

David Zhang · Guangming Lu
Lei Zhang

Advanced Biometrics

Advanced Biometrics

David Zhang • Guangming Lu • Lei Zhang

Advanced Biometrics



Springer

David Zhang
The Hong Kong Polytechnic University
Hong Kong, China

Guangming Lu
Shenzhen Graduate School
Harbin Institute of Technology
Shenzhen, China

Lei Zhang
The Hong Kong Polytechnic University
Hong Kong, China

ISBN 978-3-319-61544-8 ISBN 978-3-319-61545-5 (eBook)
DOI 10.1007/978-3-319-61545-5

Library of Congress Control Number: 2017946764

© Springer International Publishing AG 2018

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Biometric technologies, such as fingerprint, face, iris, and palmprint, have been well studied and introduced in many research books. However, these technologies have their own advantages and disadvantages, and there is not one kind of biometric technology that can be fit for different applications. Many new biometric technologies have been developed in recent years, especially for some new applications.

This book describes some new biometric technologies, such as High-Resolution fingerprint, Finger-Knuckle-Print, Hand back skin texture, 3D fingerprint, Tongueprint, and 3D ear. There are 15 chapters, and except the overview chapter (Chap. 1) and recapitulation chapter (Chap. 15), the other 13 chapters are divided into four parts, including High-Resolution Fingerprint Recognition, Finger-Knuckle-Print Verification, Other Hand-Based Biometrics, and Some New Head-Based Biometrics. Many efficient feature extraction, matching, and fusion algorithms are introduced, and some potential systems have been developed in this book.

A comprehensive introduction to both theoretical issues and practical implementation in biometric authentication is given in this book. It will serve as a textbook or as a useful reference for graduate students and researchers in the fields of computer science, electrical engineering, systems science, and information technology. Researchers and practitioners in industry and R&D laboratories working on security system design, biometrics, immigration, law enforcement, control, and pattern recognition will also find much of interest in this book.

The work is supported by the NSFC funds under project No. 61271344 and 61332011, Shenzhen Fundamental Research fund JCYJ20150403161923528, JCYJ20140508160910917, and Key Laboratory of Network Oriented Intelligent Computation, Shenzhen, China.

Hong Kong, China
Harbin, China
Hong Kong, China
Feb 2017

David Zhang
Guangming Lu
Lei Zhang

Contents

1	Overview	1
1.1	Why Biometrics	1
1.2	History of Biometrics	2
1.3	Biometrics Systems	3
1.4	Recent Biometrics Development	7
1.4.1	Multi-biometrics	7
1.4.2	3D Biometrics	8
1.4.3	Multispectral Biometrics	9
1.5	Arrangement of This Book	10
	References	10

Part I High Resolution Fingerprint Recognition

2	High Resolution Partial Fingerprint Alignment	15
2.1	Introduction	15
2.1.1	High Resolution Partial Fingerprint	17
2.1.2	Fingerprint Alignment	18
2.1.3	Partial Fingerprint Alignment Based on Pores	19
2.2	Feature Extraction	20
2.2.1	Ridge and Valley Extraction	20
2.2.2	Pore Extraction	21
2.2.3	Pore-Valley Descriptors	22
2.3	PVD-Based Partial Fingerprint Alignment	24
2.4	Experiments	27
2.4.1	The Dataset	28
2.4.2	The Neighborhood Size and Sampling Rate	29
2.4.3	Corresponding Feature Point Detection	30
2.4.4	Alignment Transformation Estimation	32

2.4.5	Partial Fingerprint Recognition	34
2.4.6	Computational Complexity Analysis	37
2.5	Summary	38
	References	39
3	Fingerprint Pore Modeling and Extraction	41
3.1	Introduction	41
3.2	Review of Existing Pore Extraction Methods	43
3.3	Dynamic Anisotropic Pore Model (DAPM)	44
3.4	Adaptive Pore Extraction	46
3.4.1	DAPM Parameter Estimation	46
3.4.2	Implementation Issues	47
3.4.3	The Pore Extraction Algorithm	49
3.5	Experiments and Performance Evaluation	52
3.5.1	Pore Detection Accuracy	53
3.5.2	Pore Based Partial-Fingerprint Recognition	55
3.5.3	Pore Based Full-Size Fingerprint Recognition	58
3.5.4	Computational Complexity Analysis	59
3.6	Summary	61
	References	62
4	The Resolution for Fingerprint Recognition	65
4.1	Introduction	65
4.2	Multiresolution Fingerprint Images Collecting	67
4.2.1	Acquisition Device	67
4.2.2	Fingerprint Samples	68
4.2.3	Implementation of Multiresolution	68
4.3	Selecting Resolution Criteria Using Minutiae and Pores	71
4.4	Experiments and Analysis	75
4.4.1	Minimum Resolution Required for Pore Extraction	75
4.4.2	Resolution Based on the Established Criteria	76
4.4.3	Analysis of Ridge Width	77
4.4.4	Fingerprint Recognition Accuracy	78
4.5	Summary	80
	References	80

Part II Finger-Knuckle-Print Verification

5	Finger-Knuckle-Print Verification	85
5.1	Introduction	85
5.2	The Finger-Knuckle-Print (FKP) Recognition System	87
5.3	ROI (Region of Interest) Extraction	90

5.4	FKP Feature Extraction and Matching	96
5.4.1	Improved Competitive Coding (ImCompCode)	96
5.4.2	Magnitude Coding (MagCode)	98
5.4.3	FKP Matching	98
5.5	Experimental Results	99
5.5.1	Database Establishment	99
5.5.2	Selection of the Image Resolution	100
5.5.3	FKP Verification	101
5.5.4	Speed	104
5.5.5	Discussion	105
5.6	Summary	107
	References	107
6	Local Features for Finger-Knuckle-Print Recognition	111
6.1	Introduction	111
6.2	Analysis of Local Features	114
6.3	Extraction and Matching of Local Features	117
6.3.1	Step 1: Phase Congruency (PC)	117
6.3.2	Local Feature Extraction and Coding	118
6.3.3	Matching of Local Feature Maps	120
6.3.4	Integration of Local Features	121
6.4	Experimental Results and Discussions	121
6.4.1	FKP Database and the Test Protocol	121
6.4.2	Determination of Parameters	122
6.4.3	Performance of Local Features	123
6.4.4	Further Discussions	124
6.5	Summary	128
	References	128
7	Global Information for Finger-Knuckle-Print Recognition	131
7.1	Introduction	131
7.2	Local Feature Extraction and Matching	134
7.3	Global Feature Extraction and Matching	137
7.3.1	From Local to Global	137
7.3.2	Phase-Only Correlation (POC)	138
7.3.3	Band-Limited Phase-Only Correlation (BLPOC)	139
7.4	Local-Global Information Combination	141
7.5	Experimental Results	143
7.5.1	FKP Database and Test Protocol	143
7.5.2	Determination of Parameters	144
7.5.3	FKP Verification Results	144
7.5.4	Speed	145
7.6	Summary	147
	References	147

8 Finger-Knuckle-Print Verification with Score Level Adaptive Binary Fusion	151
8.1 Introduction	151
8.2 Competitive Coding Based FKP Verification	153
8.3 Recognition via Reconstruction	154
8.3.1 Motivation	154
8.3.2 Reconstruction with l_1 -Norm Sparse Regularization	156
8.3.3 Reconstruction with l_2 -Norm Regularization	158
8.3.4 Patch Based Reconstruction	159
8.4 Verification by Binary Score Level Fusion	161
8.4.1 Popular Score Level Fusion Methods	162
8.4.2 Adaptive Binary Fusion	163
8.5 Summary of the Verification Algorithm	165
8.6 Experimental Results	165
8.6.1 Comparison Between Different Fusion Rules	166
8.6.2 Experiment Settings and Parameter Selection	166
8.6.3 FKP Verification Results	168
8.6.4 Integrating with Global Features	170
8.6.5 Discussions	171
8.7 Summary	172
References	173

Part III Other Hand-Based Biometrics

9 3D Fingerprint Reconstruction and Recognition	177
9.1 Introduction	177
9.2 3D Fingerprint Reconstruction System	180
9.3 Fingerprint Feature Correspondences Establishment	182
9.3.1 Correspondences Establishment Based on SIFT	182
9.3.2 Correspondences Establishment Based on Ridge Map	182
9.3.3 Correspondences Establishment Based on Minutiae	188
9.4 Finger Shape Model Estimation	189
9.5 Curvature Features Extraction and Matching	192
9.6 Experimental Results and Analysis	197
9.6.1 3D Fingerprint Reconstruction System Error Analysis	197
9.6.2 Comparison and Analysis of Reconstruction Results	197
9.6.3 Validation of Estimated Finger Shape Model	199
9.6.4 Reconstruction System Computation Time Analysis	203
9.6.5 3D Fingerprint Recognition	204
9.7 Summary	207
References	210

10 Hand Back Skin Texture for Personal Identification	213
10.1 Introduction	213
10.2 Hand Back Skin Texture Imaging System	216
10.3 HBST Feature Extraction and Classification	218
10.3.1 MR8 Filter Bank	220
10.3.2 Texton Learning Based on SR	221
10.3.3 Feature Extraction and Classification	223
10.4 Experimental Results	224
10.4.1 Database Establishment	224
10.4.2 Personal Identification	226
10.4.3 Gender Classification	229
10.4.4 Discussion	230
10.5 Summary	231
References	231
11 Line Scan Palmprint Recognition System	235
11.1 Introduction	235
11.2 Existing System	237
11.2.1 Flatbed Scanner	237
11.2.2 Web Camera Based Systems	237
11.2.3 Palmprint Systems with Pegged Flat Platen Surface	238
11.3 Line Scan Palmprint System Design	239
11.3.1 Line Scan Imaging Scheme	239
11.3.2 System Framework	242
11.3.3 Line Scan Image Sensor-CIS Module	243
11.3.4 Synchronizing Unit	246
11.3.5 Controller Board	248
11.3.6 The Device	249
11.3.7 ROI Extraction	249
11.4 Experiment and Comparison	251
11.4.1 Line Scan Palmprint Database	251
11.4.2 Verification Experiment	252
11.4.3 Comparisons with Current Palmprint Systems	253
11.5 Summary	255
References	255
12 Door Knob Hand Recognition System	259
12.1 Introduction	259
12.2 Ergonomic Biometrics Design	260
12.2.1 Development of Biometric Systems	260
12.2.2 Ergonomics Studies in Biometrics	261
12.2.3 Ergonomic Biometrics Design Model	263

12.3	Door Knob Hand Recognition System	264
12.3.1	Concept and Framework	264
12.3.2	Imaging	265
12.3.3	Feature Extraction and Classification	269
12.4	Experiments	274
12.4.1	Data Set	274
12.4.2	Experimental Results	275
12.4.3	Comparison with Conventional Biometrics	277
12.4.4	Discussion	278
12.5	Summary	280
	References	280

Part IV Some New Head-Based Biometrics

13	Dynamic Tongueprint Recognition	287
13.1	Introduction	287
13.2	Background	290
13.3	Tongue Squirm and Its Applications	292
13.3.1	Liveness Detection	292
13.3.2	Squirm Features Extraction	295
13.4	Extraction of Static Physiological Features	296
13.4.1	Geometric Features	296
13.4.2	Crack Features	298
13.4.3	Texture Features	300
13.5	Experimental Results	302
13.5.1	Database	302
13.5.2	Identification Experiments	303
13.5.3	Verification Experiments	304
13.5.4	Liveness Detection Experiments	305
13.6	Summary	306
	References	307
14	Online 3D Ear Recognition	309
14.1	Introduction	309
14.2	Scanner Design for Online 3D Ear Acquisition	311
14.3	3D Ear Global and Local Feature Classes	313
14.3.1	Global Feature Class	313
14.3.2	Local Feature Class	316
14.4	Experimental Results and Discussion	321
14.4.1	Feature Optimization	321
14.4.2	Matching Using Local Features	322
14.4.3	Recognition with Global Feature Indexing	324
14.4.4	Performance Analysis	325
14.5	Summary	326
	References	327

Contents	xiii
15 Book Review and Future Work	329
15.1 Book Recapitulation	330
15.2 Future Work	332

Chapter 1

Overview

Abstract This chapter gives an overview of biometrics technology, its development history, typical biometrics systems, and current research hotspots. Then, the chapter arrangement is introduced.

Keywords Biometrics • 3D biometrics • Multispectral biometrics

1.1 Why Biometrics

With the increasing attention on security vulnerabilities and transactional fraudulent in industries and societies, highly reliable and accessible personal authentication and identification techniques become an inevitable demand for human societies. Biometrics has emerged to meet this need and even has developed to the science combining biology technology and information technology to utilize physiological or behavioral characteristics in human body to deal with identification of individuals. In particular, biometrics technologies focus on the technologies to automatically authenticate the still traits of the human such as the DNA, ears, palmprint, hand and finger geometries, fingerprint, faces, irises, footprint, retina and tooth or the dynamic traits of the human such as voice, gait, keystroke and signature. What's more, it seems that biometrics will and is being a dominant component of the world and remarkably increasing number of biometrics systems have been developed to satisfy the research and commercial need. Biometrics systems have been widely applied to a variety of government and private domains as a technology with respect to security and convenience. Furthermore, biometrics has shown its overwhelming superiority to replace or enhance traditional identification methods, such as token-based approaches and knowledge-based approaches. This book presents a literature review of the biometrics area. And subsequently some new biometrics technologies and systems are discussed and highlighted.

1.2 History of Biometrics

The term biometric comes from the Greek words bios and metrikos. Biometrics deals with identification of individuals based on their biological or behavioral characteristics (Jain and Ross 2002). Biometrics combines biology technology and information technology to exploit physical features or behavioral features in human body to identify a person's identity so as to replace or enhance traditional personal identification methods.

With the increasing concerns on security breaches and transaction fraud, highly reliable and convenient personal verification and identification technologies are more and more requisite in our social activities and national services. Biometrics is implemented to two main aspects applications, identity verification and identity recognition. Identity verification is to require that the system has a binary option, acceptance or rejection, in response to the person's claim when the person claims to an identity. However, identity recognition is to require that the system to retrieve the pre-existing database of characteristics and identify the one that matched the characteristics of the unknown individual being presented. Historically, the development of biometric technologies is originated from different history background.

Personal identification is to associate the identity to a particular individual. Identification can be viewed as the form of recognition or verification which is known as authentication (Jain et al. 2000). Knowledge-based and token-based personal identification techniques have been treated as the two traditional widely used techniques (Miller 1994). The knowledge-based approaches authenticate the identity of an individual according to what he/she knows. Any individuals with certain secret knowledge, such as personal identification numbers or a password for telephone calling, membership or credit cards, and then answers to questions, would receive the associated service. In the token-based approach, the identity of a person is verified according to what he/she has. Anyone possessed a certain physical object (token), e.g., keys or ID cards, is authorized to receive the associated service.

Both the token-based and the knowledge-based approaches, however, have some inherent limitations, because they are not based on any intrinsic biological nature from each individual for personal identification. These drawbacks usually are fatal important for personal identification. For example, tokens may be stolen, lost, forgotten, effaced or misplaced, and even tokens are easy to fool. Knowledge-based techniques also exists flaws. For example, the personal identification number (PIN) may be forgotten by a valid user or guessed by a fraud.

Because the characteristics of knowledge-based and token-based personal identification techniques are unable to be unique, distinctive and distinguishable, they cannot meet the security requirements of electronic interconnection of the information society on account of the fraudulent vulnerability. In order to make breakthrough on verification systems, powerful identification systems have never been more in demand. Biometric recognition is an emerging personal recognition technology developed to overcome the inherent limitations of the traditional personal

recognition approaches (Jain et al. 1999; Zhang 2000, 2002). Biometrics which exploits the instinct of distinctive physiological characteristics has emerged to enhance the identification techniques by automatically verifying or recognizing the identity of a living person. Compared with the token-based and the knowledge-based methods, biometric identifiers cannot be easily forged, shared, forgotten, or lost, and thus can provide better security, higher efficiency, and increased user convenience. The appearance of biometrics makes authentication easier, faster and more accurate. Moreover, biometrical identifiers are more competent and reliable compare to the traditional identification techniques and also have gained great reputation.

1.3 Biometrics Systems

Generally, a biometric system is a computer system implemented by exploiting corresponding biometric identification methods, techniques, and technologies. Biometric systems can be regarded as pattern recognition systems, where a feature set is first extracted from the acquired data, and then compared with the stored template set to make a decision on the identity of an individual. A biometric system can be applied to two fields, verification and identification. In verification mode, the decision is whether a person is “who he claims to be?” In identification mode, the decision is “whose biometric data is this?” A biometric system is thus formalized into a two-class or multi-class pattern recognition system.

A biometric system usually includes four major modules: data acquisition, feature extraction, matching, and system database (Jian et al. 2007). In the data acquisition module, the biometric data of an individual is acquired using a capture sensor such as fingerprint sensor and digital camera for face. In the feature extraction module, the acquired data is processed to extract a set of discriminative features. In the matching module, the features are compared with the stored template set to make a decision on the identity of an individual. In the system database module, a database is built and maintained to store the biometric templates of the enrolled users. Feature extraction and matching are two of the most challenging problems in biometric recognition research, and have attracted researchers from different backgrounds: biometrics, computer vision, pattern recognition, signal processing, and neural networks.

Fingerprint Identification System

Fingerprints identification system has been used for many centuries on basis of the essence in its uniqueness and distinguishability. The application of finger identification provides solid and infallible methods of personal identification because of its high accuracy (Maio et al. 2002). Fingerprint identification mainly depends on the pattern of ridges and valleys on the surface of fingertip, whereas these patterns are determined during the first 7 months of each person. For decades, law enforcement has been classifying and determining identity by matching key points of ridge

endings and bifurcations. The uniqueness of the fingerprints is that no two fingerprints have been found alike in the many billions of human, even for the identical twins or the prints on each finger of the same person (Maio et al. 2002).

Facial Identification System

Face recognition started from computer vision and research on face recognition goes back to the earliest days of artificial intelligence and computer vision. Images of a human face are highly suitable for use as a biometric trait for personal authentication because they can be acquired non-intrusively, hands-free, continuously, and usually in a way that is acceptable to most users (Zhao et al. 2003). A general definition of face recognition problem is to verify or recognize one or more persons from a stored image and face database by comparing facial distinguishing characteristics. The authentication of a person by their facial image can be done in a number of different ways, such as by capturing an image of the face in the visible spectrum using an inexpensive camera or by using the infrared patterns of facial heat emission. Face recognition or verification mainly depends on characteristics by the overall facial structure, distance between mouth, nose, eyes and jaw edges, and then to find a match by comparing these features from a database of face images.

Hand and Finger Geometry Identification System

Hand and finger geometry-base verification and recognition is an important branch of biometrics to automatically verify and recognize individuals based on the distinguishing hand geometric characteristics. Hand and finger geometry recognition technique exploits a number of characteristics from the human hand such as finger length, finger width, finger area, finger thickness, palm width and curvature of the fingers at certain points to make personal verification and recognition (Zhang and Kanhangad 2011). Hand and finger geometry recognition technique possesses dramatical merits. For example, hand-based system just utilizes simple imaging requirement by virtue of extracting features from low-resolution hand images. What's more, hand-based system is competent with the capability of operating and unaffected under turbulent and harsh environmental conditions such as dry weather, or individual anomalies such as dirt on the hand, dry skin. These external factors do not appear to have any negative effects on the verification and recognition accuracy of hand and finger geometry technique. Furthermore, the low data-storage requirement is additional superiority. The biometric system based on hand and finger geometry has been used in physical access control in commercial and residential applications, in time and attendance systems. Additionally, hand and finger geometry technique can be effectively implemented and inexpensive. Moreover, acquisition and authentication system are efficiently obtained.

Palmprint Identification System

Palm is the inside part of human hand from the wrist to the end of our fingers and palmprint is the skin patterns of a palm, composed of the physical characteristics of the skin patterns of a palm. The palmprint recognition implements the matching characteristics by the pattern of ridges and valleys much like the fingerprints. The palms of the human possess much larger area compare to finger, and as a result,

palmprints are expected to be much more distinctive and palmprint identification are anticipated to be more robust and effective. Furthermore, additional distinguishing characteristics of palmprints also can be provided as supplements for identification, such as wrinkle, texture and principles lines. Point features also can be used to palmprint identification systems, such as minutiae points, delta points and datum points. In the early stage, palm print recognition techniques have been investigated to extract and match the singular points and minutia points from high resolution palm print images. High resolution palm print scanner, however, is expensive, and is time consuming to capture a palm print image, which restricts the potential applications of online palm print recognition systems. Subsequently, online capture device has been developed to collect real time low resolution palm print image, and low resolution palm print recognition has gradually received considerable recent interest in biometric community (Zhang et al. 2003; Zhang and Lu 2013; Xu et al. 2015). A great many different low resolution palmprint recognition algorithms have been developed, which can be roughly grouped into three categories: holistic-based, feature-based, and hybrid methods. To this end, palmprint identification system usually combines all the features of the palm such as hand geometry, ridge and valley features, principle lines, texture, wrinkles and point features to build a highly accurate system.

Iris Identification System

The patterns of the iris, the colored area that surrounds the pupil, are thought to be unique. Iris recognition is to recognize a person by utilizing and analyzing the random pattern of the iris. Iris features can be more easily obtained than other features from eyes, like retina. It is shown that an iris has more details than a fingerprint or palmprint. Feature set extracted using iris detailed and unique texture will remain stable and immutable over decades of life. Iris recognition system exploits textures with striations, contraction furrows, pits, collagenous fibers, filament, crypts (darkened areas and resembling excavations), serpentine vasculature, rings and freckles for personal identification. Iris patterns can be obtained through a video-based image acquisition system. Iris scanning devices have been used in personal authentication applications. Furthermore, these visual textures of the iris are established during the first 2 years of life which is the fetal development period for human. More complex textures of iris can contribute more useful distinctive identification features for personal authentication. It has been demonstrated that iris-based biometric system can work with individuals without regard to ethnicity or nationality. Moreover, iris identification technique can recognize the identity of claimed person with dramatic efficiency and this technology can be embraced in large-scale identification system.

Ear Identification System

Earprints identification is to utilize the salient characteristics of the ear for personal identification by exploiting the shape of the ear and the structure or contours of the congenital human tissue of the auricle. The ear identification technique matches the distance of distinguishing points on the auricle of the ear. There is evidence to show that the shape of the ear and the structure of the cartilaginous tissue of the pinna are

distinctive. As a result, the ear-based biometric system can be used for authenticating personal identity. Some scientists have demonstrated that the human ear is differentiated enough to each individual and can be practically used as a biometric by its appearance. Moreover, some researcher has claimed that the ear is fully formed when you are born and it will stay similar except for a litter descendent of your lobe (Iannarelli 1989). Furthermore, some scientists have been trying to use the ear's appearance in 2D intensity and 3D ear shape for personal identification (Liu et al. 2015a, b; Abaza et al. 2013).

Gait Identification System

From ancient time, human can recognition your familiars by one's walk. Gait identification is an outstanding biological or behavioral identification technique to analyze the walking ability of humans and animals. The aim of gain recognition is to detect, classify and identify humans by the fastest speed. Gait identification usually resort to some distinctive characteristics such as, stop length, stride length, speed, dynamic base, progression line, foot angle and hip angle. Moreover, some factors also can influence the gain identification, for instance, external factors (such as footwear and clothing), physical or intrinsic factors (such as weight and height, male or female and age), pathological factors, emotional factors and etc. Gait can be supposed to be dramatic distinctive and the stride of each person still be unique enough to be used for identification even among a group of people. Thus, gait recognition technique is to use the peculiar way of one's walk for personal identification.

Voice Identification System

Voice recognition is also one of the oldest approaches for identification by humans. Voice recognition is also commonly referred to as speaker recognition. The voice identification is the process for recognizing a person by analyzing both the behavioral features of someone and the physical structure of someone's vocal tract. That is, voice recognition is developed by analyzing the shape and size of the appendages such as vocal tracts, mouth, nasal cavities and lips. Voice authentication uses the acoustic features of voice, which have been found to be different between individuals. These acoustic patterns reflect both anatomic (e.g., size and shape of the throat and mouth) and behavioral patterns (e.g., voice pitch, speaking style) of an individual. The incorporation of learned patterns into the voice templates (the latter called "voiceprints") has allowed speaker identification to be recognized as a "behavioral biometric". Voice-based personal authentication systems employ three styles of spoken input: text-dependent, text-prompted and text-independent. Most voice authentication applications use text-dependent input, which involves selection and enrollment of one or more voice passwords. Text-prompted input is used whenever there is concern about imposters.

Signature Identification System

Signature identification is categorized to the behavioral biometric technique. General signature identification concludes two modes, that is, static signature and dynamic signature identification techniques. Static signature identification

technique mainly refers to make identification on still signature, i.e. recognizing by measuring and analyzing the shape of the pre-captured signature by photo scanner or optical scanner. On the other hand, the mode of dynamic signature identification is to recognition the signature sign on live using the characteristics of signature such as the stroke order, azimuth, the pressure applied, inclination, pen up/down, the speed of writing, horizontal and vertical coordinates of inflection point. However, the core point required by the signature biometric technique is not the image of the signature but behavioral patterns, i.e. how to sign. This technology has potential applications in e-business, where signatures can be an accepted method of personal authentication.

Keystroke Identification System

Keystroke identification technology is an important member of the biometrics and is categorized to physiological and behavioral biometrics by different measurements. Keystroke identification is considered to be the best way to authenticate a person on account of minimizing the impact on privacy. Keystroke identification technique is the process of measuring and analyzing human's typing rhythm on digital devices such as, computer keyboard, mobile phone or touch screen panel. The way and the manner in which we type on our computer keyboard varies from individual to individual and is considered to be a unique behavioral biometric. Keystroke identification is probably one of the easiest biometrics forms to implement and manage, because all we needed for this identification is the existing computer and keyboard that is already in place and use.

1.4 Recent Biometrics Development

1.4.1 *Multi-biometrics*

Verification or identification accuracy is always the first-of-all objective for biometric systems. Unibiometric system, the biometric system using a single biometric characteristic, usually suffers from some limitations and cannot provide satisfactory recognition performance. For example, manual workers with damaged or dirty hands may not be able to provide high-quality fingerprint images, and thus failure to enrol would happen for single fingerprint recognition system.

Multi-biometric systems, which integrate information from multiple biometric traits, provide some effective means to enhance the performance and reliability of the biometric system. To combine information from individual biometric traits, there are three categories of fusion strategies, feature level fusion, matching score level fusion, and decision level fusion. In feature level fusion, the data obtained from each sensor is used to compute a feature vector. As the feature extracted from one biometric trait is independent of that extracted from the other, it is reasonable to concatenate the two vectors into a single new vector for performing multi-biometric based personal authentication. Note that the new feature vector now has a higher

dimensionality than the original feature vector generated from each sensor. Feature reduction techniques may be employed to extract useful features from the set of the new feature vector. In matching score level fusion, each subsystem using one biometric trait of the multi-biometric system provides a matching score indicating the proximity of the feature vector with the template vector. These scores can be combined to assert the veracity of the claimed identity. In decision level fusion each sensor first acquire one of multiple biometric traits and the resulting feature vectors are individually classified into the two decisions—accept or reject the claimed identity. Then a scheme that exploits the known decisions to make the final decision is used. In the field of multi-biometrics, a great number of studies of feature level fusion, matching score level fusion and decision level fusion have been made. Though fusion of multi-biometrics are generally recognized as three classes as described above, in real-world applications of multi-modal biometric it is possible that the “Fusion Process” may be simultaneously involved in different levels such as in both the matching score level and the decision level.

Multi-biometric system is designed to overcome the limitations of any single biometric systems by fusing information from multiple biometric traits. The fusion can be implemented in either of three levels, feature level, matching score level, and decision level. In feature level fusion, a new feature vector is constructed using the concatenation rule (Ross and Govindarajan 2005), the parallel rule (Yang et al. 2003; Yang and Yang 2002), or the competitive rule (Kong et al. 2006). In matching score level fusion, a number of transformation-based (Jain et al. 2005; Zuo et al. 2007), classifier-based (Fierrez-Aguilar et al. 2005), and density-based (Ulery et al. 2006; Nandakumar et al. 2006) score fusion methods have been used to combine scores of multiple scores. In decision level fusion, boolean conjunctions, weighted decision methods, classical inference, Bayesian inference, Dempster-Shafer method, and voting have been proposed to make the final recognition decision (Gokberk et al. 2003).

1.4.2 3D Biometrics

In the past decade, biometrics recognition has been growing rapidly, and many biometrics systems have been widely used in various applications. However, most of the biometrics recognition techniques are based on 1D signal or 2D images. There are many limitations of 1D and 2D biometrics technologies until now:

- Fingerprints may be distorted and unreadable or unidentifiable if the person’s fingertip has dirt on it, or if the finger is twisted during the process of finger-printing. In an ink fingerprint, twisting could cause the ink to blur, distorting the shape of the fingerprint and potentially making it unreadable.
- It is found that with age, the voice of a person differs. Also when the person has flu or throat infection the voice changes, or if there are too much noise in the environment this method may not authenticate correctly.

- For Iris recognition, if people affected with diabetes, the eyes get affected resulting in differences.
- The conventional 2D palmprint recognition is a fast and effective personal authentication method, but 2D palmprint images can be easily counterfeited.

Although 2D biometrics recognition techniques can achieve high accuracy, the 2D features can be easily counterfeited and much 3D feature structural information is lost. Therefore, it is of high interest to explore new biometrics recognition techniques: 3D Biometrics.

With the development of 3D techniques, it is possible to capture 3D characteristics in real time. Recently, 3D techniques have been used in biometrics authentication, such as 3D face, 3D fingerprint, 3D palmprint and 3D ear recognition, and shown many advantages, such as:

- 3D biometrics is much more robust to illumination and pose variations than 2D biometrics.
- 3D range data may offer a richer information source for feature extraction. And usually it also can fuse with 2D biometrics to enhance the system accuracy.
- 3D biometrics systems are more robust to attack, since 3D information is more difficult to be duplicated or counterfeited.

3D biometrics technologies have been the new trend in this research field. There are some commercial devices which can obtain the 3D information of an object, such as Konica Minolta Vivid 9i/910, Cyberware whole body color 3D scanner, and so on. These commercial 3D scanners have high speed and accuracy, and can be used for 3D biometrics information collection.

1.4.3 Multispectral Biometrics

Versatility, usability, and security are some of the required characteristics of biometric system. Such system must have the capability to acquire and process biometric data at different times of day and night, in a variety of weather and environmental conditions, and be resistant to spoofing attempts. Multispectral biometrics is one of the few technologies shown to solve many of the aforementioned issues (Li and Jain 2009).

Multispectral imaging is possible to simultaneously capture images of an object in the visible spectrum and beyond. It has been extensively used in the field of remote sensing, medical imaging, and compute vision to analyze information in multiple bands of the electromagnetic spectrum. Multispectral imaging could provide more information than single modality. Usually, complementary features could be extracted. Thus, better recognition and spoof detection ability are easy to be achieved.

1.5 Arrangement of This Book

In this book, we would like to introduce some new biometric technologies. There are 15 chapters, and except this chapter and recapitulation chapter (Chap. 15), the other 13 chapters are in four parts, including High Resolution Fingerprint Recognition, Finger-Knuckle-Print Verification, Other Hand-Based Biometrics, and Some New Head-Based Biometrics. The biometrics devices, feature extraction and matching algorithms of each biometrics system are presented.

Part I

This part focuses on high resolution fingerprint recognition technology. Chapter 2 introduces the high resolution partial fingerprint alignment using pore-valley descriptors. Adaptive fingerprint pore modeling and extraction method is introduced in Chap. 3, and Chap. 4 discusses the reference high resolution of fingerprint image by using minutiae and pores.

Part II

This part discusses Finger-Knuckle-Print verification technologies. Chapter 5 introduces an online Finger-Knuckle-Print verification system. Chapter 6 studies phase congruency induced local features for FKP verification. Chapter 7 presents ensemble of local and global information for Finger-Knuckle-Print verification. In Chap. 8, FKP verification method with score level adaptive binary fusion is introduced.

Part III

This part has four chapters, focusing on some new hand-based biometric technologies. Chapter 9 introduces 3D fingerprint reconstruction and recognition method. Chapter 10 reports a new method for personal identification based on hand back skin texture patterns. Chapter 11 presents a novel line scan palmprint recognition system. In Chap. 12, another hand-based biometric system, door knob hand recognition system, is proposed.

Part IV

This part contains two chapters. Chapter 13 introduces a new biometric identifier: dynamic tongueprint. Chapter 14 gives a novel line scan palmprint recognition system.

References

- Abaza A, Ross A, Hebert C, MAF H, Nixon MS (2013) A survey on ear biometrics. ACM Comput Surv 45(2):94–111
- Fierrez-Aguilar J, Ortega-Garcia J, Gonzalez-Rodriguez J, Bigun J (2005) Discriminative multi-modal biometric authentication based on quality measure. Pattern Recogn 38(5):777–779
- Gokberk B, Salah A, Akarun L (2003) Rank-based decision fusion for 3D shape-based face recognition. Lect Notes Comput Sci 3546:1019–1028

- Iannarelli A (1989) Ear identification. Paramont, Fremont
- Jain A, Ross A (2002) Introduction to biometrics. In: Biometrics. Springer, New York, pp 1–41
- Jain A, Bolle A, Pankanti S (1999) Biometrics: personal identification in networked society. Kluwer Academic, Norwell
- Jain A, Hong L, Pankanti S (2000) Biometric identification. Commun ACM 43(2):90–98
- Jain A, Nandakumar K, Ross A (2005) Score normalization in multimodel biometric systems. Pattern Recogn 38(5):2270–2285
- Jain A, Flynn P, Ross A (2007) Handbook of biometrics. Springer, New York
- Kong A, Zhang D, Kamel M (2006) Palmprint identification using feature-level fusion. Pattern Recogn 39:478–487
- Li S, Jain A (2009) Encyclopedia of biometrics. Springer, New York
- Liu Y, Lu G, Zhang D (2015a) An effective 3D ear acquisition system. PLoS One 10(6):e0129439
- Liu Y, Zhang B, Zhang D (2015b) Ear-parotic face angle: a unique feature for 3D ear recognition. Pattern Recogn Lett 53:9–15
- Maio D, Maltoni D, Cappelli R, Wayman JL, Jain A (2002) FVC2002: fingerprint verification competition. In: Proceedings of international conference on pattern recognition (ICPR), Quebec City, QC, Canada, pp 744–747
- Miller B (1994) Vital signs of identity [biometrics]. IEEE Spectr 31(2):22–30
- Nandakumar K, Chen Y, Jain A, Dass S (2006) Quality-based score level fusion in multibiometric systems. In: Proceedings of international conference on pattern recognition, Hong Kong, pp 473–476
- Ross A, Govindarajan R (2005) Feature level fusion using hand and face biometrics. In: Proceedings of SPIE conference on biometric technology for human identification, pp 196–204
- Ulery B, Hicklin A, Watson C, Fellner W, Hallinan P (2006) Studies of biometric fusion. NIST technical report IR7346
- Xu Y, Fei L, Zhang D (2015) Combining left and right palmprint images for more accurate personal identification. IEEE Trans Image Process 24(2):549–559
- Yang J, Yang JY (2002) Generalized K-L transform based combined feature extraction. Pattern Recogn 35(1):295–297
- Yang J, Yang JY, Zhang D, JFL (2003) Feature fusion: parallel strategy vs. serial strategy. Pattern Recogn 36(6):1369–1381
- Zhang D (2000) Automated biometrics: technologies and systems. Kluwer Academic, Boston
- Zhang D (2002) Biometrics solutions for authentication in an e-world. Kluwer Academic, Dordrecht
- Zhang D, Kanhangad V (2011) Hand geometry recognition. In: Encyclopedia of cryptography and security. Springer, New York
- Zhang D, Lu G (2013) 3D biometrics. Springer, New York
- Zhang D, Kong WK, You J, Wong M (2003) On-line palmprint identification. IEEE Trans Pattern Anal Mach Intell 25(9):1041–1050
- Zhao W, Chellappa R, Rosenfeld A (2003) Face recognition: a literature survey. ACM Comput Sur 36(4):399–458
- Zuo W, Wang K, Zhang D, Zhang H (2007) Combination of two novel LDA-based methods for face recognition. Neurocomputing 70(4–6):735–742

Part I

High Resolution Fingerprint Recognition

Chapter 2

High Resolution Partial Fingerprint Alignment

Abstract This chapter discusses the alignment of high resolution partial fingerprints, which is a crucial step in partial fingerprint recognition. The previously developed fingerprint alignment methods, including minutia-based and non-minutia feature based ones, are unsuitable for partial fingerprints because small fingerprint fragments often do not have enough features required by these methods. In this chapter, we propose a new approach to aligning high resolution partial fingerprints based on pores, a type of fingerprint fine ridge features that are abundant on even small fingerprint areas. Pores are first extracted from the fingerprint images by using a difference of Gaussian filtering approach. After pore detection, a novel Pore-Valley Descriptor (PVD) is proposed to characterize pores based on their locations and orientations, as well as the ridge orientation fields and valley structures around them. A PVD-based coarse-to-fine pore matching algorithm is then developed to locate pore correspondences. Once the corresponding pores are determined, the alignment transformation between two partial fingerprints can be estimated. The proposed method is compared with representative minutia based and orientation field based methods using the established high resolution partial fingerprint dataset and two fingerprint matchers. The experimental results show that the PVD-based method can more accurately locate corresponding feature points, estimate the alignment transformations, and hence significantly improve the accuracy of high resolution partial fingerprint recognition.

Keywords Fingerprint alignment • Partial fingerprints • High resolution fingerprints • Pores

2.1 Introduction

Automatic fingerprint recognition systems (AFRS) have been nowadays widely used in personal identification applications such as access control (Ratha and Bolle 2004; Maltoni et al. 2003). Roughly speaking, there are three types of fingerprint matching methods: minutia-based, correlation-based, and image-based (Maltoni

et al. 2003; Nanni and Lumini 2008). In minutia-based approaches, minutiae (i.e. endings and bifurcations of fingerprint ridges) are extracted and matched to measure the similarity between fingerprints (Jain et al. 1997; Tico and Kuosmanen 2003; Jiang and Yau 2000; Kovacs-Vajna 2000; Feng 2008). These minutia-based methods are now the most widely used ones (Ratha and Bolle 2004; Maltoni et al. 2003). Different from the minutia-based approaches, both correlation-based and image-based methods compare fingerprints in a holistic way. The correlation-based methods spatially correlate two fingerprint images to compute the similarity between them (Bazen et al. 2000), while the image-based methods first generate a feature vector from each fingerprint image and then compute their similarity based on the feature vectors (Jain et al. 2000; Ross et al. 2003; Teoh et al. 2004; Nanni and Lumini 2007, 2008). No matter what kind of fingerprint matchers are used, the fingerprint images usually have to be aligned when matching them. Later in this section, we will discuss more about the fingerprint alignment methods.

In order to further improve the accuracy of AFRS, people are now exploring more features in addition to minutiae on fingerprints. The recently developed high resolution fingerprint scanners make it possible to reliably extract level-3 features such as pores. Pores have been used as useful supplementary features for a long time in forensic applications (Bazen et al. 2000; CDEFFS 2008). Researchers have also studied the benefit of including pores in AFRS and validated the feasibility of pore based AFRS (Kryszczuk et al. 2004a, b; Roddy and Stosz 1997; Stosz and Alyea 1994; Jain et al. 2007; Parthasaradhi et al. 2005). Using pores in AFRS has two advantages. First, pores are more difficult to be damaged or mimicked than minutiae (Parthasaradhi et al. 2005). Second, pores are abundant on fingerprints. Even a small fingerprint fragment could have a number of pores (refer to Fig. 2.1). Therefore, pores are particularly useful in high resolution partial fingerprint recognition where the number of minutiae is very limited. In this chapter, we focus on the alignment of high resolution partial fingerprints and investigate the methods for high resolution fingerprint image processing.

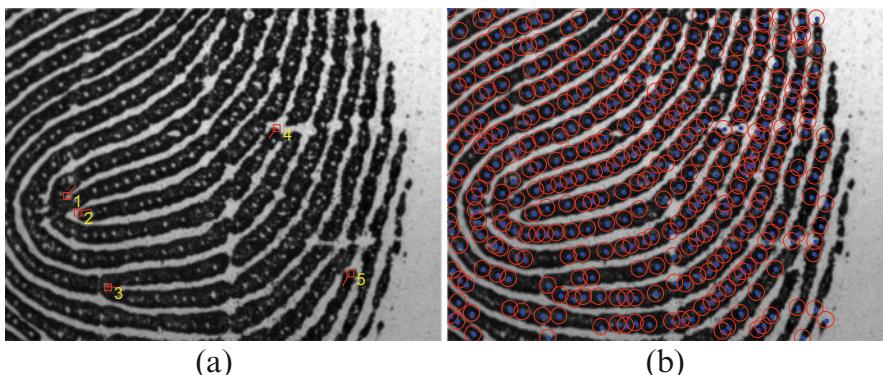


Fig. 2.1 An example high resolution partial fingerprint. It has only five minutiae as marked in (a), but hundreds of pores as marked in (b)

2.1.1 High Resolution Partial Fingerprint

In a live-scan AFRS, a user puts his/her finger against the prism and the contact fingerprint region will be captured in the resulting image. A small contact region between the finger and the prism will lead to a small partial fingerprint image. On such small fingerprint region, there could be very limited minutiae available for recognition. A natural way to solve the partial fingerprint recognition problem is to make full use of other fine fingerprint features abundant on the small fingerprint fragments. Sweat pores are such kind of features and high resolution fingerprint imaging makes it possible to reliably extract the sweat pores on fingerprints (CDEFSS 2008).

Most existing high resolution fingerprint recognition methods use full-size fingerprint images which capture large fingerprint areas. However, to capture the full fingerprints, high resolution fingerprint images should have much bigger sizes than conventional low resolution fingerprint images. As a result, much more computational resources are required to process the images. Considering the increasing demand of AFRS on mobile devices and other small portable devices, small fingerprint scanners and limited computational resources are very common (Jea and Govindaraju 2005). Consequently, the algorithms for aligning and matching partial fingerprint images are becoming important. Therefore, this chapter, different from previous study of high resolution fingerprint recognition, uses high resolution partial fingerprint images to study the partial fingerprint image alignment problem and a feasible algorithm will be proposed.

Although some methods have been proposed to construct full fingerprint templates from a number of partial fingerprint images (Choi et al. 2007), it is expensive or even impossible to collect sufficient fingerprint fragments to construct a reliable full fingerprint template. Moreover, some errors (e.g. spurious features) could be introduced in the construction process. Thus, it is meaningful and very useful if algorithms can be developed for aligning and matching partial fingerprints to partial fingerprints.

Some researchers have studied the problem of matching a partial fingerprint to full template fingerprints. In (Jea and Govindaraju 2005), Jea and Govindaraju proposed a minutia-based approach to matching incomplete or partial fingerprints with full fingerprint templates. Their approach uses brute-force matching when the input fingerprints are small and few minutiae are presented, and uses secondary feature matching otherwise. Since this approach is based on minutiae, it is very likely to produce false matches when there are very few minutiae, and it is not applicable when there are no minutiae on the fingerprint fragments. Kryszczuk et al. (2004a, b) proposed to utilize pore locations to match fingerprint fragments. Using high resolution fingerprint images (approx. 2000 dpi in (Kryszczuk et al. 2004a, b)), they studied how pores might be used in matching partial fingerprints and showed that the smaller the fingerprint fragments, the greater the benefits of using pores. In their method, Kryszczuk et al. aligned fingerprints by searching for the transformation parameters which maximize the correlation between the input fingerprint

fragment and the candidate part on the full fingerprint template. Very recently, Chen and Jain (2007). 2005 employed minutiae, dots, and incipient ridges, to align and match partial fingerprints with full template fingerprints.

One drawback of most of the above approaches in aligning fragmental fingerprints is that they are mainly based on the features which are probably very few (e.g. minutiae) or even do not exist (e.g. dots and incipient ridges) on small fingerprint fragments (refer to Fig. 2.1). When the template fingerprints are also small fingerprint fragments, it will become difficult to get correct results due to the lack of features. In (Kryszczuk et al. 2004a, b), Kryszczuk et al. proposed a correlation-based blind searching approach to fragmental fingerprint alignment. As we will show later, however, this method has limited accuracy because it has to discretize the transformation parameter space.

