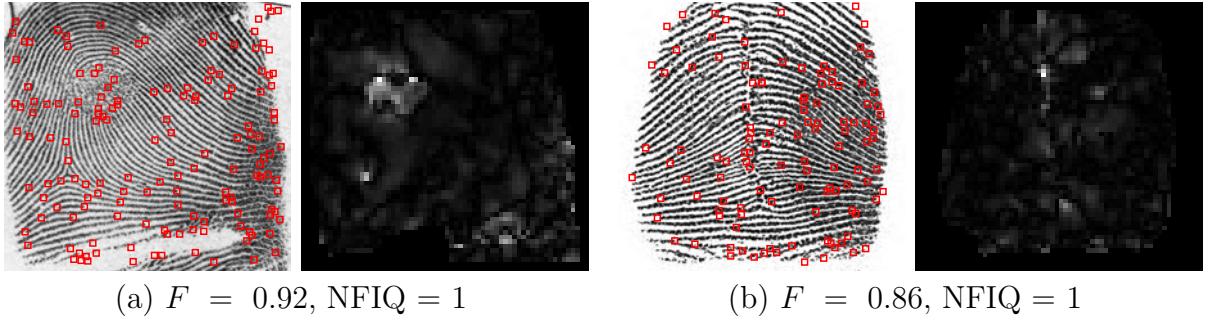
Figure 5.17: False negative examples of the proposed algorithm. Minutiae and orientation field discontinuities of each example are shown. (a) Fingerprint with small altered area and (b) imitated fingerprint. Note that NFIQ also fails to detect these two altered fingerprints.

(a) F = 0.33, NFIQ=5   (b) F = 0.57, NFIQ=1   (c) F = 0.58, NFIQ=2   (d) F = 0.52, NFIQ=3
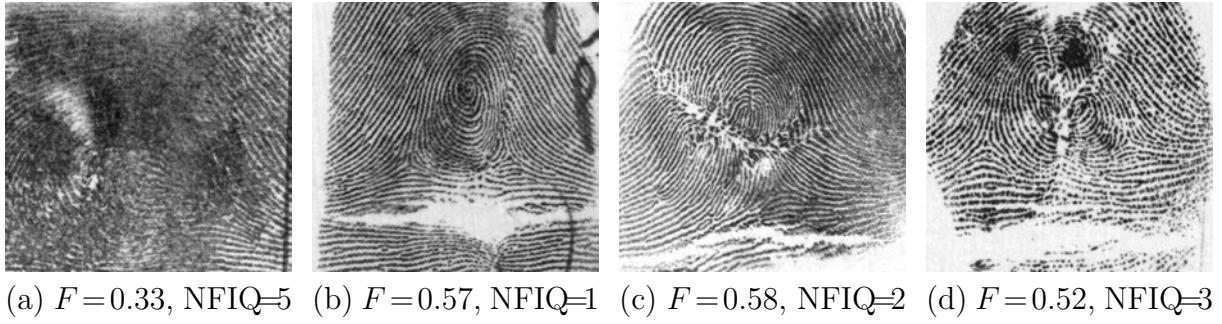
Figure 5.18: False positive examples of the proposed algorithm. Poor quality ridge patterns: (a) NIST SD14, F0010811; and possibly altered fingerprints: (b) F0019979, (c) F0002962, and (d) F0018103.

quality images need to be manually checked and (ii) criminals may purposely present poor quality fingerprints to the fingerprint system to evade identification [133]. All three false positive cases with NFIQ = 1 or 2 appear to have been altered (two of them are shown in Figures 5.18(b) and (c)).

5.3.3.2 Subject-Level Evaluation

In our altered fingerprint database, we observed that when a person resorts to fingerprint alteration, he tries to alter as many fingers as possible (Figure 5.3). This makes sense since large-scale AFIS applications typically use a fusion of match scores from all ten fingerprints for identification. So, altering just one or two fingerprints is not likely to change the identification decision. Based on this observation, we use the following decision-level fusion rule to
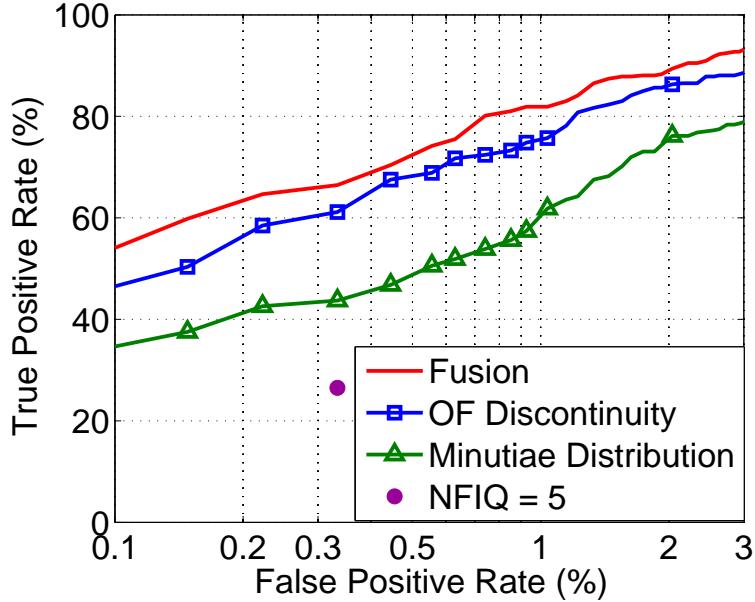
Figure 5.19: ROC curves of the proposed algorithm (including three approaches) and NFIQ criterion for detecting altered fingerprints at subject level.

perform the subject-level detection for altered fingerprints. When six or more fingerprints are detected as altered, the subject is claimed to have altered fingerprints. Subjects with fewer than six altered fingerprints are not considered as a threat to the AFIS since even five (out of ten) natural fingerprints are generally sufficient for reliable identification.

For the subject-level evaluation, 453 tenprint cards with more than five altered fingerprints and 2,700 tenprint cards in NIST SD14 are used. Figure 5.19 shows the ROC curves of the proposed algorithm (including three approaches) as well as the NFIQ criterion for detecting subjects with altered fingerprints. At a false positive rate of 0.3%, where the NFIQ criterion determines subjects with six or more fingerprints of NFIQ = 5 in NIST SD14 as persons who altered their fingerprints, the proposed algorithm attains a true positive rate of 66.4% while the NFIQ criterion obtains a 26.5% true positive rate.

Figure 5.20 shows an example of a tenprint card where the subject-level decision is successful. Even though one altered finger is not correctly detected due to the smoothness of the orientation field and the absence of abnormality in minutiae distribution in the altered region, our subject-level fusion algorithm still flags this person because as many as nine

Figure 5.20: True positive example of detection at subject level by the proposed algorithm. Although one of the altered fingerprints was not detected, this subject is still detected as having altered fingerprints with high confidence since the other nine fingerprints (boxed fingerprints) are correctly detected as altered. None of the ten fingerprints is detected as altered using the NFIQ criterion.
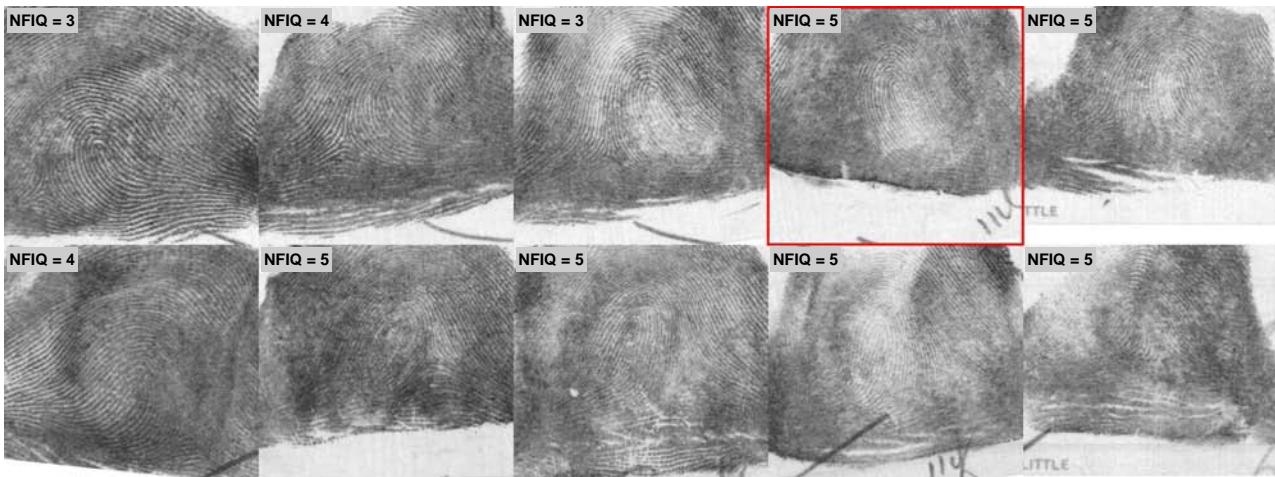
Figure 5.21: True negative example at subject level identified by the proposed algorithm (NIST SD14, F0000121–F0000130). This subject can pass our alteration detector since the nine fingerprints are determined to be natural fingerprints by the proposed algorithm. However, the NFIQ criterion raises a false alarm for this subject since six of the fingerprints have the NFIQ value of 5.

fingers are determined to be altered.

Fusion of multiple fingerprints also helps to reduce the false positive for a person who either did not alter his fingerprints or simply has one or two fingerprints that appear to have been altered due to accidents or occupational reasons. Figure 5.21 shows one such example.

In this case, however, the NFIQ criterion will falsely raise an alarm for this subject since six of the ten fingerprints are assigned the NFIQ value of 5.

We also have access to a small altered fingerprint database (254 images) from another government agency. This database has larger variance in terms of image format such as compression method, image resolution, and image type (single finger impressions, slap impressions, and tenprint cards). As a result, we report the detection performance on this database separately. We trained an SVM using all the 4,433 images in our first altered fingerprint database and tested on this second small database. The same NFIQ criterion was also used as a comparison. At the false positive rate of 2.1%, the proposed algorithm shows a 33.1% true positive rate compared to 9.4% for the NFIQ criterion.

5.4 Matching of Altered Fingerprints

Fingerprint alteration does not always fulfill its intended purpose, namely masking one's identity, for the following reasons. (i) Permanent modification or removal of friction ridge pattern often fails since the ridge details in the epidermis layer reappear on the surface of the skin after a few months [46]. (ii) Although the local fingerprint information in the altered region is lost, the true identity can still be established based on ridge details in the unaltered area. In 1941, Roscoe Pitts, a habitual criminal, had plastic surgery performed to remove the skin of his fingertips and replace it with skin grafts from his chest [14]. After he was arrested by the police, his true identity was revealed as Robert Philipps by comparing the second joints of his fingers with the original fingerprint card.

There is also a reported case where altered fingerprints were submitted as a query to AFIS. In 1995, a man named Alexander Guzman was arrested by Florida officials for possessing a false passport. His fingerprints were found to have been altered by a 'Z' shaped cut on the fingertips: two triangular skin patches from the 'Z' cut were lifted, switched, and then stitched back (see Figure 5.1(b)). Through a manual restoration of his altered fingerprints

156

and a search against the FBI database, the restored fingerprints of Alexander Guzman were linked to the fingerprints of Jose Izquierdo, an absconding drug criminal [134].

Altered fingerprint matching is a challenging problem due to the following reasons: (i) friction ridge structure can be severely damaged by abrading, cutting, burning, or applying strong chemicals on fingertips, resulting in a number of unreliable minutiae (Figure 5.8); (ii) even if the ridge structure is well defined in local regions, orientation field can be highly unusual during the procedure of switching skin patches in cases of 'Z'-cut prints (Figure 5.9(a)); and (iii) minutiae in a well-defined ridge area may not belong to the fingerprint of interest if a portion of skin on the fingertip was transplanted from other parts of the body (e.g., palm or sole) (Figure 5.9(b)).

The matching phase can be divided into two parts: (i) altered fingerprint restoration and (ii) altered fingerprint matching. Among altered fingerprint types, 'Z'-cut cases are of special interest since the original ridge structure of the finger is still retained in the finger, but in different positions. Once the 'Z'-cut prints are detected, the ridge structure in the triangular patches can be restored by reversing the transposing procedure. The restored 'Z'-cut fingerprint and all other altered fingerprints are submitted to a special matcher which is robust to a large amount of skin distortion and utilizes local minutiae information.

We investigate the feasibility of a state-of-the-art commercial fingerprint matcher to link altered fingerprints to their pre-altered mates [141] by (i) refining the minutiae that are automatically extracted by the matcher and (ii) restoring the altered region in the 'Z' shaped cut by swapping two triangular skin patches with each other.

5.4.1   Restoration of Altered Fingerprints

Minutiae distribution in altered fingerprints is significantly affected by severe skin distortion introduced during the process of alteration. Restoration of altered fingerprints attempts to relocate the minutiae to their original positions by undoing the alteration process, so that skin distortion can be relaxed. However, this is a very challenging problem because (i) each
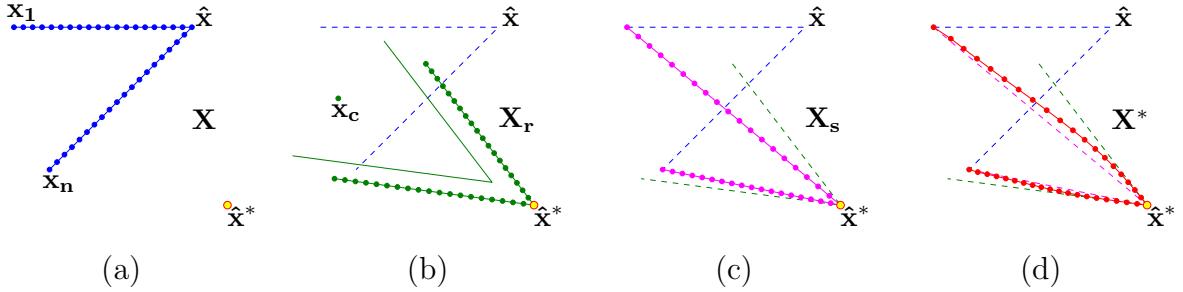
Figure 5.22: Transformation of a skin patch in 'Z'-cut fingerprint. (a) Points on a boundary ($\mathbf{X}$) including a vertex ($\hat{\mathbf{x}}$) and a new position of the vertex ($\hat{\mathbf{x}}^*$), (b) rigid transformation of $\mathbf{X}$ ($\mathbf{X_r}$), (c) scaling of $\mathbf{X}$ ($\mathbf{X_s}$), and (d) weighted sum of $\mathbf{X_r}$ and $\mathbf{X_s}$ ($\mathbf{X}^*$).

application of 'Z'-cut alteration leads to a different outcome in ridge pattern, and (ii) the alteration process often discards a portion of the fingerprint.

The following procedure is proposed to restore 'Z'-cut fingerprints:

1. Mark the edges of each skin patch along the scars and determine a vertex of each patch, $\hat{\mathbf{x}}$ in Figure 5.22(a), and its new position, $\hat{\mathbf{x}}^*$ in Figure 5.22(a), to define the rigid transformation of the skin patch (Figure 5.23(b)). One edge of the triangular region is opened and connected to the rest of the fingerprint.

2. Normalize each skin patch to make the boundaries of skin patches flat (Figure 5.23(c)).

3. Swap the two skin patches based on the thin-plate spline (TPS) model [35] (Figure 5.23(d)).

Normalizing and switching a skin patch follow a nonrigid transformation due to skin elasticity. The skin distortion is modeled by TPS for smooth transition of skin. The boundary of a skin patch is represented by a set of control points, $\mathbf{X}$; two ending points of the boundary are denoted as $\mathbf{x_1}$ and $\mathbf{x_n}$, and one of the points in $\mathbf{X}$ is selected as a vertex, $\hat{\mathbf{x}}$ (see Figure 5.22(a)). Correspondences in the TPS model are built by combining rigid transformation for rotation and translation with nonlinear scaling along two edges while preserving the

following constraints: (i) the vertex point, $\hat{\mathbf{x}}$, is mapped to a new vertex position, $\hat{\mathbf{x}}^*$; and (ii) the opened edge to the unaltered region stays in the same position.

The first constraint gives rigid transformation parameters (i.e., rotation matrix $\mathbf{R}$ and translation vector $\mathbf{t}$ in Equation (5.9)). For each point $\mathbf{x}$ in $\mathbf{X}$, the rigid transformation is applied as follows:

$$\mathbf{x_r} = \mathbf{R}(\mathbf{x} - \mathbf{x_c}) + \mathbf{x_c} + \mathbf{t}, \tag{5.9}$$

where $\mathbf{R}$ is the rotation matrix with angle θ,

$$\theta = \tan^{-1}\left(\frac{\hat{y}^*}{\hat{x}^*}\right) - \tan^{-1}\left(\frac{\hat{y}}{\hat{x}}\right), \tag{5.10}$$

$\hat{\mathbf{x}} = (\hat{x}, \hat{y})^T$, $\hat{\mathbf{x}}^* = (\hat{x}^*, \hat{y}^*)^T$, $\mathbf{x_c}$ is the center of the rotation defined by the center of two boundary ending points ($\mathbf{x_1}$ and $\mathbf{x_n}$), and $\mathbf{t}$ is the translation parameter to relocate the rotated patch to meet the new vertex, $\mathbf{t} = \hat{\mathbf{x}}^* - \hat{\mathbf{x}}_\mathbf{r}$.

Projection of the control points in $\mathbf{X}$ onto the new edges determined by $\overline{\mathbf{x_1}\hat{\mathbf{x}}^*}$ and $\overline{\mathbf{x_n}\hat{\mathbf{x}}^*}$ gives the scaling factor in smooth transition. For a point $\mathbf{x}$ in $\mathbf{X}$,

$$\mathbf{x_s} = s(\mathbf{x} \cdot \mathbf{e})\mathbf{e}, \tag{5.11}$$

where s is a scaling parameter defined as the ratio of $\overline{\mathbf{x_1}\hat{\mathbf{x}}}$ to $\overline{\mathbf{x_1}\hat{\mathbf{x}}^*}$ for the control points between $\mathbf{x_1}$ and $\hat{\mathbf{x}}$ or the ratio of $\overline{\mathbf{x_n}\hat{\mathbf{x}}}$ to $\overline{\mathbf{x_n}\hat{\mathbf{x}}^*}$ for the control points between $\mathbf{x_n}$ and $\hat{\mathbf{x}}$, and $\mathbf{e}$ is the unit vector directing from an ending point to the new vertex.