2.1.2 *Fingerprint Alignment*

Fingerprint alignment or registration is a crucial step in fingerprint recognition. Its goal is to retrieve the transformation parameters between fingerprint images and then align them for matching. Some non-rigid deformation or distortion could occur in fingerprint image acquisition. It is very costly to model and remedy such distortions in fingerprint registration, and they can be compensated to some extent in subsequent fingerprint matching. Thus, the majority of existing fingerprint alignment methods considers only translation and rotation, although some deformable models (Cappelli et al. 2001; Ross et al. 2005) have been proposed. According to the features used, existing fingerprint alignment methods can be divided into two categories, minutia based and non-minutia feature based methods. Minutia based methods are now the most widely used ones (Huvanandana et al. 2000; Jain et al. 1997; Tico and Kuosmanen 2003; Jiang and Yau 2000; Kovacs-Vajna 2000; Feng 2008; Ratha et al. 1996; Chang et al. 1997; Chen et al. 2006a, b; Nanni and Lumini 2007). Non-minutia feature based methods (Bazen et al. 2000; Zhang and Wang 2002; Jain et al. 2000; Yager and Amin 2005, 2006; Liu et al. 2006; Ross et al. 2002) include those using image intensity values, orientation fields, cores, etc. One problem in applying these methods to partial fingerprints is that the features required by them could be very few on the fragments. Consequently, they will either lead to incorrect results or be not applicable.

There are roughly two kinds of methods for estimating alignment transformations. The first kind of methods quantizes the transformation parameters into finite sets of discrete values and searches for the best solution in the quantized parameter space (Kryszczuk et al. 2004a, b; Bazen et al. 2000; Ratha et al. 1996; Chang et al. 1997; Yager and Amin 2005, 2006; Liu et al. 2006; Ross et al. 2002). The alignment accuracy of these methods is thus limited due to the quantization. The second kind of methods first detects corresponding feature points (or reference points) on fingerprints and then estimates the alignment transformation based on the detected corresponding points (Huvanandana et al. 2000; Jain et al. 1997; Tico and

Kuosmanen 2003; Jiang and Yau 2000; Kovacs-Vajna 2000; Feng 2008; Zhang and Wang 2002; Jain et al. 2000; Chen et al. 2006a, b). Most of such methods make use of minutiae as the feature points. As discussed before, however, it is problematic to align partial fingerprints based on minutiae because of the lack of such features on the fingerprint fragments.

2.1.3 *Partial Fingerprint Alignment Based on Pores*

Following the second kind of alignment methods, we need to find some reference points other than minutiae on fingerprints for the purpose of aligning partial fingerprints. One possible solution is to use sufficiently densely sampled points on ridges as such reference points. However, it is hard, or even impossible, to ensure that identical points are sampled on different fingerprint images, and a too dense sampling of points will make the matching computationally prohibitive. On the contrary, sweat pores (as well as minutiae) are unique biological characteristics and are persistent on a finger throughout the life. Compared with minutiae, they are much more abundant on small partial fingerprints. Therefore, the pores can serve as reliable reference points in aligning partial fingerprint images. Although pore shapes and sizes are also important and biophysically distinctive features (Bindra et al. 2000), they cannot be reliably captured on fingerprint images because they are greatly affected by the pressure of the fingertip against the scanner. On the other hand, the pore statuses can change between open and close from time to time. Therefore, in general only the locations of pores are used in recognizing the pores and the fingerprints (Bazen et al. 2000).

Considering the plenty of pores on partial fingerprints, in this chapter we introduce, to the best of our knowledge, for the first time an approach to aligning partial fingerprints based on the pores reliably extracted from high resolution partial fingerprint images. This approach, by making use of the pores on fingerprints as reference feature points, can effectively align partial fingerprints and estimate the transformation between them even when there is a small overlap and large translation and rotation. We first propose an efficient method to extract pores, and then present a descriptor of pores, namely the Pore-Valley Descriptor (PVD), to determine the correspondences between them. The PVD describes a pore using its location and orientation, the ridge orientation inconsistency in its neighborhood, and the structure of valleys surrounding it. The partial fingerprints are first matched based on their PVDs, and the obtained pore correspondences are further refined using the global geometrical relationship between the pores. The transformation parameters are then calculated from the best matched pores. The experiments demonstrate that the proposed PVD-based alignment method can effectively detect corresponding pores and then accurately estimate the transformation between partial fingerprints. It is also shown that the proposed alignment method can significantly improve the recognition accuracy of partial fingerprint recognition.

2.2 Feature Extraction

The fingerprint features, including pores, ridges and valleys, will be used in the proposed method. The extraction of ridge orientations and frequencies and ridge maps has been well studied in the literature (Ratha and Bolle 2004; Maltoni et al. 2003). In this chapter, we use the classical methods proposed by Jain et al. (Jain et al. 1997; Hong et al. 1998) to extract ridge orientations, frequencies and ridge maps. Because ridges and valleys are complementary on fingerprints, it is a simple matter to get skeleton valley maps by thinning the valleys on the complement of ridge maps. To extract pores, we divide the fingerprint into blocks and use Gaussian matched filters to extract them block by block. The scales of Gaussian filters are adaptively determined according to the ridge frequencies on the blocks. After extracting orientation fields, valleys, and pores, we can then generate the Pore-Valley descriptor for each pore. Next we describe the feature extraction methods in detail.

2.2.1 Ridge and Valley Extraction

The considered partial fingerprint image has a higher resolution (approx. 1200 dpi in this chapter) than the conventional fingerprints (about 500 dpi) so that level-3 features such as pores can be reliably extracted from them. To extract ridges and valleys, it is not necessary to directly work on images of such a high resolution. In order to save computational cost, we smooth the image and down-sample it to half of its original resolution, and use the method in (Hong et al. 1998) to calculate the ridge orientations and frequencies. Based on local ridge orientations and frequencies, a bank of Gabor filters are used to enhance the ridges on the fingerprint. The enhanced fingerprint image is then binarized to obtain the binary ridge map.

On fingerprints, valleys and ridges are complementary to each other. Therefore, we can easily get the binary valley map as the complement of the binary ridge map. In order to exclude the effect of background on complement calculation, the fingerprint region mask (Hong et al. 1998) is employed to filter out the background if any. The binary valley map is then thinned to make all valleys be single-pixel lines. On the resulting skeleton valley map, there could be some false and broken valleys due to scars and noise. Thus we post-process it by connecting valley endings if they are very close and have opposite directions, and by removing valley segments between valley endings and/or valley bifurcations if they are very short or their orientations differ much from the local ridge orientations. Finally, we up-sample the obtained ridge orientation and frequency images, binary ridge map and skeleton valley map to the original resolution. Figure 2.2b shows the skeleton valley map extracted from the original fingerprint fragment in Fig. 2.2a.

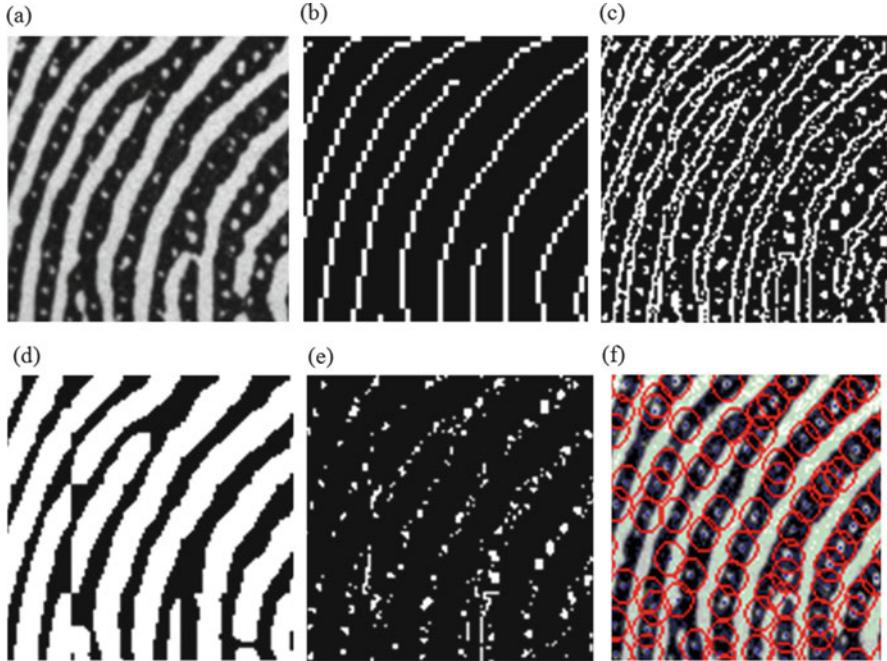


Fig. 2.2 (a) Original fingerprint image; (b) extracted skeleton valley map; Gaussian filtering output (c) at a small scale and (d) at a large scale; (e) difference between (c) and (d); (f) extracted pores after post-processing (pores are marked by red circles) (Color figure online)

2.2.2 Pore Extraction

Referring to Figs. 2.1 and 2.2a, on the fingerprint images captured using an optical contact fingerprint sensor, ridges (valleys) appear as dark (bright) lines, whereas pores are bright blobs on ridges, either isolated (i.e. closed pores) or connected with valleys (i.e. open pores). In general pores are circle-like structures and their spatial distributions are similar to 2-D Gaussian functions. Meanwhile, the cross-sections of valleys are 1-D Gaussian-like functions with different scales. To be specific, valleys usually have bigger scales than pores. Based on this observation, we use two 2-D Gaussian filters, one with a small scale and the other with a large scale, to enhance the image. The difference between their outputs can then give an initial pore extraction result. This procedure is basically the DoG (difference of Gaussian) filtering, which is a classic blob detection approach. The difficulty here is how to estimate the scales of the Gaussian filters.

Considering that the scale of either pores or valleys is usually not uniform across a fingerprint image and different fingerprints could have different ridge/valley frequencies, we partition the fingerprint into a number of blocks and estimate adaptively the scales of Gaussian filters for each block. Take a block image I_B as

an example. Suppose the mean ridge period over this block is p . It is a good measure of the scale in its corresponding fingerprint block. Thus, we set the standard deviations of the two Gaussian filters to $k_1 p$ and $k_2 p$ respectively ($0 < k_1 < k_2$ are two constants). The outputs of them are

$$F_1 = G_{k_1 p} * I_B, F_2 = G_{k_2 p} * I_B \quad (2.1)$$

$$G_\sigma(x, y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2+y^2}{2\sigma^2}} - m_G, |x|, |y| \leq 3\sigma \quad (2.2)$$

where ‘*’ denotes convolution and m_G is used to normalize the Gaussian filter to be zero-mean. Note that the settings of k_1 and k_2 should take into consideration the ridge and valley widths and the size of pores. In our experiments, we empirically chose the values for them based on the fingerprint database we used. The filtering outputs F_1 and F_2 are further normalized to [0, 1] and binarized, resulting in B_1 and B_2 . The small scale Gaussian filter $G_{k_1 p}$ will enhance both pores and valleys, whereas the large scale filter $G_{k_2 p}$ will enhance valleys only. Therefore, subtracting B_2 from B_1 , we obtain the initial result of pore extraction: $P_B = B_1 - B_2$.

To remove possible spurious pores from the initial pore extraction result P_B , we apply the following constraints to post-process the result. (1) Pores should reside on ridges only. To implement this constraint, we use the binary ridge map as a mask to filter the extracted pores. (2) Pores are circle-like features. We require that for a true pore, the eccentricity of its region should be less than a threshold. From Fig. 2.2e, f, it can be seen that this operation can successfully remove the spurious pores caused by valley contours, i.e. those line-shaped features in Fig. 2.2e. (3) Pores should be within a range of valid sizes. We measure the size of a pore by counting the pixels inside its region. In our experiments, we set the size between 3 and 30. (4) The mean intensity of a true pore region should be large enough and its variance should be small. Otherwise, the detected pores are viewed as false ones caused by noise. Finally, we get the extracted pore image. Figure 2.2c–f illustrate the pore extraction process of the fingerprint in Fig. 2.2a. It is worth mentioning that some other methods based on similar assumption (i.e. pores are circle-like features) have also been proposed in the literature (Jain et al. 2007; Ray et al. 2005). Compared with those methods, the pore extraction method proposed here takes into consideration the varying pore scales and thus has better pore extraction accuracy according to our experiments. Since it is out of the scope of this chapter, we do not make further discussion on this topic here due to the limit of space.

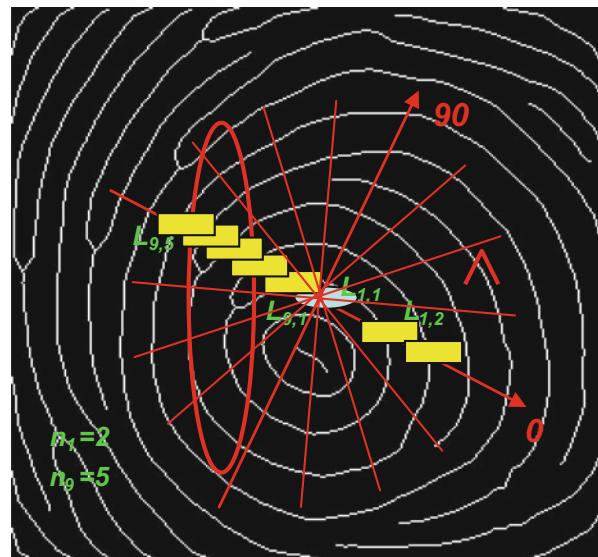
2.2.3 Pore-Valley Descriptors

In order to use pores to align fingerprints, a descriptor is needed to describe the pore features so that the correspondences between pores can be accurately determined. A good descriptor should be invariant to the deformations of rotation and translation,

which are very common when capturing fingerprints. Most previous studies on pore based fingerprint recognition (Kryszczuk et al. 2004a, b; Stosz and Alyea 1994; Jain et al. 2007) describe a pore simply by its location because they compare the pores on two fingerprints with the alignment between the two fingerprints known or estimated beforehand. However, if the alignment is not given, it is not sufficient to tell one individual pore from others by using only the location feature. Thus, it is necessary to employ some other information which can be useful in distinguishing pores. According to recent work on minutia-based fingerprint recognition methods, the ridge and valley structures and the ridge orientation field surrounding minutiae are also very important in minutia matching (Tico and Kuosmanen 2003; Feng 2008). Thus in this section we describe pores by using the neighboring valley structures and ridge orientation field. We call the resulting descriptor the Pore-Valley Descriptor (PVD).

The basic attribute of a pore is its location (X, Y) , which is defined as the column and row coordinates of the center of its mass. In this chapter, for the purpose of alignment, we introduce the orientation feature θ for a pore. It is defined as the ridge orientation at (X, Y) . Referring to Fig. 2.3, in order to sample the valley structures in the pore's neighborhood, we establish a local polar coordinate system by setting the pore's location as origin and the pore's orientation as the polar axis pointing to the right/bottom side. The polar angle is set as the counterclockwise angle from the polar axis. A circular neighborhood, denoted by N_p , is then chosen. It is centered at the origin with radius being $R_n = k_n p_{\max}$, where p_{\max} is the maximum ridge period on the fingerprint and k_n is a parameter to control the neighborhood size. Some radial lines are drawn starting from $\varphi_1 = 0^\circ$ with a degree step θ_s until $\varphi_m = m \cdot \theta_s$, where $m = \lfloor 360^\circ / \theta_s \rfloor$ is the total number of radial lines.

Fig. 2.3 Illustration of a Pore-Valley Descriptor with $k_n = 4$ and $\theta_s = 22.5$



For each line, we find where it intersects with valleys in the neighborhood. These intersections together with the pore give rise to a number of line segments. We number these segments from inside to outside and calculate their lengths. As shown in Fig. 2.3, a degree of 22.5 is taken as the step and hence 16 lines are employed. Taking the 0 degree and 180 degree lines as examples, the former has two segments and the latter has five segments. The ridge orientation field in the pore's neighborhood is another important feature. We define the ridge orientation inconsistency (OIC) in N_p as follows to exploit this information:

$$OIC(N_p) = \frac{1}{|N_p|} \sum_{(i,j) \in N_p} \left\{ [\cos(2 \cdot OF(i,j)) - m_{\cos}]^2 + [\sin(2 \cdot OF(i,j)) - m_{\sin}]^2 \right\} \quad (2.3)$$

where OF is the ridge orientation field, $|N_p|$ denotes the number of pixels in N_p , $m_{\cos} = \sum_{(i,j) \in N_p} \cos(2 \cdot OF(i,j)) / |N_p|$ and $m_{\sin} = \sum_{(i,j) \in N_p} \sin(2 \cdot OF(i,j)) / |N_p|$. With the above mentioned features, we define the PVD as the following feature vector Θ :

$$\Theta = \left[X, Y, \theta, OIC(N_p), \vec{S}_1, \vec{S}_2, \dots, \vec{S}_m \right] \quad (2.4)$$

$$\vec{S}_k = [n_k, L_{k,1}, L_{k,2}, \dots, L_{k,n_k}], \quad k = 1, 2, \dots, m \quad (2.5)$$

where n_k is the number of line segments along the k^{th} line, and $L_{k,n}$ is the length of the n^{th} segment ($1 \leq n \leq n_k$) along the k^{th} line.

The OIC component and the sampled valley structure features in the proposed PVD are invariant to rotation and translation because they are calculated in circular neighborhood of the pore which is intrinsically rotation-invariant and they are defined with respect to the local coordinate system of the pore. The OIC component is a coarse feature which captures the overall ridge flow pattern information in the neighborhood of a pore on a very coarse level. It will be used as an initial step to roughly match the pores. The sampled valley structure features are fine features. They will be used as the second step to accurately match pores. The pore locations and orientations will be used to double check pore correspondences. Finally, the transformation between fingerprints will be estimated based on the locations and orientations of their corresponding pores. In the next section, we will present the proposed PVD-based alignment algorithm.

2.3 PVD-Based Partial Fingerprint Alignment

This chapter aims to align partial fingerprints by using pores. To this end, we need to first identify pore correspondences on fingerprints. However, even a small fingerprint fragment can carry many pores (hundreds in the $6.24 \times 4.68 \text{ mm}^2$ fragments used in our experiments), making it very time consuming to match

pores in pairs directly using their surrounding valley structures (i.e. the segment lengths recorded in the PVD). Therefore, a coarse-to-fine matching strategy is necessary. The OIC components in the PVD can serve for the coarse matching. Given two pores, we first compare their OIC features. If the absolute difference between their OIC features is larger than a given threshold T_{oic} , they will not be matched; otherwise, proceed to the next fine matching step.

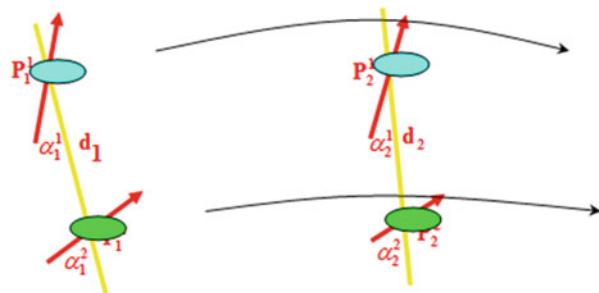
Coarse matching will eliminate a large number of false matches. In the subsequent fine matching, we compare the valley structures in the two pores' neighborhoods. According to the definition of PVD, each pore is associated with several groups of line segments which capture the information of its surrounding valleys. We compare these segments group by group. When comparing the segments in the k^{th} group, where there are n_k^1 and n_k^2 segments in the two pores' descriptors, we first find the common segments in the group, i.e. the first $\hat{n}_k = \min\{n_k^1, n_k^2\}$ segments. The dissimilarity between the two pores is then defined as

$$\sum_{k=1}^m \left(\sum_{n=1}^{\hat{n}_k} \frac{|L_{k,n}^1 - L_{k,n}^2|}{\hat{n}_k} + \frac{(n_k^1 - n_k^2)^2}{n_k^1 \cdot n_k^2} \right) \quad (2.6)$$

The first term in the formula calculates the mean absolute difference between all common segments in each group, and the second term is to penalize the missing segments. The smaller the dissimilarity is, the more similar the two pores are. After comparing all possible pairs of pores which pass coarse matching, each pair of pores is assigned with a dissimilarity calculated by Eq. 2.6. They are then sorted ascendingly according to the dissimilarities, producing the initial correspondences between the pores.

The top K initial pore correspondences (i.e. those with the smallest degree of dissimilarity) are further double checked to get the final pairs of corresponding pores for alignment transformation estimation. The purpose of double checking is to calculate the supports for all pore correspondences based on the global geometrical relationship between the pores. At the beginning of double checking, the supports to all pore correspondences are initialized to zero. Figure 2.4 illustrates the relevant measures we use.

Fig. 2.4 Illustration of the relevant measures used in pore correspondence double checking



Assume $\{P_1^1, P_2^1\}$ and $\{P_1^2, P_2^2\}$ are two pairs of corresponding pores among the top ones. To check them, we compare (1) the distances, denoted by d_1 and d_2 , between the pores on the two fingerprints; and (2) the angles, denoted by $\{\alpha_1^1, \alpha_2^1\}$ and $\{\alpha_1^2, \alpha_2^2\}$, between their orientations and the lines connecting them. If both the distance differences and the angle differences are below the given thresholds T_d and T_α , i.e.

$$|d_1 - d_2| \leq T_d, |\alpha_1^1 - \alpha_2^1| \leq T_\alpha, |\alpha_1^2 - \alpha_2^2| \leq T_\alpha \quad (2.7)$$

the supports for these two correspondences are increased by 1; otherwise, the support for the correspondence with higher dissimilarity is decreased by 1, whereas the support for the other one stays the same. After checking all the top K correspondences two by two, those with a non-negative support are taken as the final pore correspondences. If none of the correspondences has non-negative support, the two fingerprints cannot be aligned.

If some corresponding pores are found, we can then estimate the transformation between the two fingerprints. Here, we consider rotation and translation (since all the fingerprints are captured by the same type of scanner, we assume that the scaling factor is one) as follows

$$\begin{bmatrix} \tilde{X}_2 \\ \tilde{Y}_2 \end{bmatrix} = \begin{bmatrix} \cos \beta & -\sin \beta \\ \sin \beta & \cos \beta \end{bmatrix} \begin{bmatrix} X_2 \\ Y_2 \end{bmatrix} + \begin{bmatrix} \Delta X \\ \Delta Y \end{bmatrix} = R \begin{bmatrix} X_2 \\ Y_2 \end{bmatrix} + t \quad (2.8)$$

where (X_2, Y_2) are the coordinates of a pore on the second fingerprint and $(\tilde{X}_2, \tilde{Y}_2)$ are its transformed coordinates in the first fingerprint's coordinate system, $R = \begin{bmatrix} \cos \beta & -\sin \beta \\ \sin \beta & \cos \beta \end{bmatrix}$ is the rotation matrix and $t = \begin{bmatrix} \Delta X \\ \Delta Y \end{bmatrix}$ is the translation vector. Our goal is to estimate the transformation parameters $(\beta, \Delta X, \Delta Y)$, where β is the rotation angle and ΔX and ΔY are the column and row translations respectively.

If there is only one pair of corresponding pores found on the two fingerprints, we directly estimate the transformation parameters by the locations and orientations of the two pores, (X_1, Y_1, θ_1) and (X_2, Y_2, θ_2) , as follows:

$$\beta = \begin{cases} \beta_1 & \text{if } \text{abs}(\beta_1) \leq \text{abs}(\beta_2) \\ \beta_2 & \text{else} \end{cases} \quad (2.9)$$

$$\Delta X = X_1 - X_2 \cos \beta + Y_2 \sin \beta \quad (2.10)$$

$$\Delta Y = Y_1 - X_2 \sin \beta - Y_2 \cos \beta \quad (2.11)$$

where $\beta_1 = \theta_1 - \theta_2$ and $\beta_2 = \text{sgn}(\beta_1) \cdot (|\beta_1| - \pi)$.

If there are more than one pairs of corresponding pores, we employ the method similar to (Haralick et al. 1989) to estimate the rotation and translation parameters

based on the locations of the corresponding pores. Let $\{(X_1^i, Y_1^i) | i = 1, 2, \dots, C\}$ and $\{(X_2^i, Y_2^i) | i = 1, 2, \dots, C\}$ be C pairs of corresponding pores. We determine R and t by minimizing

$$\frac{1}{C} \sum_{i=1}^C \left\| \begin{bmatrix} X_1^i \\ Y_1^i \end{bmatrix} - R \begin{bmatrix} X_2^i \\ Y_2^i \end{bmatrix} - t \right\|^2 \quad (2.12)$$

Following the proof in (Haralick et al. 1989), it is easy to show that

$$t = \begin{bmatrix} \bar{X}_1 \\ \bar{Y}_1 \end{bmatrix} - R \begin{bmatrix} \bar{X}_2 \\ \bar{Y}_2 \end{bmatrix} \quad (2.13)$$

where $\bar{X}_j = (\sum_{i=1}^C X_j^i)/C$, $\bar{Y}_j = (\sum_{i=1}^C Y_j^i)/C$, $j = 1, 2$. Let

$$B = \frac{1}{C} \begin{bmatrix} \sum_{i=1}^C (X_1^i - \bar{X}_1)(X_2^i - \bar{X}_2) & \sum_{i=1}^C (Y_1^i - \bar{Y}_1)(X_2^i - \bar{X}_2) \\ \sum_{i=1}^C (X_1^i - \bar{X}_1)(Y_2^i - \bar{Y}_2) & \sum_{i=1}^C (Y_1^i - \bar{Y}_1)(Y_2^i - \bar{Y}_2) \end{bmatrix} \quad (2.14)$$

and its singular value decomposition be $B = UDV^T$, then $R = VU^T$ and $\beta = \arcsin(R_{21})$, R_{21} where is the entry at the second row and first column of R .

2.4 Experiments

In general, the feature of pores can only be reliably extracted from fingerprints with a resolution of at least 1000 dpi (CDEFFS 2008). So far there is no such free fingerprint image database available in the public domain. Therefore, we established a set of high resolution partial fingerprint images by using a custom-built fingerprint scanner of approximate 1200 dpi (refer to Fig. 2.5 for example images). With the established high resolution partial fingerprint image dataset, we evaluate the proposed fingerprint alignment method in comparison with a minutia-based method and an orientation field-based method. Next in subsection 2.4.1 we first introduce the collected dataset of high resolution partial fingerprint images. In subsection 2.4.2 we investigate the two parameters involved in the method. Subsection 2.4.3 compares the proposed method with the minutia based method in corresponding feature point detection. Subsection 2.4.4 compares the proposed method with the orientation field based method in alignment transformation estimation. In subsection 2.4.5 we compare the three methods in terms of fingerprint recognition accuracy. Finally, in subsection 2.4.6, we analyze the computational complexity of the method.

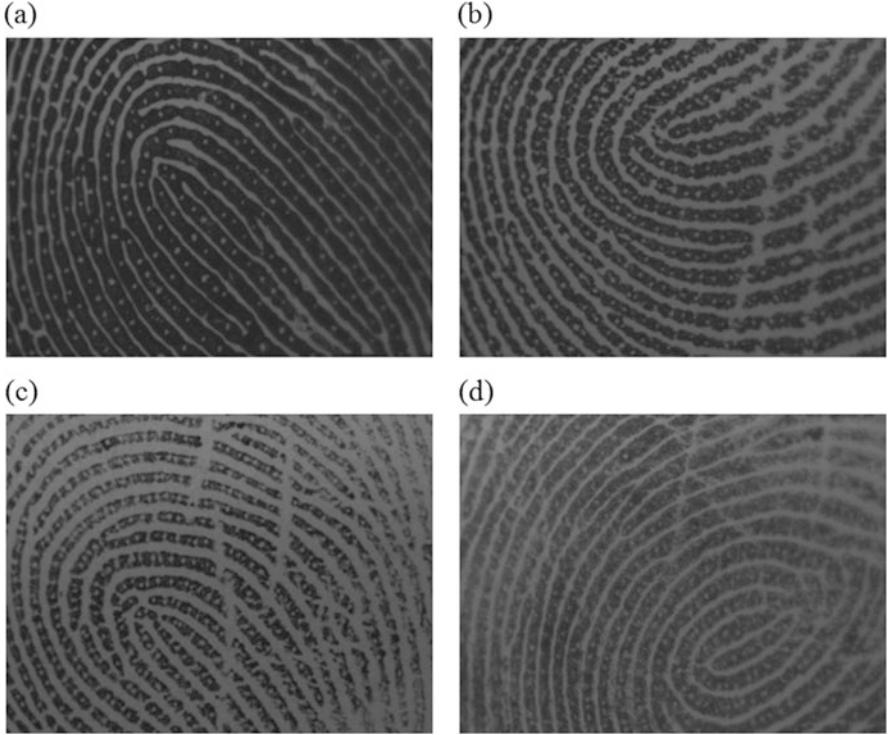


Fig. 2.5 Example fingerprint images used in the experiments. Their quality indexes are (a) 0.8777, (b) 0.7543, (c) 0.6086, and (d) 0.5531 according to the frequency domain quality index defined in Chen et al. (2005)

2.4.1 The Dataset

We first collected 210 partial fingerprint images from 35 fingers as the training set for parameter selection and evaluation, and then collected 1480 fingerprint fragments from 148 fingers (including the fingers in the training set) as the test set for performance evaluation. The data were collected in two sessions (about 2 weeks apart). Most of the participants are students and staff in our institute, whose ages are between 20 and 50 years old. In the training set, there are three images captured from each finger in each session; whereas in the test set, each finger has five images scanned in each of the two sessions.

The resolution of these fingerprint images is approximately 1200 dpi and their spatial size is 320 pixels in width and 240 pixels in height. Therefore, they cover an area of about 6.5 mm by 4.9 mm on fingertips. When capturing the fingerprint images, we simply asked the participants to naturally put their fingers against the prism of the scanner without any exaggeration of fingerprint deformation. As a result, typical transformations between different impressions of the same finger in

the dataset include translations with tens of pixels and rotations by around eight degrees. The maximal translations and rotations are respectively about 200 pixels and 20 degrees. Hence, the minimal overlap between a finger's different impressions is about one fourth of the fingerprint image area. In subsequent experiments, we will give representative examples of these cases.

The resolution of these fingerprint images is approximately 1200 dpi and their spatial size is 320 pixels in width and 240 pixels in height. Therefore, they cover an area of about 6.5 mm by 4.9 mm on fingertips. When capturing the fingerprint images, we simply asked the participants to naturally put their fingers against the prism of the scanner without any exaggeration of fingerprint deformation. As a result, typical transformations between different impressions of the same finger in the dataset include translations with tens of pixels and rotations by around eight degrees. The maximal translations and rotations are respectively about 200 pixels and 20 degrees. Hence, the minimal overlap between a finger's different impressions is about one fourth of the fingerprint image area. In subsequent experiments, we will give representative examples of these cases.

2.4.2 *The Neighborhood Size and Sampling Rate*

The proposed alignment method uses the valley structures in the neighborhood of pores. The valley structures are sampled along a number of different directions, determined by the degree step θ_s . Here we refer to θ_s as the sampling rate of directions. Obviously, the neighborhood size and sampling rate are two critical parameters in the proposed alignment method. We set the neighborhood size as k_n times the maximum ridge period. Intuitively, a small k_n or large θ_s will cost less computational resource but will make the resulting PVDs less discriminative, whereas a large k_n or small θ_s will lead to more noise-sensitive and costly PVDs. We evaluated the effect of k_n and θ_s on the accuracy of corresponding feature point detection using 50 pairs of fingerprints which were randomly chosen from the training set. Each pair is from the same finger but taken at different sessions. These fingerprints show different quality. Some example images are shown in Fig. 2.5.

We used two measures to evaluate the accuracy: the percentage of correct top one pore correspondence ($M1$) and the average percentage of correct correspondences among the top five pore correspondences ($M2$). Let N be the total number of pairs of fingerprint images, and N_{T5} the number of pairs of fingerprints on which the top one pore correspondence is correct. We also counted the number of correct pore correspondences among the top five correspondences on each pair of fingerprints. Denote by N_{T5}^i the number of correct pore correspondences among the top five correspondences on the i^{th} pair of fingerprints. Then the two measures $M1$ and $M2$ are defined as

$$M1 = N_{T1}/N \quad (2.15)$$

$$M2 = \frac{1}{N} \sum_{i=1}^N N_{T5}^i / 5 \quad (2.16)$$

We investigated several combinations of different values for k_n and θ_s , i.e. $k_n \in \{3, 3.5, 4, 4.5, 5\}$ and $\theta_s \in \{15^\circ, 18^\circ, 20^\circ, 22.5^\circ, 30^\circ, 45^\circ\}$. Table 2.1 lists the results on the 50 pairs of fingerprints. From the results, we can see that the best accuracy is obtained at a sampling rate of $\theta_s = 20^\circ$ or $\theta_s = 22.5^\circ$, and no significant difference is observed between these two different sampling rates. With respect to the neighborhood size, it appears that $k_n = 4$ is a good choice. Furthermore, it was observed that neither too small nor too large neighborhoods can produce the best accuracy. In our following experiments, considering both the accuracy and the computational cost, we set $k_n = 4$ and $\theta_s = 22.5^\circ$. Note that the settings of these two parameters should be dependent on the resolution of fingerprint images and the population from which the fingerprint images are captured. If a different fingerprint image dataset is used, the above training process has to be done again by using a subset of the fingerprint images in that dataset.

2.4.3 Corresponding Feature Point Detection

Detecting feature point correspondences is an important step in the proposed alignment method as well as in many state-of-the-art minutia-based methods. The optimal alignment transformation is estimated based on the detected corresponding feature points (i.e. pores or minutiae). Considering the significance of corresponding feature point detection, we carried out experiments to compare the proposed method with a representative minutia-based method (Chen et al. 2006a) in terms of corresponding feature point detection accuracy. In the experiments, we used 200 pairs of partial fingerprints randomly chosen from the training set for evaluation. In each pair, the two fingerprints are from the same finger but were captured at different sessions.

Table 2.1 Accuracies (%) of corresponding pore detection on 50 pairs of fingerprints under different settings of k_n and θ_s

		θ_s					
		15°	18°	20°	22.5°	30°	45°
k_n	3	32/45.2	44/51.6	50/57.5	50/58.2	46/52.5	40/49.1
	3.5	52/60.1	62/75.2	74/80.2	78/82.5	70/69.6	58/62.5
	4	66/74.6	80/80.5	96/95.1	98/94.7	94/88.5	80/78.2
	4.5	76/80.2	84/86	88/90.5	86/89.1	80/78.1	72/70.6
	5	54/49	62/56.7	66/60.5	60/61.7	54/59.2	52/52.1

Figure 2.6 shows some example pairs of fingerprint fragments with the detected corresponding minutiae (left column) or pores (right column). When there are more than five pairs of corresponding minutiae or pores, we show only the first five pairs. In Fig. 2.6a, b, both methods can correctly find the top five feature point correspondences. However, when the fingerprint quality changes between sessions, for example because of perspiration, the minutiae based method will tend to detect false minutiae and hence false minutia correspondences. In Fig. 2.6c, broken valleys occur on the second fingerprint. As a result, the detected two minutia correspondences are incorrect. Instead, the proposed PVD-based method is more robust and can correctly detect the corresponding pores as shown in Fig. 2.6d.

The fingerprint fragments in Fig. 2.6e have large deformation and small overlap. Consequently, few (fewer than ten) minutiae can be found in their overlapping region. In this case, the minutia-based method fails again because there lack sufficient minutiae. Actually, even when two partial fingerprints overlap much, there could still be very few minutiae available on them because of the small

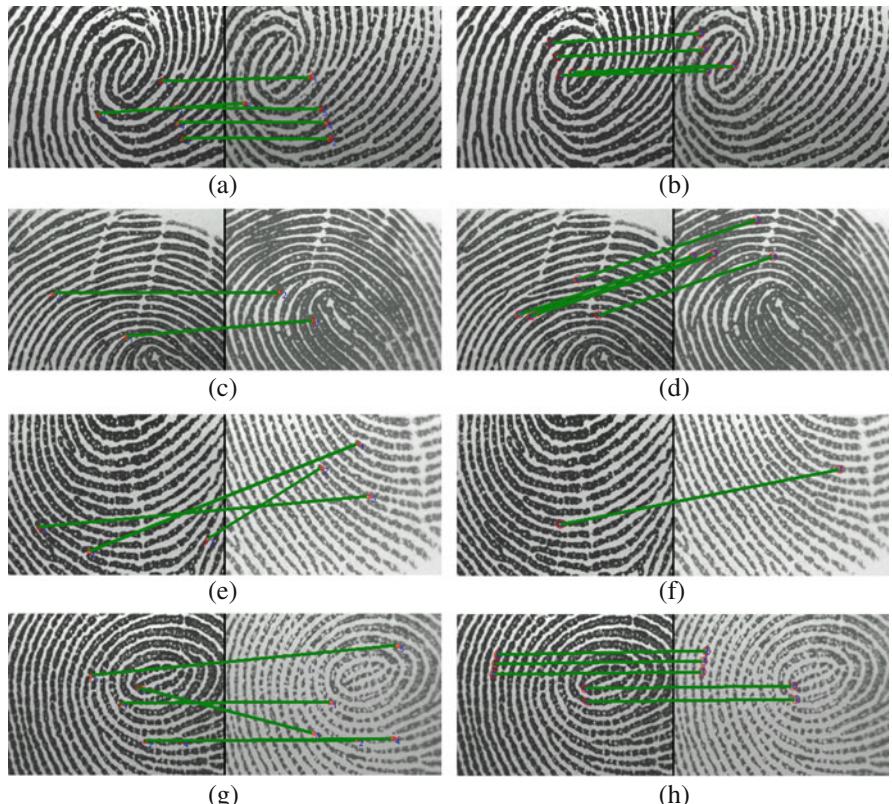


Fig. 2.6 Examples of corresponding feature point detection results using minutia based (*left column*) and PVD based (*right column*) methods

Table 2.2 Accuracies of corresponding feature point detection by the two methods

	M1	M2
Minutia based method (Chen et al. 2006b)	40%	35.1%
PVD based method	98%	95.5%

fingerprint areas. As can be seen in Fig. 2.6g, some false correspondences are detected on the two fragments due to insufficient minutiae. In contrast, as shown in Fig. 2.6f, h, the results by the proposed PVD-based method on these partial fingerprints are much better.

We calculated the two measures, M1 and M2, for the two methods on all the 200 pairs of partial fingerprints. The results are listed in Table 2.2. It can be seen that the minutia-based method works poorly whereas the proposed PVD-based method can detect the corresponding feature points with a very high accuracy, achieving significant improvements over the minutia-based method. This demonstrates that the PVD-based alignment method can cope with various fingerprint fragments more accurately than the minutia based method, largely thanks to the abundance and distinctiveness of pores on fingerprints. Since the alignment transformation estimation is based on the detected corresponding feature points, it is obvious that the PVD-based method will also estimate the alignment transformation more accurately than the minutia-based method. Next, we compare the PVD-based method with an orientation field-based method in terms of alignment transformation estimation accuracy.

2.4.4 Alignment Transformation Estimation

After obtaining the pore correspondences on two fingerprints, we can then estimate the alignment transformation between them based on the corresponding pores. To quantitatively evaluate the performance of the proposed method in alignment transformation estimation, we need some ground truth fingerprint fragment pairs. To this end, we randomly chose ten pairs of fingerprints from the test set (each pair was captured from the same finger but in two different sessions), and manually computed their transformations as the ground truth. Because we consider only translation and rotation here, we need at least two pairs of corresponding feature points on a pair of fingerprints to calculate the transformation between them. Therefore, we first manually marked two pairs of corresponding feature points on each of the ten pairs of fingerprints. Based on the coordinates of the two pairs of corresponding feature points, we then directly computed the transformation between the pair of fingerprints by solving a set of equations. The obtained ground truth on the ten pairs of fingerprints is given in 2.3. The first three pairs of fingerprints are shown in Fig. 2.7. From Table 2.3, we can see that these chosen fingerprint pairs display translations from less than 10 pixels to about 180 pixels and rotations from less than five degrees to more than 10 degrees. In our experiments,

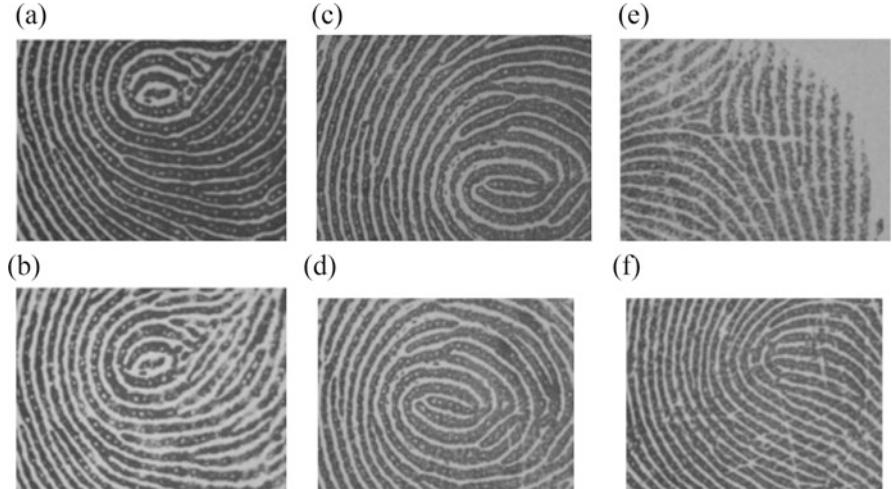


Fig. 2.7 Three of the ten chosen pairs of fingerprint fragments. (a) and (b), (c) and (d), (e) and (f) are the first three pairs of fingerprints listed in Table 2.3

Table 2.3 Alignment transformation ground truth and estimation results by the two methods

Index of fingerprint pair	Ground truth			OF based method (Yager and Amin 2006)			PVD based method		
	ΔY	ΔX	β	ΔY	ΔX	β	ΔY	ΔX	β
	01	-56	23	11.01	-11	3	-2.00	-61	33
02	90	33	-15.95	43	51	-1.00	93	26	-16.89
03	100	-181	-2.87	-8	72	-31.02	91	-176	2.93
04	-11	-8	-3.16	-1	1	0.00	-11	-6	-1.98
05	90	8	4.19	100	0	1.00	89	6	3.88
06	69	-74	-5.44	23	-21	0.00	72	-78	-9.33
07	78	-137	-3.45	45	-24	0.00	76	-142	-6.40
08	-87	2	0.96	-74	1	-1.00	-93	7	2.23
09	-73	39	-4.40	-69	47	-3.00	-79	50	-0.60
10	19	-4	-11.25	-4	-2	-1.00	12	-1	-8.61

we have observed that typical transformations in the dataset are translations by tens of pixels and rotations by around eight degrees. In this part, we compared our proposed method with the steepest descent orientation field (OF) based alignment method (Yager and Amin 2006) in terms of alignment transformation estimation accuracy using the chosen fingerprint pairs.

Table 2.3 lists the estimation results by the OF based method (the step sizes of translation and rotation are set as one pixel and one degree respectively) and the proposed method on the chosen fingerprint pairs. Figure 2.8 illustrates the aligned fingerprint images by overlaying the first image with the transformed second image in the pair shown in Fig. 2.7. Obviously, the PVD-based method estimates the

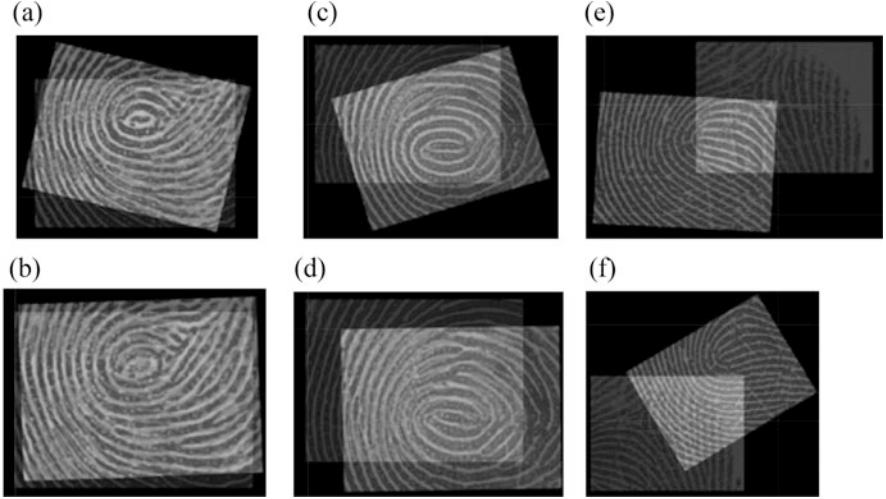


Fig. 2.8 Alignment results of PVD based method (a, c, e) and OF based method (b, d, f) on the fingerprint pairs shown in Fig. 2.7a, b, Fig. 2.7c, d, and Fig. 2.7e, f

Table 2.4 Average absolute errors (AAE) by the two methods

	AAE (ΔY)	AAE (Δx)	AAE (β)
OF based method (Yager and Amin 2006)	33.9	48.5	8.5
PVD based method	4.2	5.4	2.5

transformation parameters much more accurately and it does not have the initialization and quantization problems, which will affect greatly the performance of OF based method. Moreover, there is no guarantee that the OF based method will always converge to the global optimal solution. In fact, it can be easily trapped at local minima, for example the third pair of fingerprints which has small overlap (refer to the last column in Figs. 2.7 and 2.8). In Table 2.4, we list the average absolute errors of the two methods over the chosen ten fingerprint pairs. These results clearly demonstrate that the PVD-based method can recover the transformation between partial fingerprints more accurately.