These two transformations are combined by their weighted sum:

$$\mathbf{x}^* = w(|\mathbf{x} - \hat{\mathbf{x}}|)\mathbf{x_r} + (1 - w(|\mathbf{x} - \hat{\mathbf{x}}|))\mathbf{x_s}, \tag{5.12}$$

where $w(r)$ is a weight function with respect to the distance from the vertex and defined as

$$w(r) = \frac{1}{1 + e^{-ar}}, \tag{5.13}$$

159
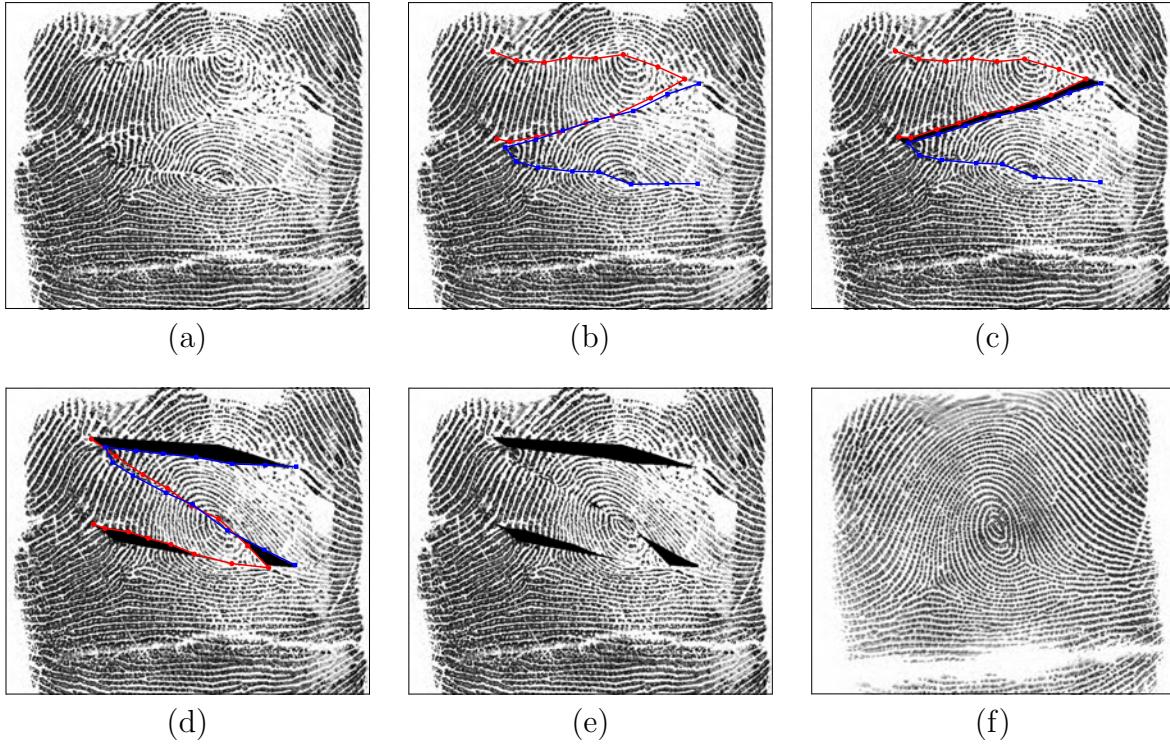
Figure 5.23: Restoration process. (a) Altered fingerprint image, (b) markups along the edges of each skin patch, (c) normalization of the boundaries, (d) swapping of two skin patches, (e) restored fingerprint, and (f) pre-altered fingerprint of (a).

where parameter a determines the slope of the sigmoid function that adjusts the transition rate between $\mathbf{X_r}$ and $\mathbf{X_s}$ along an edge.

Now, the TPS deformation model is specified by the correspondences between $\mathbf{X}$ and $\mathbf{X}^*$, in addition to the constraint that the open edge of a skin patch stays the same. Figure 5.23 shows the restoration procedure of a 'Z'-cut fingerprint and its pre-altered mate for comparison.

5.4.2 Matching of Altered Fingerprints

The number of valid minutiae can vary a lot according to the area of valid fingerprint region (i.e., the fingerprint region is unaltered and located in the ROI). The altered fingerprints in Figures 5.24(a) and 5.24(b) have very few minutiae that can be useful in matching. In this

160
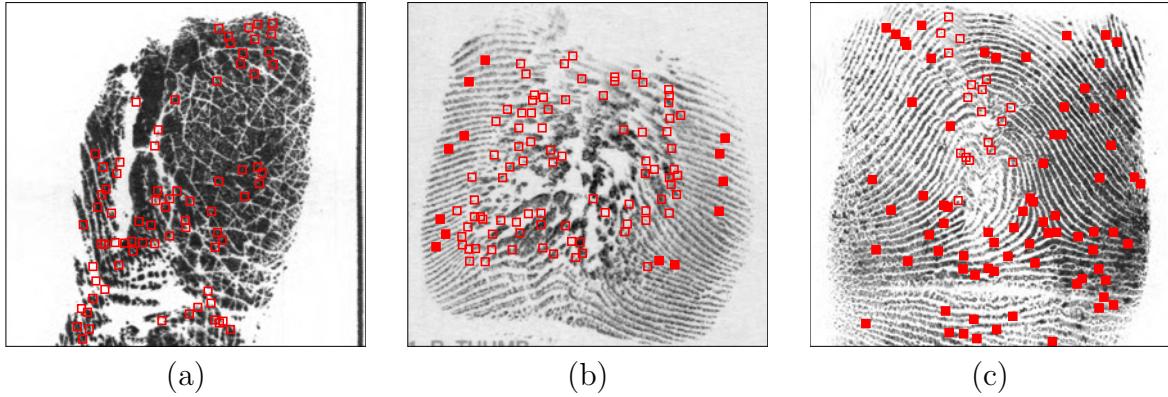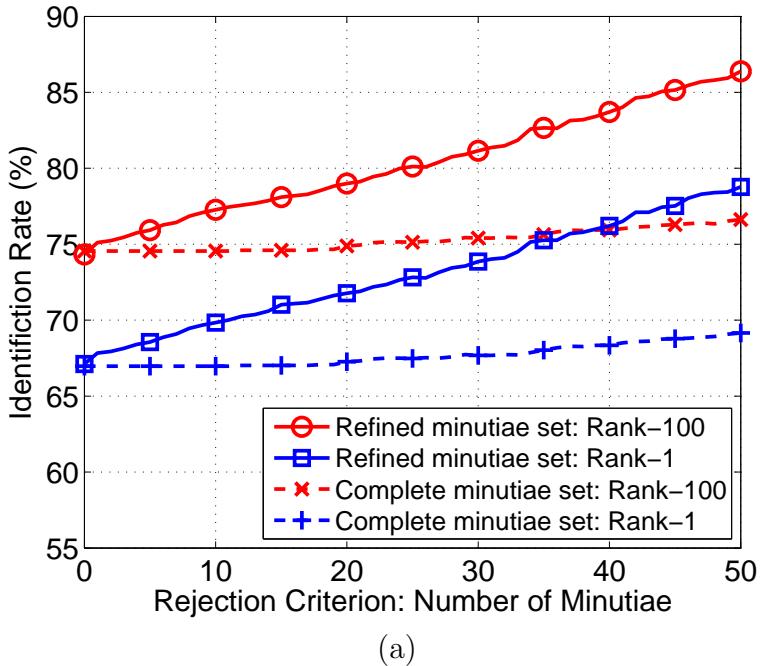
Figure 5.24: Complete minutiae set automatically extracted by the matcher (red squares) and minutiae in the valid fingerprint region (red-filled squares). (a) No valid minutiae, (b) very few minutiae in the valid region, and (c) abundant minutiae present in the valid region.
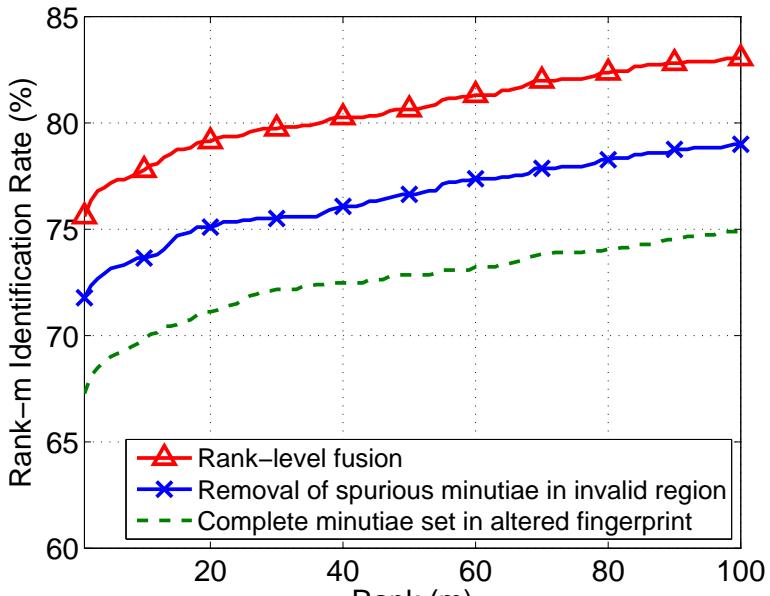
case, fingerprint matching based on minutiae may not be successful in finding corresponding mates in the database. On the other hand, fingerprints with large valid area contain a number of valid minutiae (see Figure 5.24(c)).

Three minutiae sets are evaluated: (i) all the minutiae extracted from the altered fingerprints, (ii) a subset of minutiae obtained from the altered fingerprints by removing spurious minutiae in the invalid fingerprint region, and (iii) a subset of minutiae obtained from the restored fingerprint image by removing spurious minutiae in the invalid fingerprint region. Note that all the minutiae in altered fingerprints are automatically extracted by the matcher, and then spurious minutiae in the invalid region are masked out.

Matching performance of altered fingerprints is evaluated by the CMC curves. We view altered fingerprint matching in the same spirit as latent fingerprint matching in the sense that these are high profile cases where a forensic examiner needs to examine the top N candidates retrieved from the background database. As a fingerprint matcher, Neurotechnology VeriFinger SDK 6.3 [13] is used to extract minutiae and match fingerprints. The altered fingerprint database consists of 1,332 pre/post-altered fingerprint pairs from 382 unique fingers; among them, 200 pairs are of 'Z'-cut type. If multiple pre-altered impressions of a finger exist, the best quality fingerprint image assessed by NFIQ software [124] is selected as

Figure 5.25: Matching performance of altered fingerprints. (a) Rejection criterion (i.e., the number of valid minutiae in the minutiae set) versus rank-1 and rank-100 identification rate, and (b) CMC curves at the fingerprint rejection threshold of 20 minutiae in the valid fingerprint region.
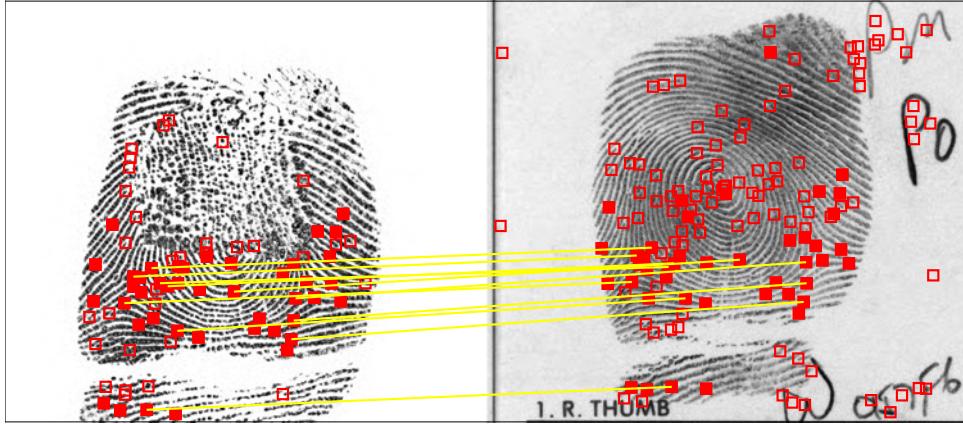
Figure 5.26: Example where the altered fingerprint matching is successful with all the minutiae extracted by the matcher. The pre-altered mate is retrieved at rank 1 from the database with 27,382 fingerprints.

an exemplar fingerprint. To enlarge the exemplar database size, 27,000 fingerprints in NIST SD14 were added.

A query fingerprint is rejected if the total number of minutiae in the valid fingerprint region is smaller than a threshold. Figure 5.25(a) shows the rank-1 and rank-100 identification rate with respect to the threshold for fingerprint rejection. The minutiae sets refined by removing spurious minutiae in the invalid fingerprint region significantly improve the matching performance compared to the minutiae sets containing all the minutiae from the altered fingerprints. At the threshold of 20 (i.e., minutiae sets with fewer than 20 minutiae in the valid region are rejected), the CMC curves for the complete minutiae set from the altered fingerprints, the refined minutiae set obtained by discarding minutiae in the invalid region, and their rank-level fusion are shown in Figure 5.25(b). The highest rank method is used for rank-level fusion.

Fingerprint alteration is not always successful in lowering the genuine match scores. Furthermore, the severity of the alteration does not predict degradation in matching performance. Figure 5.26 shows an example where the fingerprint alteration appears to be severe due to the skin transplantation over a large area. However, it can be successfully matched to its pre-altered mate; the match score with its true mate is sufficiently high to be correctly

163
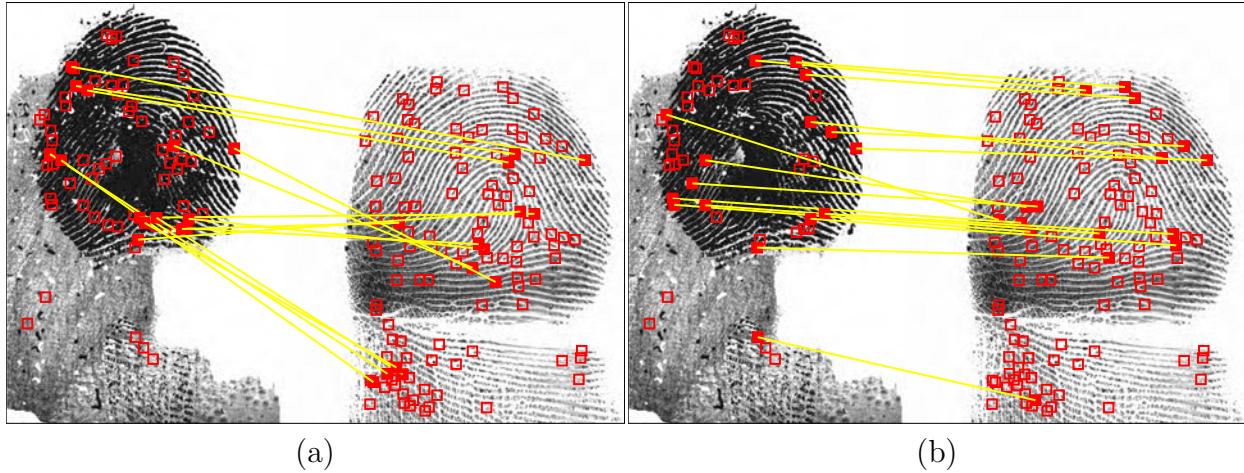
Figure 5.27: Example where removing spurious minutiae in the invalid region improves the matching performance. (a) Matching with the complete minutiae set in the altered fingerprint (match score is 9, and the pre-altered mate is retrieved at rank 10,093), and (b) matching with the refined minutiae set by removing spurious minutiae in the altered region (match score is 29, and the mate is retrieved at rank 1).
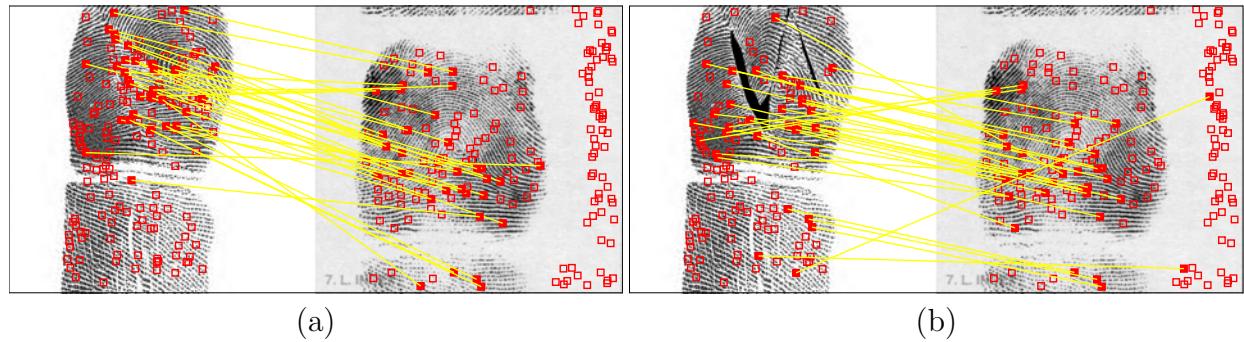
Figure 5.28: Example where restoration of the 'Z'-cut fingerprint improves the matching performance. (a) Matching result of the altered fingerprint with its pre-altered mate (match score is 5 and retrieval rank is 12,525), and (b) matching result of the restored image with the mate (match score is now 51 and retrieval rank is 1).

identified at the top rank.