2.4.5 Partial Fingerprint Recognition

We have also evaluated the proposed alignment method in partial fingerprint recognition by setting up a simple partial fingerprint recognition system as shown in Fig. 2.9. In this system, the alignment transformation is first estimated between an input fingerprint image and a template fingerprint image by using one of the three methods: minutia-based method, orientation field based method, and the

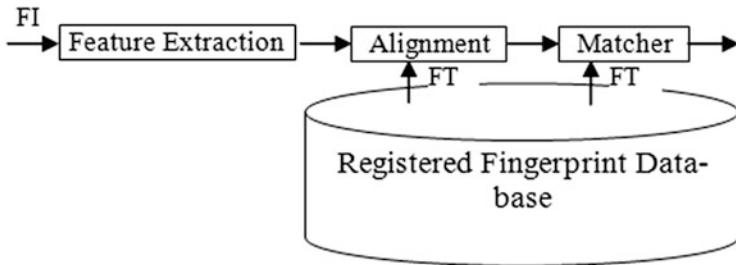


Fig. 2.9 The simple partial fingerprint recognition system used in the experiments. FI and FT denote the input and template fingerprints respectively

proposed PVD based method. As for the matcher, we employed two different approaches. The first one is a minutia and pore based matcher (called MINU-PORE matcher). It matches the minutiae and pores on the fingerprints, and then fuses the match scores of minutiae and pores to give a similarity score between the fingerprints. The second approach is an image-based matcher called GLBP matcher based on Gabor and Local Binary Patterns (LBP), recently proposed by Nanni and Lumini (2007). Please note that the purpose of the experiments here is to compare the contributions of the three different alignment methods to a fingerprint matcher. Therefore, we did not do any optimization on the matchers but considered only the relative improvement between the three alignment methods.

The MINU-PORE matcher we implemented works as follows. The minutiae and pores on the input fingerprint image are transformed into the coordinate system of the template fingerprint according to the estimated transformation. Minutiae and pores on the two fingerprint images are then matched separately. Two minutiae are thought to be matched if the difference between their locations and the difference between their directions are both below the given thresholds (15 pixels for location differences and 30 degrees for direction differences in our experiments). As for two pores, if the difference between their locations is below a given threshold (15 pixels in our experiments), they are matched. The minutia matching score is defined as the ratio between the number of matched minutiae to the total number of minutiae, and the pore matching score is defined similarly. The final matching score is obtained by fusing the minutia and pore matching scores using the summation rule.

By using the MINU-PORE matcher on the test set, we conducted the following matches. (1) Genuine matches: each of the fingerprint images in the second session was matched with all the fingerprint images of the same finger in the first session, resulting in 3700 genuine match scores. (2) Imposter matches: the first fingerprint image of each finger in the second session was matched with the first fingerprint images of all the other fingers in the first session, resulting in 21,756 imposter match scores. Based on the obtained match scores, we calculated the equal error rate (EER) of each of the three alignment methods: 29.5% by the PVD based alignment method, 38.66% by the Minutia based alignment method, and 41.03% by the OF based alignment method. The receiver operating characteristic (ROC) curves of the

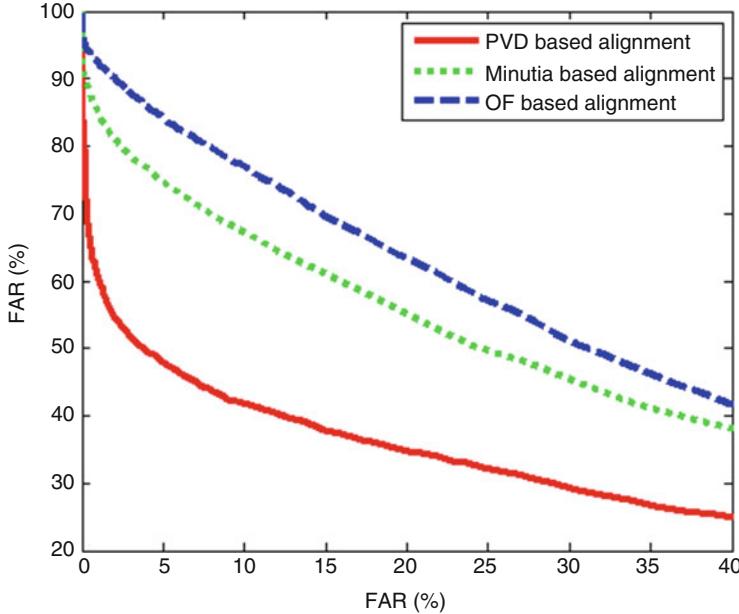


Fig. 2.10 The ROC curves of the MINU-PORE matcher by using the three alignment methods

three methods are plotted in Fig. 2.6. It can be clearly seen that the proposed PVD based alignment method makes much improvement in EER, specifically 23.69% over the minutia based method and 28.1% over the OF based method (Fig. 2.10).

As for the GLBP matcher, we first transform the input fingerprint image into the coordinate system of the template fingerprint image according to the estimated alignment transformation, then extract the Gabor-LBP feature vectors from the transformed input fingerprint image and the template fingerprint image (we directly took the configuration parameters from (Nanni and Lumini 2007)), and finally calculate the Euclidean distance between the Gabor-LBP feature vectors of the input fingerprint image and the template fingerprint image. By using the GLBP matcher, we carried out the same matching scheme as in the MINU-PORE matcher. As a result, the PVD-based alignment method leads to the EER of 34.85%, the minutia based alignment method 39.98%, and the OF based alignment method 45.11%. Figure 2.11 shows the corresponding ROC curves. Compared with the other two methods, the proposed PVD based alignment method achieves 12.83% and 22.74% improvement in EER respectively. In all the experiments, it is observed that matching errors are largely caused by inaccurate alignments. This validates that the proposed alignment method is more suitable for partial fingerprints and can significantly improve the accuracy of partial fingerprint recognition. Although the EER obtained here is relatively high, this is because the recognition of partial fingerprint images is itself very challenging due to the limited feature.

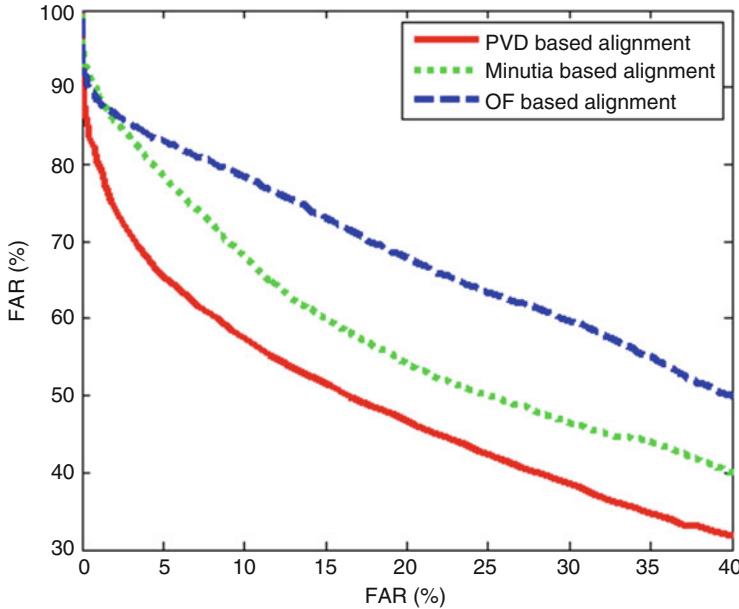


Fig. 2.11 The ROC curves of the GLBP matcher by using the three alignment methods

2.4.6 Computational Complexity Analysis

The proposed PVD based alignment method has the following main steps for each pair of fingerprint images to be aligned: (A) ridge orientation and frequency estimation; (B) ridge and valley extraction; (C) pore extraction; (D) PVD generation; (E) PVD comparison; (F) pore correspondence refinement; and (G) transformation estimation. The first two steps (A) and (B) are common to most automatic fingerprint recognition systems. The last step (G) involves a singular value decomposition of a 2×2 matrix, which can be implemented very efficiently. We have implemented the method by using Matlab and executed it on a PC with a 2.13 GHz Intel(R) Core(TM) 2 6400 CPU and RAM of 2GB. It takes about 0.02 ms to estimate the transformation from a set of corresponding feature points. The step F, pore correspondence refinement, needs to calculate some Euclidean distances and angles, which can also be done in about 0.02 ms. The step C, pore extraction, is a little bit more time-consuming. The pore extraction method we used in this chapter is a filtering based approach, which extracts pores by some linear filter operations. In our experiments, it takes about two seconds to extract the pores from a fingerprint image.

The most time-consuming steps are PVD generation (Step D) and comparison (Step E). Although it does not take much time to generate the PVD for one pore (about 0.02 s) or to compare the PVD of two pores (about 0.02 ms), processing the whole set of pores on fingerprints takes more time because of the large quantity of

pores. With regard to the fingerprint images used in our experiments, there are averagely around 500 pores on a fingerprint fragment. Therefore, it takes in average about 10 s and 5 s respectively to generate the PVD for the pores on a fingerprint fragment and to compare the PVD on two fingerprint fragments. Considering that we did not optimize the code and that the Matlab code itself has low efficiency, we expect that the computational cost can be much reduced after optimization and the speed can be significantly improved by using languages like C/C++. Compared with the proposed method, the minutia based method is more efficient, taking usually less than one second for either extracting or matching the minutiae (but using C/C++ implementation). As for the OF based method, the time needed to align two fingerprints depends on a number of factors, such as the amount of transformation between the fingerprints, the initial estimation of the transformation, and the step sizes used in the search process. Therefore, it is difficult to draw a conclusion on its efficiency. In our experiments, the OF based method can sometimes converge in less than one second, but sometimes converge after more than one minute. Generally speaking, the proposed method achieves much higher alignment accuracy than the other two approaches with an acceptable computational cost.

2.5 Summary

A new approach was proposed in this chapter to aligning partial high resolution fingerprints using pores. After pore detection, a novel descriptor, namely Pore-Valley Descriptor (PVD), was defined to describe pores based on their local characteristics. Then a coarse-to-fine pore matching method was used to find the pore correspondences based on PVD. With the detected corresponding pores, we estimated the alignment transformation between the fingerprint fragments. To evaluate the performance of the proposed PVD based high resolution partial fingerprint alignment method, we established a set of partial fingerprint images and used them to compare the proposed method with state-of-the-art minutia-based and orientation field-based fingerprint alignment methods. The experimental results demonstrated that the PVD-based method can more accurately detect the corresponding feature points and hence estimate better the alignment transformation. It was also shown in our experiments that the accuracy of partial fingerprint recognition can be significantly improved by using the PVD based alignment method.

One important issue in high resolution fingerprint recognition is the stability of pores. Despite that not all pores will appear in the fingerprint images of the same person but captured at different times, we experimentally found that usually there will be enough corresponding pores that can be detected on the fingerprint images from the same person. It is interesting and very important to further investigate the statistical characteristics of pores on fingerprint images.

Although the PVD based alignment method proposed in this chapter is designed for high resolution partial fingerprint recognition, it is not limited to partial fingerprints. It can also be applied to full fingerprint images. One problem may be the expensive computational cost caused by the large amount of pore features. One solution could be to perform coarse registration first by using OF based schemes and then apply the PVD based method for a fine estimation of the alignment transformation. It also deserves to do more investigation of the discriminative power of pores.

References

- Bazen A, Verwaaijen G, Gerez S, Veelenturf L, Zwaag B (2000) A correlation-based fingerprint verification system. In: Proceedings of the workshop on circuits systems and signal processing, pp 205–213
- Bindra B, Jasuja O, Singla A (2000) Poroscopy: a method of personal identification revisited. Internet J Forensic Med Toxicol 1(1)
- Cappelli R, Maio D, Maltoni D (2001) Modeling plastic distortion in fingerprint images. In: Proceedings of ICAPR, Rio de Janeiro
- CDEFFS (2008) Data format for the interchange of extended fingerprint and palmprint features, Working Draft Version 0.2. <http://fingerprint.nist.gov/standard/cdeffs/index.html>
- Chang S, Hsu W, Wu G (1997) Fast algorithm for point pattern matching: invariant to translations, rotations and scale changes. Pattern Recogn 30:311–320
- Chen Y, Jain A (2007) Dots and incipients: extended features for partial fingerprint matching. Paper presented at biometric symposium, BCC, Baltimore, Sept 2007
- Chen Y, Dass S, Jain A (2005) Fingerprint quality indices for predicting authentication performance. In: Proceedings of AVBPA, pp 160–170
- Chen X, Tian J, Yang X (2006a) A new algorithm for distorted fingerprints matching based on normalized fuzzy similarity measure. IEEE Trans Image Process 15:767–776
- Chen X, Tian J, Yang X, Zhang Y (2006b) An algorithm for distorted fingerprint matching based on local triangle feature set. IEEE Trans Inf Forensics Secur 1:169–177
- Choi K, Choi H, Lee S, Kim J (2007) Fingerprint image mosaicking by recursive ridge mapping. IEEE Trans Syst Man Cybern B 37:1191–1203
- Feng J (2008) Combining minutiae descriptors for fingerprint matching. Pattern Recogn 41:342–352
- Haralick R, Joo H, Lee C, Zhuang X, Vaidya V, Kim M (1989) Pose estimation from corresponding point data. IEEE Trans Syst Man Cybern 19:1426–1446
- Hong L, Wan Y, Jain A (1998) Fingerprint image enhancement: algorithms and performance evaluation. IEEE Trans Pattern Anal Mach Intell 20(8):777–789
- Huvanandana S, Kim C, Hwang J (2000) Reliable and fast fingerprint identification for security applications. Proc Int Conf Image Process 2:503–506
- Jain A, Hong L, Bolle R (1997) On-line fingerprint verification. IEEE Trans Pattern Anal Mach Intell 19:302–314
- Jain A, Prabhakar S, Hong L, Pankanti S (2000) Filterbank-based fingerprint matching. IEEE Trans Image Process 9:846–859
- Jain A, Che Y, Demirkus M (2007) Pores and ridges: fingerprint matching using level 3 features. IEEE Trans Pattern Anal Mach Intell 29:15–27
- Jea T, Govindaraju V (2005) A minutia-based partial fingerprint recognition system. Pattern Recogn 38:1672–1684

- Jiang X, Yau WY (2000) Fingerprint minutiae matching based on the local and global structures. *Proc Int Conf Pattern Recogn* 2:1042–1045
- Kovacs-Vajna Z (2000) A fingerprint verification system based on triangular matching and dynamic time warping. *IEEE Trans Pattern Anal Mach Intell* 22:1266–1276
- Kryszczuk K, Drygajlo A, Morier P (2004a) Extraction of level 2 and level 3 features for fragmentary fingerprints. In: Proceedings of the Second COST Action 275 workshop, Vigo, Spain, pp 83–88
- Kryszczuk K, Morier P, Drygajlo A (2004b) Study of the distinctiveness of level 2 and level 3 features in fragmentary fingerprint comparison, BioAW2004. LNCS 3087:124–133
- Liu L, Jiang T, Yang J, Zhu C (2006) Fingerprint registration by maximization of mutual information. *IEEE Trans Image Process* 15:1100–1110
- Maltoni D, Maio D, Jain A, Prabhakar S (2003) Handbook of fingerprint recognition. Springer, New York
- Nanni L, Lumini A (2007) A hybrid wavelet-based fingerprint matcher. *Pattern Recogn* 40 (11):3146–3151
- Nanni L, Lumini A (2008) Local binary patterns for a hybrid fingerprint matcher. *Pattern Recogn* 41(11):3461–3466
- Parthasaradhi S, Derakhshani R, Hornak L, Schuckers SAC (2005) Time-series detection of perspiration as a liveness test in fingerprint devices. *IEEE Trans Syst Man Cybern C* 35:335–343
- Ratha N, Bolle R (2004) Automatic fingerprint recognition systems. Springer, New York
- Ratha N, Karu K, Chen S, Jain A (1996) A real-time matching system for large fingerprint databases. *IEEE Trans Pattern Anal Mach Intell* 18:799–813
- Ray M, Meenen P, Adhami R (2005) A novel approach to fingerprint pore extraction. Proceedings of the 37th south-eastern symposium on system theory, pp 282–286
- Roddy A, Stosz J (1997) Fingerprint features – statistical analysis and system performance estimates. *Proc IEEE* 85:1390–1421
- Ross A, Reisman J, Jain A (2002) Fingerprint matching using feature space correlation. In: Proceedings of post-European conference on computer vision workshop on biometric authentication, LNCS 2359, pp 48–57
- Ross A, Jain A, Reisman J (2003) A hybrid fingerprint matcher. *Pattern Recogn* 36:1661–1673
- Ross A, Dass S, Jain A (2005) A deformable model for fingerprint matching. *Pattern Recogn* 38:95–103
- Stosz JD, Alyea L (1994) Automated system for fingerprint authentication using pores and ridge structure. In: Proceedings of SPIE conference on automatic systems for the identification and inspection of humans, San Diego, vol 2277, pp 210–223
- Teoh A, Ngo D, Song O (2004) An efficient fingerprint verification system using integrated wavelet and Fourier-Mellin invariant transform. *Image Vis Comput* 22(6):503–513
- Tico M, Kuosmanen P (2003) Fingerprint matching using an orientation-based minutia descriptor. *IEEE Trans Pattern Anal Mach Intell* 25:1009–1014
- Yager N, Amin A (2005) Coarse fingerprint registration using orientation fields. *EURASIP J Appl Signal Process* (13):2043–2053
- Yager N, Amin A (2006) Fingerprint alignment using a two stage optimization. *Pattern Recogn* 27:317–324
- Zhang W, Wang Y (2002) Core-based structure matching algorithm of fingerprint verification. *Proc Int Conf Pattern Recogn* 1:70–74

Chapter 3

Fingerprint Pore Modeling and Extraction

Abstract Sweat pores on fingerprints have proven to be discriminative features and have recently been successfully employed in automatic fingerprint recognition systems (AFRS), where the extraction of fingerprint pores is a critical step. Most of the existing pore extraction methods detect pores by using a static isotropic pore model; however, their detection accuracy is not satisfactory due to the limited approximation capability of static isotropic models to various types of pores. This chapter presents a dynamic anisotropic pore model to describe pores more accurately by using orientation and scale parameters. An adaptive pore extraction method is then developed based on the proposed dynamic anisotropic pore model. The fingerprint image is first partitioned into well-defined, ill-posed, and background blocks. According to the dominant ridge orientation and frequency on each foreground block, a local instantiation of appropriate pore model is obtained. Finally, the pores are extracted by filtering the block with the adaptively generated pore model. Extensive experiments are performed on the high resolution fingerprint databases we established. The results demonstrate that the proposed method can detect pores more accurately and robustly, and consequently improve the fingerprint recognition accuracy of pore-based AFRS.

Keywords Automatic fingerprint recognition • Pore extraction • Pore models

3.1 Introduction

Fingerprint is the most widely used biometric characteristic for personal identification because of its uniqueness and stability over time (Zhang 2000; Jain et al. 2007a, b; Maltoni et al. 2003). Most of the existing automatic fingerprint recognition systems (AFRS) use the minutia features on fingerprints, i.e. the terminations and bifurcations of fingerprint ridges (Ratha and Bolle 2004; Ratha et al. 1996; Jain et al. 1997), for recognition. Although they can achieve good recognition accuracy and have been used in many civil applications, their performance still needs much improvement when a large population is involved or a high security level is required. One solution to enhancing the accuracy of AFRS is to employ more

features on fingerprints other than only minutiae (Jain et al. 2007a, b; Ross et al. 2003; He et al. 2006).

Generally, fingerprint features can be divided into three levels (CDEFFS 2009). Level 1 features (e.g. overall fingerprint ridge patterns) and Level 2 features (e.g. minutiae) have been extensively studied and they are employed in most existing AFRS. Level 3 features, however, are ignored in many AFRS even though they are also very distinctive and have been used for a long time in the forensic community (Ashbaugh 1999; Bindra et al. 2000). Level 3 features refer to ridge dimensional attributes such as ridge contours and pores, which are fine details on ridges and require high resolution imaging techniques to reliably capture (CDEFFS 2009). Such requirements limit the use of Level 3 features in conventional AFRS. Thanks to the advances in imaging techniques and the demand for more secure biometric systems, recently researchers have been paying more and more attention to using Level 3 features in AFRS. Roddy and Stosz (1997), and Parsons et al. (2008) statistically analyzed the discriminative power of pores and validated the effectiveness of pore configuration in personal identification. The first AFRS using pores was developed by Stosz and Alyea (1994). They combined minutiae and pores to recognize persons. Subsequently, Kryszczuk et al. (2004a, b) investigated the contribution of pores to fragmentary fingerprint recognition and showed that the smaller the fingerprint fragments, the greater the benefit of using pores. Recently, Jain et al. (2006, 2007a, b) proposed a high resolution fingerprint recognition system which uses features from all the three levels (i.e. ridge orientation fields, minutiae, ridge contours, and pores).

A critical step in the pore based AFRS is the extraction of pores from fingerprint images. Existing methods extract pores by using skeleton-tracking-based (Roddy and Stosz 1997; Stosz and Alyea 1994; Kryszczuk et al. 2004a, b) or filtering-based approaches (Jain et al. 2006, 2007a, b; Ray et al. 2005; Parsons et al. 2008). The skeleton-tracking-based approaches are quite time-consuming and work well only with very high quality fingerprint images (Jain et al. 2006, 2007a, b). The filtering-based approaches are more efficient and more robust. They use static isotropic pore models to detect pores. As we will see later, however, the pores on real fingerprint images appear anisotropic and vary greatly in scale from fingerprint to fingerprint and from region to region.

In this chapter we will first propose a novel dynamic anisotropic pore model which describes the pores more flexibly and accurately than the previous models. With the proposed pore model, we will then develop an adaptive pore extraction method to accurately and robustly extract pores. To evaluate the proposed method, we established two sets of high resolution fingerprint images and conducted extensive experiments on them. The results demonstrated that the proposed pore model and pore extraction method can detect pores more accurately and robustly than the previous methods, and the extracted pore features can consequently improve the recognition accuracy of pore based AFRS.

3.2 Review of Existing Pore Extraction Methods

Existing pore extraction methods can be classified into two categories, skeleton-tracking-based methods and filtering-based methods. All earlier works (Roddy and Stosz 1997; Stosz and Alyea 1994; Kryszczuk et al. 2004a, b) are skeleton-tracking-based methods. They first binarize and skeletonize the fingerprint image and then track the fingerprint skeletons. A pore is detected when certain criteria are met during the tracking. As pointed out in (Jain et al. 2006, 2007a, b), however, skeletonization is computationally expensive and very sensitive to noise and it works well only on very high resolution fingerprint images of high quality. For example, the fingerprint images used in (Roddy and Stosz 1997; Stosz and Alyea 1994; Kryszczuk et al. 2004a, b) are all at least 2000 dpi. Recently proposed approaches are filtering-based methods that detect pores by using pore models to filter fingerprint images. Figure 3.1 shows three typical isotropic pore models: Ray's model (Ray et al. 2005), Jain's model (Jain et al. 2006, 2007a, b), and the DoG (difference of Gaussian) model (Parsons et al. 2008).

Ray et al. (2005) proposed an approach to extracting pores from fingerprint images based on the pore model in Fig. 3.1, which is a modified 2-dimensional Gaussian function. They first calculated an error map for the fingerprint image, with each entry in this map being the sum of the squared errors between the pore model and the local area surrounding the pixel. The error map is then binarized such that only areas of high pore probability (i.e. low error) are retained. In these areas, the pores are detected as the local minima in a $[(2r_m)] \times (2r_m)$ neighborhood. In Ray et al. (2005), the authors used unitary parameters r (the variance of the Gaussian) and r_m to detect pores. However, the pore scales and ridge/valley widths could vary greatly from one fingerprint to another fingerprint or from one region to another region in the same fingerprint (referring to Fig. 3.2 for examples). Moreover, Ray's pore model is isotropic, yet as we can see from Fig. 3.2 that the appearance of open pores on real fingerprint images is not isotropic.

More recently, Jain et al. (2006, 2007a, b) proposed to use the Mexican hat wavelet transform to extract pores based on the observation that pore regions typically have a high negative frequency response as intensity values change abruptly from bright to dark at the pores. The Mexican hat wavelet actually serves as the pore model, and its scale parameter is experimentally set for specific datasets. Figure 3.1b shows the Mexican hat wavelet. Obviously, it is also isotropic. This

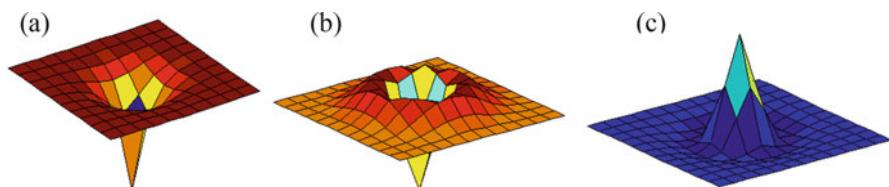


Fig. 3.1 Three typical pore models: (a) Ray's model (Ray et al. 2005), (b) Jain's model (Jain et al. 2006, 2007a, b), and (c) the DoG model (Parsons et al. 2008), which are all isotropic

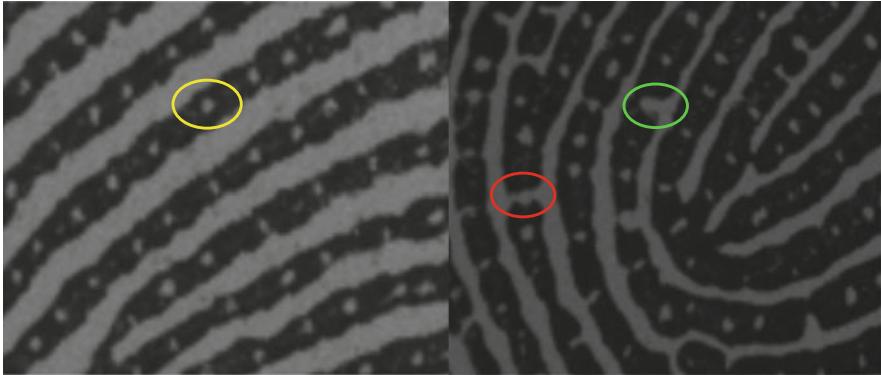


Fig. 3.2 Two fingerprint images with very different ridge and valley widths. A closed pore is marked on the *left* image and two open pores are marked on the *right* image

pore model is also limited in that it cannot adapt itself to different fingerprints or different regions on a fingerprint.

Another pore extraction method was proposed by Parsons et al. (2008). Its basic idea is to use a band-pass filter to detect circle-like features. In other words, the method assumes that pores appear as circular objects on fingerprint images, and the pore is thus modeled by the DoG filter. Figure 3.1c shows this pore model. In Parsons et al. (2008), the authors did not consider the variation of pore scales in fingerprint images but simply used a unitary scale in their model. To deal with the limitations caused by unitary scale, we have recently proposed in Zhao et al. (2010) an adaptive DoG-based pore extraction method. It divides a fingerprint image into blocks and defines for each block a DoG filter according to the local ridge period on the block. One limitation of the DoG-based methods is that the pore models are isotropic. The underlying assumption that pores are circular features does not hold well on real fingerprint images. In this chapter, we will propose another novel pore model and extraction method, which can well solve the problems with existing pore models and extraction methods. Next, we introduce the new pore model first.

3.3 Dynamic Anisotropic Pore Model (DAPM)

Sweat pores reside on finger ridges and may be either closed or open (Ashbaugh 1999). As can be seen in Fig. 3.2, a closed pore looks like an isolated bright blob on the dark ridge, whereas an open pore, which is perspiring, is connected with its neighboring bright valleys. To investigate the spatial appearances of pores on fingerprint images, we manually marked and cropped hundreds of pores on many fingerprint images, including both open and closed pores. We generalized three types of representative pore structures, which are illustrated in Fig. 3.3. It can be seen that the two open pores (b) and (c) are not isotropic. Along the ridge direction,

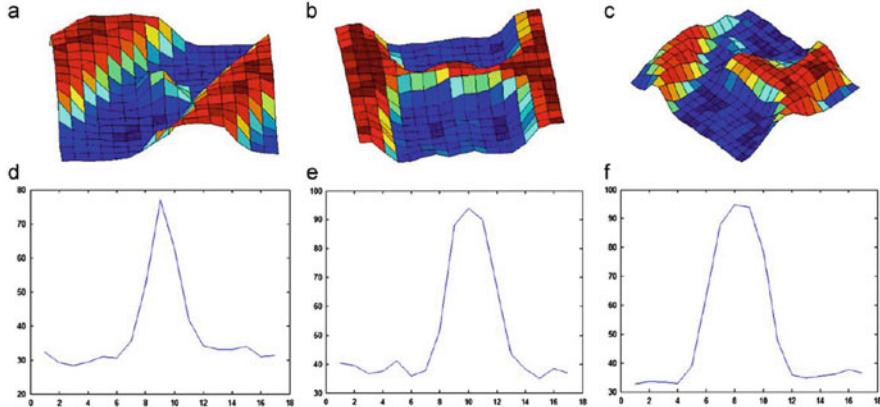


Fig. 3.3 The appearance of three typical pores on real fingerprint images. (a) is a closed pore, and (b) and (c) are open pores. (d)–(f) are the corresponding intensity profiles across them along the ridge orientation. All the profiles are Gaussian-shaped

all the three types of pores appear with Gaussian-shaped profiles. Furthermore, the width of Gaussian profile will vary from one pore to another.

These observations clearly show that the previously proposed pore models (refer to Sect. 3.2) are not accurate enough to model the various pores because they are isotropic and static (i.e. using a unitary scale). In order to represent the pores more accurately, we propose here a new pore model which has two parameters to adjust scale and orientation. When applying this model to a real pore, these two parameters are adaptively determined according to the local ridge features (i.e. ridge orientation and frequency). Therefore, we name the proposed model as the dynamic anisotropic pore model (DAPM), which is defined as follows:

$$\begin{cases} P_0(i,j) = e^{-\frac{j^2}{2\sigma^2}} \cdot \cos\left(\frac{\pi}{3\sigma}i\right) \\ -3\sigma \leq i, j \leq 3\sigma \end{cases} \quad (3.1)$$

$$\begin{cases} P_\theta(i,j) = \text{Rot}(P_0, \theta) = e^{-\frac{\hat{j}^2}{2\sigma^2}} \cdot \cos\left(\frac{\pi}{3\sigma}\hat{i}\right) \\ \hat{i} = i \cos(\theta) - j \sin(\theta), \hat{j} = i \sin(\theta) + j \cos(\theta) \\ -3\sigma \leq i, j \leq 3\sigma \end{cases} \quad (3.2)$$

Equation (3.1) is the reference model (i.e. the zero-degree model) and Eq. (3.2) is the rotated model. Here, σ is the scale parameter which is used to control the pore size. It can be determined by the local ridge frequency. θ is the orientation parameter which is used to control the direction of the pore model. It can be estimated by the local ridge orientation. Figure 3.4 shows some example instances of the proposed DAPM. With the proposed DAPM, next we present an adaptive pore extraction method in Sect. 3.4.

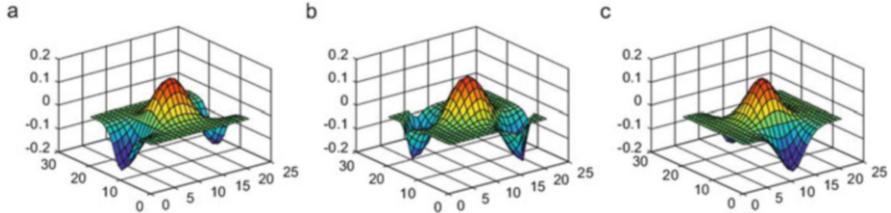


Fig. 3.4 Example instances of the dynamic anisotropic pore model. (a) $\theta = 0$, (b) $\theta = 45$, and (c) $\theta = 90$

3.4 Adaptive Pore Extraction

Pore extraction is essentially a problem of object detection. Generally, given a model of an object, we can detect the object by using the model as a matched filter. When convoluting an image with a matched filter describing the desired object, strong responses will be obtained at the locations of the object on the image. The techniques of matched filters have been successfully used in many applications, for example, vessel detection on retinal images (Sofka and Stewart 2006). In this section, we will first estimate the parameters in the DAPM to instantiate the pore model, and then discuss the important implementation issues in using the instantiated pore models to extract pores. Finally, the adaptive pore extraction algorithm will be presented.

3.4.1 DAPM Parameter Estimation

The matched filters for pore extraction can be generated by instantiating the pore models. In order to instantiate the DAPM in Eqs. (3.1) and (3.2), it is necessary to initialize two parameters, orientation θ and scale σ . As for the orientation parameter θ , an intuitive way is to set it as the local fingerprint ridge orientation. To estimate the ridge orientation field on the fingerprint, we first smooth the fingerprint image by using a smoothing kernel and then calculate the gradients along the x and y directions by using some derivative operator (e.g. the Sobel operator). Let $G_x(i,j)$ and $G_y(i,j)$ be the gradients at the pixel (i,j) , and the squared gradients be $G_{xx}(i,j) = G_x(i,j) \times G_x(i,j)$, $G_{xy}(i,j) = G_x(i,j) \times G_y(i,j)$, and $G_{yy}(i,j) = G_y(i,j) \times G_y(i,j)$. The squared gradients are then smoothed by using a Gaussian kernel, resulting in \bar{G}_{xx} , \bar{G}_{xy} , and \bar{G}_{yy} . The ridge orientation at (i,j) is estimated by

$$O(i,j) = \frac{\pi}{2} + \frac{1}{2} \cdot \arctan\left(\frac{\bar{G}_{xx}(i,j) - \bar{G}_{yy}(i,j)}{2 \cdot \bar{G}_{xy}(i,j)}\right) \quad (3.3)$$

which is in the range of $[0, \pi]$. For more details on fingerprint ridge orientation field estimation, please refer to Bazen and Gerez (2002).

With regard to the scale parameter σ , if we can estimate the range of pore scales, we can then use a bank of multi-scale matched filters to detect the pores; however, this is very time-consuming. Therefore, we estimate and use the maximum valid pore scale when designing the matched filters in this chapter. As shown in Sect. 3.2, the pores are located on ridges. Consequently, the pore scales should be restricted by the ridge widths. This motivates us to associate the maximum pore scale with the local fingerprint ridge period by a ratio, i.e. $\sigma = \tau/k$, where τ is the local ridge period (or the reciprocal of local ridge frequency) and k is a positive constant. In this chapter, we empirically set $k = 12$. The local ridge frequency is estimated in a local window by using the projection-based method in Hong et al. (1998).

3.4.2 Implementation Issues

With the estimated parameters θ and σ in Sect. 3.4.1, an adaptive pore model can be instantiated for each pixel and then we can apply it as a matched filter to extracting pores from the fingerprint image. However, there will be two problems if directly applying the matched filters in a pixel-wise way. Next, we discuss these issues in detail and present the solutions to practical implementation.

The first problem is the computational cost. Obviously, it will be very expensive to calculate the DAPM in a pixel-wise way. Noting that in a local region on the fingerprint, the ridges run nearly parallel with each other, and the intervals between them vary slightly, we could therefore calculate a common DAPM in a local region to detect pores. The second problem is that on some parts of a fingerprint image it is difficult to get an accurate estimate of the local ridge orientation and frequency, which is needed in order to initialize an accurate instance of DAPM. For example, on the image shown in Fig. 3.5a, the region highlighted by the red circle is mashed and no dominant orientation can be obtained. The sharp change of ridge orientation at the singular points of a fingerprint will also raise difficulties in estimating the ridge orientation and frequency surrounding the singular points.

To deal with these issues, we propose a block-wise approach to implementing the matched filters for pore extraction. This approach defines three kinds of blocks on fingerprint images: well-defined blocks, ill-posed blocks, and background blocks. Well-defined and ill-posed blocks are both foreground fingerprint regions. On a well-defined block, it is able to directly estimate a dominant ridge orientation and a ridge frequency. On an ill-posed block, there is not a dominant ridge orientation but the ridge frequency can be estimated by interpolation of the frequencies on its neighboring blocks.

The block partition and classification is performed in a hierarchical way. First, a large block size is applied to the image. For each block B , the following structure tensor is calculated

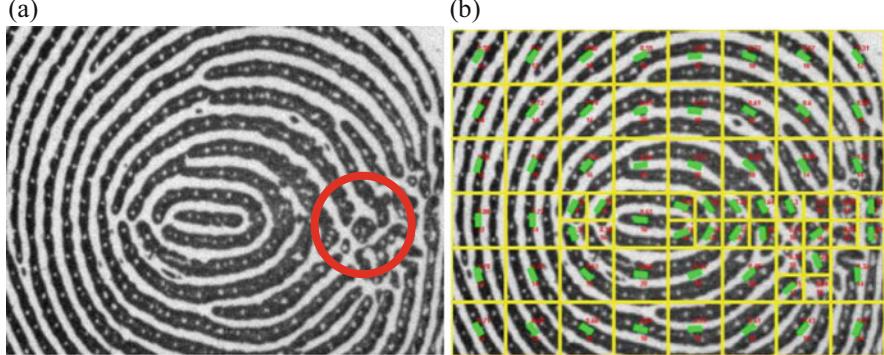


Fig. 3.5 (a) A fingerprint image. The ridge orientation and frequency cannot be accurately estimated on the region marked by the *red circle*. (b) The partition result. The dominant ridge orientations of the well-defined blocks are shown by the *green lines* (Color figure online)

$$J = \frac{1}{N_B} \sum_{i \in B} \nabla B_i \nabla B_i^T = \begin{bmatrix} j_{11} & j_{12} \\ j_{21} & j_{22} \end{bmatrix} \quad (3.4)$$

where N_B denotes the number of pixels in the block, $\nabla B_i = (\partial B_i / \partial x, \partial B_i / \partial y)^T$ is the gradient vector at pixel i , and ‘ T ’ represents the transpose operator. The structure tensor J contains information of ridge orientation in the block and the eigenvalues of J can be used to measure the consistency of ridge orientation. Specifically, we use the orientation certainty (OC) defined as follows (Chen et al. 2005):

$$OC = \frac{(\lambda_1 - \lambda_2)^2}{(\lambda_1 + \lambda_2)^2} = \frac{(j_{11} - j_{22})^2 + 4j_{12}^2}{(j_{11} + j_{22})^2} \quad (3.5)$$

where λ_1 and λ_2 are the two eigenvalues of 2×2 structure tensor J and we assume $\lambda_1 \geq \lambda_2$. This quantity of OC can indicate how strongly the energy is concentrated along the ridge orientation. If there is a dominant ridge orientation, then $\lambda_1 \geq \lambda_2$ and OC will be close to 1. Otherwise, λ_1 and λ_2 will not differ much and consequently OC will be close to 0.

We also calculate a measurement related to the intensity contrast (IC) of the block as follows:

$$IC = std(B) \quad (3.6)$$

where std denotes the standard deviation. The purpose of this is to exclude the background from the foreground fingerprint. The two measurements, OC and IC , are evaluated with pre-specified thresholds. If both of them are above the thresholds, the block is recorded as a well-defined block and will not be further partitioned. Otherwise, the block larger than the minimum size is evenly partitioned into four equal sub-blocks, each of which is further examined. Suppose a block has

reached the minimum size, this block will be marked as a well-defined block if its *OC* and *IC* measures are above the thresholds; it is marked as an ill-posed block if its *OC* measure is less than the threshold but the *IC* measure is above the threshold; otherwise, it is marked as a background block. Figure 3.5b shows the partition result of the image in Fig. 3.5a. The dominant ridge orientations of the well-defined blocks are shown by the green lines.

After partitioning the fingerprint image into the three kinds of blocks, the pores can be extracted from each of the foreground (well-defined or ill-posed) blocks. For a well-defined block, the dominant ridge orientation and the mean ridge frequency on it can be calculated directly, and hence the DAPM can be consequently instantiated. For an ill-posed block, there is no dominant ridge orientation but the mean ridge frequency can be calculated by interpolating the mean ridge frequencies of its neighboring blocks. Hence, as a compromise, we apply to the ill-posed blocks the adaptive DoG based pore models (Zhao et al. 2010). Next we discuss on how to calculate the dominant ridge orientation of a well-defined block and the mean ridge frequency on a foreground block.

The dominant orientation of a well-defined block is defined as the average orientation of the ridge orientation field on the block. To average the orientation field of block B , denoted by B_{OF} , we first multiply the orientation angle at each pixel by 2, and then calculate its cosine and sine values. Finally, the dominant orientation of the block is calculated as

$$B_{DO} = \frac{1}{2} \arctan \left(\frac{\text{aver}(\sin(2 \cdot B_{OF}))}{\text{aver}(\cos(2 \cdot B_{OF}))} \right) \quad (3.7)$$

where $\text{aver}(F)$ denotes the average of the elements in F .

For each well-defined block, the average ridge frequency on the block is calculated by using the method in Hong et al. (1998). The ridge frequencies on the ill-posed blocks are estimated by interpolating their surrounding blocks whose ridge frequencies have already been calculated. Specifically, after the ridge frequencies on well-defined blocks have been calculated, we iteratively check the fingerprint image until all the ridge frequencies of the foreground blocks have been calculated. If the ridge frequency of a foreground block has not been calculated, we take the mean of the ridge frequencies of its neighboring blocks as its ridge frequency. Finally, all foreground fingerprint blocks, no matter with or without dominant orientation, are assigned with ridge frequencies.

3.4.3 The Pore Extraction Algorithm

We now summarize the complete adaptive fingerprint pore extraction algorithm. As shown in Fig. 3.6, the proposed pore extraction algorithm consists of five main steps. Take the fingerprint fragment in Fig. 3.7a, which is a part of Fig. 3.5a, as an

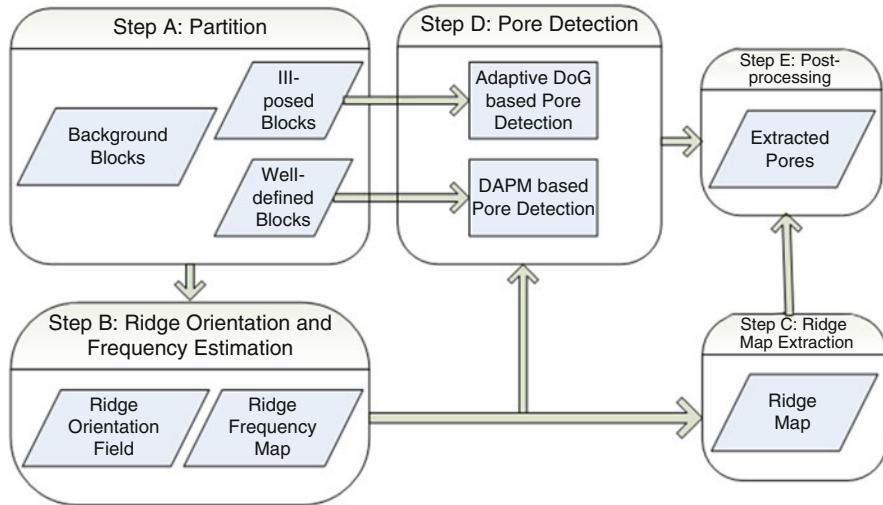


Fig. 3.6 The main steps of the proposed pore extraction method

example. The first step is to partition the fingerprint image into a number of blocks, each being a well-defined block, an ill-defined block or a background block (see Fig. 3.7b). In the second step, the ridge orientation field of the fingerprint image is calculated. Meanwhile, the mean ridge frequencies on all foreground blocks are estimated, which form the ridge frequency map of the fingerprint image (see Fig. 3.7c). It then proceeds to the third step, in which the binary ridge map of the fingerprint image is calculated as follows. Based on the estimated ridge orientation field and ridge frequency map, the fingerprint image is first enhanced by using a bank of Gabor filters (Hong et al. 1998) to enhance the bright valleys and suppress dark ridges. In order to extract the ridges from the fingerprint image, we binarize the enhanced image and calculate its complement, where the ridge pixels have value ‘1’. With this complement image, we can readily obtain the binary ridge map by setting the corresponding ridge pixels in the foreground fingerprint blocks to ‘1’ and the other pixels to ‘0’. This binary ridge map (see Fig. 3.7d) will be used in the post-processing step to remove spurious pores because pores can only locate on ridges.