Removal of spurious minutiae in the altered region can improve the matching performance (see Figure 5.27). In most altered fingerprint matching, it is observed that minutiae pairing results are globally inconsistent due to a number of spurious minutiae from scars and mutilated regions. By removing spurious minutiae, the matcher is able to find more

164
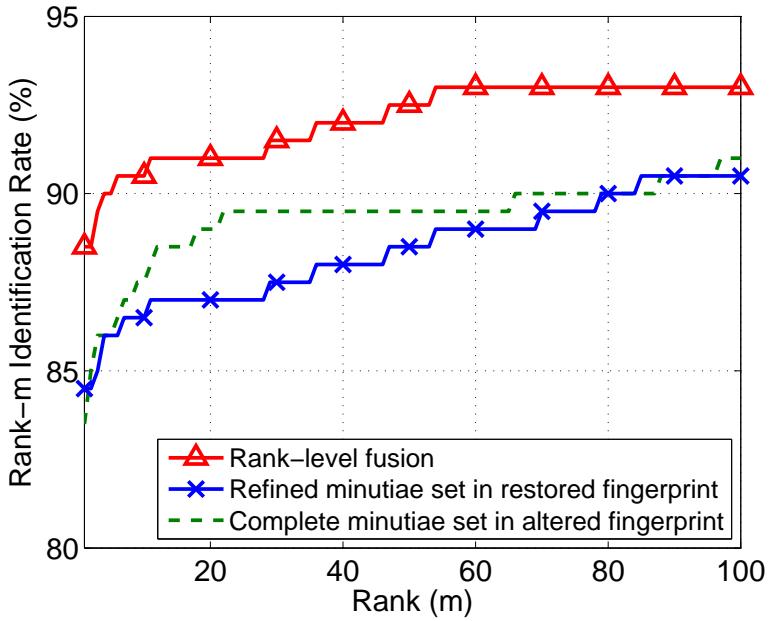
Figure 5.29: CMC curves for 'Z'-cut fingerprints at the fingerprint rejection threshold of 20 minutiae in the valid region. While matching with the restored images alone is still challenging, the rank-level fusion of the altered fingerprint matching and the restored fingerprint matching significantly improves the matching performance.

consistent mates in minutiae pairings, which results in a higher genuine match score.

For 'Z'-cut prints, the restoration of the distorted skin patches is helpful to improve the matching performance. Figure 5.28 shows an example where ridge structures of two local patches are successfully relocated. The restored image of the altered fingerprint shows much better consistency in minutiae pairing. Further, a significantly larger number of minutiae contribute to correct matchings. The CMC curves in Figure 5.29 show that the rank-level fusion of the minutiae from an altered fingerprint and the minutiae from its restored fingerprint helps to improve the overall matching performance.

Altered fingerprint matching fails mainly due to: (i) insufficient number of minutiae in the unaltered region which leads to rejection from minutiae-based matching, and (ii) a large amount of skin distortion that changes the structure of the neighboring minutiae (Figure 5.30).
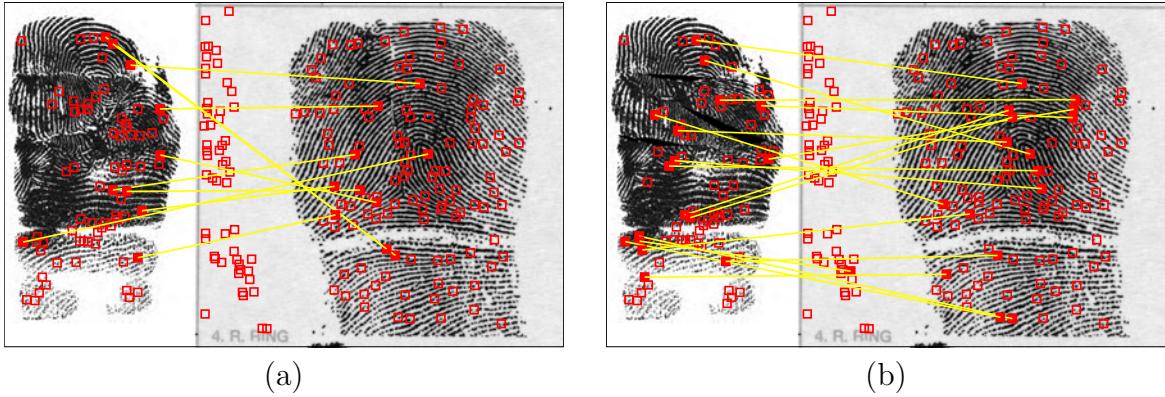
Figure 5.30: Example where the altered fingerprints cannot be correctly matched with their true mates using any refined minutiae sets. (a) Refined minutiae set from the altered image and (b) refined minutiae set from its restored image. Match scores for each pair in (a) and (b) are 3 and 9, respectively. Due to the large amount of skin distortion, none of these methods are successful in finding the correct mate within a practical size of the candidate list.

5.5   Summary

The success of AFIS and their extensive deployment all over the world have prompted some individuals to take extreme measures to evade identification by altering their fingerprints. The problem of fingerprint alteration or obfuscation is very different from that of fingerprint spoofing, where an individual uses a fake fingerprint in order to adopt the identity of another individual. While the problem of spoofing has received substantial attention in the literature, the problem of obfuscation has not been addressed in the biometric literature, in spite of numerous documented cases of fingerprint alteration for the purpose of avoiding identification. Large-scale fingerprint identification systems are facing a growing threat due to fingerprint alteration since, compared to other biometric modalities (e.g., face and iris), it is relatively easier to obfuscate or alter fingerprints.

We have introduced the problem of fingerprint alteration and conducted a quantitative analysis of the threat of altered fingerprints to a commercial fingerprint matcher. We also evaluated the capability of a well-known fingerprint image quality assessment software, NFIQ, for detecting altered fingerprints. Since the NFIQ has limited capability in distin-

guishing altered fingerprints from natural fingerprints, we developed an algorithm to automatically detect altered fingerprints based on the characteristics of the fingerprint orientation field and minutiae distribution. The proposed algorithm based on the features extracted from the orientation field and minutiae satisfies the three essential requirements for an alteration detection algorithm: (i) fast operational time, (ii) high true positive rate at low false positive rate, and (iii) easy integration into any AFIS. The proposed algorithm and the NFIQ criterion were tested on a large public domain fingerprint database (NIST SD14) as natural fingerprints and altered fingerprint databases provided by law enforcement agencies. At a false positive rate of 0.3%, the proposed algorithm can correctly detect 66.4% of the subjects with altered fingerprints, while 26.5% of such subjects are detected by the NFIQ algorithm.

Once altered fingerprints are detected, the subject will be sent to a secondary inspection for further investigation and for matching his altered fingerprints against an exemplar database. We show that some of the altered fingerprints still have sufficient ridge information to be identified even though the severity of the alteration performed to the fingerprint appears to be large. Given an exemplar fingerprint database consisting of 27,000 images from NIST SD14 and 382 images from a pre-altered mated fingerprint set, 86% of 1,332 altered fingerprints find their true mates within rank 100 by simply removing spurious minutiae in the invalid fingerprint region, while the minutiae sets including all the minutiae extracted from the altered fingerprint achieve a 77% identification rate at rank 100. In addition, restoration of 'Z'-cut fingerprints can boost the matching performance significantly by fusing the minutiae sets from the altered fingerprints and their restored versions at rank level.

This study can be further extended along the following directions:

1. Reconstruct altered fingerprints. For some types of altered fingerprints where the ridge patterns are damaged locally or the ridge structure is still present on the finger but possibly at a different location, reconstruction is indeed possible.

2. Match altered fingerprints to their pre-altered mates. A matcher specialized for altered fingerprints can be developed to link them to pre-altered mates in the database, which

is robust to skin distortion and that maximally uses local ridge structure in the valid fingerprint region.

3. Use multi-biometrics [115] to combat the growing threat of individuals evading AFIS. Federal agencies in the United States have adopted or are planning to adopt multi-biometrics in their identity management systems (FBI's NGI [11] and DoD's Automated Biometric Identification System (ABIS) [15]). However, other biometric traits can also be altered successfully. It has been reported that plastic surgery can significantly degrade the performance of face recognition systems [123] and that cataract surgery can reduce the accuracy of iris recognition systems [114]. To effectively deal with the problem of evading identification by altering biometric traits, a systematic study of possible alteration approaches for each major biometric trait is needed.

Chapter 6

Conclusions and Future Research

6.1 Conclusions

Friction ridge skin on fingers and palms has been purportedly known to be a unique physical characteristic of an individual that does not change over time and can be used as a person's "seal" or "signature" since ancient times. With the high and growing demand for reliable person recognition in modern era, fingerprint recognition has become one of the most successful solutions that achieves high recognition accuracy, ease of use, and high throughput. Fingerprints are now routinely collected not only from apprehended criminals, but also for the purpose of border crossing, entrance to amusement parks, and accessing laptops and mobile phones. Fingerprint recognition technology enables authorities to find suspect(s) involved in a crime and check the background of government job applicants.

Despite the widespread use and success of fingerprint recognition technology, there remain certain limitations and concerns about fingerprint recognition. This dissertation has made an attempt to address some of these challenging issues. A summary of our contributions is listed below.

1. *Persistence of fingerprints* Although fingerprint pattern is supposedly known to be permanent during one's life time and unique to an individual, no statistical study to

demonstrate the stability of fingerprint recognition accuracy over time has been reported. Based on a longitudinal fingerprint database collected from repeat offenders by a law enforcement agency, we have analyzed the genuine match scores obtained by two state-of-the-art fingerprint matchers. Our analysis based on multilevel statistical models demonstrated the impact of time interval between two fingerprints being matched on the corresponding genuine match score. While the genuine match score tends to decrease with increase in the time interval between two different impressions of the same finger and the subject's age, our analysis showed that fingerprint quality is a better predictor (covariate) of the variation in genuine match scores. Furthermore, an analysis of the binary identification decisions on genuine fingerprint pairs showed that true acceptance rates of fingerprint matchers tend to be stable even as the time interval between two fingerprints being compared increases.

2. *Fingerprint modeling* Fingerprint orientation field models are useful in distinguishing fingerprints from any other textured pattern, representing the fingerprint orientation field in a compact form, and estimating the orientation field from noisy or partially missing fingerprints. We have developed a global orientation field model that captures the distinctive characteristics of fingerprint ridge flow pattern, namely the number and type of singular points. The model has been evaluated by testing its capability to filter out non-fingerprint images in a fingerprint database.

3. *Latent fingerprint enhancement* The high accuracy of fingerprint recognition is conditioned on the high quality of fingerprint images being compared. However, latent fingerprints, unintentionally left on the surface of an object, often have very poor quality possibly with severe distortion. NIST evaluations of state-of-the-art latent fingerprint matchers show a large gap between exemplar fingerprint matching and latent fingerprint matching performances. The latent fingerprint enhancement algorithm that we have developed not only improves the matching accuracy, but also facilitates manual

170

markup by latent examiners.

4. *Fingerprint obfuscation* The widespread deployment of AFIS, especially at border crossings, has prompted some individuals with prior criminal record to purposely alter or change their fingerprint pattern in order to conceal their true identity. In order to identify such individuals, it is necessary to develop a method that detects "abnormal" fingerprint patterns. We have developed a simple and computationally efficient algorithm for this task. An input fingerprint is suspected to have been altered if (i) its orientation field does not fit well the fingerprint orientation field model, and (ii) its minutiae distribution deviates from the empirical distribution of normal fingerprints. Furthermore, we have developed methods to restore and match altered fingerprints in order to reveal the person's true identity from the unaltered ridge pattern remained after fingerprint alteration.

6.2   Future Research

Fingerprint recognition research presented here can be further advanced as follows:

- A statistical analysis of impostor match scores needs to be conducted to determine the temporal tendency of the impostor match scores and the false acceptance rate of the fingerprint matchers.

- Fingerprint orientation field modeling should be extended by integrating arch-like global orientation field in the model and designing a convex objective function for fast model fitting.

- Latent fingerprint enhancement algorithm can be further improved by utilizing a fingerprint-specific model in the orientation field estimation.

- Restoration and matching algorithms for altered fingerprints that can handle severe distortion need to be developed.

BIBLIOGRAPHY

Bibliography

[1] NIST Special Database 29: Plain and Rolled Images from Paired Fingerprint Cards. `http://www.nist.gov/srd/nistsd29.cfm`.

[2] NIST Special Database 4: NIST 8-Bit Gray Scale Images of Fingerprint Image Groups (FIGS). `http://www.nist.gov/srd/nistsd4.cfm`.

[3] NIST Special Database 27: Fingerprint Minutiae from Latent and Matching Tenprint Images. `http://www.nist.gov/srd/nistsd27.cfm`.

[4] Enhanced Spoof-proofing of Fingerprint Readers by Optical Coherence Tomography. `http://www0.egr.uh.edu/bol/biometric.html`.

[5] Liveness Detection. `http://www.lumidigm.com/liveness-detection/`.

[6] Surgically Altered Fingerprints. `http://www.clpex.com/images/FeetMutilation/L4.JPG`.

[7] The Federal Bureau of Investigation, Integrated Automated Fingerprint Identification System (IAFIS). `http://www.fbi.gov/about-us/cjis/fingerprints_biometrics/iafis/iafis`.

[8] The U.S. Department of Homeland Security, US-VISIT. `http://www.dhs.gov/us-visit-office`.

[9] Unique Identification Authority of India. `http://uidai.gov.in/`.

[10] NIST Special Database 14: NIST Mated Fingerprint Card Pairs 2 (MFCP2). `http://www.nist.gov/srd/nistsd14.cfm`.

[11] The FBI's Next Generation Identification (NGI). `http://www.fbi.gov/about-us/cjis/fingerprints_biometrics/ngi`.

[12] NIST Minutiae Interoperability Exchange Test (MINEX). `http://www.nist.gov/itl/iad/ig/minex04.cfm`.

[13] Neurotechnology Inc., VeriFinger. `http://www.neurotechnology.com/vf_sdk.html`.

[14] History of Fingerprint Removal. `http://jimfisher.edinboro.edu/forensics/fire/print.html`.

[15] DoD Biometrics Task Force. `http://www.biometrics.dod.mil/`.

[16] Frye v. United States. 54 App. D.C. 46, 293 F. 1013, 1923.

[17] Daubert v. Merrell Dow Pharmaceuticals Inc. 509 U.S. 579, 1993.

[18] Men in Black, 1997. `http://www.imdb.com/title/tt0119654/`.

[19] Doubt Cast on Fingerprint Security, May 2002. `http://news.bbc.co.uk/2/hi/science/nature/1991517.stm`.

[20] Sweden Refugees Mutilate Fingers, April 2004. `http://news.bbc.co.uk/2/hi/europe/3593895.stm`.

[21] Is That a Finger or a Jell-O Mold?, December 2005. `http://www.nytimes.com/2005/12/20/science/20find.html?_r=0`.

[22] EURODAC: A European Union-Wide Electronic System for the Identification of Asylum-Seekers, September 2006. `http://europa.eu/rapid/press-release_MEMO-06-334_en.htm`.

[23] Criminals Cutting Off Fingertips To Hide IDs, 2008. `http://www.thebostonchannel.com/news/15478914/detail.html`.

[24] Asylum Seekers Torch Skin off Their Fingertips So They Can't Be ID'd by Police, 2008. http://www.mirror.co.uk/sunday-mirror/2008/06/29/asylum-seekers-torch-skin-off-their-fingertips-so-they-can-t-be-id-d-by-police-98487-20624559/.

[25] Surgically Altered Fingerprints Help Woman Evade Immigration, December 2009. `http://abcnews.go.com/Technology/GadgetGuide/surgically-altered-fingerprints-woman-evade-immigration/story?id=9302505`.

[26] Altered Fingerprints Detected in Illegal Immigration Attempts, June 2009. `http://www.japantoday.com/category/crime/view/altered-fingerprints-detected-in-illegal-immigration-attempts`.

[27] Three Charged with Conspiring to Mutilate Fingerprints of Illegal Aliens, July 2010. `http://www.eagletribune.com/local/x739950408/Three-charged-with-conspiring-to-mutilate-fingerprints-of-illegal-aliens`.

[28] Lumidigm Wins Another Brazilian Bank for ATM Fingerprint Access, June 2012. `http://www.lumidigm.com/lumidigm-wins-another-brazilian-bank-for-atm-fingerprint-access/`.

[29] Authentec Laptop Fingerprint Sensor/Reader, April 2012. `http://support.authentec.com/KnowledgeBase/KBview/tabid/843/ArticleId/504/Laptop-Fingerprint-Sensor-Reader.aspx`.

[30] W. W. Adams and P. Loustaunau. *An Introduction to Gröbner Bases (Graduate Studies in Mathematics, Vol 3).* American Mathematical Society, 1994.

[31] F. Alonso-Fernandez, J. Fierrez, J. Ortega-Garcia, J. Gonzalez-Rodriguez, H. Fronthaler, K. Kollreider, and J. Bigun. A Comparative Study of Fingerprint Image-Quality Estimation Methods. *IEEE Transactions on Information Forensics and Security,* 2(4):734–743, 2007.

174

[32] A. Antonelli, R. Cappelli, D. Maio, and D. Maltoni. Fake Finger Detection by Skin Distortion Analysis. *IEEE Transactions on Information Forensics and Security*, 1(3):360–373, 2006.

[33] D. R. Ashbaugh. *Quantitative-Qualitative Friction Ridge Analysis: An Introduction to Basic and Advanced Ridgeology*. CRC Press, 1999.

[34] A. M. Bazen and S. H. Gerez. Systematic Methods for the Computation of the Directional Fields and Singular Points of Fingerprints. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7):905–919, 2002.

[35] F. L. Bookstein. Principal Warps: Thin Plate Splines and the Decomposition of Deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6):567–585, 1989.

[36] J. W. Burks. The Effect of Dermabrasion on Fingerprints: A Preliminary Report. *Archives of Dermatology*, 77(1):8–11, 1958.

[37] R. Cappelli, A. Lumini, D. Maio, and D. Maltoni. Fingerprint Image Reconstruction from Standard Templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(9):1489–1503, 2007.

[38] R. Cappelli, D. Maio, and D. Maltoni. Synthetic Fingerprint-Database Generation. In *Proceedings of International Conference on Pattern Recognition*, pages 744–747, 2002.

[39] R. Cappelli, D. Maio, D. Maltoni, J. L. Wayman, and A. K. Jain. Performance Evaluation Of Fingerprint Verification Systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(1):3–18, 2006.