It is worth mentioning that the first three steps can be performed on a down-sampled small image of the original high resolution fingerprint image because they do not depend on the Level 3 features. In our experiments, we down-sampled the images to half of their original resolution and then carried out steps A, B, and C. Afterwards, the obtained image partition result, ridge orientation field, ridge frequency map, and the ridge map were all up-sampled to the original resolution. They will be used in the subsequent pore detection and post-processing. Working on the down-sampled images can reduce a lot the computational cost. The pore detection and post-processing are performed on the original fingerprint images because the Level 3 pore features can hardly be reliably extracted in the down-sampled low resolution fingerprint images.

In the pore detection step, the foreground fingerprint blocks are processed one by one to detect pores on them. A local instantiation of the DAPM is established for each well-defined block based on the local ridge orientation and frequency on that block, and a local instantiation of the adaptive DoG based pore model is established for each ill-posed block based on the local ridge frequency on the block. Applying the adaptively instantiated pore model to the block as a matched filter will enhance the pores while suppressing valleys and noise. A threshold is then applied to the filtering response to segment out the candidate pores on the block. After processing all the blocks, we obtain a binary image where the candidate pores have the value ‘1’ and other pixels have value ‘0’. This binary image gives the initial pore extraction result (pore map). Figure 3.7e shows an example. We can see that there could have some spurious and false pores in this map.

The last step is to remove the spurious and false pores from the initial pore extraction result. In previous work, most methods remove false pores by applying the restraint that pores should reside only on ridges (Jain et al. 2006, 2007a, b; Ray et al. 2005, Ratha et al. 1996) and that the size of pores should be within a valid range (Jain et al. 2006; 2007a, b; Ratha et al. 1996). Some researchers also propose to refine the pore extraction result based on the intensity contrast (Parsons et al. 2008). In Parsons et al. (2008), a PCA method is applied to a set of extracted

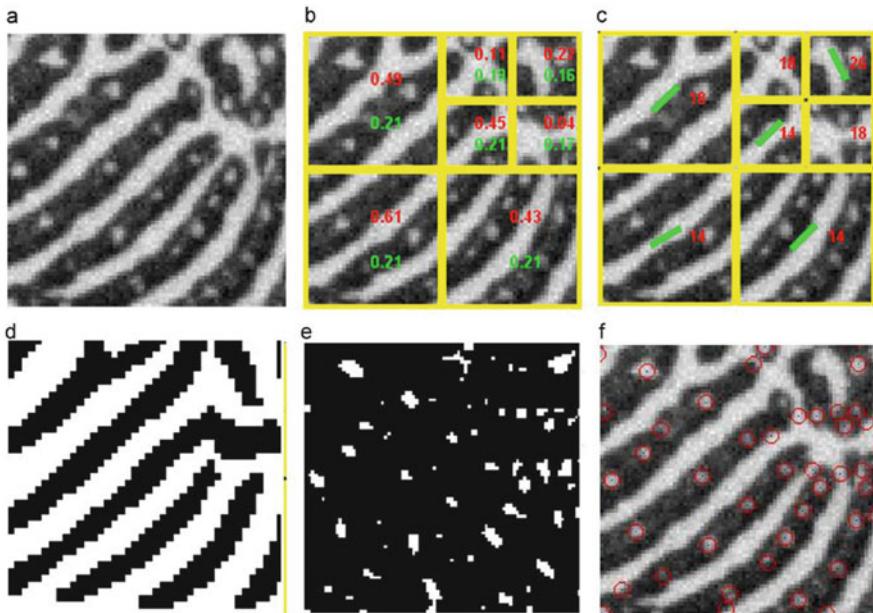


Fig. 3.7 (a) is a fingerprint fragment; (b) shows the blocks on it (the value in red is the orientation certainty and the value in green is the intensity contrast); (c) displays the estimated dominant ridge orientations and periods (the green lines denote the orientations on well-defined blocks, and if there is no orientation shown on a block, the block is an ill-posed block); (d) is the ridge map; (e) is the initial pore map; and (f) shows the final detected pores (marked by circles) (Color figure online)

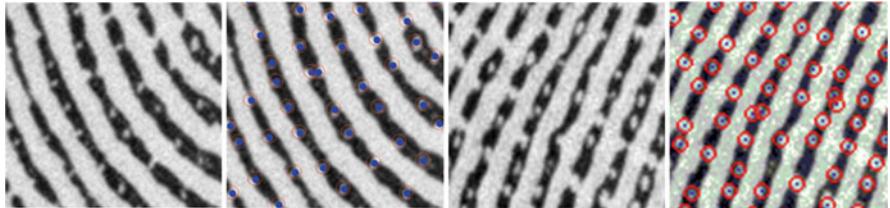


Fig. 3.8 Some example pore extraction results obtained by using the proposed method

putative pores to estimate a model for the gray level distribution over pores. This model is then used to exclude falsely detected pores based on a method of minimizing squared error. This method is however greatly affected by the chosen putative pores.

In this chapter, we take the following steps to post-process the extracted candidate pores. First, we use the binary ridge map as a mask to filter the pore map. In this step, the pixels which are not on ridges are removed. Second, we sort the remaining candidate pore pixels according to their gray level values descendingly and then discard the last 5% pixels because they are more probably caused by noise. Third, we identify all the connected components on the pore map, and each component is taken as a candidate pore. We check the size of each component, i.e. the number of pixels it has. If the size is out of the pre-specified range of valid pore size (from 3 to 30 in our experiments), the candidate pore is removed from the pore map. The final pore map is obtained after the above refinement. We record the extracted pores by recording the coordinates of their mass centers. See Fig. 3.7f for an example. More examples of different fingerprint fragments are given in Fig. 3.8.

3.5 Experiments and Performance Evaluation

To evaluate the proposed fingerprint pore extraction method, a high resolution fingerprint image dataset is required. It is well accepted that the fingerprint image resolution should be at least 1000 dpi to reliably capture the Level 3 features such as pores (CDEFFS 2009). Unfortunately, so far there is no such high resolution fingerprint image database free available in the public domain. So we built an optical fingerprint scanner by ourselves, which could collect the fingerprint images at a resolution about 1200 dpi. Figure 3.9 shows the scanner we developed. It uses a CCD camera (Lumenera Camera LU135M) to capture the fingerprint image when the finger touches against the prism of the scanner.

Two databases have been established by using the scanner we developed. The first database (denoted as DBI) is a partial fingerprint image database where the image size is 320 pixels in width and 240 pixels in height. The second database (denoted as DBII) contains full-size fingerprint images which are 640 pixels in width and 480 pixels in height. Both databases have 1480 fingerprint images taken

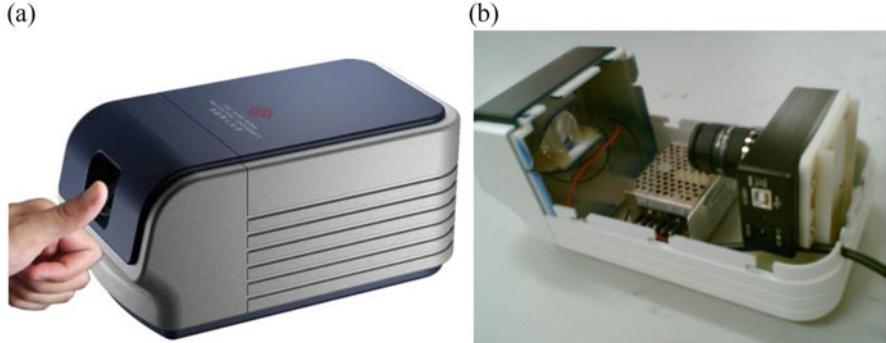


Fig. 3.9 (a) The high-resolution fingerprint scanner we developed and (b) its inner structure

from 148 fingers with each finger having ten samples in two sessions. Five images were captured for each finger in each of the two sessions which were about two weeks apart.

Using the established databases, we have conducted extensive experiments to evaluate the proposed pore extraction method in comparison with three state-of-the-art methods [Jain’s method (Jain et al. 2006, 2007a, b), Ray’s method (Ray et al. 2005), and the adaptive DoG based method (Zhao et al. 2010)]. Three types of experiments were conducted. First, we compared the proposed method with its counterparts in terms of pore detection accuracy using a set of fingerprint images chosen from DBI. Second, using the partial fingerprint image database DBI and a minutia-pore-based fingerprint matcher, we evaluated the fingerprint recognition performance by using the pores extracted by the proposed method and the other three methods. Third, we evaluated the fingerprint recognition performance of the four methods on the full-size fingerprint image database DBII. In the following, we present the experiments in detail.

3.5.1 Pore Detection Accuracy

We first assess the pore detection accuracy of the proposed method. For this purpose, we chose a set of 24 fingerprint images from DBI. The chosen images have relatively good quality so that the pores on them can be easily marked. Figure 3.10 shows two example fingerprint images. We manually marked the pores on these fingerprint images as the ground truth for our experiments. We then used our proposed method, Jain’s method, Ray’s method, and the adaptive DoG based method to extract pores on them. Figure 3.11 shows the pore extraction results of the four methods on the fingerprint image in Fig. 3.10b. On this fingerprint fragment, the ridges on the left hand side are thinner than those on the right hand side, and both open and closed pores can be observed. From Fig. 3.11a, b, we can

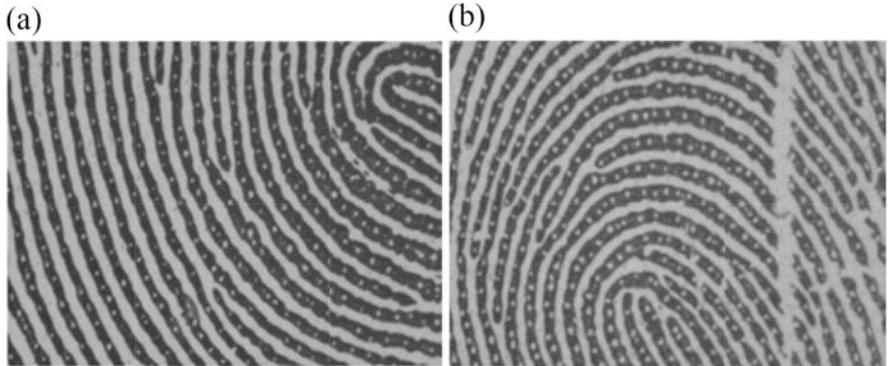


Fig. 3.10 Two example fingerprint images used in the experiments

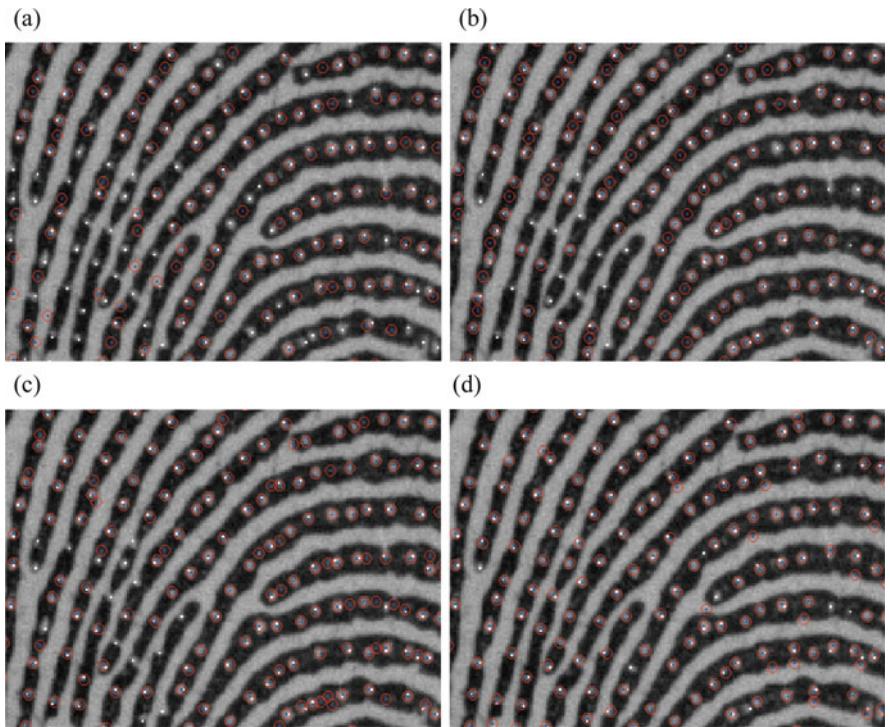


Fig. 3.11 Example pore extraction results of (a) Ray's method, (b) Jain's method, (c) the adaptive DoG based method, and (d) the proposed method. The pores are manually marked by *bright dots* and the detected pores are marked by *red circles* (Color figure online)

see that due to the unitary scale they use, Ray's method and Jain's method cannot work consistently well on the left and the right parts of the fingerprint image because of the varying ridge widths and pore sizes. In addition, all the three comparison methods miss many open pores because their isotropic pore models

Table 3.1 The average pore detection accuracy (%) and the standard deviation of the four methods on the 24 fingerprint images

	Ray's method	Jain's method	Adaptive DoG based method	The proposed method
R_T	60.6 (11.9)	75.9 (7.5)	80.8 (6.5)	84.8 (4.5)
R_F	30.5 (10.9)	23.0 (8.2)	22.2 (9.0)	17.6 (6.3)

cannot accurately handle open pores. In contrast, the proposed method successfully detects most of the pores on both the left and the right parts of the fingerprint image no matter they are open or closed. This demonstrates that the proposed DAPM model can better adapt to varying ridge widths and pore sizes, and can better cope with both closed and open pores.

In addition to the visual evaluation of the pore detection results, we calculated the average detection accuracy on the 24 fingerprint images by using two metrics: R_T (true detection rate) and R_F (false detection rate). R_T is defined as the ratio of the number of detected true pores to the number of all true pores, while R_F is defined as the ratio of the number of falsely detected pores to the total number of detected pores. A good pore extraction algorithm should have a high R_T and a low R_F simultaneously. Table 3.1 lists the average detection accuracy and the standard deviation of detection accuracy of the four methods. According to the average detection accuracy listed in Table 3.1, the proposed method achieves not only the highest true detection rate but also the lowest false detection rate. With regard to the standard deviation, as shown in Table 3.1, the proposed method again achieves the smallest deviation over the whole image set for both true detection rate and false detection rate. As for the other three methods, none beats its counterparts in all cases. From these results, we can see that the proposed method can detect pores on fingerprint images more accurately and more robustly.

3.5.2 Pore Based Partial-Fingerprint Recognition

Since the purpose of pore extraction is to introduce new features for fingerprint recognition, it is necessary to test how the pores extracted by the methods will contribute to a fingerprint recognition system. According to (Kryszczuk et al. 2004a, b; Zhao et al. 2009), the fingerprint recognition benefits more from the pores when the used fingerprint images cover small fingerprint area. Therefore, in order to emphasize the contribution of pores, we evaluated in this sub-section the improvement of fingerprint recognition accuracy made by the extracted pores based on the partial fingerprint image database DBI.

We implemented an AFRS like the one in Zhao et al. (2009) which is based on minutiae and pores. The block diagram of the AFRS is shown in Fig. 3.12. It consists of five main modules, i.e. minutia extraction, pore extraction, minutia matching, pore matching, and match score fusion. We use the methods in Feng (2008) for minutia extraction and matching modules. The pore matching is

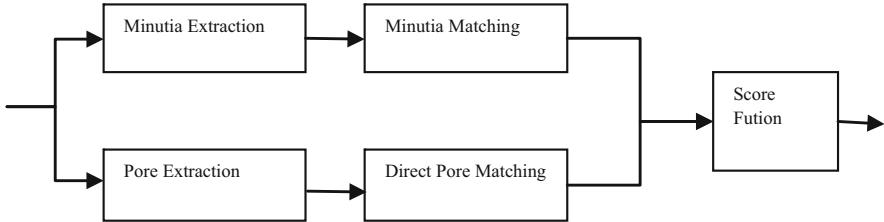


Fig. 3.12 Block diagram of the AFRS used in partial fingerprint recognition experiments

accomplished by using the direct pore matching method in Zhao et al. (2009). It firstly establishes initial correspondences between the pores on two fingerprints based on their local features, and then uses the RANSAC (Random Sample Consensus) algorithm (Hartley and Zisserman 2003) to refine the pore correspondences, and finally calculates a pore-based similarity score between the two fingerprints based on the number of corresponding pores. The pore matching is independent of minutia matching in this method. This method is very suitable for small partial fingerprint recognition where the minutia matching results are often unreliable due to the limited number of minutiae on the small fingerprint fragments (Zhao et al. 2009). The pore match score and the minutia match score are finally fused by using a simple weighted summation scheme to give the final match score between two fingerprint images (before fusion, both match scores are normalized to the range between 0 and 1) as follows

$$MS = \omega \cdot MS_{\text{minu}} + (1 - \omega) \cdot MS_{\text{pore}} \quad (3.8)$$

where ω is the weight of minutiae with respect to pores.

By using database DBI and the above described AFRS, we evaluated the fingerprint recognition performance of the four pore extraction methods. Considering the expensive computational cost, the following matches were carried out: (1) Genuine matches: Each of the fingerprint images in the second session was matched with all the fingerprint images in the first session, leading to 3700 genuine matches, and (2) Imposter matches: the first fingerprint image of each finger in the second session was matched with the first fingerprint image of all the other fingers in the first session, resulting in 21,756 imposter matches. Figure 3.13 shows the equal error rates (EER) obtained by the four methods on DBI under different weights. By using only minutiae, the EER is 17.67%. The EERs when using only pores (i.e. $\omega = 0$) are respectively 21.96% by Ray's method, 21.53% by Jain's method, 22.99% by the adaptive DoG based method, and 20.49% by the proposed method. By fusing minutiae and pores, the best results are 12.41% ($\omega = 0.9$), 12.4% ($\omega = 0.8$), 14.18% ($\omega = 0.9$), and 11.51% ($\omega = 0.8$) by the four methods respectively. Figure 3.14 shows their receiver operating characteristics (ROC) curves when the best EERs are obtained. It is seen that the proposed method leads to the best recognition results. The improvement of recognition accuracy made by fusing the pore features over using only minutia features are 29.77%, 29.82%, 19.75% and 34.86% respectively by the four methods.

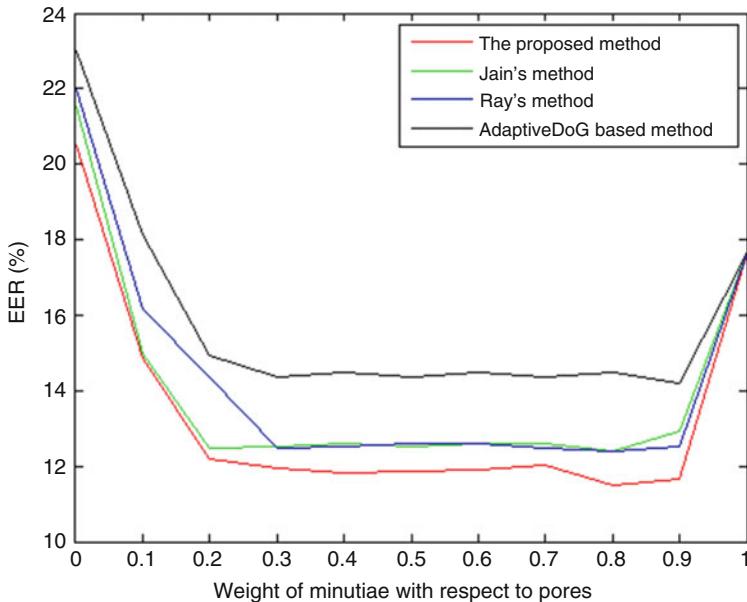


Fig. 3.13 The EERs of the four methods on DBI with different fusing weights

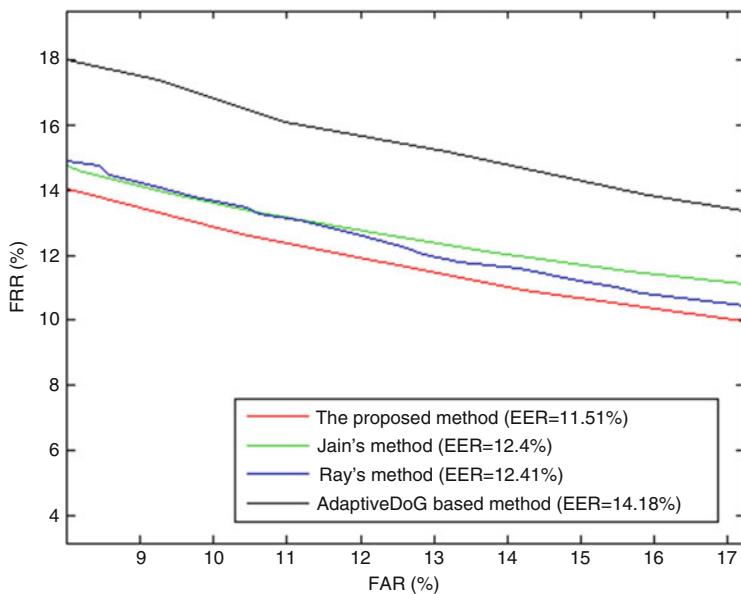


Fig. 3.14 The ROC curves of the four methods on DBI when the lowest EERs are obtained

3.5.3 Pore Based Full-Size Fingerprint Recognition

The experiments in this sub-section were to evaluate the contribution of the extracted pores to full-size fingerprint recognition. We compared the four pore extraction methods by using a different AFRS in Jain et al. (2006), which is appropriate for full-size fingerprint images, and using the full-size fingerprint image database DBII. Figure 3.15 shows the block diagram of the AFRS. We used the same minutia extraction and matching modules and the same score fusion module as in the last sub-section, but implemented the pore matching module by using the ICP (iterative closest point) based method as in (Jain et al. 2006, 2007a, b). This is because on fingerprint images covering large fingerprint area, there are sufficient minutiae to provide reliable minutiae match results. We can thus compare the pores locally in the neighborhoods of matched minutiae. In this way, the pores can be matched much more efficiently. Specifically, after matching the minutiae on two fingerprints, the pores lying in the neighborhoods of each pair of matched minutiae are matched by using the ICP algorithm (Jain et al. 2006, 2007a, b), resulting in N match scores (N is the number of pairs of matched minutiae), which are defined as the summation of two terms: the mean distance between all matched pores and the percentage of unmatched pores. The pore match score between the two fingerprints is finally defined as the average of the first three smallest match scores.

By using the above AFRS, we matched pair-wise all the fingerprint images in DBII (avoiding symmetric matches), generating 6660 genuine match scores and 1,087,800 imposter match scores. Figure 3.16 presents the EERs obtained by the four methods on DBII. Because the fingerprint images in DBII are full-size fingerprint images and have more minutiae, it can be seen that the EER of using only minutiae is 0.61%, which is much better than that obtained on DBI (17.67%, referring to Sect. 3.5.2). When only using pores, the EER of Ray’s method is 9.45%, Jain’s method 8.82%, the adaptive DoG based method 10.85%, and the proposed method 7.81%. The best results of these methods after fusion with minutia match scores are 0.59% ($\omega = 0.9$), 0.6% ($\omega = 0.9$), 0.56% ($\omega = 0.8$), and 0.53% ($\omega = 0.7$). Figure 3.17 shows their corresponding ROC curves when the best results are obtained. The proposed method improves on the best EERs of Ray’s method, Jain’s method, and the adaptive DoG based method by 10.17%, 11.67%, and 5.36% respectively.

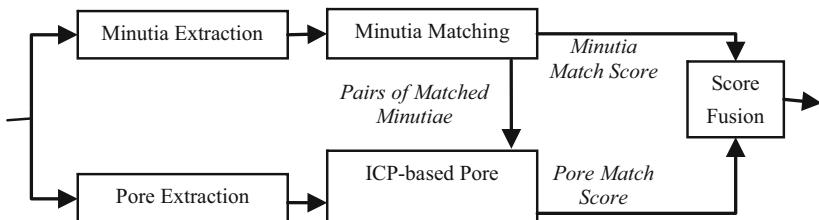


Fig. 3.15 Block diagram of the AFRS used in full-size fingerprint recognition experiments

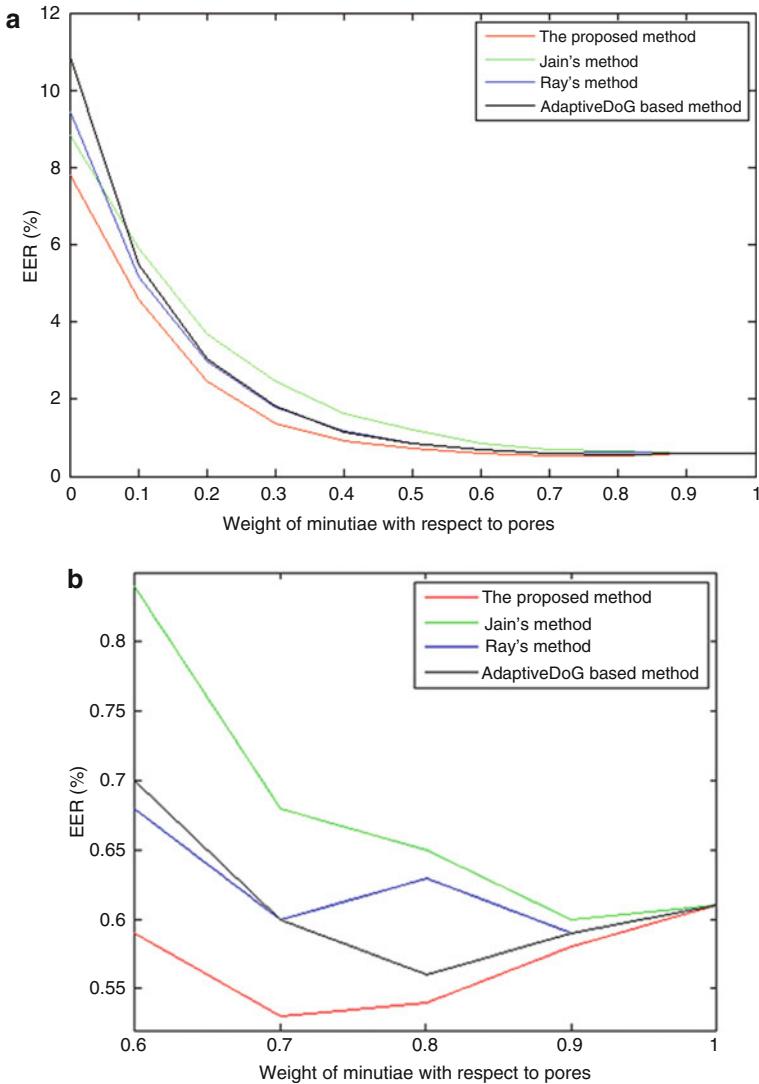


Fig. 3.16 The EERs of the four methods on DBII when different weights are used. (b) is the zoom-in of (a) when the weight is from 0.6 to 1

3.5.4 Computational Complexity Analysis

Before closing this section, we would like to briefly analyse the computational complexity of the methods. As shown in Fig. 3.6, the proposed pore extraction method has five main steps: (A) partition, (B) ridge orientation and frequency estimation, (C) ridge map extraction, (D) pore detection, and (E) post-processing.

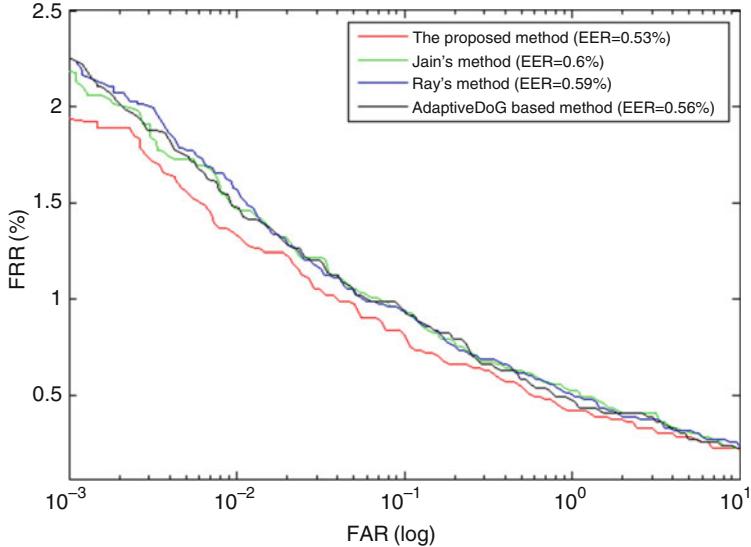


Fig. 3.17 The ROC curves of the four methods on DBII when the best results are obtained

Among these steps, the steps (B) and (C) are common to most automatic fingerprint recognition systems. The step (A) needs to calculate the two measurements, OC and IC , for each of the blocks, which can be done very efficiently. On a PC with 2.13 GHz Intel(R) Core(TM) 2 6400 CPU and RAM of 2GB, the Matlab implementation of the method took about 0.05 ms in average to conduct the step (A) for one partial fingerprint image used in the experiments (note that down-sampling was applied in the experiments). In the pore detection step (D), the main operation is the convolution of pore models with the fingerprint image. This is essentially common to all filtering-based pore extraction methods including the three counterpart methods considered in the experiments here. However, because Jain's and Ray's methods both apply a single pore filter to the whole fingerprint image, they are more efficient than the Adaptive DoG-based method and the method proposed here which apply different pore filters to different blocks. Specifically, Jain's and Ray's methods took less than 0.1 ms to detect the pores on one partial fingerprint image, whereas the other two methods used about 0.5 ms. The last post-processing step (E) is also common to all pore extraction methods. Using our Matlab implementation, it took about 0.4 ms to carry out all the post-processing operations defined in Sect. 3.4.3. From the above analysis, we can see that the proposed pore extraction method is a little more complex than Jain's and Ray's methods due to the more elaborated pore models on which it is based. However, considering the gain of accuracy by the proposed method, the increased computational complexity is deserved. More importantly, its computational cost is still acceptable (with the Matlab implementation, it took about 1–2 s to extract the pores on one partial fingerprint image), and we expect that the computational cost can be much reduced by using languages like C/C++ and after optimization.

3.6 Summary

This chapter presented a dynamic anisotropic pore model (DAPM). It differs from previous pore models in that it is anisotropic and dynamic so that it can more accurately represent pores by using orientation and scale parameters. A novel adaptive pore extraction method was then developed based on the DAPM. The fingerprint image was partitioned into well-defined, ill-posed and background blocks according to the orientation certainty and intensity contrast on the blocks. For each well-defined block, the ridge orientation and frequency were estimated directly, while for each ill-posed block, the ridge frequency was estimated by interpolating the ridge frequencies on its neighboring blocks. The local instances of the DAPM were then instantiated for the well-defined blocks based on the estimated orientations and frequencies of them. The instantiated pore models were taken as the matched filters and applied to the blocks to detect the pores thereon. In the post-processing step, some constraints were used to remove possible spurious and false pores in the detection result.

We have established two high-resolution fingerprint databases to evaluate the proposed method in comparison with three state-of-the-art pore extraction methods. The proposed method obtained a true detection rate as 84.8% and a false detection rate as 17.6%, and corresponding deviations 4.5% and 6.3%. In contrast, the best true detection rate and false detection rate of existing methods were respectively 80.8% and 22.2%, and their deviations were 6.5% and 8.2%. These experimental results demonstrated that the proposed DAPM is more accurate and robust than the previous models. We consequently evaluated the pore based fingerprint recognition systems. The experiments show that by using pores as additional features to the conventional minutia features, higher recognition accuracy can be obtained. Since the proposed DAPM achieves higher pore detection accuracy, it obtains the best fingerprint recognition accuracy among the state-of-the-art pore extraction methods on both partial and full-size fingerprint image databases. On the partial fingerprint image database, by fusing minutia and pore match scores it improves the recognition accuracy of using only minutiae by 34.86%, whereas the improvements made by the other methods are all below 30%. On the full-size fingerprint image database, again, the proposed method achieves the best EER at 0.53%, which improves the other methods by over 5% to about 12%.

In this chapter, we experimentally set the scale parameter of the proposed pore model according to the local ridge periods by a ratio. In our future work, we are going to determine it in a more adaptive way based on the automatic scale selection theory (Sofka and Stewart 2006; Lindeberg 1998). Currently, the proposed pore extraction method is mainly for live-scan fingerprints. In the future, we are also going to extend the proposed method to more challenging latent fingerprints.

References

- Ashbaugh D (1999) Quantitative–qualitative friction ridge analysis: an introduction to basic and advanced ridgeology. CRC Press, Boca Raton
- Bazen A, Gerez S (2002) Systematic methods for the computation of the directional fields and singular points of fingerprints. *IEEE Trans Pattern Anal Mach Intell* 24(7):905–919
- Bindra B, Jasuja O, Singla A (2000) Poroscopy: a method of personal identification revisited. *Internet J Forensic Med Toxicol* 1(1)
- CDEFFS (2009) Data format for the interchange of extended fingerprint and palmprint features. Working draft version 0.4, June. <http://fingerprint.nist.gov/standard/cdeffs/index.html>
- Chen Y, Dass S, Jain A (2005) Fingerprint quality indices for predicting authentication performance. In: Proceedings of AVBPA2005, pp 160–170
- Feng J (2008) Combining minutiae descriptors for fingerprint matching. *Pattern Recogn* 41:342–352
- Hartley R, Zisserman A (2003) Multiple view geometry in computer vision, 2nd edn. Cambridge University, Cambridge
- He Y, Tian J, Li L, Chen H, Yang X (2006) Fingerprint matching based on global comprehensive similarity. *IEEE Trans Pattern Anal Mach Intell* 28(6):850–862
- Hong L, Wan Y, Jain A (1998) Fingerprint image enhancement: algorithms and performance evaluation. *IEEE Trans Pattern Anal Mach Intell* 20(8):777–789
- Jain A, Hong L, Bolle R (1997) On-line fingerprint verification. *IEEE Trans Pattern Anal Mach Intell* 19(4):302–314
- Jain A, Chen Y, Demirkus M (2006) Pores and ridges: fingerprint matching using level 3 features. In: Proceedings of ICPR06, vol 4, pp 477–480
- Jain A, Chen Y, Demirkus M (2007a) Pores and ridges: fingerprint matching using level 3 features. *IEEE Trans Pattern Anal Mach Intell* 29(1):15–27
- Jain A, Flynn P, Ross A (2007b) Handbook of biometrics. Springer, New York
- Kryszczuk K, Drygajlo A, Morier P (2004a) Extraction of level 2 and level 3 features for fragmentary fingerprints. In: Proceedings of the 2nd COST Action 275 workshop, pp 83–88
- Kryszczuk K, Morier P, Drygajlo A (2004b) Study of the distinctiveness of level 2 and level 3 features in fragmentary fingerprint comparison. In: BioAW2004, LNCS 3087, pp 124–133
- Lindeberg T (1998) Edge detection and ridge detection with automatic scale selection. *Int J Comput Vis* 30:117–156
- Maltoni D, Maio D, Jain A, Prabhakar S (2003) Handbook of fingerprint recognition. Springer, New York
- Parsons N, Smith JQ, Thonnes E, Wang L, Wilson RG (2008) Rotationally invariant statistics for examining the evidence from the pores in fingerprints. *Law Probab Risk* 7:1–14
- Ratha N, Bolle R (2004) Automatic fingerprint recognition systems. Springer, New York
- Ratha N, Karu K, Chen S, Jain A (1996) A real-time matching system for large fingerprint databases. *IEEE Trans Pattern Anal Mach Intell* 18(8):799–813
- Ray M, Meenen P, Adhami R (2005) A novel approach to fingerprint pore extraction. In: Proceedings of the 37th south-eastern symposium on system theory, pp 282–286
- Roddy A, Stosz J (1997) Fingerprint features – statistical analysis and system performance estimates. *Proc IEEE* 85(9):1390–1421
- Ross A, Jain A, Reisman J (2003) A hybrid fingerprint matcher. *Pattern Recogn* 36:1661–1673
- Sofka M, Stewart C (2006) Retinal vessel centerline extraction using multiscale matched filters, confidence and edge measures. *IEEE Trans Med Imaging* 25(12):1531–1546
- Stosz J, Alyea L (1994) Automated system for fingerprint authentication using pores and ridge structure. In: Proceedings of SPIE conference on automatic systems for the identification and inspection of humans, San Diego, vol 2277, pp 210–223

- Zhang D (2000) Automated biometrics: technologies and systems. Kluwer and Academic, Dordrecht
- Zhao Q, Zhang L, Zhang D, Luo N (2009) Direct pore matching for fingerprint recognition. In: Proceedings of ICB2009, pp 597–606
- Zhao Q, Zhang D, Zhang L, Luo N (2010) High resolution partial fingerprint alignment using pore-valley descriptors. Pattern Recogn 43(3):1050–1061

Chapter 4

The Resolution for Fingerprint Recognition

Abstract High-resolution automated fingerprint recognition systems (AFRS) offer higher security because they are able to make use of level 3 features, such as pores, that are not available in lower-resolution (<500 dpi) images. One of the main parameters affecting the quality of a digital fingerprint image and issues such as cost, interoperability, and performance of an AFRS is the choice of image resolution. In this chapter, we identify the optimal resolution for an AFRS using the two most representative fingerprint features, minutiae and pores. We first designed a multi-resolution fingerprint acquisition device to collect fingerprint images at multiple resolutions and captured fingerprints at various resolutions but at a fixed image size. We then carried out a theoretical analysis to identify the minimum required resolution for fingerprint recognition using minutiae and pores. After experiments on our collected fingerprint images and applying three requirements for the proportions of minutiae and pores that must be retained in a fingerprint image, we recommend a reference resolution of 800 dpi. Subsequent tests have further confirmed the proposed reference resolution.

Keywords High-resolution AFRS • Selecting resolution criteria • Minutiae • Pores • Fingerprint recognition accuracy

4.1 Introduction

As one of the most popular biometric traits, fingerprints are widely used in personal authentication, particularly with the availability of a variety of fingerprint acquisition devices and the advent of thousands of advanced fingerprint recognition algorithms. Such algorithms make use of distinctive fingerprint features that can usually be classified at three levels of detail (Ashbaugh 1999), as shown in Fig. 4.1 and referred to as level 1, level 2, and level 3. Level-1 features are the macro details of fingerprints, such as singular points and global ridge patterns, e.g., deltas and cores (indicated by red triangles in Fig. 4.1). They are not very distinctive and are thus mainly used for fingerprint classification rather than recognition. The level-2 features (red rectangles) primarily refer to the Galton features or minutiae, namely,

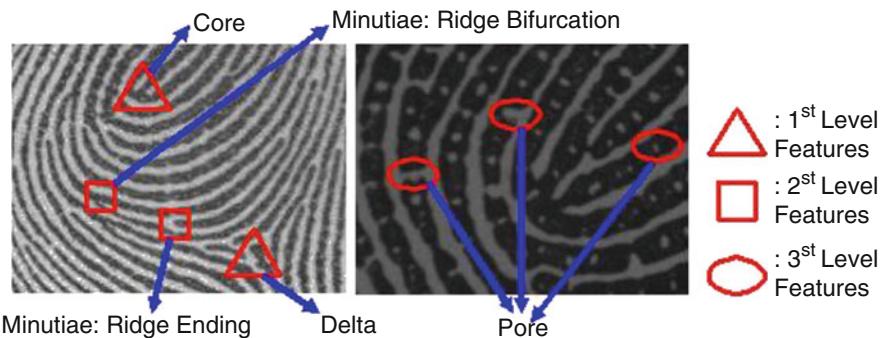


Fig. 4.1 Three levels of fingerprint features

ridge endings and bifurcations. Level-2 features are the most distinctive and stable features, which are used in almost all automated fingerprint recognition systems (AFRSs) (Ashbaugh 1999; Maltoni et al. 2009; Ratha and Bolle 2004) and can reliably be extracted from low-resolution fingerprint images (~ 500 dpi). A resolution of 500 dpi is also the standard fingerprint resolution of the Federal Bureau of Investigation for AFRSs using minutiae (Jain et al. 2007). Level-3 features (red circles) are often defined as the dimensional attributes of the ridges and include sweat pores, ridge contours, and ridge edge features, all of which provide quantitative data supporting more accurate and robust fingerprint recognition. Among these features, pores have most extensively been studied (Jain et al. 2006, 2007; Stosz and Alyea 1994; Roddy and Stosz 1997; Kryszczuk et al. 2004a, b; Zhao et al. 2008, 2009a, b; CDEFFS 2009; International Biometric Group 2008; Ray et al. 2005; Parsons et al. 2008; Chen 2009) and are considered to be reliably available only at a resolution higher than 500 dpi.

Resolution is one of the main parameters affecting the quality of a digital fingerprint image, and so, it has an important role in the design and deployment of AFRSs and impacts both their cost and recognition performance. Despite this, the field of AFRS does not currently have a well-proven reference resolution or standard resolution for high-resolution AFRS that can be used interoperably between different AFRSs. For example, Stosz and Alyea extracted pores at a resolution of approximately 1270 dpi in the vertical direction and 2400 dpi in the horizontal direction (1270×2400 dpi) (Stosz and Alyea 1994). Jain et al. chose a resolution of 1000 dpi based on the 2005 ANSI/NIST fingerprint standard update workshop (Jain et al. 2007). The Committee to Define an Extended Fingerprint Feature Set (CDEFFS 2009) defined level-3 features at a resolution of 1000 dpi. Zhao et al. proposed some pore extraction and matching methods at a resolution of 902×1200 dpi (Zhao et al. 2008, 2009a, b). Finally, the International Biometric Group analyzed level-3 features at a resolution of 2000 dpi (International Biometric Group 2008).

In this chapter, we take steps toward establishing such a reference resolution, assuming a fixed image size and making use of the two most representative fingerprint features, i.e., minutiae and pores, and providing a minimum resolution for pore extraction that is based on anatomical evidence. The use of a fixed image

size is determined by the fact that the quality of a digital fingerprint image is mainly determined by three factors, the resolution, the number of pixels in a fingerprint image, and the measured area of the fingerprint, with it being possible to uniquely determine the value of any one given the other two. In analyzing the influence of resolution on AFRS, it was thus necessary to fix one of the other two parameters. Here, we choose to fix the image size. We conducted experiments on a set of fingerprint images of different resolutions (from 500 to 2000 dpi). By evaluating these resolutions in terms of the number of minutiae and pores, the results have shown that 800 dpi would be a good choice for a reference resolution. Finally, we applied state-of-the-art automated fingerprint recognition algorithms to our collected fingerprint images. Via cross validation experiments, we found the recognition precision under resolution 700–1000 dpi is one order of magnitude higher than that under other considered resolutions. The highest recognition accuracy in different fingerprint groups is almost always obtained under 800 dpi. These results validate our proposed resolution from the point of view of automated fingerprint recognition accuracy.