[40] R. Cappelli and D. Maltoni. On the Spatial Distribution of Fingerprint Singularities. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(4):742–748, 2009.

[41] C. C. Chang and C. J. Lin. *LIBSVM: A Library For Support Vector Machines*, 2001. `http://www.csie.ntu.edu.tw/~cjlin/libsvm/`.

[42] Y. Chen and A. K. Jain. Beyond Minutiae: A Fingerprint Individuality Model with Pattern, Ridge and Pore Features. In *Proceedings of International Conference on Biometrics*, pages 523–533, 2009.

[43] S. Chikkerur, A. N. Cartwright, and V. Govindaraju. Fingerprint Enhancement Using STFT Analysis. *Pattern Recognition*, 40(1):198–211, 2007.

[44] S. A. Cole. *Suspect Identities: A History of Fingerprinting and Criminal Identification*. Harvard University Press, 2001.

[45] H. Cummins. Attempts to Alter and Obliterate Finger-prints. *Journal of American Institute of Criminal Law and Criminology*, 25:982–991, 1935.

[46] H. Cummins and M. Midlo. *Finger Prints, Palms and Soles: An Introduction to Dermatoglyphics*. Dover Publications, New York, 1961.

[47] N. Dalal and B. Triggs. Histograms of Oriented Gradients for Human Detection. In *Proceedings of Computer Vision and Pattern Recognition*, pages 886–893, 2005.

[48] J. Deng, W. Dong, R. Socher, L-J. Li, K. Li, and F-F. Li. Imagenet: A Large-Scale Hierarchical Image Database. In *Proceedings of Computer Vision and Pattern Recognition*, pages 248–255, 2009.

[49] P. J. Diggle, P. Heagerty, K-Y. Liang, and S. Zeger. *Analysis of Longitudinal Data*. Oxford University Press, 2003.

[50] G. Doddington, W. Liggett, A. Martin, M. Przybocki, and D. Reynolds. Sheep, Goats, Lambs and Wolves: A Statistical Analysis of Speaker Performance in the NIST 1998 Speaker Recognition Evaluation. In *International Conference on Spoken Language Processing*, 1998.

[51] H. Faulds. On the Skin-furrows of the Hand. *Nature*, page 605, 1880.

[52] Federal Bureau of Investigation (FBI). *The Science of Fingerprints : Classification and Uses*. United States Department of Justice, 1984.

[53] Federal Office for Information Security (BSI). Evaluation of Fingerprint Recognition Technology–BioFinger, 2004.

[54] J. Feng, A. K. Jain, and A. Ross. Detecting Altered Fingerprints. In *Proceedings of International Conference on Pattern Recognition*, pages 1622–1625, 2010.

[55] J. Feng, J. Zhou, and A. K. Jain. Orientation Field Estimation for Latent Fingerprint Enhancement. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013. To appear.

[56] S. P. Fenker and K. W. Bowyer. Experimental Evidence of a Template Aging Effect in Iris Biometrics. In *IEEE Workshop on Applications of Computer Vision*, pages 232–239, 2011.

[57] S. P. Fenker and K. W. Bowyer. Analysis of Template Aging in Iris Biometrics. In *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 45–51, 2012.

[58] M. A. Fischler and R. C. Bolles. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Communications of the ACM*, 24(6):381–396, 1981.

[59] G. Fitzmaurice, M. Davidian, G. Verbeke, and G. Molenberghs. *Longitudinal Data Analysis*. Boca Raton: CRC Press, 2009.

[60] F. Galton. Personal Identification and Description. *Nature*, pages 173–177, 201–202, 1888.

[61] F. Galton. *Finger Prints*. Macmillan, 1892.

[62] H. Goldstein. *Multilevel Statistical Models (Fourth Edition)*. Wiley, 2010.

[63] C. Gottschlich, T. Hotz, R. Lorenz, S. Bernhardt, M. Hantschel, and A. Munk. Modeling the Growth of Fingerprints Improves Matching for Adolescents. *IEEE Transactions on Information Forensics and Security*, 6(3):1165–1169, 2011.

[64] P. Grother, J. R. Matey, E. Tabassi, G. W. Quinn, and M. Chumakov. IREX VI: Temporal Stability of Iris Recognition Accuracy. NISTIR 7948, 2013.

[65] P. Grother and E. Tabassi. Performance of Biometric Quality Measures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(4):531–543, 2007.

[66] J. Gu, J. Zhou, and D. Zhang. A Combination Model for Orientation Field of Fingerprints. *Pattern Recognition*, 37(3):543–553, 2004.

[67] M. Hall. Criminals Go to Extremes to Hide Identities, November 2007. `http://www.usatoday.com/news/nation/2007-11-06-criminal-extreme_N.htm`.

[68] R. M. Haralick, K. Shanmugam, and I. Dinstein. Textural Features for Image Classification. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-3(6):610–621, 1973.

[69] M. R. Hawthorne. *Fingerprints: Analysis and Understanding*. CRC Press, 2009.

[70] D. Hedeker, R. D. Gibbons, and B. R. Flay. Random-effects Regression Models for Clustered Data with an Example from Smoking Prevention Research. *Journal of Consulting and Clinical Psychology*, 62(4):757–765, 1994.

[71] E. R. Henry. *Classification and Uses of Finger Prints*. George Routledge and Sons, 1900.

[72] W. J. Herschel. Finger-Prints. *Nature*, 51(1308):77–78, 1894.

[73] W. J. Herschel. *The Origins of Finger-Printing*. Oxford University Press, 1916.

[74] R. A. Hicklin, J. Buscaglia, M. A. Roberts, S. B. Meagher, W. Fellner, M. J. Burge, M. Monaco, D. Vera, L. R. Pantzer, C. C. Yeung, and T. N. Unnikumaran. Latent Fingerprint Quality: A Survey of Examiners. *Journal of Forensic Identification*, 61(4):385–419, 2011.

[75] M. W. Hirsch, S. Smale, and R. L. Devaney. *Differential Equations, Dynamical Systems, and an Introduction to Chaos (Second Edition)*. Elsevier Academic Press, 2004.

[76] L. Hong, Y. Wan, and A. K. Jain. Fingerprint Image Enhancement: Algorithm and Performance Evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(8):777–789, 1998.

[77] S. Huckemann, T. Hotz, and A. Munk. Global Models for the Orientation Field of Fingerprints: An Approach Based on Quadratic Differentials. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(9):1507–1519, 2008.

177

[78] M. Indovina, V. Dvornychenko, R. A. Hicklin, and G. I. Kiebuzinski. Evaluation of Latent Fingerprint Technologies: Extended Feature Sets [Evaluation #2]. NISTIR 7859, 2012.

[79] M. Indovina, V. Dvornychenko, E. Tabassi, G. Quinn, P. Grother, S. Meagher, and M. Garris. ELFT Phase II - An Evaluation of Automated Latent Fingerprint Identification Technologies. NISTIR 7577, 2009.

[80] A. K. Jain, S. Pankanti, S. Prabhakar, L. Hong, A. Ross, and J. L. Wayman. Biometrics: A Grand Challenge. In *Proc. International Conference on Pattern Recognition*, volume 2, pages 935–942, 2004.

[81] A. K. Jain, S. Prabhakar, and S. Pankanti. On the Similarity of Identical Twin Fingerprints. *Pattern Recognition*, 35(11):2653–2663, 2002.

[82] X. Jiang and W. Ser. Online Fingerprint Template Improvement. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(8):1121–1126, 2002.

[83] M. Kawagoe and A. Tojo. Fingerprint Pattern Classification. *Pattern Recognition*, 17(3):295–303, 1984.

[84] M. Kücken. Models for Fingerprint Pattern Formation. *Forensic Science International*, 171:85–96, 2007.

[85] N. M. Laird and J. H. Ware. Random-Effects Models for Longitudinal Data. *Biometrics*, 38(4):963–974, 1982.

[86] A. Lanitis. A Survey of the Effects of Aging on Biometric Identity Verification. *International Journal of Biometrics*, 2(1):34–52, 2010.

[87] H. C. Lee and R. E. Gaensslen. *Advances in Fingerprint Technology*. CRC Press, 2001.

[88] R. Leeden, F. Busing, and E. Meijer. Bootstrap Methods for Two-Level Models. In *Multilevel Conference*, 1997.

[89] J. Li, W-Y. Yau, and H. Wang. Constrained Nonlinear Models of Fingerprint Orientations with Prediction. *Pattern Recognition*, 39(1):102–114, 2006.

[90] C. J. M. Maas and J. J. Hox. The Influence of Violations of Assumptions on Multilevel Parameter Estimates and Their Standard Errors. *Computational Statistics and Data Analysis*, 46(3):427–440, 2004.

[91] D. Maltoni, D. Maio, A. K. Jain, and S. Prabhakar. *Handbook of Fingerprint Recognition (Second Edition)*. Springer-Verlag, 2009.

[92] A. J. Mansfield and J. L. Wayman. Best Practices in Testing and Reporting Performance. Center for Mathematics and Scientific Computing, National Physical Laboratory Teddington, U.K., 2002.

[93] J. Matas and O. Chum. Randomized RANSAC with T$_{d,d}$ Test. *Image and Vision Computing*, 22(10):837–842, 2004.

[94] T. Matsumoto, H. Matsumoto, K. Yamada, and S. Hoshino. Impact of Artificial Gummy Fingers on Fingerprint Systems. In *Proceedings of SPIE, Optical Security and Counterfeit Deterrence Techniques IV*, volume 4677, pages 275–289, 2002.

[95] K. R. Moses, P. Higgins, M. McCabe, S. Probhakar, and S. Swann. Chapter 6: Automated Fingerprint Identification System (AFIS). In *Fingerprint Sourcebook*. National Institute of Justice/NCJRS, 2010.

[96] K. Nandakumar, A. K. Jain, and A. Ross. Fusion in Multibiometric Identification Systems: What about the Missing Data? In *Proceedings of International Conference on Biometrics*, pages 743–752, 2009.

[97] National Institute of Justice (NIJ). *The Fingerprint Sourcebook*. 2012.

[98] National Institute of Standards and Technology (NIST). Fingerprint Vendor Technology Evaluation 2012. `http://www.nist.gov/itl/iad/ig/fpvte2012.cfm`.

[99] National Institute of Standards and Technology (NIST) and National Institute of Justice (NIJ). Latent Print Examination and Human Factors: Improving the Practice through a Systems Approach, 2012.

[100] National Research Council. Strengthening Forensic Science in the United States: A Path Forward, 2009.

[101] C. Neumann, C. Champod, R. Puch-Solis, N. Egli, A. Anthonioz, and A. Bromage-Griffiths. Computation of Likelihood Ratios in Fingerprint Identification for Configurations of Any Number of Minutiae. *Journal of Forensic Science*, 52(1):54–64, 2007.

[102] K. A. Nixon and R. K. Rowe. Multispectral Fingerprint Imaging for Spoof Detection. In *Proceedings of SPIE, Biometric Technology for Human Identification II*, volume 5779, pages 214–225, 2005.

[103] Office of the Inspector General, Oversight and Review Division. A Review of the FBI's Handling of the Brandon Mayfield Case, 2006.

[104] Office of the Inspector General, Oversight and Review Division. A Review of the FBI's Progress in Responding to the Recommendations in the Office of the Inspector General Report on the Fingerprint Misidentification in the Brandon Mayfield Case, 2011.

[105] S. Pankanti, S. Prabhakar, and A. K. Jain. On the Individuality of Fingerprints. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(8):1010–1025, 2002.

[106] S. T. V. Parthasaradhi, R. Derakhshani, L. A. Hornak, and S. A. C. Schuckers. Time-series Detection of Perspiration as a Liveness Test in Fingerprint Devices. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 35(3):335–343, 2005.

[107] J. Patten. Savvy Criminals Obliterating Fingerprints To Avoid Identification, March 2008. `http://www.eagletribune.com/punews/local_story_062071408.html`.

[108] P. Pedersen, M-F. Roy, and A. Szpirglas. Counting Real Zeros in the Multivariate Case. In F. Eyssette and A. Galligo, editors, *Computational Algebraic Geometry*, volume 109 of *Progress in Mathematics*, pages 203–224. Birkhäuser Boston, 1993.

[109] S. Petitjean. Algebraic Geometry and Computer Vision: Polynomial Systems, Real and Complex Roots. *Journal of Mathematical Imaging and Vision*, 10:191–220, 1999.

[110] H. Plotnick and H. Pinkus. The Epidermal vs. the Dermal Fingerprint: An Experimental and Anatomical Study. *Archives of Dermatology*, 77(1):12–17, 1958.

[111] S. Ram, H. Bischof, and J. Birchbauer. Modelling Fingerprint Ridge Orientation Using Legendre Polynomials. *Pattern Recognition*, 43(1):342–357, 2010.

[112] S. W. Raudenbush and A. S. Bryk. *Hierarchical Linear Models: Applications and Data Analysis Methods (Second Edition)*. SAGE Publications, 2002.

[113] J. A. Rice. *Mathematical Statistics and Data Analysis*. Cengage Learning, 2007.

[114] R. Roizenblatt, P. Schor, F. Dante, J. Roizenblatt, and R. Belfort. Iris Recognition as a Biometric Method after Cataract Surgery. *American Journal of Ophthalmology*, 140(5):969, 2005.

[115] A. Ross, K. Nandakumar, and A. K. Jain. *Handbook of Multibiometrics*. Springer Verlag, 2006.

[116] A. Rutherford. *ANOVA and ANCOVA: A GLM Approach (Second Edition)*. Wiley, 2012.

[117] Science Working Group on Friction Ridge Analysis, Study and Technology (SWG-FAST). Standards for Examining Friction Ridge Impressions and Resulting Conclusions (Latent/Tenprint), 2013.

[118] D. R. Setlak. Fingerprint Sensor Having Spoof Reduction Features and Related Methods, 1999.

[119] B. G. Sherlock and D. M. Monro. A Model for Interpreting Fingerprint Topology. *Pattern Recognition*, 26(7):1047–1055, 1993.

[120] C-F. Shu and R. C. Jain. Vector Field Analysis for Oriented Patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(9):946–950, 1994.

[121] J. D. Singer and J. B. Willett. *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence*. Oxford University Press, 2003.

[122] K. Singh. Altered Fingerprints, 2008. `https://secure.interpol.int/Public/Forensic/fingerprints/research/alteredfingerprints.pdf`.

[123] R. Singh, M. Vatsa, H. S. Bhatt, S. Bharadwaj, A. Noore, and S. S. Nooreyezdan. Plastic Surgery: A New Dimension to Face Recognition. *IEEE Transactions on Information Forensics and Security*, 5(3):441–448, 2010.

[124] E. Tabassi, C. Wilson, and C. Watson. Fingerprint Image Quality. NISTIR 7151, 2004.

[125] B. T. Ulery, R. A. Hicklin, J. Buscaglia, and M. A. Roberts. Accuracy and Reliability of Forensic Latent Fingerprint Decisions. *Proceedings of the National Academy of Sciences*, 108(19):7733–7738, 2011.

[126] B. T. Ulery, R. A. Hicklin, J. Buscaglia, and M. A. Roberts. Repeatability and Reproducibility of Decisions by Latent Fingerprint Examiners. *PLoS ONE*, 7(3), 2012.

[127] U. Uludag, A. Ross, and A. K. Jain. Biometric Template Selection and Update: A Case Study in Fingerprints. *Pattern Recognition*, 37(7):1533–1542, 2004.

[128] P. R. Vizcaya and L. A. Gerhardt. A Nonlinear Orientation Model for Global Description of Fingerprints. *Pattern Recognition*, 29(7):1221–1231, 1996.

[129] Y. Wang and J. Hu. Global Ridge Orientation Modeling for Partial Fingerprint Identification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(1):72–87, 2010.

[130] Y. Wang, J. K. Hu, and D. Phillips. 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):573–585, 2007.

[131] M. V. Water. Can Fingerprints Be Forged? *The Science News Letter*, 29(774):90–92, 1936.

[132] C. Watson, M. Garris, E. Tabassi, C. Wilson, R. M. McCabe, S. Janet, and K. Ko. NIST Biometric Image Software. `http://www.nist.gov/itl/iad/ig/nbis.cfm`.

[133] L. M. Wein and M. Baveja. Using Fingerprint Image Quality to Improve the Identification Performance of the U.S. Visitor and Immigrant Status Indicator Technology Program. *Proceedings of the National Academy of Sciences of the U.S.A.*, 102(21):7772–7775, 2005.

[134] K. Wertheim. An Extreme Case of Fingerprint Mutilation. *Journal of Forensic Identification*, 48(4):466–477, 1998.

[135] C. Wilson, R. A. Hicklin, M. Bone, H. Korves, P. Grother, B. Ulery, R. Micheals, M. Zoepfl, S. Otto, and C. Watson. Fingerprint Vendor Technology Evaluation 2003: Summary of Results and Analysis Report. NISTIR 7123, 2004.

[136] M. Wong, S. P. Choo, and E. H. Tan. Travel Warning with Capecitabine. *Annals of Oncology*, 2009.

[137] S. Yoon, J. Feng, and A. K. Jain. On Latent Fingerprint Enhancement. In *Proceedings of SPIE, Biometric Technology for Human Identification VII*, volume 7667, pages 766707_01–766707_10, 2010.

[138] S. Yoon, J. Feng, and A. K. Jain. Latent Fingerprint Enhancement via Robust Orientation Field Estimation. In *Proceedings of International Joint Conference on Biometrics*, pages 1–8, 2011.

[139] S. Yoon, J. Feng, and A. K. Jain. Altered Fingerprints: Analysis and Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(3):451–464, 2012.

[140] S. Yoon and A. K. Jain. Is There a Fingerprint Pattern in the Image? In *Proceedings of International Conference on Biometrics*, 2013.

[141] S. Yoon, Q. Zhao, and A. K. Jain. On Matching Altered Fingerprints. In *Proceedings of International Conference on Biometrics*, pages 222–229, 2012.

[142] Y. Zhu, S. C. Dass, and A. K. Jain. Statistical Models for Assessing the Individuality of Fingerprints. *IEEE Transactions on Information Forensics and Security*, 2(3):391–401, 2007.