4.2 Multiresolution Fingerprint Images Collecting

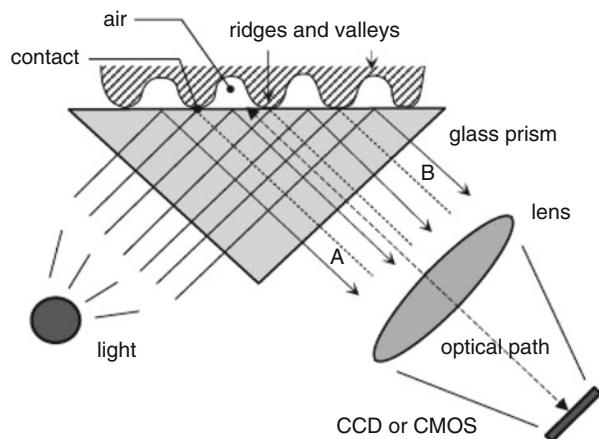
According to our knowledge, there is no data set of multiresolution fingerprint images publicly available. We therefore collected a multiresolution fingerprint image database by using our custom-built fingerprint image acquisition device. In this section, we introduce the fingerprint acquisition device and the established multiresolution fingerprint image database.

4.2.1 Acquisition Device

A multiresolution fingerprint acquisition device (or sensor) must be cost-effective but should particularly be able to acquire fingerprint images at multiple resolutions without any negative impact on the quality of the image (Maltoni et al. 2009). There are generally three kinds of fingerprint sensors: solid state, ultrasound, and optical (Maltoni et al. 2009; Ross and Jain 2004). Solid-state sensors are small and inexpensive but cannot capture high-resolution images (Han et al. 2005). Ultrasound sensors can capture high-resolution images but are usually bulky and expensive (Bicz et al., online). Optical sensors can capture a variety of different image resolutions, varying in a range of sizes and prices. They are easy to implement and have been found to have a high degree of stability and reliability (Xia and O’Gorman 2003). Our system is thus equipped with an optical fingerprint sensor.

While there are also several different ways to implement optical fingerprint sensors, the oldest and most widely used way (Maltoni et al. 2009) and the way we have chosen to implement our sensor is frustrated total internal reflection (FTIR). As shown in Fig. 4.2, an FTIR-based fingerprint sensor consists of a light

Fig. 4.2 Operation of an FTIR-based fingerprint sensor (Maltoni et al. 2009)



source, a glass prism, a lens, and a charge-coupled device (CCD) or complementary metal-oxide-semiconductor camera. When users put their fingers on the surface of the glass prism, ridges absorb light, and so, they appear dark, whereas valleys and the fine details on ridges reflect light and thus appear bright. Different resolutions can be obtained by simply adjusting the distance between the glass prism and the lens and the distance between the lens and the camera.

4.2.2 *Fingerprint Samples*

The most commonly used fingers in fingerprint recognition are the thumb, index finger, and middle finger. These are also the fingers that we use for the images used in our experiments. We collected fingerprint images from both males and females. This is pertinent because male and female fingers are, on the average, different in terms of area and ridge width (or pore size). A total of 25 males and 25 females contributed to our database. Four fingerprint images were captured from each of their six fingers (i.e., thumb, index, and middle fingers on right and left hands) under each of the following resolutions: 500, 600, 700, 800, 900, 1000, 1200, 1600, and 2000. As a result, there are totally 1200 fingerprint images for each of the considered resolutions in the database. Figure 4.3 shows some example fingerprint images collected from a male and a female.

4.2.3 *Implementation of Multiresolution*

Three factors among others can affect the quality of a fingerprint image: its resolution, the measured area of the fingerprint that is captured or sensed, and the



Fig. 4.3 Example 800-dpi fingerprint images in our established database. **(a)** From a female. **(b)** From a male. *From left to right: thumb, index finger, and middle finger*

Table 4.1 The values of H and W at various r when h and w are set as 640 and 480

(h, w) (pixel)	r (dpi)	(H, W) (mm)
(640, 480)	500	(32.5, 24.4)
	600	(27.1, 20.3)
	700	(23.2, 17.4)
	800	(20.3, 15.2)
	900	(18.1, 13.5)
	1000	(16.3, 12.2)
	1200	(13.6, 10.2)
	1600	(10.2, 7.6)
	2000	(8.1, 6.1)

size of the image (the number of pixels). These factors are essentially not independent but related with each other as follows:

$$H = 25.4 \times h/r, \quad W = 25.4 \times w/r \quad (4.1)$$

where r denotes the resolution, h and w denote the height and width of the image, and H and W denote the height and width of the captured area (in millimeters). To generate fingerprint images of different resolutions, one of the other two parameters must be fixed. Table 4.1 shows the values of H and W according to (4.1) at different resolutions when h and w are set as 640 and 480 pixels. It can be seen that, at a fixed image size, the area captured by the image decreases as the resolution increases. Different resolutions can easily be obtained by adjusting the distances between the

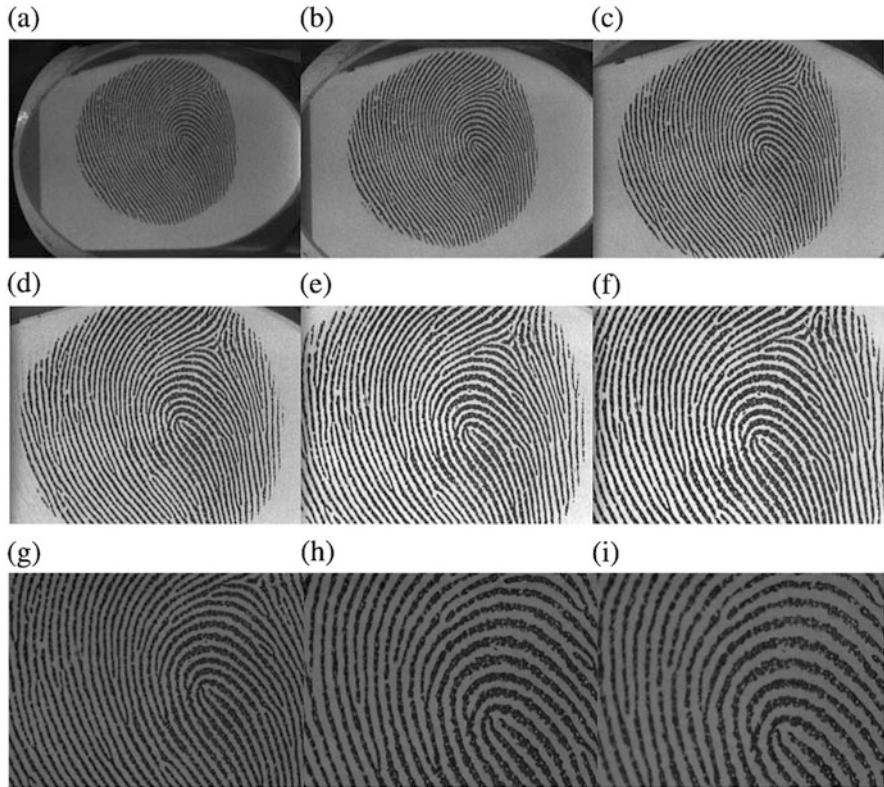


Fig. 4.4 Example fingerprint images at different resolutions when using a fixed image size of 640×480 pixels. (a) 500 dpi. (b) 600 dpi. (c) 700 dpi. (d) 800 dpi. (e) 900 dpi. (f) 1000 dpi. (g) 1200 dpi. (h) 1600 dpi. (i) 2000 dpi

glass prism, the lens, and the CCD. Figure 4.4 shows some example fingerprint images at different resolutions.

It should be noted that the resolution of our device is not identical along the vertical and horizontal directions. This is because the CCD camera has a vertical resolution of 1040 lines and a horizontal resolution of 1394 lines. At 500 dpi, this is not a large difference, and researchers usually ignore it. However, as the resolution increases, the difference between the vertical and horizontal resolutions becomes more obvious. For example, at a vertical resolution of 800 dpi, the horizontal resolution is 1064 dpi, but at a vertical resolution of 1200 dpi, the horizontal resolution is 1596 dpi. The ratio between the horizontal resolution and vertical resolution is equal to the one between the horizontal resolution and vertical resolution of the CCD camera. Thus, given both the vertical or horizontal resolution of fingerprint images and the parameters of CCD camera, we can calculate the resolution of fingerprint images along the other direction. For simplicity, in this chapter, we refer just to the vertical resolution.

4.3 Selecting Resolution Criteria Using Minutiae and Pores

Generally, people may think that higher recognition accuracy can be achieved by increasing the resolution. It is true if the whole fingerprint region is covered. However, in practical AFRSs, the fingerprint image size is usually confined to a relatively small one for the purpose of miniaturization and reducing the computational complexity. Until now, the most widely used image size in most chapters (Roddy and Stosz 1997; Kryszczuk et al. 2004a, b; Jain et al. 2006; Zhao et al. 2008, 2009a, b; CDEFFS 2009; International Biometric Group 2008; Ray et al. 2005; Parsons et al. 2008; Chen 2009) or in most public fingerprint image databases such as the fingerprint verification competition (FVC) databases (e.g., FVC2000, FVC2002, FVC2004, and FVC2006) is 640×480 pixels. With a limited image size, the larger the resolution is, the smaller the captured fingerprint region will be. Although increasing the fingerprint image resolution can provide more fine details on fingerprints for fingerprint matching, it would degrade the fingerprint recognition accuracy if the loss of useful discriminative information (e.g., minutiae) due to decreased fingerprint areas dominates the newly emerged fingerprint details (e.g., pores). For instance, the fingerprint images of a fixed size might cover the whole fingerprint regions at low resolution but capture only few ridges on the fingers at high resolution (see Fig. 4.4). Thus, in this chapter aiming at a balance between various fingerprint features (in particular, minutiae and pores) available on high-resolution fingerprint images, we investigate the fingerprint distinctiveness and recognition accuracy at different resolutions when a fixed image size is adopted. It is also worth mentioning that noise caused by the skin condition or the amount of pressure applied by the finger (Maltoni et al. 2009) also plays an important role in the recognition performance of AFRS due to its influence on the quality of fingerprint images. However, it is a common issue to fingerprint images at all resolutions and is thus out of the scope of the resolution selection work in this chapter.

Since 500-dpi minutiae-based AFRSs were taken as the baseline systems, we chose the fingerprint image size so that as many minutiae as possible are captured by the 500-dpi fingerprint image, or in other words, it can cover the full fingerprint region. By experience, we used an initial image size of 640×480 pixels. As can be seen in Fig. 4.4, this size can actually capture the full fingerprint region at resolutions of 500 and 600 dpi as well. Thus, we cropped the foreground fingerprint regions on these 500-dpi fingerprint images by using rectangles. The maximum width and height of these rectangles observed in the database are 380 and 360 pixels, which were finally taken as the image size for the fingerprint images captured under higher resolutions (i.e., 600–2000 dpi in the experiments in this chapter). Such an image size, which may be comparable with the templates stored in most of existing minutiae-based AFRSs, will be very helpful to realize the interoperability between different AFRSs, which is one motivation of this chapter.

In order to utilize the minutiae and pores on fingerprints, it is necessary that we be able to robustly extract both of these features. Minutiae can be robustly extracted

from images of 500 dpi or above but pore extraction requires higher resolution images according to investigating most of the chapters about fingerprints' studies (Ashbaugh 1999; Maltoni et al. 2009; Ratha and Bolle 2004; Jain et al. 1997, 2006, 2007; Stosz and Alyea 1994; Roddy and Stosz 1997; Kryszczuk et al. 2004a, b; Zhao et al. 2008, 2009a, b; CDEFFS 2009; International Biometric Group 2008; Ray et al. 2005; Parsons et al. 2008; Chen 2009; Ross and Jain 2004). It thus became necessary to figure out what would be the minimum resolution needed to extract pore features. Intuitively, such a figure can be arrived at based on anatomical evidence, i.e. the possible smallest physical size of pores on fingers. We will discuss this in detail in Sect. 4.4.1.

We finally raised three criteria to select the image resolution for high-resolution AFRSs by considering the following. (1) Given a fixed image size, retain as many minutiae as possible while pores begin to be available. (2) The number of pores begins to decrease, and no other useful information but the position of pores will be conveyed when resolution reaches a certain value. (3) Minutiae are more discriminative than pores if the same number of them is considered. Retain as many minutiae as possible while also retaining an acceptable number of pores.

We can better understand the rationale for the criteria by considering the images of an example finger shown in Fig. 4.5, whose image size is 380×360 pixels and resolution increases from 500 to 2000 dpi. The minutiae are the features of interest and are marked with red circles. The availability of pores can also be seen on these images. One may clearly observe the change of available minutiae and pores across these fingerprint images of different resolutions. Next, we introduce the three selection criteria in detail.

Criterion 1 Given a Fixed Image Size, Retain as Many Minutiae as Possible While Pores Begin to Be Available. A Lower Limit Image Resolution Can Be Obtained.

Most minutiae based AFRSs judge whether two fingerprints are from the same finger by counting the number of matched minutiae; basically, the larger the number of minutiae is, the higher the possibility of making correct judgment will be. Thus, we should try to retain as many minutiae as possible. Table 4.2 lists the number of minutiae and pores in an image at different resolutions. As expected, the number of minutiae decreases as resolution increases. On the other hand, the number of pores first increases and then decreases as resolution increases. According to the analysis of Stosz and Alyea (1994), there is a minimum resolution (larger than 500 dpi) for robust pore extraction. As a consequence, criterion 1 is established to determine the lower limit of resolution.

Criterion 2 The Number of Pores Begins to Decrease, and No Other Useful Information but the Position of Pores Will Be Conveyed When Resolution Reaches a Certain Value. An Upper Limit Image Resolution Can Be Obtained.

As can be seen in Fig. 4.5, the size and shape of pores become more visible at higher resolutions. However, according to Ashbaugh (1999) and Roddy and Stosz (1997), usually, only the location of pores is reliable discriminative information for fingerprint recognition; on the contrary, the size and shape of one pore can significantly vary from one impression to another. The two 2000-dpi images in Fig. 4.6 are from the same finger but collected at different times. Clearly, the pores'

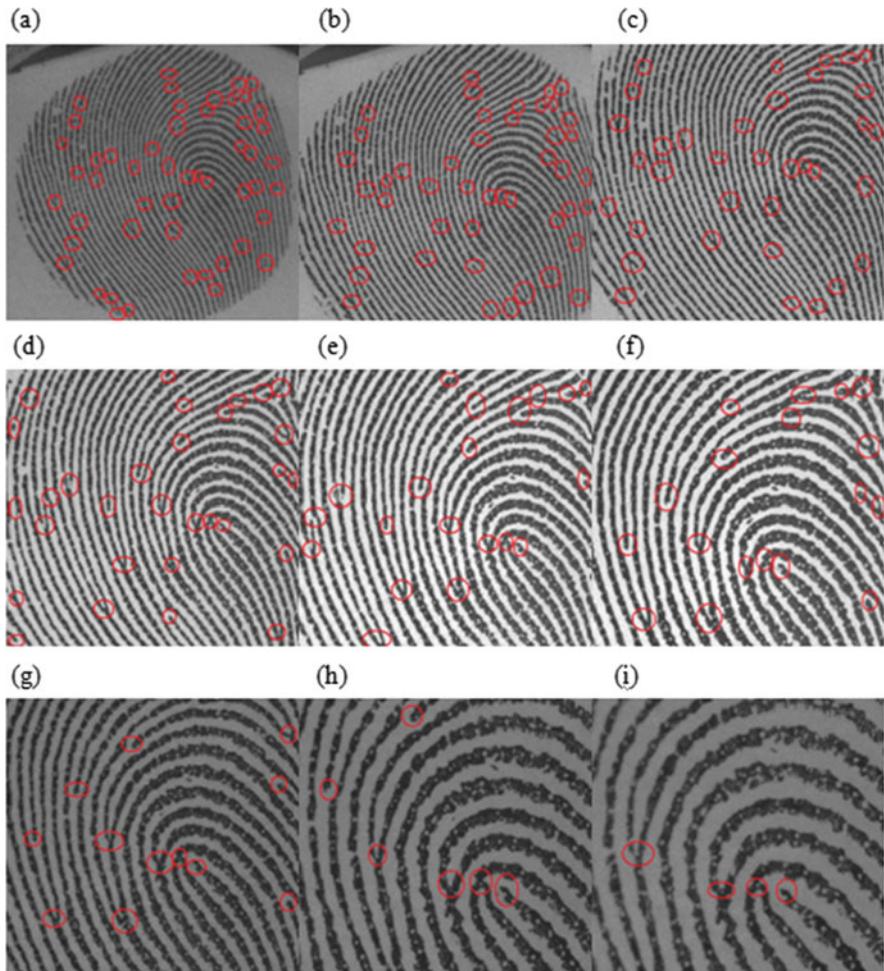


Fig. 4.5 Minutiae and pores on fingerprint images of 380×360 pixels at different resolutions. (a) 500 dpi. (b) 600 dpi. (c) 700 dpi. (d) 800 dpi. (e) 900 dpi. (f) 1000 dpi. (g) 1200 dpi. (h) 1600 dpi. (i) 2000 dpi

Table 4.2 Number of minutiae and pores in Fig. 4.5 at different resolutions

r (dpi)	500	600	700	800	900	1000	1200	1600	2000
Num_minu	51	46	35	30	20	18	12	6	4
Num_pore	0	85	617	683	710	609	356	172	140

size and shape (see the pores marked by red circles) are corrupted by noise or influenced by the condition of pores (open or closed). We thus set another criterion for resolution selection based on the number of pores at different resolutions, which can offer us the upper limit resolution.

Criterion 3 Minutiae Are More Discriminative than Pores If the Same Number of Each Is Considered. Retain as Many Minutiae as Possible While Also Retaining an Acceptable Number of Pores. A Reference Image Resolution Is Then Proposed.

Criteria 1 and 2 put emphasis on the number of minutiae and pores, respectively, which just offer the lower resolution and upper resolution for high-resolution AFRSs. However, it is obvious that this will at times also require us to make some kind of tradeoff between the two. In this tradeoff, the bias will be toward retaining minutiae because the distribution of minutiae is more random than that of pores, and so, the number of minutiae in an image will have a greater influence on fingerprint recognition. The blue line on Fig. 4.7 links ten adjacent minutiae on a fingerprint image, while the red line links adjacent pores. We can see that the blue

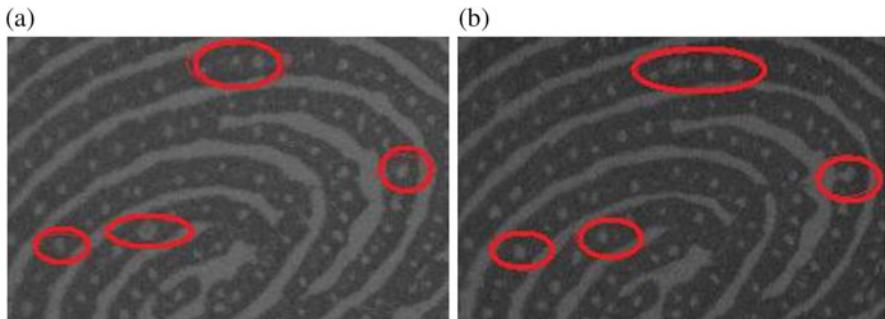


Fig. 4.6 Two prints of one finger under 2000 dpi captured at different times

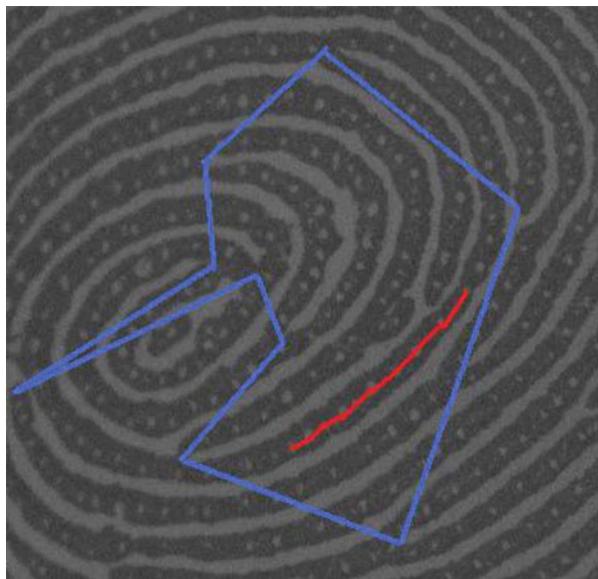


Fig. 4.7 Distribution of a similar number of minutiae and pores on a fingerprint image of 380×360 pixels

line traverses approximately 1/3 of the entire fingerprint image while the red line is concentrated in just one area of about 1/100 of the fingerprint image. From this, it would seem that if one or the other, i.e., minutiae or pores, must be traded off, then we lose less discriminative power if we bias toward retaining minutiae in the selection of a suitable resolution. We thus set our last criterion for resolution selection as retaining as many minutiae as possible while an acceptable number of pores are available.

Note that all the aforementioned three criteria are about the number of minutiae and pores with a fixed image size. However, the ridge width, which differs between different kinds of fingers (e.g., thumb, index finger, and middle finger) (Cummins et al. 1941) and between different genders (female and male) (Acree 1999), also has some effect on the number of minutiae and pores for a fixed image size and would consequently affect the selection of resolution. To make the reference resolution we selected based on the established criteria being universal to all fingers, it is necessary to study the relationship between the ridge width and the resolution. An analysis of ridge width on different kinds of fingers (e.g., thumb, index finger, and middle finger) and on fingers from different genders (female and male) is conducted with respect to the resolution selected based on the established criteria. Section 4.4.3 will report the analysis result.

4.4 Experiments and Analysis

To get a reference resolution based on our established criteria and to verify it, some analysis and experiments are organized as follows. First, a theoretical analysis of the minimum resolution for pore extraction is given. Second, the statistical number of minutiae and pores manually counted is offered. Third, an analysis of the ridge width on different kinds of fingers (thumb, index finger, and middle finger) and on fingers of different genders (female and male) is given. Finally, the automated fingerprint recognition results of different resolution fingerprint images are provided.

4.4.1 *Minimum Resolution Required for Pore Extraction*

There is a minimum resolution that is required to be able to extract pores well. In 1994, Stosz and Alyea (1994) automatically extracted pores using a high-resolution fingerprint sensor. They noted that pores could range in size from 60 to 250 μm in one dimension and that the smallest detectable pores, i.e., 60 μmin one dimension, determined the minimum resolution required by a sensor. They assumed a sampling period half the size of the smallest pore and concluded that the minimum required are solution of 800 dpi. In a later chapter, in 1997, Roddy and Stosz (1997) talked

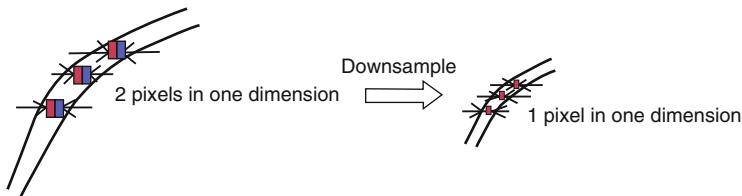


Fig. 4.8 Rule used to choose the minimum resolution for pore extraction

Table 4.3 The minimum value of h of different resolutions

Resolutions (dpi)	500	600	700	800	900	1000	1200	1600	2000
Minimum value of h (pixel)	1.5	1.7	2.0	2.3	2.6	2.9	3.5	4.7	5.8

about a range of pore sizes of 88–220 μm . Taking these two figures into account, in this chapter, we use the average of these two minimum pore sizes.

To determine the minimum required resolution, we take the size of pores and the resolution and apply (4.1) to calculate the number of pixels in a pore. Then, based on the rule that the size of the smallest pores in one dimension can be down sampled (Stosz and Alyea 1994), we know that the minimum resolution for pore extraction should guarantee that there are at least 2 pixels of the smallest pores in one dimension, as illustrated in Fig. 4.8. Table 4.3 shows the minimum values of height h for different resolutions. We can see that the minimum resolution required for pore detection is 700 dpi when assuming a sampling period half the size of the smallest pores.

4.4.2 Resolution Based on the Established Criteria

Given a fixed image size, as resolution increases, the number of minutiae decreases and pores become more visible. We manually counted the numbers of minutiae and pores in the 120 fingerprint images at each resolution (500×2000 dpi) at an image size of 380×360 pixels and then averaged these numbers. Figure 4.9 shows the relationship between the numbers of minutiae and pores. We have exaggerated the number of minutiae tenfold for the purpose of display. We can see that the number of minutiae is monotonically decreasing but within an acceptable range from 500 to 1000 dpi and that a relatively large number of pores (statistical number by counting manually) are retained at resolutions in the range of 700–1000 dpi. It would appear that the best choice of resolution for fingerprint recognition is 700 dpi. However, given that 700 dpi is the minimum resolution for pore extraction, we decided that to make the system more robust to noise, 800 dpi would be a better choice.

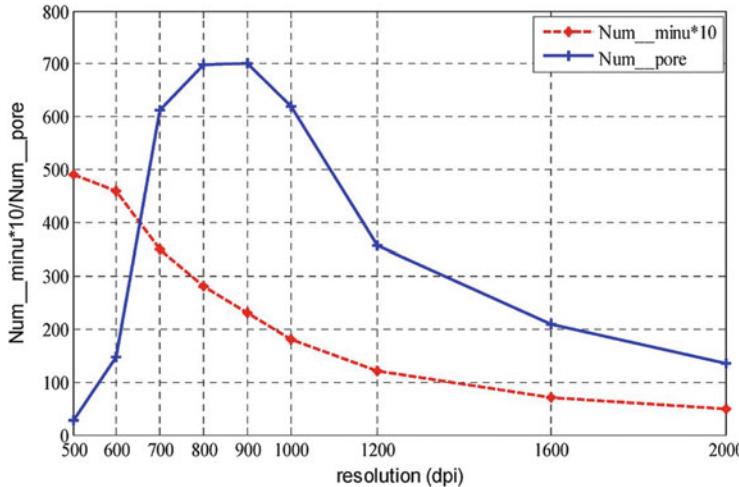


Fig. 4.9 Average numbers of minutiae and pores in 120 selected images in our database at different resolutions

4.4.3 Analysis of Ridge Width

Since minutiae and pores are both related to fingerprint ridges, there is some influence of ridge width on the number of minutiae and pores. We thus did some analysis about the ridge width for different groups of fingers. Ridge width has been studied in Cummins et al. (1941) and Acree (1999). In Cummins et al. (1941), the ridge width was determined by counting the ridges transversely crossing a line of 1 cm. In the chapter, the authors concluded that ridge width has little to do with the body weight, stature, hand length, and so on. They also summarized that the ridge width is different for different finger seven though they do not greatly differ. However, they did not discuss the relationship between ridge width and gender, for the reason that all the samples used in their chapter are from males. The relationship between ridge width and gender was studied in Acree (1999). Ridge width in that chapter was decided by the ridge density, which counted the epidermal ridges on fingerprints with a 5×5 mm square drawn on transparent film. The authors of Acree (1999) concluded that women tend to have a statistically significant greater ridge density. Getting aware of the variation of ridge width, we also studied the ridge width on different kinds of fingers (e.g., thumb, index finger and middle finger) and on fingers from different genders (female and male) by using our collected fingerprint image database at the selected resolution 800 dpi. The ridge density used in Acree (1999) is adopted here to determine the ridge width. In our database, there are 150 female fingers, 150 male fingers, 100 thumbs, 100 index fingers, and 100 middle fingers. Some descriptive statistics of dermal ridge densities as mentioned in Acree (1999) and the corresponding ridge width represented in terms of micrometers [calculated by the following (4.2)] and pixels [calculated by formula (4.1)] are given for different groups of fingers in Table 4.4. The

Table 4.4 Descriptive statistics comparisons of ridge density and their corresponding ridge width on different group of fingers.

	Females	Males	Thumb	Index finger	Middle finger
Number of fingers	150	150	100	100	100
Mean (ridges/25 mm ²)	19.13	17.67	18	18.23	18.67
Corresponding ridge width (μm, pixel)	(185, 5.8)	(200, 6.3)	(196, 6.2)	(194, 6.1)	(189, 6.0)
Minimuma (ridges/25 mm ²)	16.83	15.50	16.13	16.00	16.67
Corresponding ridge width (μm, pixel)	(210, 6.6)	(228, 7.2)	(219, 6.9)	(220, 6.9)	(212, 6.7)
Maximuma (ridges/25 mm ²)	22.67	21.67	21.17	21.77	22

corresponding ridge width represented in terms of micrometers is calculated as follows:

$$\text{ridge width} = \sqrt{5^2 + 5^2} / (\text{ridge density} \times 2) \quad (4.2)$$

Here, the diagonal length of the 5 × 5 mm square is considered as the overall length of all ridge-valley period.

Table 4.4 shows the standard variation (SD), mean value (Mean), minimum value per person (Minimum), and maximum value per person (Maximum) of the ridge density, as well as their corresponding ridge width on different groups of fingers. The results of different kinds of fingers (thumb, index finger, and middle finger) in Table 4.4 show that there is little difference in ridge width between them, which agrees with the conclusion made in Cummins et al. (1941). The results in Table 4.4 also show that the ridge width of females is generally smaller than that of males by 15 μm or 0.5 pixels. However, this difference is not significant (i.e., of sub pixel level). Thus, we conclude that, under a resolution of 800 dpi, the ridge width had little influence on the number of minutiae and pores. It makes our proposed reference resolution universal to all fingers.

4.4.4 Fingerprint Recognition Accuracy

To verify our choice of resolution and its relationship to accurate fingerprint recognition, we conducted a series of experiments using the fusion strategy presented in Zhao et al. (2009a, b) by combining the state-of-the-art minutia-based method proposed in Jain et al. (1997) and the pore-based method proposed in Zhao et al. (2009a, b), evaluating recognition accuracy according to the equal error rate (EER). Specifically, we did cross-validation experiments by dividing all fingers into three groups according to the types of fingers (i.e., thumb, index finger, and middle finger, respectively), as well as by dividing all fingers into female and male groups. The recognition results by considering all the fingers included in our

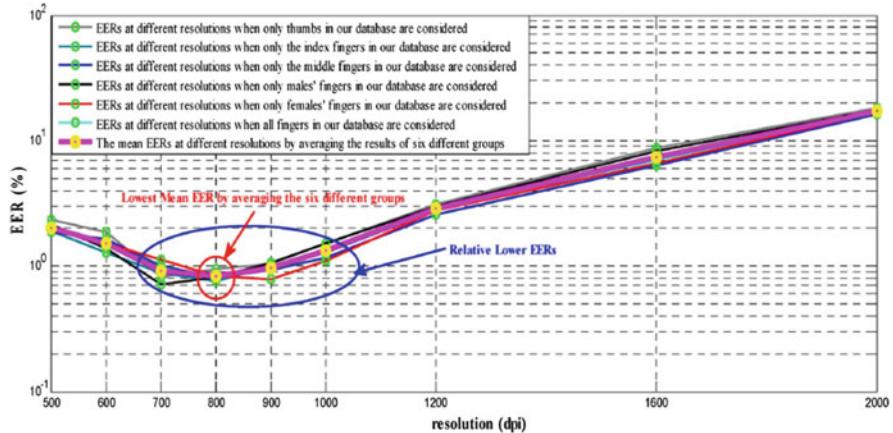


Fig. 4.10 EERs obtained at different resolutions on the six different groups of fingers and the mean EERs by averaging those EERs at different resolutions

database were also given. The lower the value of EER is, the higher the recognition accuracy will be. Figure 4.10 shows the EERs obtained at different resolutions on the six different groups of fingers and the mean EERs by averaging those EERs at different resolutions. For the thumb, index finger, and middle finger groups, the EERs were obtained from 600 genuine scores (generated from 100 fingers, 4 pictures of each finger) and 4950 imposter scores (generated from 100 fingers, comparing the first images of different fingers). For the female and male groups, the EERs were obtained from 900 genuine scores (generated from 150 fingers, four pictures of each finger) and 11,175 imposter scores (generated from 150 fingers, comparing the first images of different fingers). When considering all the fingers, the EERs were obtained from 1800 genuine scores (generated from 300 fingers, four pictures of each finger) and 44,850 imposter scores (generated from 300 fingers, comparing the first images of different fingers).

Figure 4.10 shows the EERs on different groups of finger sat different resolutions by fusing the state-of-the-art minutia based method proposed in Jain et al. (1997) and the pore-based method proposed in Zhao et al. (2009a, b). Specifically, the black line in Fig. 4.10 shows the recognition results when only the males' fingers in our database are considered. The lowest EER is obtained when resolution is 700 dpi. The red line in Fig. 4.10 shows the EER values at different resolutions when only females' fingers in our database are involved. The lowest EER is obtained when the resolution is 900 dpi. The gray line that represents the EERs when only thumbs are considered shows that the lowest EER can be obtained at the resolution of 700 dpi. The rest of the lines in Fig. 4.10 all show that the lowest EER is achieved at the resolution of 800 dpi. However, all of the results in Fig. 4.10 show that a relatively lower EER can be obtained when the resolution is between 700 and 1000 dpi. A resolution of 800 dpi can achieve the lowest EER in most cases and the lowest mean EER of the sixfold experiment (pink line). This result further confirms our proposed reference resolution.

4.5 Summary

This chapter has proposed a method for selecting a reference resolution for use in high-resolution AFRSs based on minutiae and pores. We have initially found that, based on anatomical evidence, a minimum resolution of 700 dpi would give good results, but further analysis based upon an analysis of the number of minutiae and pores and the ridge width on different kinds of fingers and on fingers of different genders, as well as tests of comparative accuracy, has led us to recommend a reference resolution of 800 dpi. While we have regarded this as an advance, we must point out that the image size also has an important role in high-resolution AFRSs. In this chapter, we limited images to a size of 380×360 pixels to allow us to investigate only the impact of resolution. In future work, we will investigate how to best make the tradeoff between the influences of resolution and image size within a certain range on high-resolution AFRS and to figure out whether there exists a dynamic resolution to different image sizes for high-resolution AFRSs.

References

- Acree M (1999) Is there an gender difference in fingerprint ridge density? *Forensic Sci Int* 102(1):35–44
- Ashbaugh D (1999) Quantitative–qualitative friction ridge analysis: an introduction to basic and advanced ridgeology. CRC Press, Boca Raton, FL
- Bicz W, Banasiak D, Bruciak P, Gumienny S, Gumulinski D, Keysiak A, Kuczynski W, Pluta M, Rabiej G. Fingerprint structure imaging based on an ultrasound camera [online]. <http://www.optel.com.pl/article/english/article.htm>
- CDEFFS (2009) Data format for the interchange of extended fingerprint and palmprint features, version 0.4
- Chen Y (2009) Extended feature set and touchless imaging for fingerprint matching. Michigan State University
- Cummins H, Waits W, McQuitty J (1941) The breadths of epidermal ridges on the finger tips and palms: a study of variation. *Am J Anat* 68(1):127–150
- Han J, Tan Z, Sato K, Shikida M (2005) Thermal characterization of micro heater arrays on a polyimide film substrate for fingerprint sensing applications. *J Micromech Microeng* 15(2):282–289
- International Biometric Group (2008) Analysis of level 3 features at high resolutions (phase II)
- Jain A, Hong L, Bolle R (1997) On-line fingerprint verification. *IEEE Trans Pattern Anal Mach Intell* 19(4):302–314
- Jain A, Chen Y, Demirkus M (2006) Pores and ridges: fingerprint matching using level 3 features. In: Proceedings of the 18th international conference on pattern recognition, pp 477–480
- Jain A, Chen Y, Demirkus M (2007) Pores and ridges: high-resolution fingerprint matching using level-3 features. *IEEE Trans Pattern Anal Mach Intell* 29(1):15–27
- Kryszczuk K, Drygajlo A, Morier P (2004a) Extraction of level 2 and level 3 features for fragmentary fingerprints. In: Proceedings of the 2nd COST Action 275 workshop, Vigo, Spain, pp 83–88
- Kryszczuk K, Morier P, Drygajlo A (2004b) Study of the distinctiveness of the level 2 and level 3 features in fragmentary fingerprint comparison. In: Proceedings of the biometrics authentication, ECCV 2004 international workshop, pp 124–133

- Maltoni D, Maio D, Jain A, Prabhakar S (2009) Handbook of fingerprint recognition. Springer, New York
- Parsons N, Smith J, Thonnes E, Wang L, Wilson R (2008) Rotationally invariant statistics for examining the evidence from the pores in fingerprints. *Law Probab Risk* 7(1):1–14
- Ratha N, Bolle R (2004) Automatic fingerprint recognition systems. Springer, New York
- Ray M, Meenen P, Adhami R (2005) A novel approach to fingerprint pore extraction. In: Proceedings of the thirty-seventh southeastern symposium on system theory, pp 282–286
- Roddy A, Stosz J (1997) Fingerprint features-statistical analysis and system performance estimates. *Proc IEEE* 85(9):1390–1421
- Ross A, Jain A (2004) Biometric sensor interoperability: a case study in fingerprints. In: Proceedings of the international ECCV workshop on biometric authentication, LNCS 3087, pp 134–145
- Stosz J, Alyea L (1994) Automated system for fingerprint authentication using pores and ridge structure. *Proc SPIE* 2277:210–223
- Xia X, O'Gorman L (2003) Innovation in fingerprint capture devices. *Pattern Recogn* 36(2):361–369
- Zhao Q, Zhang L, Zhang D, Luo N, Bao J (2008) Adaptive pore model for fingerprint pore extraction. In: Proceedings of the 18th international conference on pattern recognition, pp 1–4
- Zhao Q, Zhang L, Zhang D, Luo N (2009a) Direct pore matching for fingerprint recognition. In: ICB 2009, pp 597–606
- Zhao Q, Zhang D, Zhang L, Luo N (2009b) High resolution partial fingerprint alignment using pore-valley descriptors. *Pattern Recogn* 43(3):1050–1061

Part II

Finger-Knuckle-Print Verification

Chapter 5

Finger-Knuckle-Print Verification

Abstract Biometric based personal authentication is an effective method for automatically recognizing, with a high confidence, a person's identity. By observing that the texture pattern produced by bending the finger knuckle is highly distinctive, in this chapter we present a new biometric authentication system using finger-knuckle-print (FKP) imaging. A specific data acquisition device is constructed to capture the FKP images, and then an efficient FKP recognition algorithm is presented to process the acquired data in real time. The local convex direction map of the FKP image is extracted, based on which a local coordinate system is established to align the images and a region of interest is cropped for feature extraction. For matching two FKPs, a feature extraction scheme which combines orientation and magnitude information extracted by Gabor filtering is proposed. An FKP database, which consists of 7920 images from 660 different fingers, is established to verify the efficacy of the proposed system and promising results are obtained. Compared with the other existing finger-back surface based biometric systems, the proposed FKP system achieves much higher recognition rate and it works in real time. It provides a practical solution to finger-back surface based biometric systems and has great potentials for commercial applications.

Keywords Biometrics • Finger-knuckle-print • Personal authentication

5.1 Introduction

Personal authentication is a common concern to both industries and academia due to its numerous applications such as physical access control, computer security, banking and law enforcement, etc. Biometrics, which refers to the unique physiological or behavioral characteristics of human beings, can be used to distinguish between individuals and hence can serve as an ideal solution to this problem. With the rapid development of computer techniques, in the past three decades researchers have exhaustively investigated the use of a number of biometric characteristics, including fingerprint (Jain and Flynn 2007; Maltoni et al. 2003; Ratha and Bolle 2004), face (Delac and Grgic 2007; Wechsler 2006), iris (Daugman 1993, 2004),

retina (Hill 1999; Borgen et al. 2008), palmprint (Guo et al. 2009; Zhang et al. 2003; Kong and Zhang 2004; Kong et al. 2006; Sun et al. 2005; Huang et al. 2008; Jia et al. 2008), hand geometry (Jain et al. 1999; Jain and Duta 1999; Sanchez-Reillo et al. 2000), hand vein (Wang et al. 2008; Kumar and Prathyusha 2008), finger surface (Woodard and Flynn 2005a, b; Ravikanth and Kumar 2007; Kumar and Zhou 2009a, b), inner-knuckle-print (Li et al. 2004; Nanni and Lumini 2009a, b), voice (Hollien 2002), ear (Burge and Burger 1999), gait (Nixon et al. 2006) and signature (Plamondona and Loretteb 1989; Liu et al. 1979), etc. Although many biometric techniques are still under the stage of research and development, some biometric systems have been used in a large scale; for example, the Hong Kong government has been using the fingerprint recognition system as the automated passenger clearance system (e-channel) since 2004 (E-channel System of the Hong Kong government 2004).

Among various kinds of biometric identifiers, hand-based biometrics has been attracting considerable attention over recent years. Fingerprint (Jain and Flynn 2007; Maltoni et al. 2003; Ratha and Bolle 2004), palmprint (Guo et al. 2009; Zhang et al. 2003; Kong and Zhang 2004; Kong et al. 2006; Sun et al. 2005; Huang et al. 2008; Jia et al. 2008), hand geometry (Jain et al. 1999; Jain and Duta 1999; Sanchez-Reillo et al. 2000), hand vein (Wang et al. 2008; Kumar and Prathyusha 2008), and inner-knuckle-print (Li et al. 2004; Nanni and Lumini 2009a, b) have been proposed and well investigated in the literature. The popularity of hand-based biometrics should be attributed to its high user acceptance. In fact, the image pattern in the finger knuckle surface is highly unique and thus can serve as a distinctive biometric identifier. Woodard and Flynn (2005a, b) are among the first scholars who exploit the use of finger knuckle surface in biometric systems. They set up a 3-D finger back surface database with the Minolta 900/910 sensor. For feature extraction, they used the curvature based shape index to represent the finger back surface. Woodard's work makes a good effort to validate the uniqueness of outer finger surface as a biometric characteristic. However, their work did not provide a practical solution to establishing an efficient system using the outer finger surface features. The cost, size and weight of the Minolta 900/910 sensor limit the use of it in a practical biometric system, and the time-consuming 3-D data acquisition and processing limit its use in real-time applications. In addition, they did not fully exploit the finger knuckle texture information in feature extraction. These systems are not only different in the imaging technologies, but also different in the feature extraction methods. In the following parts, the imaging technique and feature extraction will be presented separately.

Later, Kumar and Ravikanth (Ravikanth and Kumar 2007; Kumar and Ravikanth 2009) proposed another approach to personal authentication by using 2-D finger-back surface imaging. They developed a system to capture hand-back images and then extracted the finger knuckle areas by some preprocessing steps. The subspace analysis methods such as PCA, LDA and ICA were combined to do feature extraction and matching. With Kumar's design, the acquisition device is doomed to have a large size because nearly the whole hand back area has to be captured, despite the fact that the finger knuckle area only occupies a small portion of the acquired image. Furthermore, subspace analysis methods may be effective

for face recognition but they may not be able to effectively extract the distinctive line and junction features from the finger knuckle surface. In Kumar's later work (Kumar and Zhou 2009a, b), they used the robust line orientation code proposed in (Jia et al. 2008) to extract the orientation of the finger-back surface images.

In this chapter, a new hand-based biometric technique, namely finger-knuckle-print (FKP), is developed for online personal authentication. FKP refers to the image pattern of the outer surface around the phalangeal joint of one's finger, which is formed by bending slightly the finger knuckle. A specially designed acquisition device is constructed to collect FKP images. Unlike the systems in (Woodard and Flynn 2005a, b) and (Kumar and Ravikanth 2009) which first capture the image of the whole hand and then extract the finger or finger knuckle surface areas, the proposed system captures the image around the finger knuckle area of a finger directly, which largely simplifies the following preprocessing steps. Meanwhile, with such a design the size of the imaging system can be greatly reduced, which improves much its applicability. Since the finger knuckle will be slightly bent when being imaged in the proposed system, the inherent finger knuckle print patterns can be clearly captured and hence the unique features of FKP can be better exploited.

After an FKP image is captured, a region of interest (ROI) needs to be cropped from the original image for the following feature extraction. An efficient FKP ROI extraction algorithm is proposed based on the intrinsic characteristics of FKP images. For matching two FKP ROI images, we propose a new feature extraction scheme which combines orientation and magnitude information extracted by Gabor filters. Experimental results show that it outperforms the other state-of-the-arts coding based feature extraction methods, such as CompCode (Kong and Zhang 2004), OrdinalCode (Sun et al. 2005), RLOC (Jia et al. 2008; Kumar and Zhou 2009a, b) and BOCV (Guo et al. 2009), in FKP recognition. To evaluate the performance of the proposed technique, an FKP database was established using our prototype system, which consists of 7920 images from 660 different fingers. Experimental results demonstrated that the proposed FKP based authentication system can verify the personal identity in real time with a high recognition rate. Compared with the other existing finger knuckle surface based biometric systems (Woodard and Flynn 2005a, b; Ravikanth and Kumar 2007; Kumar and Ravikanth 2009), the proposed one performs much better in terms of both the recognition accuracy and the speed.

5.2 The Finger-Knuckle-Print (FKP) Recognition System

The schematic diagram of our FKP based personal authentication system is shown in Fig. 5.1. The system is composed of a data acquisition module and a data processing module. The data acquisition module is composed of a finger bracket, a ring LED light source, a lens, a CCD camera and a frame grabber. The captured FKP image is inputted to the data processing module, which comprises three basic steps: ROI (region of interest) extraction, feature extraction and coding, and

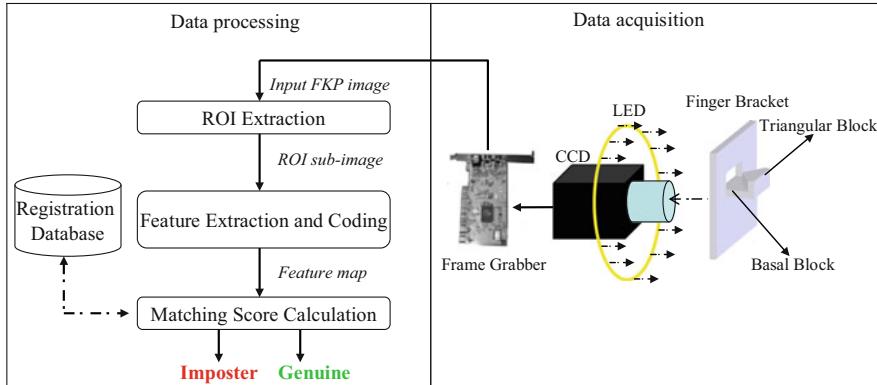


Fig. 5.1 Structure of the proposed FKP-based personal authentication system. The whole system is composed of a data acquisition module and a data processing module

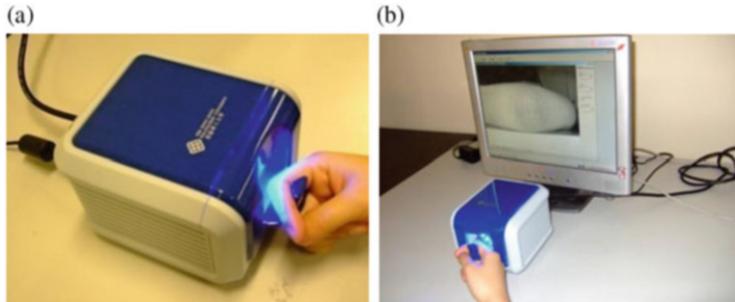


Fig. 5.2 (a) The outlook of the developed FKP image acquisition device; (b) The device is being used to collect FKP samples

matching. Figure 5.2 shows the outlook of our FKP image acquisition device whose overall size is 160 mm × 125 mm × 100 mm.

A critical issue in data acquisition is to make the data collection environment as stable and consistent as possible so that variations among images collected from the same finger can be reduced to the minimum. In general, a stable image acquisition process can effectively reduce the complexity of the data processing algorithms and improve the image recognition accuracy. Meanwhile, we want to put as little constraint as possible on the users in order for high user friendliness of the system. With the above considerations, a semi-closed data collection environment is designed in our system. The LED light source and the CCD camera are enclosed in a box so that the illumination is nearly constant. One difficult problem is how to make the gesture of the finger be nearly constant so that the captured FKP images from the same finger are consistent. In our system, the finger bracket is designed for this purpose.

Refer to Fig. 5.1, a basal block and a triangular block are used to fix the position of the finger joint. In data acquisition, the user can easily put his/her finger on the basal block with the middle phalanx and the proximal phalanx touching the two slopes of the triangular block. Such a design aims at reducing the spatial position variations of the finger in different capturing sessions. The triangular block is also used to constrain the angle between the proximal phalanx and the middle phalanx to a certain magnitude so that line features of the finger knuckle surface can be clearly imaged.

After the image is captured, it is sent to the data processing module for preprocessing, feature extraction and matching (refer to Sects 3 and 4 for details). The size of the acquired FKP images is 768×576 under a resolution about 400 dpi. Figure 5.3 shows four sample images acquired by the developed device. Two images in the first row are from one finger and images in the second row are from another finger. Example images for the same finger were captured at two different collection sessions with an interval of 56 days. We see that by using the developed system, images from the same finger but collected at different times are very similar to each other. Meanwhile, images from different fingers are very different, which implies that FKP has the potential for personal identification.

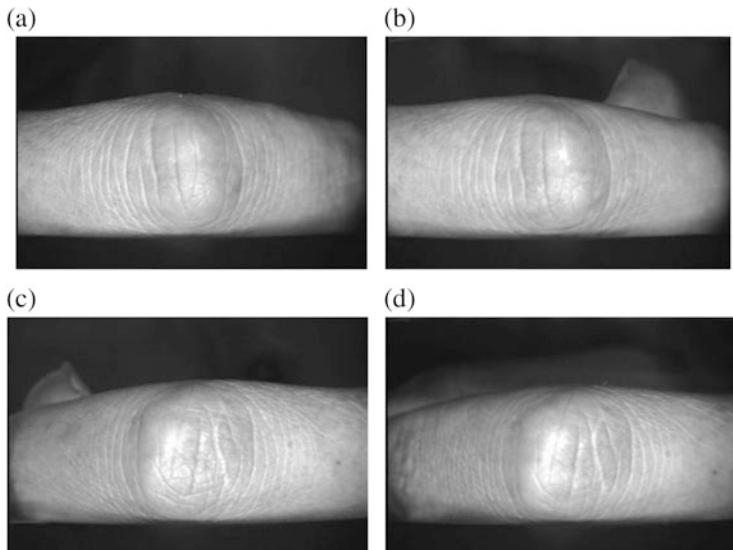


Fig. 5.3 Sample FKP images acquired by the developed system. (a) and (b) are from one finger while (c) and (d) are from another finger. Images from the same finger are taken at two different sessions with an interval of 56 days

5.3 ROI (Region of Interest) Extraction

FKP images collected from different fingers are very different. On the other hand, for the same finger, images collected at different collection sessions will also vary because of the variation of spatial locations of the finger. Therefore, it is necessary and critical to align FKP images by adaptively constructing a local coordinate system for each image. With such a coordinate system, an ROI can be cropped from the original image in order for reliable feature extraction and matching. In this section, we will propose an algorithm for the local coordinate system determination and ROI sub-image extraction.

Because the finger is always put flatly on the basal block when the FKP image is captured, the bottom boundary of the finger is stable in every image and can be taken as the X -axis of the ROI coordinate system. However, the Y -axis is much more difficult to determine. Intuitively, we want to locate the Y -axis in the center of the phalangeal joint so that most of the useful features in the FKP image can be preserved within the ROI. It can be observed that line features on the two sides of the phalangeal joint have different convex directions. Taking this fact as a hint, we propose to code line pixels based on their convex directions and then make use of the convex direction codes to determine the Y -axis. Figure 5.4 illustrates the main steps of the coordinate system determination and the ROI extraction. In the following, we describe these steps in detail.

Step 1: Image down-sampling

The size of the captured FKP image is 768×576 under a resolution of 400 dpi. Based on our experiments, it is not necessary to use such a high resolution for feature extraction and matching. Therefore, we apply a Gaussian smoothing operation to the original image, and then down-sample the smoothed image to about 150 dpi (see Sect. 5.2 for the discussion of resolution selection). The down-sampling operation has two advantages. First it can significantly reduce the computational cost by reducing the data amount. Second, the Gaussian smoothing will suppress the noise in the original image, which can benefit the following feature extraction and matching steps. We denote by I_D the down-sampled image and Fig. 5.4a shows such an image.

Step 2: Determine the X-axis of the coordinate system

Refer to Fig. 5.4b, the bottom boundary of the finger can be easily extracted by a Canny edge detector. Actually, this bottom boundary is nearly consistent to all FKP images because all the fingers are put flatly on the basal block in data acquisition. By fitting this boundary as a straight line, the X -axis of the local coordinate system is determined.

Step 3: Crop a sub-image I_S from I_D

Useful information which can be used for personal identification only resides in a portion of the whole FKP image. Therefore, we first crop a sub-image I_S from the original image for the convenience of later processing. The left and right boundaries of I_S are two fixed values evaluated empirically. The top and bottom boundaries are

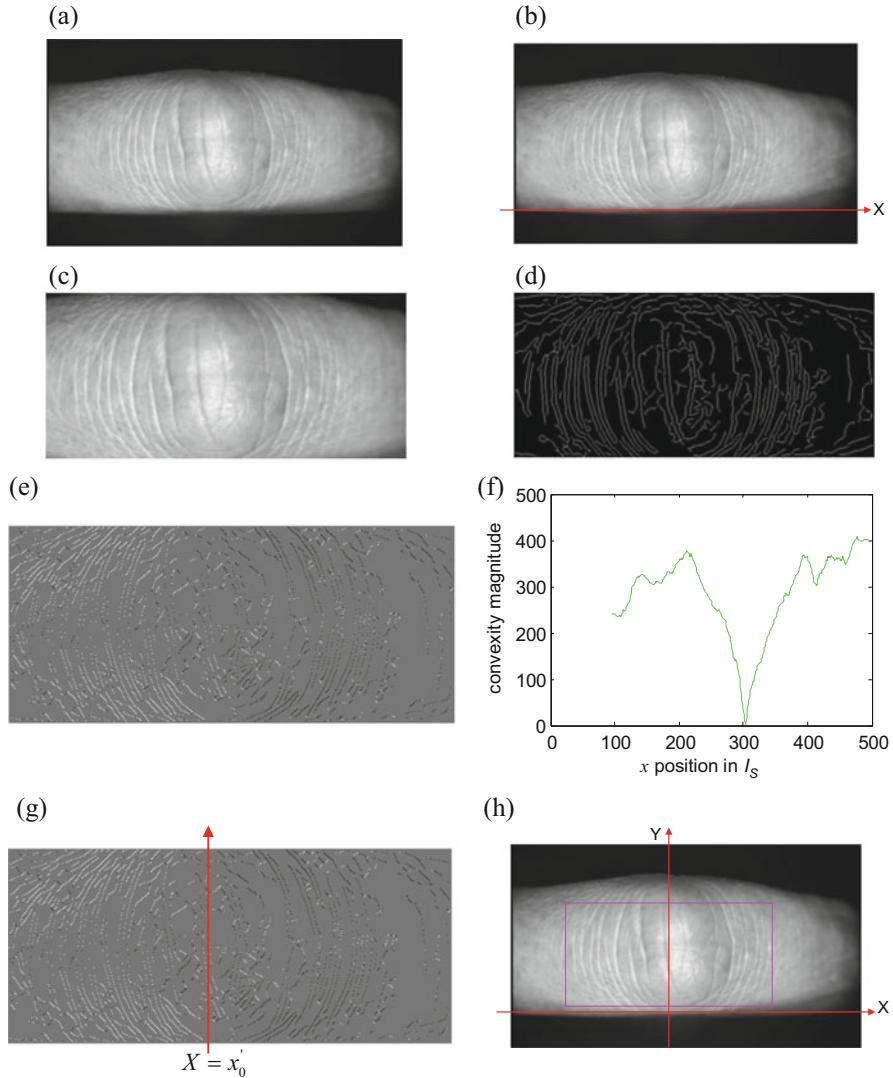


Fig. 5.4 Illustration for the ROI extraction process. (a) I_D image which is obtained by a down-sampling operation after a Gaussian smoothing; (b) X -axis of the coordinate system, which is the line $Y = Y_0$, fitted from the bottom boundary of the finger; (c) I_S image extracted from I_D ; (d) I_E image obtained by applying a Canny edge detector on I_S ; (e) I_{CD} image obtained by applying the convex direction coding scheme to I_E ; (f) plot of $conMag(x)$ for a typical FKP image; (g) line $X = x'_0$, where $x'_0 = \arg \min_x (conMag(x))$; (h) ROI coordinate system and the rectangle indicates the area of the ROI sub-image that will be extracted

estimated according to the boundary of real fingers. Figure 5.4c shows an example I_S image. This roughly cropped sub-image will be used to calculate the Y-axis so that an accurate ROI image can be cropped. Modulated light projector projects a continually changing light at the subject. Usually the light source simply cycles its amplitude in a [sinusoidal](#) pattern. A camera detects the reflected light and the changes in the brightness of each pixel of the image and the light phase to get the distance of the projected light. The structure is shown in Fig. 5.4.

Step 4: Canny edge detection

By applying Canny edge detector to I_S , the corresponding edge map I_E can be obtained. See Fig. 5.4(d) for an example.

Step 5: Convex direction coding for I_E

Based on the local convexity characteristics of the edge map I_E , we can code I_E to get the convex direction coding map I_{CD} . At this step, each pixel on I_E will be given a code to represent the local convex direction of this pixel. The underlying principle of this coding scheme is as follows. Based on the observation of FKP images, we abstract an ideal model for “curves” on an FKP image as shown in Fig. 5.5a. In this model, an FKP “curve” is either convex leftward or convex rightward. We code the pixels on convex leftward curves as “1”, the pixels on convex rightward curves as “−1”, and the other pixels not on any curves as “0”. Figure 5.5b illustrates the coding scheme. In our system, we regard the edges obtained in step 4 as “curves” and this convex direction coding is performed on I_E . The pseudo code of this algorithm is as follows (Tables 5.1 and 5.2):

After convex direction coding, each I_{CD} point is assigned a value 0, 1 or −1. Figure 5.4e shows an I_{CD} map in false color image format. White pixels on it are the ones with convexity value “1”; black pixels are the ones with value “−1”; and gray pixels are of value “0”.

Step 6: Determine the Y-axis of the coordinate system

Consider the ideal FKP curve model set up at step 5. For an FKP image, “curves” on the left part of phalangeal joint are mostly convex leftward and those on the right part are mostly convex rightward. Meanwhile, “curves” in a small area around the phalangeal joint do not have obvious convex directions. Based on this observation,

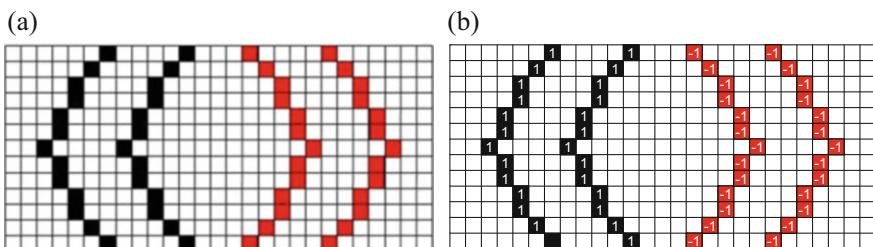


Fig. 5.5 (a) Ideal model for FKP “curves”; (b) Convex direction coding scheme

Table 5.1 Algorithm Convex_Direction_Coding (IE)

Input: I_E ($m \times n$ binary edge map computed in step 4)

Output: I_{CD} ($m \times n$ convex direction code map)

Begin module

$$y_{mid} = \frac{\text{height of } I_E}{2};$$

for each $I_E(i, j)$ **do**

if $I_E(i, j) = 0$ // it is a background pixel

$$I_{CD}(i, j) = 0;$$

else if $I_E(i+1, j-1) = 1$ and $I_E(i+1, j+1) = 1$ // it is a bifurcation pixel

$$I_{CD}(i, j) = 0;$$

else if ($I_E(i+1, j-1) = 1$ and $i \leq y_{mid}$) or ($I_E(i+1, j+1) = 1$ and $i > y_{mid}$)

$$I_{CD}(i, j) = 1;$$

else if ($I_E(i+1, j+1) = 1$ and $i \leq y_{mid}$) or ($I_E(i+1, j-1) = 1$ and $i > y_{mid}$)

$$I_{CD}(i, j) = -1;$$

end if

end for

End module

Table 5.2 Algorithm ImCompCode (I_{ROI})

```

Input:  $I_{ROI}$  (  $m \times n$  ROI sub-image)

Output:  $ImCompCode$  (  $m \times n$  integer matrix)

Begin module

for each  $I_{ROI}(x,y)$  do

 $R = \{R_j = I_{ROI}(x,y) * G_R(x,y, \theta_j)\}$ , where  $\theta_j = j\pi / 6, j = \{0,1,...5\}$  , 

 $oriMag = \frac{abs(\max(R) - \min(R))}{\max(abs(\max(R)), abs(\min(R)))}$ 

if  $oriMag < T$  // this pixel does not have a dominant orientation

 $ImCompCode(x,y) = 6$  ;

else

 $ImCompCode(x,y) = \arg \min_j \{R_j\}$  ;

end if

end for

End module

```

at a horizontal position x (x represents the column) of an FKP image, we define the “convexity magnitude” as:

$$conMag(x) = abs \left(\sum_W I_{CD} \right) \quad (5.1)$$

where W is a window that is symmetrical about the axis $X = x$ and W is of size $d \times h$ with h being the height of I_S . d is experimentally chosen as 35 in this chapter. The

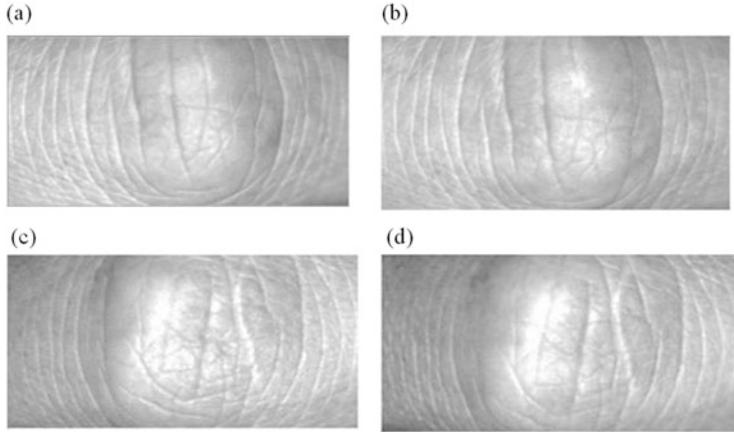


Fig. 5.6 Sample ROI images extracted by the proposed method. These four images are ROI images for the sample images shown in Fig. 5.3, respectively

“convexity magnitude” is proposed to measure how strong the dominant convex direction is in a local area on the FKP image. The characteristic of the FKP image suggests that $conMag(x)$ will reach a minimum around the center of the phalangeal joint and this position can be used to set the Y -axis of the coordinate system. Let

$$x'_0 = \arg \min_x (conMag(x)) \quad (5.2)$$

Then $X = x'_0$ can be set as the Y -axis. Figure 5.4f plots the curve $conMag(x)$ of an FKP image and Fig. 5.4g shows the vertical line $X = x'_0$, which is the Y -axis of the ROI system.

Step 7: Crop the ROI image

Now that we have fixed the X -axis and Y -axis, the local coordinate system can then be determined. Refer to Fig. 5.4h, with the constructed coordinate system, the ROI sub-image I_{ROI} can be extracted from I_D with a fixed size, which is empirically set as 160×80 in our system.

Figure 5.6 shows some examples of the extracted ROI images. We can see that the proposed coordinate system construction and ROI extraction method can effectively align the different FKP images and normalize the area for feature extraction. Such operations reduce greatly the variations caused by the various poses of the finger in data collection.

5.4 FKP Feature Extraction and Matching

The Gabor filtering technique can simultaneously extract the spatial-frequency information from the original signal (Gabor 1946). Since 1980s, it has been widely used as an effective tool to fulfill the feature extraction job in face, iris, fingerprint and palmprint systems. In (Nanni and Lumini 2009a, b), Loris and Alessandra described a Gabor feature selection technique. The Gabor filter can produce three types of features—magnitude, phase, and orientation, which can be used separately or jointly in different systems (Kong 2008). In this chapter, we propose a method combining the orientation and magnitude features for FKP recognition. Experimental results in Sect. 5 verifies that the proposed scheme performs better in FKP recognition than the other coding-based methods, such as BOCV (Guo et al. 2009), CompCode (Kong and Zhang 2004), OrdinalCode (Sun et al. 2005) and RLOC (Jia et al. 2008; Kumar and Zhou 2009a, b).

The Gabor function has several slightly different forms in the literature and here we adopt the one proposed by Lee (Lee 1996):

$$G(x, y, \omega, \theta) = \frac{\omega}{\sqrt{2\pi}\kappa} e^{-\frac{\omega^2}{8\kappa^2}(4x'^2+y'^2)} \left(e^{i\omega x'} - e^{-\frac{\kappa^2}{2}} \right) \quad (5.3)$$

where $x' = (x-x_0)\cos\theta + (y-y_0)\sin\theta$, $y' = -(x-x_0)\sin\theta + (y-y_0)\cos\theta$, (x_0, y_0) is the center of the function, ω is the radial frequency in radians per unit length and θ is the orientation of the Gabor functions in radians. κ is defined by $\kappa = \sqrt{2 \ln 2} \left(\frac{2^\delta + 1}{2^\delta - 1} \right)$, where δ is the half-amplitude bandwidth of the frequency response. ω can be determined by $\omega = \kappa/\sigma$, where σ is the standard deviation of the Gaussian envelop.

Using Gabor filtering, next we propose an improved competitive coding (ImCompCode) method to exploit the orientation information, and then we propose a magnitude coding (MagCode) method to exploit magnitude information. Finally, we fuse these two kinds of features in FKP matching.

5.4.1 Improved Competitive Coding (ImCompCode)

At each pixel $I_{ROI}(x,y)$, we extract the orientation information and represent it as an “orientation code”. This process is similar to the CompCode scheme in (Kong and Zhang 2004).

With a bank of Gabor filters, the orientation feature at each pixel $I_{ROI}(x,y)$ can be extracted. In our system, we only use the real part of the Gabor filter to perform this job. Mathematically, this orientation coding process can be represented as:

$$oriCode(x, y) = \arg \min_j \{I_{ROI}(x, y)^* G_R(x, y, \theta_j)\} \quad (5.4)$$

where symbol $*$ represents the convolution operation, G_R represents the real part of the Gabor function G , $\theta_j = j\pi/J$, $j = \{0, \dots, J-1\}$, and J represents the number of different orientations.

Here we improve the original CompCode scheme. Often on an FKP image, there are some pixels lying on relatively “plane” areas, i.e. these pixels do not reside on any lines and consequently do not have a dominate orientation. Accordingly, the J Gabor filter responses at such pixels do not have much variation. If we still assign an orientation code to it, this code may not be stable and will be sensitive to noise, making the feature representation and matching performance decreased. Therefore, those “plane” pixels should be removed from orientation coding. We define the “orientation magnitude” at a pixel as:

$$oriMag(x, y) = \frac{abs(\max(R) - \min(R))}{\max(abs(\max(R)), abs(\min(R)))} \quad (5.5)$$

where $R = \{R_j = I_{ROI}(x, y) * G_R(x, y, \theta_j)\}, j = \{0, \dots, J-1\}$ are the Gabor filtering responses at this pixel. The “orientation magnitude” $oriMag(x, y)$ can measure how likely the pixel (x, y) has a dominant orientation. If it is smaller than a threshold, we reckon that this pixel has no dominant orientation and the corresponding competitive code is assigned as J .

Based on our experimental results, using 6 Gabor filters of different orientations is enough. This is in accordance with the conclusion made by Lee in (Lee 1996) that the simple neural cells are sensitive to specific orientations with approximate bandwidths of $\pi/6$. Thus, we choose 6 orientations, $\theta_j = j\pi/6$, $j = \{0, \dots, 5\}$ for the competition. The pseudo code for our ImCompCode scheme is summarized as follows and Fig. 5.7a–d are ImCompCode maps for FKP ROI images shown in Fig. 5.6.

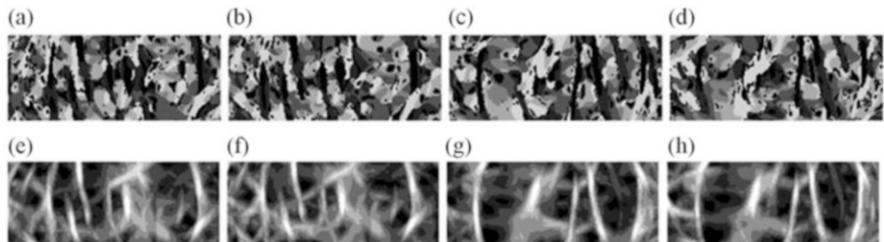


Fig. 5.7 (a)–(d) and (e)–(h) are ImCompCode maps and MagCode maps for the FKP ROI images shown in Fig. 5.6, respectively

5.4.2 Magnitude Coding (*MagCode*)

Besides orientation information, we also want to exploit magnitude information from Gabor filter responses. The magnitude of the Gabor filter response at $I_{ROI}(x, y)$ is:

$$\sqrt{(I_{ROI}(x, y)^* G_R(x, y, \omega, \theta_j))^2 + (I_{ROI}(x, y)^* G_I(x, y, \omega, \theta_j))^2} \quad (5.6)$$

where G_R and G_I represent the real part and the imaginary part of the Gabor function G respectively. However, in order to reduce the computational cost, when generating the magnitude code map, we want to make use of the temporary results generated from the “orientation coding” process. Thus, we still only use the real part of the Gabor filter and define the magnitude at $I_{ROI}(x, y)$ as:

$$mag(x, y) = \max_j (abs(I_{ROI}(x, y)^* G_R(x, y, \theta_j))) \quad (5.7)$$

Then a localized quantization is applied to $mag(x, y)$ to get the magnitude code. This process can be expressed as:

$$magCode(x, y) = ceil\left((mag(x, y) - lmin)/\left(\frac{lmax - lmin}{N}\right)\right) \quad (5.8)$$

where N is the number of quantization levels, $lmin = \min_{(x,y) \in W_m} (mag(x, y))$, and $lmax = \max_{(x,y) \in W_m} (mag(x, y))$. W_m is a $w \times w$ window centered at (x, y) . The resulting magnitude code is an integer within $1-N$. w and N can be tuned by experiments on a sub-dataset and they are experimentally set as 49 and 8 in this chapter, respectively. Figure 5.7e-h show magnitude code maps for FKP ROI images presented in Fig. 5.6.

5.4.3 FKP Matching

Suppose we are comparing two FKP ROI images, P and Q . Let P_o and Q_o be the two orientation code maps; and let P_m and Q_m be the two magnitude code maps. At first, we will calculate the matching distance between P_o and Q_o and the matching distance between P_m and Q_m respectively, and then fuse the two matching distances together as the final matching distance between P and Q .

When calculating the matching distance between P_o and Q_o , we adopt the angular distance, which is defined as:

$$angD(P, Q) = \frac{\sum_{y=1}^{Rows} \sum_{x=1}^{Cols} G(P_o(x, y), Q_o(x, y))}{(J/2) \cdot S} \quad (5.9)$$

where S is the area of the code map, and

$$G(P_o(x, y), Q_o(x, y)) = \begin{cases} 1, & P_o(x, y) = 6 \text{ and } Q_o(x, y) \neq 6 \\ 1, & P_o(x, y) \neq 6 \text{ and } Q_o(x, y) = 6 \\ 0, & P_o(x, y) = Q_o(x, y) \\ \min(P_o(x, y) - Q_o(x, y), Q_o(x, y) - (P_o(x, y) - 6)), & \text{if } P_o(x, y) > Q_o(x, y) \text{ and } P_o(x, y) \neq 6 \\ \min(Q_o(x, y) - P_o(x, y), P_o(x, y) - (Q_o(x, y) - 6)), & \text{if } P_o(x, y) < Q_o(x, y) \text{ and } Q_o(x, y) \neq 6 \end{cases} \quad (5.10)$$

The matching distance between P_m and Q_m is defined as:

$$magD(P, Q) = \frac{\sum_{y=1}^{Rows} \sum_{x=1}^{Cols} abs(P_m(x, y) - Q_m(x, y))}{(N - 1) \cdot S} \quad (5.11)$$

Then, the final matching distance between P and Q can be fused from $angD$ and $magD$ as:

$$dist(P, Q) = (1 - \lambda) \cdot angD(P, Q) + \lambda \cdot magD(P, Q) \quad (5.12)$$

where λ is used to control the contribution of $magD$ to $dist$ and it is experimentally set as 0.15 in our system.

Taking into account the possible translations in the extracted ROI sub-image (with respect to the one extracted in the enrolment), multiple matches are performed by translating one set of features in horizontal and vertical directions. And in such case, S is the area of the overlapping parts of the two code maps. The minimum of the resulting matching distances is considered to be the final distance. The ranges of the horizontal translation and the vertical translation are empirically set as -8 to 8 and -4 to 4 in this chapter, respectively.

5.5 Experimental Results

5.5.1 Database Establishment

In order to evaluate the proposed FKP-based personal authentication system, a database was established using the developed FKP image acquisition device

(refer to Figs. 5.1 and 5.2). FKP images were collected from 165 volunteers, including 125 males and 40 females. Among them, 143 subjects were 20–30 years old and the others were 30–50 years old. The database will be available in the website of Biometrics Research Center, The Hong Kong Polytechnic University.

We collected samples in two separate sessions. In each session, the subject was asked to provide 6 images for each of the left index finger, the left middle finger, the right index finger and the right middle finger. Therefore, 48 images from 4 fingers were collected from each subject. In total, the database contains 7920 images from 660 different fingers. The average time interval between the first and the second sessions was about 25 days. The maximum and minimum time intervals were 96 days and 14 days respectively. In all of the following experiments, we took images collected at the first session as the gallery set and images collected at the second session as the probe set. To obtain statistical results, each test image in the probe set was matched with all the training images in the gallery set. If the test image and the training image were from the same finger, the matching between them was counted as a genuine matching; otherwise it was counted as an imposter matching.

5.5.2 Selection of the Image Resolution

The resolution of original FKP images acquired in our system is about 400 dpi, which may not be optimal in terms of the accuracy and efficiency of FKP verification. In fact, many factors, such as the storage space, the computational cost, the employed feature extraction and matching method, and the recognition accuracy, should be considered in selecting a suitable resolution of the FKP images for a more efficient biometric system. To this end, we conducted a series of experiments to select the “optimal” resolution and set the selection criterion as: the minimum resolution with which a satisfying verification performance could be obtained. The experiments were performed on a sub-dataset of the whole FKP database. In this sub-dataset, there were 120 classes, including 1440 images. With respect to the feature extraction method, the CompCode was used (Kong and Zhang 2004). We smoothed the original images by using a Gaussian kernel and then down-sampled the images to five lower resolutions: 200 dpi, 170 dpi, 150 dpi, 120 dpi and 100 dpi. The experimental results are summarized in Table 5.3.

Table 5.3 EERs obtained under different resolutions

Resolution (dpi)	EER (%)
200	1.73
170	1.41
150	1.36
120	1.71
100	1.92

Based on the results listed in Table 5.3, it can be seen that 150 dpi is a good choice. It leads to the lowest EER, while the resolution is much smaller than the original one (400 dpi). This will reduce the computational cost and speed up the feature extraction and matching processes significantly. Therefore, in all of the following experiments, we used the FKP images with a resolution 150 dpi.

5.5.3 FKP Verification

Verification aims to answer the question of “whether the person is the one he/she claims to be”. In order to show and explain the performance of the proposed system clearly, 3 experiments were conducted. In each experiment, we evaluated and compared the performance of five coding based feature extraction methods: CompCode (Kong and Zhang 2004), OrdinalCode (Sun et al. 2005), RLOC (Jia et al. 2008), BOCV (Guo et al. 2009) and the proposed ImCompCode&MagCode. In OrdinalCode, differences between Gaussians from two orthogonal directions were used to extract the orientation fields. The scales of the 2D Gaussian function along two orthogonal directions in OrdinalCode were set as 1.7 and 4.2 in our implementation. The RLOC method is based on the modified finite Radon transform. It was originally proposed for palmprint recognition (Jia et al. 2008), and was later adopted for feature extraction of the finger-back -surface images (Kumar and Zhou 2009a, b). In our implementation of RLOC, the “line width” was set as 2 and the kernel size was 16×16 . BOCV (Guo et al. 2009) is a recently proposed method for palmprint recognition which tends to represent multiple orientations for a local region. The threshold for the binarization used in BOCV was set as 0 in this chapter. Gabor filters used in CompCode, BOCV and the proposed ImCompCode&MagCode were all of the form (5.3), and the parameters were set as: $\delta = 3.3$ and $\sigma = 4$.

Experiment 1

In the first experiment, all classes of FKPs were involved. Each image in the probe set was matched against all the images in the gallery set. Therefore, in this experiment there were 660 (165×4) classes and 3960 (660×6) images in the gallery set and the probe set each. The numbers of genuine matchings and imposter matchings are 23,760 and 7,828,920, respectively. By adjusting the matching threshold, a ROC (Receiver Operating Characteristic) curve, which is a plot of Genuine Accept Rate (GAR) against False Accept Rate (FAR) for all possible thresholds, can be created. The ROC curve can reflect the overall performance of a biometric system. Figure 5.8a shows ROC curves generated by the five different FKP recognition schemes and Table 5.4 lists the corresponding EERs, from which we can see that the proposed ImCompCode&MagCode scheme performs the best among the five methods evaluated in terms of EER. Distance distributions of genuine matchings and imposter matchings obtained by the proposed scheme are plotted in Fig. 5.8b.

Fig. 5.8 (a) ROC curves obtained by the five recognition methods in experiment 1. (b) Distance distributions of genuine matchings and imposter matchings with the proposed scheme in experiment 1

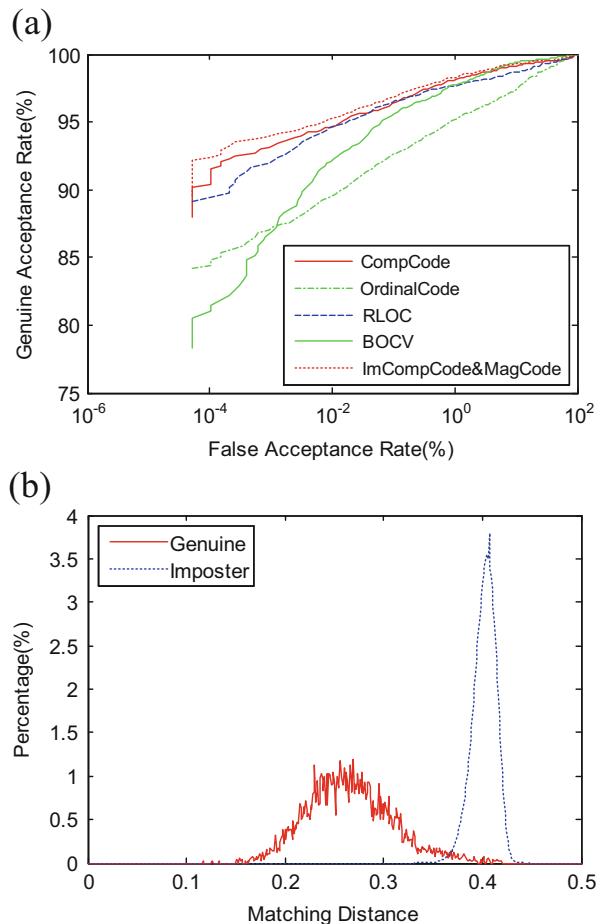


Table 5.4 EERs obtained by different methods in experiment 1

Method	EER (%)
CompCode	1.72
OrdinalCode	3.83
RLOC	1.93
BOCV	1.82
ImCompCode&MagCode	1.48

Experiment 2

As mentioned in Sect. 5.1, our database contains FKP s from four types of fingers, left index fingers, left middle fingers, right index fingers and right middle fingers. The aim of this experiment is to evaluate the performance of the proposed FKP-based personal authentication system on each type of fingers separately. For

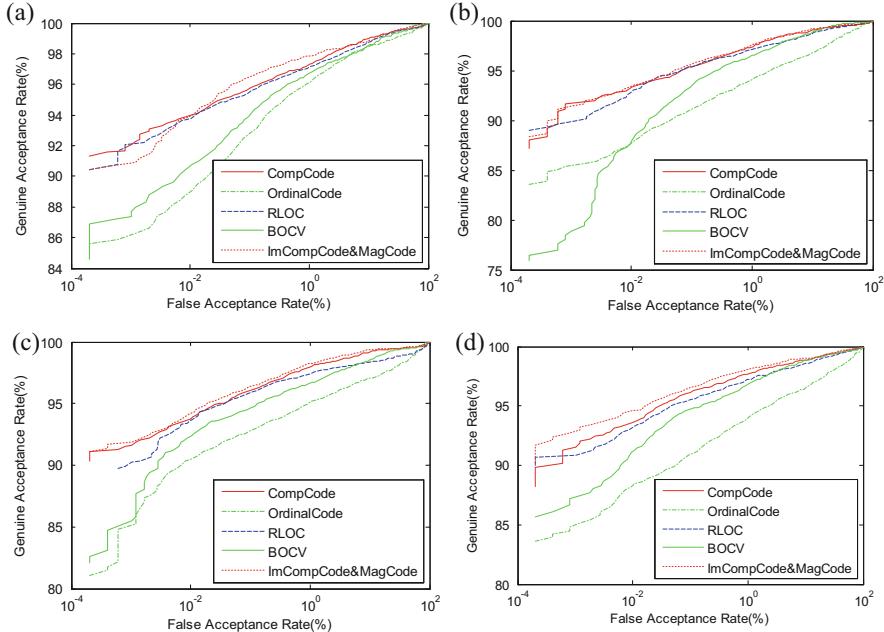


Fig. 5.9 ROC curves for FKPs from (a) left index fingers, (b) left middle fingers, (c) right index fingers, and (d) right middle fingers

Table 5.5 EERs (%) by different schemes in experiment 2

Finger type	CompCode	OrdinalCode	RLOC	BOCV	ImCompCode&Magcode
Left index	2.06	2.69	2.20	2.45	1.73
Left middle	1.96	4.07	2.27	2.42	1.78
Right index	1.82	3.66	2.07	2.43	1.44
Right middle	1.87	4.15	2.32	2.30	1.64

each type of FKPs, the gallery and the probe each contains 165 classes and 990 (165×6) sample images, and the numbers of genuine matchings and imposter matchings are 5940 and 487,080 respectively. Similarly, five FKP recognition schemes were evaluated. ROC curves for different finger types and by different recognition schemes are shown in Fig. 5.9. Experimental results in terms of EER are summarized in Table 5.5 for comparison.

The experimental results indicate that in general the right middle and index fingers perform better than their left counterparts in terms of EER. This is probably because that the majority of people who provided FKP samples in our database are right-handed. They would feel more convenience to use our image acquisition device with their right hand than with the left one, which consequently leads to a result that the variations of finger poses of right hand fingers are less severe than left hand fingers. Remarkable variations of finger poses would cause severe affine

transforms and deformations between two FKP images of the same finger, which lead to more challenge to FKP recognition.

Experiment 3

The goal of this experiment is to investigate the system's performance when we fuse information from 2 or more fingers of a person. In fact, in such a case the system works as a kind of multi-modal system with a single biometric trait but multiple units (Ross and Jain 2004). Suppose that we want to fuse information from the right index FKP and the right middle FKP. Each template in the enrolment database will be composed by a right index FKP t_{ri} and a right middle FKP t_{rm} . When matching, a client's right index FKP c_{ri} and right middle FKP c_{rm} will be matched to t_{ri} and t_{rm} respectively to get two matching distances, d_{ri} and d_{rm} . Then d_{ri} and d_{rm} will be fused according to some fusion rules to obtain the final matching distance, by which the client's identity can be identified.

With respect to fusion rules, in this chapter we simply examined the SUM rule and the MIN rule as they are easily to be implemented and can well reflect the system's performance. The SUM rule is defined as:

$$d_{sum} = \sum d_i \quad (5.13)$$

and the MIN rule is defined as:

$$d_{min} = \min(d_i) \quad (5.14)$$

where d_i is the matching distance of the client's i^{th} finger.

We tested several different fusion schemes of fingers with the two fusion rules. Results are presented in Table 5.6, from which it can be easily observed that by integrating information from more fingers the recognition performance of the system could be largely improved. We can also find that the SUM rule works better than the MIN rule in our system.

5.5.4 Speed

The FKP recognition software is implemented using Visual C#.Net 2005 on a Dell Inspiron 530s PC embedded Intel E6550 processor and 2GB of RAM. The execution time for data preprocessing and ROI extraction is 198 ms. The time for ImCompCode&MagCode-based feature extraction and matching is 75 ms and 1.2 ms, respectively. Thus, the total execution time for one verification operation is less than 0.5 s in our prototype system, which is fast enough for real-time applications. We believe that with the optimization of the implementation, the system's efficiency could be much further improved.

Table 5.6 EERs (%) obtained in experiment 3

Fingers in fusion	CompCode		OrdinalCode		RLOC		BOCV		ImCompCode& MagCode	
	S-rule	M-rule	S-rule	M-rule	S-rule	M-rule	S-rule	M-rule	S-rule	M-rule
Left index	0.33	0.67	0.64	0.72	0.26	0.68	0.41	0.45	0.20	0.63
Left middle										
Right index	0.32	0.39	0.71	0.79	0.34	0.37	0.36	0.47	0.26	0.36
Right middle										
Left index	0.36	0.67	0.84	0.76	0.42	0.91	0.63	0.63	0.26	0.64
Right index										
Left middle	0.29	0.33	0.90	0.87	0.33	0.30	0.43	0.42	0.27	0.30
Right middle										
All the four	0	0.03	0.02	0.05	0	0.16	0.01	0.09	0	0.02

Bold value represents best results

5.5.5 Discussion

FKP is a new member in the biometrics family compared with other biometric identifiers. As mentioned in the section of Introduction, Woodard and Flynn (Woodard and Flynn 2005a, b) did some salient work to validate the uniqueness of features from finger-back surfaces by using 3D imaging, and Kumar (Ravikanth and Kumar 2007; Kumar and Ravikanth 2009) integrated 2D finger knuckle surface information with the finger shape information in their system. However, such a system is doomed to have a relatively large size because they need to collect the image of the whole hand. Complex preprocessing steps are also needed to extract the finger knuckle area which only occupies a small portion of the whole acquired image. In our system, we make use of a triangular block to control the finger freedom. This gadget does not sacrifice the user convenience and it is easy to use. Such a design brings the following merits: (1) the acquisition device could be easily made to a small size; (2) image around the finger knuckle area is captured directly, which largely simplifies the following data preprocessing steps; and (3) since the finger knuckle is slightly bent when being captured, the distinctive FKP texture patterns can be clearly imaged, which makes the proposed FKP system have high accuracy.

Advantages of the proposed system could be reflected by experimental results on FKP verification. For comparison, experimental results in (Woodard and Flynn 2005a, b) and (Kumar and Ravikanth 2009) are extracted from the original chapters and listed in Table 5.7 and partial experimental results by our system are also presented. The scale of the dataset used in our experiment is much larger than the ones mentioned in (Woodard and Flynn 2005a, b) and (Kumar and Ravikanth 2009). Woodard's result (Woodard and Flynn 2005a, b) and Kumar's result

Table 5.7 Comparison of experimental results

Method	Gallery classes	Gallery samples	Probe classes	Probe samples	Finger types for fusion	EER (%)
Woodard and Flynn (2005a, b)	132	660	177	531	r-index, r-middle, r-ring	5.5
Kumar and Ravikanth (2009)	105	420	105	210	index, middle, ring, little	1.39
Ours	165	990	165	990	r-index, r-middle	0.26
Ours	165	990	165	990	r-index, r-middle, l-index, l-middle	0

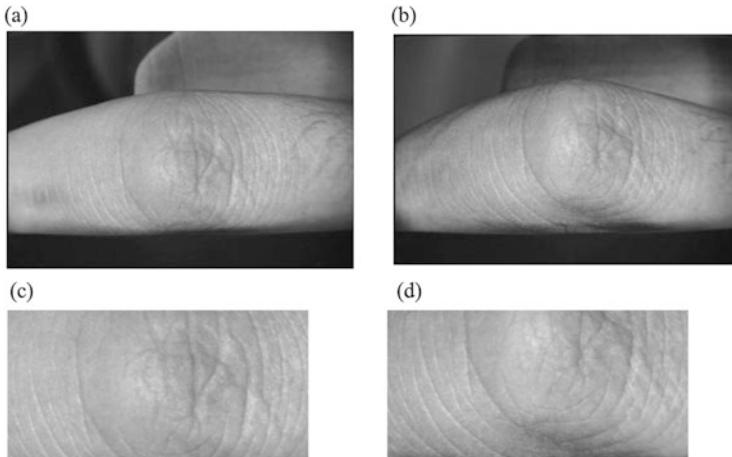


Fig. 5.10 (a) and (b) are two intra-class FKP images captured by our system. (c) and (d) are their ROI sub-images. There is an obvious pose variation between the two FKPs and they are recognized as different classes in our system

(Kumar and Ravikanth 2009) were obtained by fusing information from three and four fingers, respectively. It can be clearly seen that the proposed system performs much better even if we incorporate information from fewer fingers. Particularly, if four fingers are used, our system could achieve an EER of 0.

It should be noted that although we use a triangular block to control the finger freedom in FKP image acquisition, there are still variations for the same fingers at different collection sessions. Sometimes such variations can result in severe affine transforms or even non-elastic deformations among intra-class FKPs. And as a result, feature maps of such FKPs can have large matching distances. Figure 5.10 shows a typical example. The two FKP images in Fig. 5.10 are from the same finger but are recognized as different classes in our system. Hence, how to reduce the effects of affine transforms and deformations will be a direction of our future work.

There are two main approaches currently employed in To F technology. The first one utilizes modulated, incoherent light, and is based on a phase measurement. The second one is based on an optical shutter technology. The main details about the technology will be showed in the chapters that introduce the 3D face verification system.

5.6 Summary

This chapter presents a new approach to online personal authentication using finger-knuckle-print (FKP), which has distinctive line features. A cost-effective FKP system, including a novel image acquisition device and the associated data processing algorithms, is developed. A region of interest (ROI) extraction algorithm is proposed to align and extract the FKP sub-image for feature extraction. For efficient FKP matching, a feature extraction scheme is proposed to exploit both orientation and magnitude information extracted by Gabor filters. To evaluate the performance of the proposed system, an FKP database is established, consisting of 7920 images from 660 different fingers. Extensive experiments are conducted and promising results demonstrate the efficiency and effectiveness of the proposed technique. Compared with other existing finger back surface based systems, the proposed one has merits of high accuracy, high speed, small size and cost-effective. It has a great potential to be future improved and employed in real commercial applications. In the future, we will focus on how to deal with affine or even non-elastic deformations between FKP images from the same finger to further improve the system's performance.

References

- Borgen H, Bours P, Wolthusen S (2008) Visible-spectrum biometric retina recognition. In: Proceedings of the international conference on intelligent information hiding and multimedia signal processing, pp 1056–1062
- Burge M, Burger W (1999) Ear. In: Jain AK, Bolle R, Pankanti S (eds) Biometrics: personal identification in networked society. Kluwer Academic, Boston, pp 273–286
- Daugman J (1993) High confidence visual recognition of persons by a test of statistical independence. *IEEE Trans Pattern Anal Mach Intell* 15(11):1148–1161
- Daugman J (2004) How iris recognition works. *IEEE Trans Circuits Syst Video Technol* 14 (1):21–30
- Delac K, Grgic M (2007) Face recognition. I-Tech Education and Publishing, Vienna
- E-channel (2004) E-channel system of the Hong Kong government. <http://www.immd.gov.hk/ehtml/20041216.htm>
- Gabor D (1946) Theory of communication. *J Inst Electr Eng* 93:429–457
- Guo Z, Zhang D, Zhang L, Zuo W (2009) Palmprint verification using binary orientation co-occurrence vector. *Pattern Recogn Lett* 30(13):1219–1227

- Hill R (1999) Retinal identification. In: Jain A, Bolle R, Pankati S (eds) *Biometrics: personal identification in networked society*. Springer, Berlin
- Hollien H (2002) Forensic voice identification. Academic Press, London
- Huang D, Jia W, Zhang D (2008) Palmprint verification based on principal lines. *Pattern Recogn* 41(4):1316–1328
- Jain A, Duta N (1999) Deformable matching of hand shapes for verification. *Proc ICIP* 1999:857–861
- Jain A, Ross A, Pankanti S (1999) A prototype hand geometry-based verification system. In: *Proceedings of the 2nd international conference on audio- and video-based biometric person authentication*, pp 166–171
- Jain A, Flynn P, Ross A (2007) *Handbook of biometrics*. Springer, New York
- Jia W, Huang D, Zhang D (2008) Palmprint verification based on robust line orientation code. *Pattern Recogn* 41(5):1504–1513
- Kong A (2008) An evaluation of Gabor orientation as a feature for face recognition. *Proc ICPR* 2008:1–4
- Kong A, Zhang D (2004) Competitive coding scheme for palmprint verification. In: *Proceedings of the ICPR'04*, pp 520–523
- Kong A, Zhang D, Kamel M (2006) Palmprint identification using feature-level fusion. *Pattern Recogn* 39(3):478–487
- Kumar A, Prathyusha K (2008) Personal authentication using hand vein triangulation. *Proc SPIE Biom Technol Hum Identif* 6944:13
- Kumar A, Ravikanth C (2009) Personal authentication using finger knuckle surface. *IEEE Trans Inf Forensics Secur* 4(1):98–109
- Kumar A, Zhou Y (2009a) Human identification using knucklecodes. *Proc BTAS* 2009:1–6
- Kumar A, Zhou Y (2009b) Personal identification using finger knuckle orientation features. *Electron Lett* 45(20):1023–1025
- Lee T (1996) Image representation using 2D Gabor wavelet. *IEEE Trans Pattern Anal Mach Intell* 18(10):957–971
- Li Q, Qiu Z, Sun D, Wu J (2004) Personal identification using knuckleprint. In: *Proceedings of Sino Biometrics*, pp 680–689
- Liu C, Herbst N, Anthony N (1979) Automatic signature verification: system description and field test results. *IEEE Trans Syst Man Cybren* 9(1):35–38
- Maltoni D, Maio D, Jain A, Prabhakar S (2003) *Handbook of fingerprint recognition*. Springer, New York
- Nanni L, Lumini A (2009a) A multi-matcher system based on knuckle-based features. *Neural Comput Appl* 18(1):87–91
- Nanni L, Lumini A (2009b) On selecting Gabor features for biometric authentication. *Int J Comput Appl Technol* 35(1):23–28
- Nixon M, Tan T, Chellappa R (2006) *Human identification based on Gait*. Springer, New York
- Plamondon R, Lorette G (1989) Automatic signature verification and writer identification- the state of the art. *Pattern Recogn* 22(2):107–131
- Ratha N, Bolle R (2004) *Automatic fingerprint recognition systems*. Springer, New York
- Ravikanth C, Kumar A (2007) Biometric authentication using finger-back surface. *Proc CVPR* 2007:1–6
- Ross A, Jain A (2004) Multimodal biometrics: an overview. In: *Proceedings of the 12th European Signal Processing Conference*, pp 1221–1224
- Sanchez-Reillo R, Sanchez-Avila C, Gonzalez-Marcos A (2000) Biometric identification through hand geometry measurements. *IEEE Trans Pattern Anal Mach Intell* 22(10):1168–1171
- Sun Z, Tan T, Wang Y, Li S (2005) Ordinal palmprint representation for personal identification. *Proc CVPR* 2005:279–284
- Wang J, Yau W, Suwandy A, Sung E (2008) Personal recognition by fusing palmprint and palm vein images based on “Lapacianpalm” representation. *Pattern Recogn* 41(5):1531–1544

- Wechsler H (2006) Reliable face recognition methods – system design, implementation and evaluation. Springer, New York
- Woodard D, Flynn P (2005a) Finger surface as a biometric identifier. *Comput Vis Image Underst* 100(3):357–384
- Woodard D, Flynn P (2005b) Personal identification utilizing finger surface features. *Proc CVPR* 2:1030–1036
- Zhang D, Kong W, You J, Wong M (2003) Online palmpprint identification. *IEEE Trans Pattern Anal Mach Intell* 25(9):1041–1050

Chapter 6

Local Features for Finger-Knuckle-Print Recognition

Abstract Researchers have recently found that the finger-knuckle-print (FKP), which refers to the inherent skin patterns of the outer surface around the phalangeal joint of one's finger, has high discriminability, making it an emerging promising biometric identifier. Effective feature extraction and matching plays a key role in such an FKP based personal authentication system. This chapter studies image local features induced by the phase congruency model, which is supported by strong psychophysical and neurophysiological evidences, for FKP recognition. In the computation of phase congruency, the local orientation and the local phase can also be defined and extracted from a local image patch. These three local features are independent of each other and reflect different aspects of the image local information. We compute efficiently the three local features under the computation framework of phase congruency using a set of quadrature pair filters. We then propose to integrate these three local features by score-level fusion to improve the FKP recognition accuracy. Such kinds of local features can also be naturally combined with Fourier transform coefficients, which are global features. Experiments are performed on the PolyU FKP database to validate the proposed FKP recognition scheme.

Keywords Biometrics • Finger-knuckle-print recognition • Phase congruency

6.1 Introduction

The need for reliable automated user authentication techniques has been significantly increased in the wake of heightened concerns about security (Jain et al. 2007). Biometrics based methods, which use unique physical or behavioral characteristics of human beings, are drawing increasing attention in both academic research and industrial applications because of their high accuracy and robustness in the modern e-world. In the past decades, researchers have exhaustively investigated a number of different biometric identifiers, including fingerprint, face, iris, palm print, hand geometry, voice, and gait, etc. (Li 2009).

Among various kinds of biometric identifiers, hand-based biometrics attracts much interest because of their high user acceptance and convenience. Some

commonly used hand-based biometrics, e.g., fingerprint (Hong et al. 1998; Ross et al. 2003; Maltoni et al. 2003; Zhao et al. 2010), palm print (Han et al. 2003; Hu et al. 2007; Zhang et al. 2003, 2012; Kong and Zhang 2004; Kong et al. 2006, Kong et al. 2009; Jia et al. 2008; Struc and Pavesic 2009; Jain and Feng 2009), hand geometry (Feng et al. 2011; Duta 2009), and hand vein (Wang et al. 2008; Wilson 2010), have been well investigated in the literature. Recently, scholars have reported that finger-knuckle-print (FKP), the image pattern of skin folds and creases in the outer finger knuckle surface, is highly unique and can serve as a distinctive biometric identifier (Zhang et al. 2009a, b, 2010, 2011; Morales et al. 2011; Woodard and Flynn 2005; Kumar and Ravikanth 2009; Meraoumia et al. 2011). Compared with fingerprint, FKP is hard to be abraded since people hold stuffs with the inner side of the hand. In addition, unlike fingerprint, there is no stigma of criminal investigation associated with the finger knuckle surface, so FKP can have a higher user acceptance rate (Kumar and Ravikanth 2009). Moreover, people rarely leave FKP remains on the stuff surface, making the loss of private data less possible. Thus, FKP has a great potential to turn into a widely accepted biometric identifier.

A novel online FKP-based personal authentication system has been established in our previous works (Zhang et al. 2010), which comprises four major components: FKP image acquisition, ROI (region of interest) extraction, feature extraction, and feature matching. In our design, the finger knuckle will be slightly bent when being imaged, and hence the inherent skin patterns can be clearly captured.

As in many pattern classification tasks, feature extraction and matching plays a key role in FKP-based personal authentication system. In (Zhang et al. 2009a, b), Zhang et al. used the Gabor filter based competitive coding scheme, which was originally designed for palmprint recognition (Kong and Zhang 2004), to extract the local orientation information as FKP features. In (Zhang et al. 2010), Zhang et al. combined the orientation information and the magnitude information extracted by Gabor filters. In (Zhang et al. 2009a, b), the Fourier transform coefficients of the image were taken as the feature and the band-limited phase-only correlation technique was employed to calculate the similarity between two FKP images. In the local-global information combination (LGIC) feature extraction scheme (Zhang et al. 2011), the local orientation extracted by the Gabor filters is taken as the local feature while the Fourier coefficients are taken as a global feature. In (Morales et al. 2011), Morales et al. used a real Gabor filter to enhance the FKP image and then used the scale invariant feature transform (SIFT) to extract features; they called the proposed method as OE-SIFT (orientation enhanced-SIFT).

In our previous methods (Zhang et al. 2009a, b, 2010), real Gabor filters were used to extract the local orientation information; such an idea was inspired by the method “competitive coding” (Kong and Zhang 2004) proposed for palmprint recognition. Local orientation feature of biometric images can also be defined and extracted using other different mathematical models. For example, in (Jia et al. 2008), Jia et al. proposed a coding method to extract the local orientation of palmprints, namely robust line orientation code (RLOC), which is based on a modified finite Radon transform. In addition to the local orientation, the local phase is also widely used in the biometrics community and it is usually extracted by using band-pass complex filters, e.g., Gabor filters (Gabor 1946; Daugman 1985)

and log-Gabor filters (Field 1987). By making use of the local phase feature extracted by Gabor filters, Daugman invented the famous IrisCode (Daugman 1993); inspired by Daugman's work, Zhang et al. adopted a similar idea to match palmprint images (Zhang et al. 2003). Actually, according to (Venkatesh and Owens 1990; Sierra-Vázquez and Serrano-Pedraza 2010), the local phase reflects the type of local features. However, it is not clear whether such a feature is significant and stable. To address such an issue, we need to know whether the local phases over scales are consistent. The phase congruency (PC) model (Morrone et al. 1986; Morrone and Burr 1988; Kovesi 1999; Henriksson et al. 2009) serves as a solution to this issue. Studies of psychophysics and neurophysiology have revealed that visually discernable image features coincide with those points where Fourier waves at different frequencies have congruent phases. PC has been exploited as features by some biometrics researchers for face recognition (Gundimada and Asari 2007), iris recognition (Yuan and Shi 2005), and palmprint recognition (Struc and Pavesic 2009), and it has also been used in some object recognition applications (Verikas et al. 2012).

In fact, local orientation, local phase, and local phase congruency reflect different aspects of information embedded in a local image patch. Moreover, they are independent of each other and none of them can be covered by the others. They can provide complementary discriminating power to each other for matching biometric images. Thus, better recognition performance could be expected by combining these three local features together in some way. However, to the best of our knowledge, in the biometrics community there is no work reported to define and analyze systematically these three local features in a unified framework, and there is no attempt trying to integrate these three features to improve the performance of biometrics systems, either. Based on these considerations, in this chapter, we first define these three local features under a unified framework, and then present an efficient method to compute them using the computation framework of PC. Finally, we integrate these three local features together for FKP recognition. Experimental results demonstrate that the integration of the three local features performs better than using any of them separately. Moreover, we report the system's performance when integrating the three local features with one global feature, the Fourier transform coefficients, which leads to the best result on our benchmark FKP database. This work differs from our previous works (Zhang et al. 2010, 2011) mainly in three aspects. At first, besides the local orientation, the local phase and the local phase congruency are investigated. Secondly, the three local features are defined, analyzed, and extracted in a unified framework. And thirdly, we propose to integrate the three local features together to improve the accuracy of FKP recognition.

The remainder of this chapter is organized as follows. Section 6.2 defines and analyzes the three local features, while Sect. 6.3 presents the extraction and matching scheme for each local feature. Section 6.4 reports the experimental results and discussions. Finally, Sect. 6.5 gives the summary.

6.2 Analysis of Local Features

As stated in Sect. 6.1, in literature the three local features (the local orientation, the local phase, and the phase congruency) are extracted by using different mathematical models and their relationships are not systematically investigated. In this section, we will examine these three local features in detail under a unified framework.

To ease the following discussions, we first introduce the concept of intrinsic dimension here. The intrinsic dimension is the number of degrees of freedom necessary to describe a local image structure (Krieger and Zetsche 1996). A 2D image patch I can be classified as a local region, denoted by R , of a specific intrinsic dimension. For example, constant areas are of intrinsic dimension zero (i0D) while straight lines and edges are of intrinsic dimension one (i1D). Mathematically, such a classification can be expressed as (Krieger and Zetsche 1996)

$$I \in \begin{cases} i0D_R, I(\mathbf{x}_i) = I(\mathbf{x}_j), \forall \mathbf{x}_i, \mathbf{x}_j \in R \\ i1D_R, I(x, y) = g(x \cos \theta + y \sin \theta), \forall (x, y) \in R, I \notin i0D_R \\ i2D_R, \text{else} \end{cases} \quad (6.1)$$

where g is a 1D real-valued function. Examples of i0D, i1D, and i2D signals are shown in Fig. 6.1.

A point x in an image can be characterized by its “local features”, which are derived from a local patch centered on it. Before we define local features we need to have a model for the signal to be analyzed. In our case, we are dealing with 2D FKP images, which are actually a special kind of 2D images in that they are abundant of line-like features. And these line-like features play dominant roles in distinguishing different individuals. Thus, we assume that FKP images are locally i1D (intrinsic one dimensional) signals.

Let us consider the one dimensional (1D) real signal first. In order to analyze the local structure of the 1D real signal, analytic signal was proposed in the literature (Gabor 1946) and it has been corroborated to be quite effective (Granlund and Knutsson 1995). Analytic representation makes certain attributes of a real signal more accessible and facilitates the derivation of modulation and demodulation

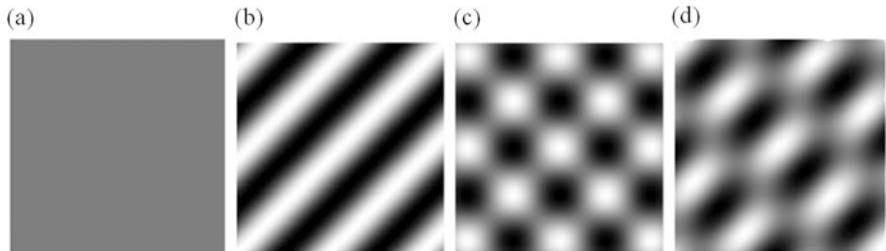


Fig. 6.1 Signals with different intrinsic dimensions: (a) i0D; (b) i1D; (c) i2D; (d) i2D

techniques. Given a 1D real signal $f(x)$, the corresponding analytic signal is defined as (Gabor 1946; Granlund and Knutsson 1995)

$$f_A(x) = f(x) + if_H(x) \quad (6.2)$$

Where $f_H(x) = f(x) * h(x)$, $i^2 = -1$, $*$ denotes the convolution operation, and $h(x) = 1/\pi x$ refers to the Hilbert transform kernel in the spatial domain. With such a complex representation, the local amplitude and the local phase of the 1D analytic signal are defined as (Gabor 1946; Granlund and Knutsson 1995)

$$a(x) = \sqrt{f^2(x) + f_H^2(x)}, \phi(x) = \arctan 2(f_H(x), f(x)) \quad (6.3)$$

The local amplitude indicates the energetic information of the signal, while the local phase can be used to distinguish between different local structures and it is independent of the local amplitude (Venkatesh and Owens 1990). In practice, since the Hilbert transform operator is an improper integral and difficult to calculate, researchers usually use a pair of spatial filters forming a quadrature pair to construct the analytic signal (Venkatesh and Owens 1990; Morrone and Owens 1987). To this end, complex Gabor (Gabor 1946; Daugman 1985) or log-Gabor (Field 1987) filters are widely used. When the 1D signal is embedded into the 2D space, its orientation should be considered. Thus, the local amplitude, the local phase, and the local orientation are three independent measures to characterize a 2D image point.

The local phase reflects the type of local structures (Venkatesh and Owens 1990). However, we do not know to what degree it is a significant feature. To address such an issue, we make use of the phase congruency (PC) (Morrone et al. 1986; Morrone and Burr 1988; Kovesi 1999), a dimensionless quantity, to measure the consistency of the local phases over scales. Based on the physiological and psychophysical evidence, it is found that visually discernable features coincide with those points having maximal phase congruency. Such a conclusion has been further corroborated by some recent studies in neurobiology using functional magnetic resonance imaging (fMRI) (Henriksson et al. 2009). Phase congruency has an intriguing property that it is almost invariant to changes in image brightness or contrast.

Thus, within the local window surrounding an image point x , four features—the local amplitude, the local phase, the local orientation and the phase congruency—can be extracted and they reflect different aspects of information contained in the local window. However, we will not use the local amplitude for recognition because it is not contrast invariant. Hence, the local phase, the local orientation, and the phase congruency will be used as three local features in this article.

For a real 2D image, these three local features can be defined and extracted using a set of 2D quadrature filter pairs, such as 2D complex Gabor or log-Gabor filters. Suppose that complex Gabor filters are adopted, which are defined as

$$G(x, y) = \exp\left(-\frac{1}{2}\left(\frac{x'^2}{\sigma_x^2} + \frac{y'^2}{\sigma_y^2}\right)\right) \exp\left(i\frac{2\pi}{\lambda}x'\right) \quad (6.4)$$

where $x' = x \cos \theta + y \sin \theta$, $y' = -x \sin \theta + y \cos \theta$. In Eq. (6.4), λ represents the wavelength of the sinusoid factor, θ represents the orientation of the normal to the parallel stripes of the Gabor function, σ_x and σ_y are the standard deviations of the 2D Gaussian envelop. It can be seen from the definition that a Gabor filter is actually a Gaussian envelop modulated by a sinusoidal plane wave. The Gaussian envelop ensures that the convolution is dominated by the image patch near the center of the filter. Therefore, the Gabor filter is a local operator and can extract information at a specific scale and a specific orientation within a local region.

To define and extract the local orientation, we make use of the competitive coding scheme which has been successfully used for palmprint (Kong and Zhang 2004) and FKP (Zhang et al. 2009a, b, 2010) recognition. Competitive coding scheme assumes that every image pixel resides on a negative “line” and it extracts the orientation of the line by using a set of real Gabor filters with different orientations. Specifically, the orientation along which the Gabor responses get the minimum is taken as the feature at this point. Denote by G_R (G_I) the real (imaginary) part of the Gabor filter G . With a series of G_R s sharing the same parameters, except the parameter of orientation, the local orientation of the image I at the position (x, y) can be extracted. Mathematically, the local orientation is defined as

$$\text{ori}(x, y) = \operatorname{argmin}_j \{I(x, y)*G_R(x, y, \theta_j)\} \quad (6.5)$$

where $\theta_j = j\pi/J$, $j = \{0, \dots, J-1\}$. J represents the number of orientations. It needs to be noted that theoretically speaking, the local orientation of ideal i1D 2D image signals can be accurately extracted by the Riesz transform-based monogenic signal model, which is a 2D extension of the classical 1D analytic signal (Felsberg and Sommer 2001; Wietzke and Sommer 2010); however, for real 2D images, multi-dimensional even-symmetric filters usually perform better for this task (Kong and Zhang 2004; Zhang et al. 2010).

The extraction of PC using quadrature pair filters will be presented in Sect. 6.3.1 in detail. Actually, PC is a 1D concept. For 2D images, we can compute PC_{θ_j} along different orientations $\{\theta_j : |j=0 \sim J-1\}$. Then the maximum of $\{PC_{\theta_j} : |j=0 \sim J-1\}$ can be taken as the PC value at the examined position:

$$PC_2(x, y) = \max_i \{PC_{\theta_j}(x, y) : |j=0 \sim J-1\} \quad (6.6)$$

We denote by θ_m the orientation along which the 1D PC takes the maximum. Then, we can apply Gabor filtering along θ_m and define the local phase as:

$$\text{phase}(x, y) = \arctan 2(I(x, y)*G_I(x, y, \theta_m), I(x, y)*G_R(x, y, \theta_m)) \quad (6.7)$$

6.3 Extraction and Matching of Local Features

In Sect. 6.2, we have defined and analyzed three local features. In practice, for the reason of computational efficiency, we do not compute the three local features separately. Instead, we present a scheme based on the computational framework of PC in (Kovesi 1999) to extract those features more efficiently. So, in the following sub-sections, the PC computation will be described first.

6.3.1 Step 1: Phase Congruency (PC)

Rather than assume a feature is a point of sharp changes in intensity, the PC model postulates that features are perceived at points where the Fourier components are maximal in phase (Morrone et al. 1986; Morrone and Burr 1988; Kovesi 1999). Phase congruency can be considered as a dimensionless measure for the significance of a structure independently of the signal amplitude. The technique to calculate PC is based on Kovesi's salient work (Kovesi 1999).

We start from the 1D signal $f(x)$. Denote by M_n^e and M_n^0 the even-symmetric and odd-symmetric filters at scale n and they form a quadrature pair. Responses of each quadrature pair to the signal will form a response vector at position x and on scale n :

$$[e_n(x), o_n(x)] = [f(x)*M_n^e, f(x)*M_n^0] \quad (6.8)$$

The local amplitude on scale n is given by

$$A_n(x) = \sqrt{e_n^2(x) + o_n^2(x)} \quad (6.9)$$

and the local phase is given by

$$\phi_n(x) = \arctan 2(o_n(x), e_n(x)) \quad (6.10)$$

These response vectors form the basis of our localized representation of the signal and the PC can be derived from them.

Let $F(x) = \sum_n e_n(x)$ and $H(x) = \sum_n o_n(x)$. Then, the 1-D PC can be computed as

$$PC(x) = \frac{E(x)}{\varepsilon + \sum_n A_n(x)} \quad (6.11)$$

where $E(x) = \sqrt{F^2(x) + H^2(x)}$ and ε is a small positive constant. We can also define the local phase as

$$Phase(x) = \arctan 2(H(x), F(x)) \quad (6.12)$$

Actually, it is the average local phase over n scales.

For 2D images, we have to apply the 1D analysis over several orientations and combine the results in some way. In such case, 2D filters with the orientation selection property can be used, such as the Gabor filters (Gabor 1946; Daugman 1985) or log-Gabor filters (Field 1987). Let $\theta_j = j\pi/J, j = \{0, \dots, J - 1\}$, denote the orientation angle of the filter, where J is the number of orientations. By modulating n and θ_j and convolving with the 2D image, we can get a set of responses at each image point x as

$$[e_{n,\theta_j}(\mathbf{x}), o_{n,\theta_j}(\mathbf{x})] \quad (6.13)$$

The local amplitude of point x on scale n and along orientation θ_j is given by

$$A_{n,\theta_j}(\mathbf{x}) = \sqrt{e_{n,\theta_j}(\mathbf{x})^2 + o_{n,\theta_j}(\mathbf{x})^2} \quad (6.14)$$

The local energy along orientation θ_j is given by

$$E_{\theta_j}(\mathbf{x}) = \sqrt{F_{\theta_j}(\mathbf{x})^2 + H_{\theta_j}(\mathbf{x})^2} \quad (6.15)$$

Where $F_{\theta_j}(\mathbf{x}) = \sum_n e_{n,\theta_j}(\mathbf{x})$ and $H_{\theta_j}(\mathbf{x}) = \sum_n o_{n,\theta_j}(\mathbf{x})$. Then, the phase congruency along orientation θ_j is computed by

$$PC_{\theta_j}(\mathbf{x}) = \frac{E_{\theta_j}(\mathbf{x})}{\varepsilon + \sum_n A_{n,\theta_j}(\mathbf{x})} \quad (6.16)$$

The average local phase along orientation θ_j is defined as

$$Phase_{\theta_j}(\mathbf{x}) = \arctan 2(H_{\theta_j}(\mathbf{x}), F_{\theta_j}(\mathbf{x})) \quad (6.17)$$

We define the 2D PC at x as

$$PC_2(\mathbf{x}) = \max_j PC_{\theta_j}(\mathbf{x}) \quad (6.18)$$

It should be noted that $PC_2(\mathbf{x})$ is a real number within 0–1.

6.3.2 Local Feature Extraction and Coding

In this section, we present the extraction and coding algorithm for each local feature. The local orientation and the local phase can be efficiently extracted by using the intermediate results of the PC computation.

Having obtained two raw PC maps of two images, we do not match them directly. Instead, we quantize them to several levels and then code them into integers. In practice, such a scheme can have three advantages: (a) it can save a lot of storage space; (b) for recognition, it works more robustly than using raw PC maps; and (c) it allows a fast matching of two maps. Therefore, we quantize PC into L levels and define the PC code as

$$pcCode(\mathbf{x}) = \left\lfloor \frac{PC_2(\mathbf{x})}{1/L} \right\rfloor \quad (6.19)$$

Where $\lfloor x \rfloor$ is the operator to return the largest integer not bigger than x . It is easy to see that each pcCode is an integer within $0 \sim L - 1$.

Although there are other kinds of methods to evaluate the local phase feature and the local orientation feature, we want to make a full use of the intermediate results in the process of computing PC in order to reduce the computational cost. It is easy to see that when calculating PC, we can get responses from a set of even-symmetric and odd-symmetric quadrature filters at different scales and different orientations. We can compute the local orientation and the local phase directly from them. For the local orientation evaluation, we borrow the idea from the competitive coding scheme (Kong and Zhang 2004; Zhang et al. 2009a, b, 2010). With the responses from the even-symmetric filters at a certain scale ζ , e.g. $\{e_{\zeta, \theta_j}(x) : j = 0, \dots, J - 1\}$, the orientation code at x can be defined as

$$oriCode(\mathbf{x}) = \operatorname{argmin}_j \{e_{\zeta, \theta_j}(\mathbf{x})\}, j = 0, \dots, J - 1 \quad (6.20)$$

Obviously, each orientation code $oriCode(x)$ is an integer within $0 \sim J - 1$.

Refer to Eq. (6.18), by our definition the 2D PC is actually the maximum of the 1D PCs along different orientations. We denote by θ_m the orientation along which the 1D PC takes the maximum value. Then, we can take the average local phase along θ_m as the local phase at x . That is

$$LP(\mathbf{x}) = Phase_{\theta_m}(\mathbf{x}) \quad (6.21)$$

The range of LP is $[0, 2\pi]$. Once again, we do not need the exact local phase angle. Instead, we quantize LP into several discrete levels to get the “phase code” as

$$phaCode(\mathbf{x}) = \lfloor LP(\mathbf{x})/(2\pi/M) \rfloor \quad (6.22)$$

where M is the number of quantization levels. Thus, each phase code is an integer within $0 \sim M - 1$.

Finally, for a given image, we can get its three code maps: pcCode, oriCode, and phaCode. Examples of them are shown in Fig. 6.2.

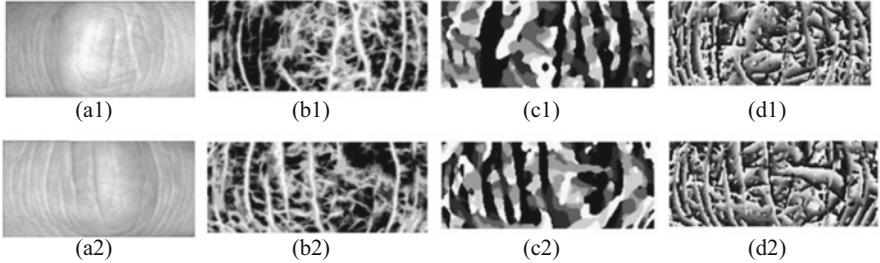


Fig. 6.2 Examples of local feature maps. (a1) and (a2) are the original FKP ROI images; (b1) and (b2) are the corresponding pcCode maps; (c1) and (c2) are the corresponding oriCode maps; (d1) and (d2) are the corresponding phaCode maps

6.3.3 Matching of Local Feature Maps

Having obtained three code maps pcCode, oriCode, and phaCode for each image, the next issue is how to match them for recognition. Since the PC is a dimensionless measure, we can use the absolute difference to measure the distance between two pcCode maps. Specifically, given two PC code maps, pcCode1 and pcCode2, we define their normalized matching distance as

$$pcD = \frac{\sum \sum abs(pcCode1(\mathbf{x}) - pcCode2(\mathbf{x}))}{(L - 1)S} \quad (6.23)$$

where S is the area of the image.

For comparing two orientation code maps, oriCode1 and oriCode2, we resort to the normalized angular distance proposed in (Kong and Zhang 2004), which is defined as

$$oriD = \frac{\sum \sum ang(oriCode1(\mathbf{x}), oriCode2(\mathbf{x}))}{SJ/2}$$

$$ang(p, q) = \begin{cases} \min(p - q, q - p + J), & p \geq q \\ \min(q - p, p - q + J), & p < q \end{cases} \quad (6.24)$$

When matching two phase code maps, we use a similar method as matching two orientation code maps. The matching distance between two phase code maps, phaCode1 and phaCode2, is given by

$$phaD = \frac{\sum \sum ang(phaCode1(\mathbf{x}), phaCode2(\mathbf{x}))}{SM/2} \quad (6.25)$$

In real implementation, it is easy to design “bitwise” representations for pcCode, oriCode, and phaCode, and accordingly, pcD, oriD, and phaD can be effectively computed.

6.3.4 Integration of Local Features

The three local features reflect different aspects of information contained in an image patch. Thus, we can expect higher recognition accuracy when assembling information from the three features together. This can be achieved by a score-level fusion and we refer to this feature integration scheme as local feature integration (LFI). Suppose that three matching distances pcD , $oriD$, and $phaD$ have been calculated by matching the three kinds of local features respectively. These three distances can be fused together to get the final matching distance. There are a couple of rules for the fusion of matching distances, such as the Simple-Sum (SS) rule, the MIn-Score (MIS) rule, the MAx-Score (MAS) rule, and the Matcher-Weighting (MW) rule (Snelick et al. 2005). In our case, pcD , $oriD$, and $phaD$ can be considered to be obtained from three different matchers and we adopt the MW rule. With the MW fusion rule, weights are assigned according to the equal error rate (EER) obtained on a training dataset by different matchers. Denote by e_k the EER of the matcher k , $k = 1, \dots, 3$. Then, the weight w_k associated with matcher k can be calculated as

$$w_k = \left(1 / \sum_{j=1}^3 \frac{1}{e_j} \right) / e_k \quad (6.26)$$

where $0 \leq w_k \leq 1$ and $\sum_{k=1}^3 w_k = 1$. It is obvious that the weights are inversely proportional to the corresponding EERs. Then, the final matching distance between two FKP images using LFI is calculated as

$$d = w_1 pcD + w_2 oriD + w_3 phaD \quad (6.27)$$

6.4 Experimental Results and Discussions

6.4.1 FKP Database and the Test Protocol

In our previous work (Zhang et al. 2010), an FKP database was established using the developed FKP image acquisition device. This database is intended to be a benchmark to evaluate the performance of various FKP recognition methods, and it is publicly online available at (PolyU Finger-Knuckle-Print Database 2010). In this database, FKP images were collected from 165 volunteers, including 125 males and 40 females. Among them, 143 subjects were 20–30 years old and the others were 30–50 years old. We collected samples in two separate sessions. In each session, the subject was asked to provide six images for each of the left index finger, the left middle finger, the right index finger, and the right middle finger. Therefore, 48 images from four fingers were collected from each subject. In total, the database

contains 7920 images from 660 different fingers. The average time interval between the first and the second sessions was about 25 days. The maximum and minimum time intervals were 96 days and 14 days, respectively. In all of the following experiments, we took images collected at the first session as the gallery set and images collected at the second session as the probe set. To obtain statistical results, each image in the probe set was matched with all the images in the gallery set. If the two images were from the same finger, the matching between them was counted as a genuine matching; otherwise it was counted as an imposter matching.

The equal error rate (EER), which is the point where the false accept rate (FAR) is equal to the false reject rate (FRR), is used to evaluate the verification accuracy. The decidability index d' (Daugman 2003) is used to measure how well the genuine and the imposter distributions are separated. d' is defined as

$$d' = \frac{|\mu_1 - \mu_2|}{\sqrt{(\sigma_1^2 + \sigma_2^2)/2}} \quad (6.28)$$

where $\mu_1(\mu_2)$ is the mean of the genuine (imposter) matching distances and $\sigma_1(\sigma_2)$ is the standard deviation of the genuine (imposter) matching distances. Besides, by adjusting the matching threshold, a detection error tradeoff (DET) curve (Martin et al. 1997), which is a plot of false reject rate (FRR) against false accept rate (FAR) for all possible thresholds, can be created. The DET curve can reflect the overall verification accuracy of a biometric system. Thus, the DET curve obtained by using each evaluated method will be provided.

6.4.2 Determination of Parameters

In real implementation, with respect to the quadrature pair filters, we utilized the log-Gabor filters whose transfer function in the frequency domain is

$$G_2(\omega, \theta_j) = \exp\left(-\frac{(\log(\omega/\omega_0))^2}{2\sigma_r^2}\right) \exp\left(-\frac{(\theta - \theta_j)^2}{2\sigma_\theta^2}\right) \quad (6.29)$$

where ω_0 is the filter's center frequency, σ_r controls the filter's radial bandwidth and σ_θ determines the filter's angular bandwidth. In the spatial domain, a log-Gabor filter has a similar shape with a Gabor filter (Kovesi 1999). However, compared with Gabor filters, log-Gabor filters have some special advantages (Field 1987; Kovesi 1999). At first, one cannot construct Gabor filters of arbitrary bandwidth and still maintain a reasonably small DC component in the even-symmetric filter, while log-Gabor filters, by definition, has no DC component. Secondly, the transfer function of the log-Gabor filter has an extended tail at the high frequency end, which makes it more capable to encode natural images than ordinary Gabor filters.

Thus, we chose to use log-Gabor filters to compute local features discussed in Sects. 6.2 and 6.3. Parameters were empirically tuned based on a sub-dataset containing images from the first 300 FKP classes and the tuning criterion was that parameter values that could lead to a lower EER would be chosen. As a result, parameters were set as the following: $n = 3, J = 6, \sigma_\theta = 0.44, L = 5, \varsigma = 3, M = 8, \omega_0^1 = 0.60, \omega_0^2 = 0.167, \omega_0^3 = 0.083$, where ω_0^1 , ω_0^2 and ω_0^3 represent the three center frequencies of the log-Gabor filters at three scales. In LFI, the weights assigned to the local orientation matcher, the local phase matcher, and the phase congruency matcher are 0.45, 0.25, and 0.30, respectively.

6.4.3 Performance of Local Features

In this experiment, we validate our claim that LFI could provide higher performance than using any of the three local features (the local orientation, the local phase, and the phase congruency) individually. In this experiment, all the classes of FKPs were involved. Therefore, there were 660 (165×4) classes and 3960 (660×6) images in the gallery set and the probe set each. Each image in the probe set was matched against all the images in the gallery set. Thus, the numbers of genuine matchings and imposter matchings were 23,760 and 15,657,840, respectively.

The verification accuracy by using each single feature, the local orientation, the local phase, or the PC, is given in Table 6.1. The performance of the LFI scheme is also reported in Table 6.1. The performance of a state-of-the-art FKP recognition method, Comp-Code (Zhang et al. 2010), is listed in Table 6.1 for comparison. The DET curves obtained by the evaluated methods are shown in Fig. 6.3.

From the experimental results shown in Table 6.1 and Fig. 6.3, we can have the following findings. At first, the local orientation can provide higher discriminability than the other two local features, the local phase and the PC, for the task of FKP verification. Secondly, “local orientation” and “CompCode” have nearly the same performance because both of them exploit the orientation information. In fact, the local orientation in this chapter is extracted by using log-Gabor filters while CompCode extracts such information by using Gabor filters. Thus, we can conclude that Gabor filters and log-Gabor filters have very similar performance for orientation feature extraction. Thirdly, the LFI scheme which integrates all the three local

Table 6.1 Performance of different FKP verification schemes based on local feature(s)

Feature type	EER (%)	d'
Local orientation	1.67	4.2847
Local phase	3.01	2.9213
Local phase congruency	2.59	3.3811
CompCode (Zhang et al. 2010)	1.66	4.2989
LFI	1.27	4.3221

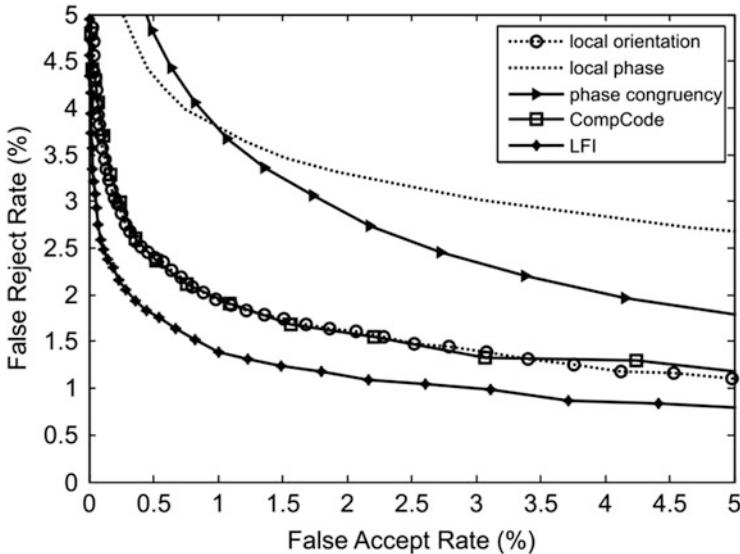


Fig. 6.3 DET curves obtained by using various FKP recognition methods based on local feature(s)

features together performs obviously better than using any of them individually, which corroborates our claim.

It should be noted that LFI has higher computational cost and needs more storage space than CompCode (Zhang et al. 2010). At the feature extraction stage, the major operations involved are convolutions. So, the number of convolutions used can roughly reflect the overall computational complexity of the feature extraction. For CompCode, 6 convolutions are needed (Zhang et al. 2010). For LFI, in order to compute the phase congruency, at each specific scale 6 log-Gabor filtering are applied and altogether 3 scales are adopted; thus, $6 \times 3 = 18$ convolutions are needed. Therefore, LFI has about 3 times the computational complexity of CompCode. In CompCode, each feature point is represented by 3-bits. In LFI, for each feature point, 3-bits are used to represent each local feature, and thus 9-bits are used to represent a feature point. So, it is easy to see that LFI needs 3 times storage space compared with CompCode.

6.4.4 Further Discussions

6.4.4.1 Robustness to Small Rotations

From the experimental results reported in Sect. 6.4.3, we can see that the local orientation has much higher discriminability than the local phase and PC for FKP recognition, and by incorporating the local phase and PC, the verification

performance could be much better than using the local orientation feature alone. The local phase and PC could provide additional discriminative information that is independent of local orientation for the FKP verification task. In addition, the local phase and PC features are more robust to small rotations than the local orientation. Due to the imperfection of the imaging device and the ROI extraction algorithm, there will be small rotations among intra-class images captured at different times, which will have negative effect to the algorithm mainly depending on the orientation information. To validate such a conjecture, we tested the robustness to small rotations of each local feature in this section.

For this purpose, we selected images from the first 400 FKP classes as the dataset and the experiment protocols were the same as described in Sect. 6.4.1. Let $\alpha = \{0, 1, 2, 3, 4, 5\}$. For each α , by rotating each image in the dataset randomly by a degree within range $[-\alpha, \alpha]$, we can get a new virtual dataset. The verification performances of each local feature in terms of EER on these 6 virtual datasets are summarized in Table 6.2. For comparison, we also list the results obtained by CompCode and LFI under the same experimental settings in Table 6.2.

As shown in Table 6.2, when α increases, the EER obtained by using each local feature increases. However, for different local features, the “accelerations” of the EER increase are different. We use the following measure to characterize the “acceleration” of the EER increase:

$$\eta = \frac{1}{5} \sum_{i=1}^5 \frac{eer_i - eer_0}{eer_0} \quad (6.30)$$

where eer_i is the EER obtained when $\alpha = i$. By using this metric η , the “acceleration” of the EER increases by using different local features while α increases can be compared. For the local orientation, $\eta = 0.2798$; for the local phase $\eta = 0.2050$; and for the PC, $\eta = 0.1477$. From this experiment, we can clearly see that with the increase of the rotation degree, the performance decrease of local phase and PC is much less than local orientation, which indicates that the local phase and PC are more robust to small rotations than the local orientation. Moreover, for CompCode, $\eta = 0.2866$ and for LFI, $\eta = 0.2171$, which indicates that the proposed local feature integration scheme LFI is much more robust to small rotations than the CompCode scheme which depends on local orientation information only.

Table 6.2 Verification performance (measured by EER) of local features on the virtual rotated datasets

	$\alpha = 0$	$\alpha = 1$	$\alpha = 2$	$\alpha = 3$	$\alpha = 4$	$\alpha = 5$
Local orientation	2.03%	2.04%	2.16%	2.40%	2.78%	3.61%
Local phase	3.63%	3.65%	3.87%	4.21%	4.64%	5.50%
Phase congruency	2.83%	2.86%	2.91%	3.23%	3.36%	3.88%
CompCode (Zhang et al. 2010)	2.01%	2.03%	2.15%	2.38%	2.77%	3.60%
LFI	1.52%	1.54%	1.64%	1.77%	1.96%	2.34%

6.4.4.2 Integrating Local Features with a Global Feature

In our previous work (Zhang et al. 2011), we presented a local-global information combination (LGIC) scheme for FKP recognition, in which the local orientation extracted by Gabor filters was taken as the local feature while the image's Fourier transform coefficients were taken as the global feature. Similarity of Fourier transform coefficients from two images were compared using the phase-only correlation (POC) technique (Kuglin and Hines 1975; Reddy and Chatterji 1996). LGIC could achieve the highest verification accuracy on our FKP database. In fact, the local features discussed can also be integrated with the global feature, i.e., the Fourier transform coefficients, using the same framework as LGIC. We call this new local-global information combination scheme as LGIC2. Compared with LGIC, LGIC2 involves two more local features, the local phase and the phase congruency. We compared the performance of LGIC and LGIC2 under the same experiment settings as described in Sect. 6.4.3. The results in terms of EER and d' are summarized in Table 6.3. Besides, the EER reported by another state-of-the-art method OE-SIFT (Morales et al. 2011) under the same experimental settings is also listed in Table 6.3 for comparison. DET curves obtained by LGIC and LGIC2 are shown in Fig. 6.4. Distance distributions of genuine matchings and imposter matchings obtained by LGIC2 are plotted in Fig. 6.5. From the experimental results,

Table 6.3 FKP verification performance of OE-SIFT, LGIC and LGIC2

Method	EER (%)	d'
OE-SIFT (Morales et al. 2011)	0.850	—
LGIC (Zhang et al. 2011)	0.402	4.5356
LGIC ₂	0.358	4.7001

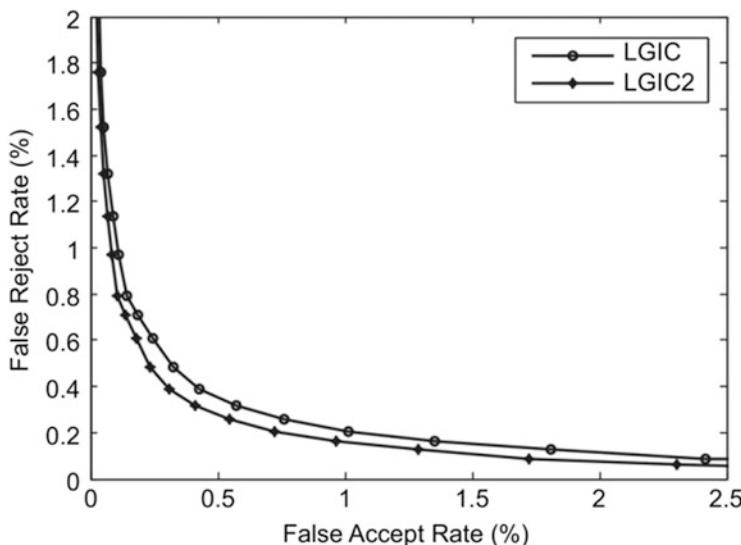


Fig. 6.4 DET curves obtained by LGIC and LGIC2

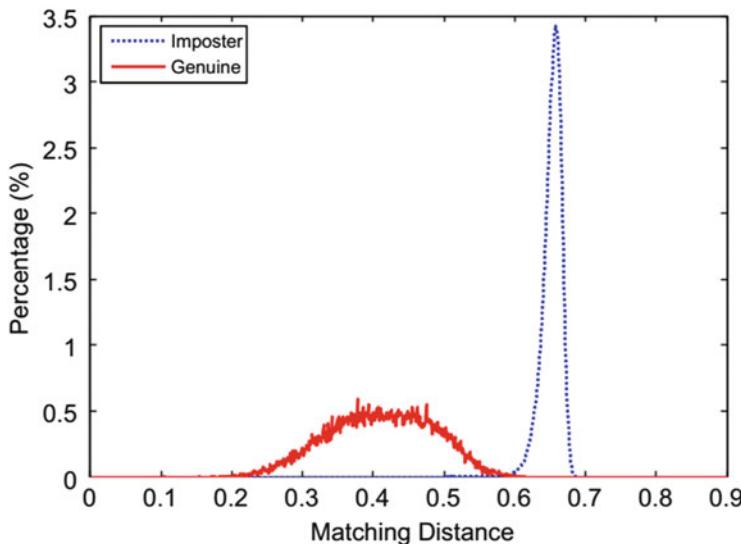


Fig. 6.5 Distance distributions of genuine matchings and imposter matchings obtained by LGIC₂

Table 6.4 Computation time for key processes of LGIC2

Operations	Time (ms)
ROI extraction	198
Local feature extraction	405
Local feature matching	0.9
Global feature matching	2.1

we can see that LGIC2 performs better than LGIC. It once again corroborates our claim that the local phase and the phase congruency could afford more discriminative information for FKP recognition.

LGIC2 is implemented using Visual C#.Net 2005 on a Dell Inspiron 530s PC embedded Intel E6550 processor and 2 GB of RAM. Computation time for the key processes is listed in Table 6.4. The total execution time for one verification operation is less than 0.7 s in our prototype system, which is fast enough for real-time applications. We believe that with the optimization of the implementation, the system's efficiency could be much further improved.

It should be noted that though LGIC2 performs the best among all the existing FKP verification methods, it cannot deal with severe intra-class pose variations. Such variations can result in severe affine transforms or even non-elastic deformations among intra-class FKP images. In fact, most of the failure cases of LGIC2 can be attributed to such large-scale intra-class pose variations. Figure 6.6 shows a typical example. Figure 6.6a, b are two FKP images captured from the same finger in different sessions. Figure 6.6c, d are the ROIs extracted from Fig. 6.6a, b, respectively. It can be seen that there is an obvious pose variation between the two FKPs. They are recognized as different classes by LGIC2. Hence, in the future, we will focus on devising high performance FKP recognition algorithms being

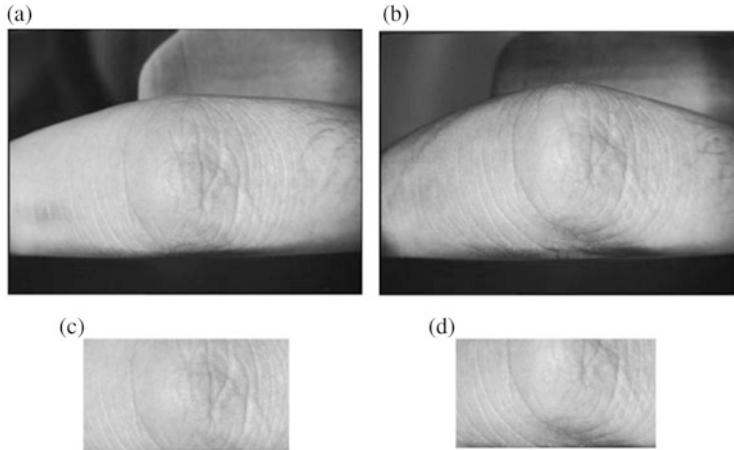


Fig. 6.6 (a) and (b) are two intra-class FKP images in PolyU FKP dataset (PolyU Finger-Knuckle-Print Database, 2010); (c) and (d) are their ROI sub-images. There is an obvious pose variation between the two FKPs and they are recognized as different classes by LGIC2

robust to such intra-class pose variations. For example, the idea proposed in Morales et al.’s work (Morales et al. 2011) can be borrowed.

6.5 Summary

In this chapter, we focused on developing new effective feature extraction and matching method for FKP recognition. To this end, we analyzed three commonly used local features, the local orientation, the local phase, and the phase congruency systematically and presented a method for computing them efficiently using the phase congruency computation framework. Coding and matching algorithm for each local feature was presented. Extensive experiments were conducted on the benchmark PolyU FKP database. The experimental results showed that the integration of all the local features together performs better than using any of them separately. The algorithm LGIC2, which integrates all the three local features and one global feature, Fourier transform coefficients, could achieve the best verification result on the benchmark FKP database, with the EER 0.358%.

References

- Daugman J (1985) Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *J Opt Soc Am A* 2(7):1160–1169
- Daugman J (1993) High confidence visual recognition of persons by a test of statistical independence. *IEEE Trans Pattern Anal Mach Intell* 15(11):1148–1161

- Daugman J (2003) The importance of being random: statistical principles of iris recognition. *Pattern Recogn* 36(2):279–291
- Duta N (2009) A survey of biometric technology based on hand shape. *Pattern Recogn* 42 (11):2797–2806
- Felsberg M, Sommer G (2001) The monogenic signal. *IEEE Trans Signal Process* 49 (12):3136–3144
- Feng Z, Yang B, Chen Y, Zheng Y, Xu T, Li Y, Zhu D (2011) Features extraction from hand images based on new detection operators. *Pattern Recogn* 44(5):1089–1105
- Field D (1987) Relations between the statistics of natural images and the response properties of cortical cells. *J Opt Soc Am A* 4(12):2379–2394
- Gabor D (1946) Theory of communication. *J Inst Electr Eng Part III Radio Commun Eng* 93 (26):429–441
- Granlund G, Knutsson H (1995) Signal processing for computer vision. Springer, Berlin
- Gundimada S, Asari V (2007) A novel neighborhood defined feature selection on phase congruency images for recognition of faces with extreme variations. *Int J Inf Technol* 3(1):25–31
- Han C, Cheng H, Lin C, Fan K (2003) Personal authentication using palm-print features. *Pattern Recogn* 36(2):371–381
- Henriksson L, Hyvärinen A, Vanni S (2009) Representation of cross-frequency spatial phase relationships in human visual cortex. *J Neurosci* 29(45):14342–14351
- Hong L, Wan Y, Jain A (1998) Fingerprint image enhancement: algorithm and performance evaluation. *IEEE Trans Pattern Anal Mach Intell* 20(8):777–789
- Hu D, Feng G, Zhou Z (2007) Two-dimensional locality preserving projections (2DLPP) with its application to palmprint recognition. *Pattern Recogn* 40(1):339–342
- Jain A, Feng J (2009) Latent palmprint matching. *IEEE Trans Pattern Anal Mach Intell* 31 (6):1032–1047
- Jain A, Flynn P, Ross A (2007) Handbook of biometrics. Springer Science & Business Media, New York
- Jia W, Huang D, Zhang D (2008) Palmprint verification based on robust line orientation code. *Pattern Recogn* 41(5):1504–1513
- Kong A, Zhang D (2004) Competitive coding scheme for palmprint verification. In: Proceedings of the 17th international conference on pattern recognition, pp 520–523
- Kong A, Zhang D, Kamel M (2006) Palmprint identification using feature-level fusion. *Pattern Recogn* 39(3):478–487
- Kong A, Zhang D, Kamel M (2009) A survey of palmprint recognition. *Pattern Recogn* 42 (7):1408–1418
- Kovesi P (1999) Image features from phase congruency. *Videre J Comput Vis Res* 1(3):1–26
- Krieger G, Zetsche C (1996) Nonlinear image operators for the evaluation of local intrinsic dimensionality. *IEEE Trans Image Process* 5(6):1026–1042
- Kuglin C, Hines D (1975), The phase correlation image alignment method. In: Proceedings of international conference on cybernetics and society, pp 163–165
- Kumar A, Ravikanth C (2009) Personal authentication using finger knuckle surface. *IEEE Trans Inf Forensics Secur* 4(1):98–110
- Li S (2009) Encyclopedia of biometrics: I-Z, vol 1. Springer Science & Business Media, New York
- Maltoni D, Maio D, Jain A, Prabhakar S (2003) Handbook of fingerprint recognition. Springer, New York
- Martin A, Doddington G, Kamm T, Ordowski M, Przybocki M (1997) The DET curve in assessment of detection task performance. In: European conference on speech communication & technology, pp 1895–1898
- Meraoumia A, Chitroub S, Bouridane A (2011) Palmprint and finger-knuckle-print for efficient person recognition based on log-gabor filter response. *Analog Integr Circ Sig Process* 69 (1):17–27
- Morales A, Travieso C, Ferrer M, Alonso J (2011) Improved finger-knuckle-print authentication based on orientation enhancement. *Electron Lett* 47(6):380–381

- Morrone M, Burr D (1988) Feature detection in human vision: a phase-dependent energy model. *Proc R Soc Lond B Biol Sci* 235(1280):221–245
- Morrone M, Owens R (1987) Feature detection from local energy. *Pattern Recogn Lett* 6 (5):303–313
- Morrone M, Ross J, Burr D, Owens R (1986) Mach bands are phase dependent. *Nature* 324 (6094):250–253
- PolyU (2010) Finger-Knuckle-Print Database. <http://www.comp.polyu.edu.hk/~biometrics>
- Reddy B, Chatterji B (1996) An FFT-based technique for translation, rotation, and scale-invariant image registration. *IEEE Trans Image Process* 5(8):1266–1271
- Ross A, Jain A, Reisman J (2003) A hybrid fingerprint matcher. *Pattern Recogn* 36(7):1661–1673
- Sierra-Vázquez V, Serrano-Pedraza I (2010) Application of Riesz transforms to the isotropic AM-PM decomposition of geometrical-optical illusion images. *J Opt Soc Am A* 27 (4):781–796
- Snelick R, Uludag U, Mink A, Indovina M, Jain A (2005) Large-scale evaluation of multimodal biometric authentication using state-of-the-art systems. *IEEE Trans Pattern Anal Mach Intell* 27(3):450–455
- Struc V, Pavescic N (2009) Phase congruency features for palm-print verification. *IET Sig Process* 3(4):258–268
- Venkatesh S, Owens R (1990) On the classification of image features. *Pattern Recogn Lett* 11 (5):339–349
- Verikas A, Gelzinis A, Bacauskiene M, Olenina I, Olenin S, Vaiciukynas E (2012) Phase congruency-based detection of circular objects applied to analysis of phytoplankton images. *Pattern Recogn* 45(4):1659–1670
- Wang J, Yau W, Suwandy A, Sung E (2008) Person recognition by fusing palmprint and palm vein images based on “Laplacianpalm” representation. *Pattern Recogn* 41(5):1514–1527
- Wietzke L, Sommer G (2010) The signal multi-vector. *J Math Imaging Vision* 37(2):132–150
- Wilson C (2010) Vein pattern recognition. A privacy-enhancing biometric. CRC, Boca Raton
- Woodard D, Flynn P (2005) Finger surface as a biometric identifier. *Comput Vis Image Underst* 100(3):357–384
- Yuan X, Shi P (2005) Iris feature extraction using 2D phase congruency. *Third Int Conf Inf Technol Appl* 2:437–441
- Zhang D, Kong W, You J, Wong M (2003) Online palmpoint identification. *IEEE Trans Pattern Anal Mach Intell* 25(9):1041–1050
- Zhang L, Zhang L, Zhang D (2009a) Finger-knuckle-print: a new biometric identifier. *Proceedings of the ICIP 2009*, pp 1981–1984
- Zhang L, Zhang L, Zhang D (2009b) Finger-knuckle-print verification based on band-limited phase-only correlation. *Proc CAIP 2009*:141–148
- Zhang L, Zhang L, Zhang D, Zhu H (2010) Online finger-knuckle-print verification for personal authentication. *Pattern Recogn* 43(7):2560–2571
- Zhang L, Zhang L, Zhang D, Zhu H (2011) Ensemble of local and global information for finger-knuckle-print recognition. *Pattern Recogn* 44(9):1990–1998
- Zhang D, Zuo W, Yue F (2012) A comparative study of palmpoint recognition algorithms. *ACM Comput Surv (CSUR)* 44(1):1–37
- Zhao Q, Zhang D, Zhang L, Luo N (2010) Adaptive fingerprint pore modeling and extraction. *Pattern Recogn* 43(8):2833–2844

Chapter 7

Global Information for Finger-Knuckle-Print Recognition

Abstract Biometrics authentication is an effective method for automatically recognizing a person's identity. Recently, it has been found that the finger-knuckle-print (FKP), which refers to the inherent skin patterns of the outer surface around the phalangeal joint of one's finger, has high capability to discriminate different individuals, making it an emerging biometric identifier. In this chapter, based on the results of psychophysics and neurophysiology studies that both local and global information is crucial for the image perception, we present an effective FKP recognition scheme by extracting and assembling local and global features of FKP images. Specifically, the orientation information extracted by the Gabor filters is coded as the local feature. By increasing the scale of Gabor filters to infinite, actually we can get the Fourier transform of the image, and hence the Fourier transform coefficients of the image can be taken as the global features. Such kinds of local and global features are naturally linked via the framework of time-frequency analysis. The proposed scheme exploits both local and global information for the FKP verification, where global information is also utilized to refine the alignment of FKP images in matching. The final matching distance of two FKPs is a weighted average of local and global matching distances. The experimental results conducted on our FKP database demonstrate that the proposed local-global information combination scheme could significantly improve the recognition accuracy obtained by either local or global information, and lead to promising performance of an FKP-based personal authentication system.

Keywords Biometrics • Finger-knuckle-print • Local-global information combination

7.1 Introduction

Recognizing the identity of a person with high confidence is a critical issue in various applications, such as e-banking, access control, passenger clearance, etc. The need for reliable user authentication techniques has significantly increased in the wake of heightened concerns about security, and rapid advancement in networking, communication and mobility (Jain et al. 2007). Biometrics based

methods, which use unique physical or behavioral characteristics of human beings, are of broad interest and have great potentials because of their high accuracy and convenience to use in the modern e-world. With the rapid development of computing techniques, in the past several decades researchers have exhaustively investigated the use of a number of biometric characteristics, including fingerprint, face, iris, palmprint, hand geometry, voice and ear.

Among various kinds of biometric identifiers, hand-based biometrics has been attracting considerable attention over recent years. Fingerprint (Maltoni et al. 2003; Ratha and Bolle 2004; Gu et al. 2006; Nikam et al. 2007; Aguilar et al. 2006), palmprint (Zhang et al. 2003; Kong and Zhang 2004; Jia et al. 2008), hand geometry (Jain et al. 1999; Sanchez-Reillo et al. 2000), hand vein (Wang et al. 2008) and inner-knuckle-print (Li et al. 2004; Nanni and Lumini 2009) have been proposed and well investigated in the literature. Recently, it has been found that the image pattern of skin folds and creases in the outer finger knuckle surface is highly unique and thus can serve as a distinctive biometric identifier. Compared with fingerprint, the finger knuckle surface has some advantages as a biometric identifier. At first, it is not easy to be abraded since people usually hold stuffs with the inner side of the hand. In addition, unlike the use of fingerprint, there is no stigma of criminal investigation associated with the finger knuckle surface, so it can have a high user acceptance (Kumar and Zhou 2009). Thus, the finger knuckle feature has a great potential to be widely accepted as a biometric identifier. Some researchers have already done salient works in this field. Woodard and Flynn (Woodard and Flynn 2005) are among the first scholars who exploited the use of finger knuckle surface in biometric systems. They set up a 3D finger back surface database with the Minolta 900/910 sensor. For feature extraction, they used the curvature based shape index to represent the finger back surface. Woodard's work makes a good effort to validate the uniqueness of outer finger surface as a biometric characteristic. However, the cost, size and weight of the Minolta 900/910 sensor limit the use of it in a practical biometric system, and the time-consuming 3D data acquisition and processing limit its use in real-time applications. Later, Kumar and Ravikanth (Kumar and Ravikanth 2009) proposed a 2D finger-back surface based personal authentication system. With respect to the feature extraction, they resorted to some subspace analysis methods such as PCA, LDA and ICA. With their design, the acquisition device is doomed to have a large size because nearly the whole hand back area has to be captured, despite the fact that the finger knuckle area only occupies a small portion of the acquired image. Furthermore, subspace analysis methods may be effective for face recognition but they may not be able to effectively extract the distinctive line and junction features from the finger knuckle surface. In Kumar's later work (Kumar and Zhou 2009), they adopted the robust line orientation code (RLOC) (Jia et al. 2008) to extract the local orientation information of the finger-back surface images.

In our previous works (Zhang et al. 2009a, b, 2010), a novel online personal authentication system using finger-knuckle-print (FKP), which refers to the inherent skin pattern of the outer surface around the phalangeal joint of one's finger, has been established. It comprises four major components: FKP image acquisition, ROI (region of interest) extraction, feature extraction and feature matching. The

proposed FKP imaging system has a small size and it simplifies the preprocessing steps, such as the finger segmentation and the ROI extraction. Since the finger knuckle will be slightly bent when being imaged, the inherent skin patterns can be clearly captured and hence the unique FKP features can be better exploited. The later feature extraction and matching are based on the extracted ROIs. As in any pattern classification task, the feature extraction and matching plays a key role in our FKP-based personal authentication system. To this end, we have developed a couple of different methods. In (Zhang et al. 2009a, b), we used the Gabor filter based competitive coding scheme, which was originally designed for palmprint recognition (Kong and Zhang 2004), to extract the local orientation information as FKP features. In (Zhang et al. 2010), we proposed to combine the orientation information and the magnitude information extracted by Gabor filters together. In (Zhang et al. 2009a, b), the Fourier transform of the whole image was taken as the feature and the band-limited phase-only correlation technique was employed to calculate the similarity between two FKP images.

In this chapter, we focus on feature extraction and matching of FKP images. Based on the area of pixels involved in feature extraction, we can label the features as “local” or “global” ones. Intuitively, a local feature is a measure computed within a local patch, encoding the detailed traits within this specific area; by contrast, a global feature is a measure derived from all (or most of) the pixels in the image, reflecting some holistic characteristic of the examined image. According to such definitions, existing FKP recognition schemes can be classified into local-based methods (Kumar and Ravikanth 2009; Woodard and Flynn 2005; Zhang et al. 2009a, b, 2010) and global-based methods (Kumar and Zhou 2009; Zhang et al. 2009a, b). However, few chapters have yet discussed the local–global information combination for FKP recognition. In the literature of psychophysics and neurophysiology, many studies have shown that both local and global information is crucial for the image perception and recognition of human beings (Su et al. 2009) and they play different but complementary roles. A global feature reflects the holistic characteristics of the image and is suitable for coarse representation, while a local feature encodes more detailed information within a specific local region and is appropriate for finer representation. Hence, better recognition accuracy can be expected if local and global information can be appropriately combined.

Such an idea has already been explored in iris recognition, palmprint recognition, face recognition and fingerprint recognition. For iris matching, Sun et al. (Sun et al. 2004, 2005) proposed a “cascade” system in which the first stage is a conventional Daugman-like classifier while the classifier at the second stage uses “global” features—areas enclosed by zero-crossing boundaries. In (Li et al. 2007), the authors described a two-level palmprint matching scheme. For coarse-level filtering, Hough transform is used to extract global features; for fine-level matching, the local information extracted from the location sand orientations of individual lines is used. Pan et al. (Pan et al. 2007) also proposed to combine the local and global features for palmprint recognition. In their work, non-negative matrix factorization with sparseness constraint and PCA are used to extract local and global features, respectively. For face recognition, Fang et al. (2002) presented a

method by combining global PCA features and component-based local features extracted by Haar wavelets. In (Su et al. 2009), Su et al. proposed a hierarchical ensemble classifier by combining global Fourier features and local Gabor features. In their method, global features are extracted from the whole face images by keeping the low-frequency Fourier coefficients while local features are exploited using Gabor filters with various scales and orientations. After that, Fisher's linear discriminant (FLD) is applied to the global Fourier features and local Gabor features. In the fingerprint recognition community, the idea of combining local and global information was also exploited (Gu et al. 2006; Nikam et al. 2007; Aguilar et al. 2006).

In this chapter, we propose a novel local-global information combination (LGIC) scheme for FKP recognition. Specifically, we take the local orientation information extracted by the Gabor filters as the local feature because local orientation has been successfully used in palmprint recognition systems (Kong and Zhang 2004; Jia et al. 2008) and FKP recognition systems (Kumar and Zhou 2009; Zhang et al. 2009a, b, 2010). By increasing the scale of the Gabor filters, more and more global information will be involved, yet the characterization of image local structures will be weakened rapidly. Particularly, if the scale of the Gabor filter is increased to infinity, the Gabor transform can be reduced to the Fourier transform of the whole image. In this case, no local information can be extracted but we can get the finest resolution for the global frequency analysis of the image. Thus, the Fourier transform coefficients are naturally taken as the global features in this chapter. With the global Fourier features, the alignment between intra-class FKP ROIs can also be refined. At the matching stage, two matching distances can be computed by comparing the local features and the global features separately. Finally, the two matching distances are fused according to some fusion rule to get the final matching distance. Extensive experiments and comparisons are conducted on our established FKP database (PolyU Finger-Knuckle-Print Database 2010) to validate the efficacy of the proposed LGIC scheme.

7.2 Local Feature Extraction and Matching

Gabor filters (Gabor 1946) have been widely used as an effective tool to fulfill the feature extraction tasks in many biometrics systems, such as face, iris, fingerprint, palmprint, etc. The frequency and orientation representations of Gabor filters are similar to those of the human visual system (Pinto et al. 2009, 2008). In the spatial domain, 2D Gabor filters can be expressed as

$$G(x, y) = \exp\left(-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right) \exp(i2\pi f x') \quad (7.1)$$

where $x' = x \cos \theta + y \sin \theta$, $y' = -x \sin \theta + y \cos \theta$. In Eq. (7.1), f represents the frequency of the sinusoid factor, y represents the orientation of the normal to the parallel stripes of the Gabor function, σ_x and σ_y are the standard deviations of the 2D Gaussian envelop.

It can be seen from the definition that a Gabor filter is actually a Gaussian envelop modulated by a sinusoidal plane wave. The Gaussian envelop ensures that the convolution is dominated by the image patch near the center of the filter. Thereby, when an image is convolved with a Gabor filter, the information near the center of the Gaussian envelop is encoded, and by contrast, the information far away from the center of the Gaussian envelop will be neglected. Therefore, the Gabor filter is a local operator and can extract the information at a specific scale and a specific orientation within a local region. Gabor filters can have a variety of different forms with different scales and orientations. Figure 7.1 shows the real part of the Gabor filters at 4 scales and along 6 orientations.

With the Gabor filters, three basic features, magnitude, phase and orientation, can be extracted (Kong 2008). However, previous studies have shown that the local orientation information is the most robust and distinctive local feature for palmprint and FKP recognition (Kong and Zhang 2004; Jia et al. 2008; Kumar and Zhou 2009; Zhang et al. 2009a, b, 2010). Hence, in this chapter, we only take the local orientation as the local feature and make use of the Gabor filter based CompCode (Kong and Zhang 2004; Zhang et al. 2009a, b) scheme to extract and code it. Such

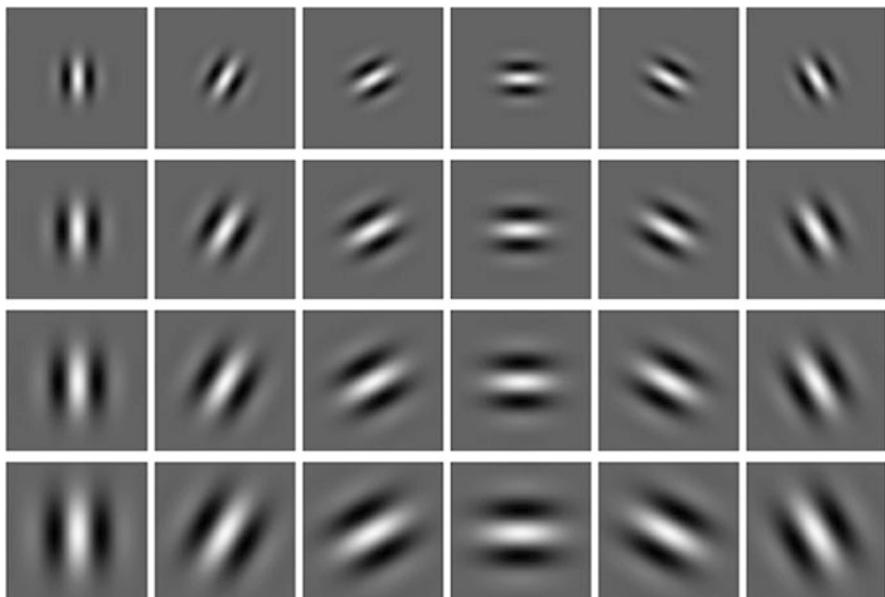


Fig. 7.1 Real parts of the 24 2D Gabor filters with four scales and six orientations

an orientation coding based feature extraction method is suitable for images containing abundant line-like structures and it has the merits of high accuracy, robustness to illumination variation and fast matching. The working principle of CompCode and its matching metric are briefly reviewed as follows.

Denote by G_R the real part of a Gabor filter. With a bank of G_R s sharing the same parameters, except the parameter of orientation, the local orientation information of the image I at the position (x, y) can be extracted and coded. Mathematically, this competitive coding process can be expressed as

$$\text{CompCode}(x, y) = \arg \min_j \{I(x, y)^* G_R(x, y, \theta_j)\} \quad (7.2)$$

where $*$ stands for the convolution operation, $\theta_j = j\pi/J$, $j = \{0, \dots, J - 1\}$, and J represents the number of orientations. Based on our previous studies (Kong and Zhang 2004; Zhang et al. 2009a, b, 2010), we set $J = 6$ in this chapter and this is in accordance with the conclusion made by Lee (Lee 1996) that the simple neural cells are sensitive to specific orientations with approximate bandwidths of $\pi/6$. Figure 7.2c, d shows two CompCode maps extracted from the FKP ROI images in Fig. 7.2a, b, respectively.

In order for real-time recognition, CompCode uses three bits to represent each orientation. When matching two CompCode maps P and Q , the angular distance based on the normalized Hamming distance is used (Kong and Zhang 2004):

$$d_L = \frac{\sum_{Y=1}^{\text{Rows}} \sum_{x=1}^{\text{Cols}} \sum_{i=0}^2 (P_i^b(x, y) \otimes Q_i^b(x, y))}{3S} \quad (7.3)$$

where $P_i^b(Q_i^b)$ is the i th bit plane of $P(Q)$, S is the area of the CompCode map and \otimes represents the bitwise “exclusive OR” operation.

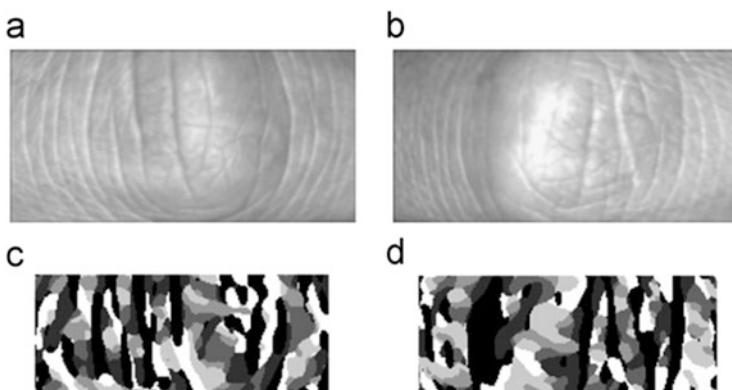


Fig. 7.2 (a) and (b) are two FKP ROI images; (c) and (d) are the CompCode maps generated from (a) and (b), respectively

7.3 Global Feature Extraction and Matching

7.3.1 From Local to Global

In Sect. 7.2, Gabor transforms are utilized to extract the local orientation information. Actually, the Gabor transform can be regarded as a windowed Fourier transform. The corresponding Gabor transform (i.e. filtering) of a function f with respect to a local window function g is (Zayed 1996)

$$G[f](w, t) = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} f(x) g(x - t) e^{-i(w \cdot x)} dx \quad (7.4)$$

where $t, w, x \in \Re^n$, $dx = dx_1 dx_2 \dots dx_n$, $x = (x_1, x_2, \dots, x_n)$ and $w \cdot x = \sum_{k=1}^n w_k x_k$. The signal $f(t)$ to be analyzed is defined in the $n-D$ spatial domain. t is the coordinate variable in the $n-D$ spatial domain and correspondingly, w is the coordinate variable in the $n-D$ frequency domain. The Gabor transform of f , $G[f](w, t)$ can give the frequency spectrum of f for a specified frequency w at a specified position t . For the convenience of discussion, we confine ourselves to the case that $n = 1$, g is a Gaussian-shaped window and $f(x)$ is of finite length $[0, T]$. Then, the Gabor transform of f is

$$G[f](\omega, x) = \int_0^T e^{-(\tau-x)^2/2\sigma^2} f(\tau) e^{-i\omega\tau} d\tau, x \in [0, T] \quad (7.5)$$

The parameter σ controls the size of the local window and the scale of the Gabor transform. Naturally, when σ goes to infinity, the whole signal $f(x)$, $x \in [0, T]$, is involved in calculating $G[f](\omega, t)$ and Eq. (7.5) is reduced to

$$G[f](\omega) = \int_0^T f(\tau) e^{-i\omega\tau} d\tau \quad (7.6)$$

It is seen that $G[f]$ does not depend on x anymore, which implies that we lose the local information in the Gabor transform. Obviously, Eq. (7.6) is the Fourier transform of f .

The above discussion on the 1D case can be easily extended to the 2D case. For 2D images, by increasing the scale of the Gabor filters, more and more global information will be involved, yet the characterization of image local structures will be rapidly weakened. Particularly, if the scale of the Gabor filter goes to infinity, the Gabor transform will degrade to the 2D Fourier transform of the whole image. In such case, though the local characterization is totally lost, we can get the finest frequency resolution for the image analysis. Therefore, in our work the Fourier transform is selected as the global feature extractor.

7.3.2 Phase-Only Correlation (POC)

Now that the Fourier transform coefficients are used as the global feature, the next problem is how to measure the similarity of two given Fourier transforms. Phase-only correlation (POC) is a classical method to this end (Kuglin and Hines 1975). In the literature, POC based methods have been widely used in image registration tasks (Reddy and Chatterji 1996). Recently, POC has also been adopted as a similarity measure in some biometrics systems (Miyazawa et al. 2008; Ito et al. 2004, 2008). Compared with the conventional POC, the band-limited phase-only correlation (BLPOC) proposed by Ito et al. (Ito et al. 2004) is more effective. Hence, in this chapter, we use BLPOC to evaluate the displacement parameters between FKP ROIs and to measure the similarity of the Fourier transforms of the aligned ROIs. In this subsection, POC will be introduced and in the next subsection BLPOC will be described.

POC is a kind of effective method to evaluate the translation parameters between two images in the Fourier domain. Its underlying principle is the translation property of the Fourier transforms (Bracewell 1965). Let f and g be the two images that differ only by a displacement (x_0, y_0) , i.e.

$$g(x, y) = f(x - x_0, y - y_0) \quad (7.7)$$

Their corresponding Fourier transforms $G(u, v)$ and $F(u, v)$ will be related by

$$G(u, v) = \sum_{m=-M_0}^{M_0} \sum_{n=-N_0}^{N_0} g(m, n) e^{-j2\pi(mu/M+nv/N)} = A_G(u, v) e^{j\Phi_F(u, v)} \quad (7.8)$$

The cross-phase spectrum $R_{GF}(u, v)$ between $G(u, v)$ and $F(u, v)$ is given by

$$R_{GF}(u, v) = \frac{G(u, v)F^*(u, v)}{|G(u, v)F^*(u, v)|} = e^{-j2\pi(ux_0+vy_0)} \quad (7.9)$$

where F^* is the complex conjugate of F . By taking inverse Fourier transform of R_{GF} back to the spatial domain, we will have a Dirac impulse centered on (x_0, y_0) .

In practice, we should consider the finite discrete representations. Consider two $M \times N$ images, $f(m, n)$ and $g(m, n)$, where the index ranges are $m = -M_0, \dots, M_0$ and $n = -N_0, \dots, N_0$ and $M = 2M_0 + 1$ and $N = 2N_0 + 1$. Denote by $F(u, v)$ and $G(u, v)$ the 2D DFTs of the two images and they are given by

$$F(u, v) = \sum_{m=-M_0}^{M_0} \sum_{n=-N_0}^{N_0} f(m, n) e^{-j2\pi(mu/M+nv/N)} = A_F(u, v) e^{j\Phi_F(u, v)} \quad (7.10)$$

$$G(u, v) = \sum_{m=-M_0}^{M_0} \sum_{n=-N_0}^{N_0} g(m, n) e^{-j2\pi(mu/M+nv/N)} = A_G(u, v) e^{j\Phi_G(u, v)} \quad (7.11)$$

where $u = -M_0, \dots, M_0$, $v = -N_0, \dots, N_0$, $A_F(u, v)$ and $A_G(u, v)$ are amplitude components, and $\Phi_F(u, v)$ and $\Phi_G(u, v)$ are phase components. Then, the cross phase spectrum $R_{GF}(u, v)$ between $G(u, v)$ and $F(u, v)$ is given by

$$\begin{aligned} w_k &= \left(1 / \sum_{k=1}^2 \frac{1}{e_k} \right) / e_k \quad 0 \leq w_k \leq 1 \sum_{k=1}^2 w_k = 1 \\ d' &= \frac{|\mu_1 - \mu_2|}{\sqrt{(\sigma_1^2 + \sigma_2^2)/2}} (\mu_2)(\sigma_y) f U_0 / M_0 = 0.25 \\ \sigma_x &= 5.0 \\ \sigma_y &= 9.0 \\ f &= 0.0435 \\ &= 1.1 \times 10^{-3} = 1.1 \times 10^{-3} \\ d_L & \\ d_G & \\ R_{GF}(u, v) &= \frac{G(u, v) F^*(u, v)}{|G(u, v) F^*(u, v)|} = e^{j\{\Phi_G(u, v) - \Phi_F(u, v)\}} \end{aligned} \quad (7.12)$$

The POC function $p_{GF}(m, n)$ is the 2D inverse DFT (IDFT) of $R_{GF}(u, v)$:

$$p_{gf}(m, n) = \frac{1}{MN} \sum_{m=-M_0}^{M_0} \sum_{n=-N_0}^{N_0} R_{GF}(u, v) e^{j2\pi(mu/M+nv/N)} \quad (7.13)$$

The peak value of p_{gf} can be calculated as $\max\{p_{gf}(m, n) | m \in [-M_0, M_0], n \in [-N_0, N_0]\}$

If the two images f and g are similar, their POC function p_{gf} will give a distinct sharp peak. If not, the peak value will drop significantly. Thus, the amplitude of the peak value can be used as a similarity measure, and the location of the peak shows the translational displacement between the two images.

7.3.3 Band-Limited Phase-Only Correlation (BLPOC)

In the POC-based image matching method, all the frequency components are involved. However, high frequency components can be prone to noise. To eliminate noisy high frequency components, Ito et al. (Ito et al. 2004) proposed the band-limited POC (BLPOC). BLPOC limits the range of the spectrum of the given FKP image. Suppose that the ranges of the inherent frequency band of FKP texture are given by $u = -U_0, \dots, U_0$ and $v = -V_0, \dots, V$ where $0 \leq U_0 \leq M_0$, $0 \leq V_0 \leq N_0$.

Thus, the effective size of spectrum is given by $L_1 = 2U_0 + 1$ and $L_2 = 2V_0 + 1$. BLPOC function is defined as

$$p_{gf}^{U_0 V}(m, n) = \frac{1}{L_1 L_2} \sum_{u=-U_0}^{U_0} \sum_{v=-V_0}^{V_0} R_{GF}(u, v) e^{j2\pi(mu/L_1 + nv/L_2)} \quad (7.14)$$

where $m = -U_0, \dots, U_0$ and $n = -V_0, \dots, V_0$. From the definition of BLPOC, we can see that U_0/M_0 and V_0/N_0 can inherent frequency distribution of the FKP images.

From the definition of BLPOC, it can be seen that the BLPOC function between two images f and g can be considered as the POC function between their low-pass filtered versions. Thus, the BLPOC function can maintain the properties of the POC function. Specifically, if two images are similar, their BLPOC function will have a distinct sharp peak. At the same time, the translational displacement between the two images can be estimated by the location of the peak. Experiments indicate that the BLPOC function provides a much higher discrimination capability than the original POC function in FKP recognition. This can be reflected in the matching examples shown in Fig. 7.3. Figure 7.3a, b are two FKP ROI images from the same finger (captured indifferent collection sessions), whose POC function and BLPOC function are shown in Fig. 7.3c, d, respectively; Fig. 7.3e, f are two FKPROI images from different fingers, whose POC function and BLPOC function are shown in

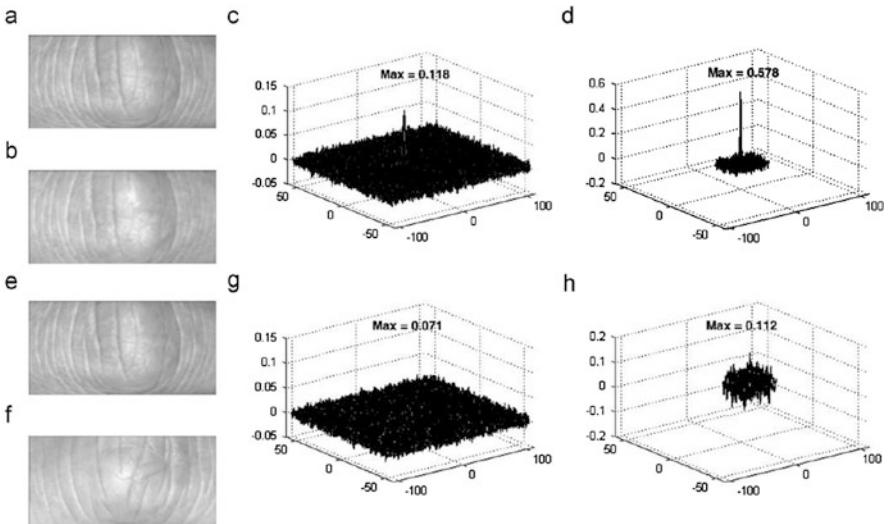


Fig. 7.3 Examples of a genuine matching and an imposter matching using POC and BLPOC, respectively: (a) and (b) are two FKPROI images from the same finger, (c) is their POC function and (d) is their BLPOC function; (e) and (f) are two FKPROI images from different fingers, (g) is their POC function and (h) is their BLPOC function

Fig. 7.3g, h, respectively. These examples indicate that in the case of a genuine matching (a matching performed between a pair of FKP images from the same finger), the BLPOC will exhibit a much sharper peak than POC; however, for an imposter matching (a matching performed between a pair of FKP images from different fingers), neither BLPOC nor POC will show a distinct sharp peak. Hence, in this chapter, we adopt the BLPOC to align the displacement between FKP ROI images and then to measure the similarity between Fourier transforms of the aligned ROIs.

7.4 Local-Global Information Combination

In this section, we present our local-global information combination (LGIC) based FKP recognition algorithm. The entire process of our LGIC-based FKP matching is illustrated in Fig. 7.4. Given two FKPROI images f and g , the following four steps will be taken to compute their similarity.

Step 1: Translation Alignment by Global Features with BLPOC

Although the FKP image acquisition device and the ROI extraction algorithm can reduce the geometric transformations between intra-class ROIs much, it is still inevitable that there is some displacement between intra-class ROIs. This will weaken the genuine matching scores. In our previous coding-based works (Zhang et al. 2009a, b, 2010), this problem was addressed by translating one set of features in horizontal and vertical directions several times and the minimum of the resulting matching distances was considered to be the final matching distance. In this chapter, we solve this problem in a different way by evaluating the translation parameters between the two ROIs using the BLPOC function. Then we crop the common regions, based on which the feature matching is performed. The translation parameters (t_1, t_2) between f and g can be estimated from the peak location of the global BLPOC of them. Then, we can align f and g based on (t_1, t_2) and extract the common regions f_c and g_c . It should be noted that in our system, we will check the ratio between the common region area and the area of the original ROI. If $\text{area}(f_c)/\text{area}(f) < t$ (or $\text{area}(g_c)/\text{area}(g) < t$) where t is a threshold, f_c and g_c will be simply set as f and g . Generally, this will happen when the two FKP images are from different fingers, i.e. inter-classes.

Step 2: Local Feature Extraction and Matching

After alignment and common area cropping, two CompCode maps C_f and C_g are constructed from f_c and g_c . Then, by C_f matching and C_g , we could get the matching distance d_L . For technical details at this step, please refer to Sect. 7.2.

Step 3: Global Feature Extraction and Matching

We use the peak value of the BLPOC function $p_{f_c g_c}^{U_0 V_0}$ between f_c and g_c to measure the similarity of their Fourier transforms. Denote by $pocS$ the peak value of $p_{f_c g_c}^{U_0 V_0}$, then the matching distance is defined as: $d_G = 1 - pocS$.

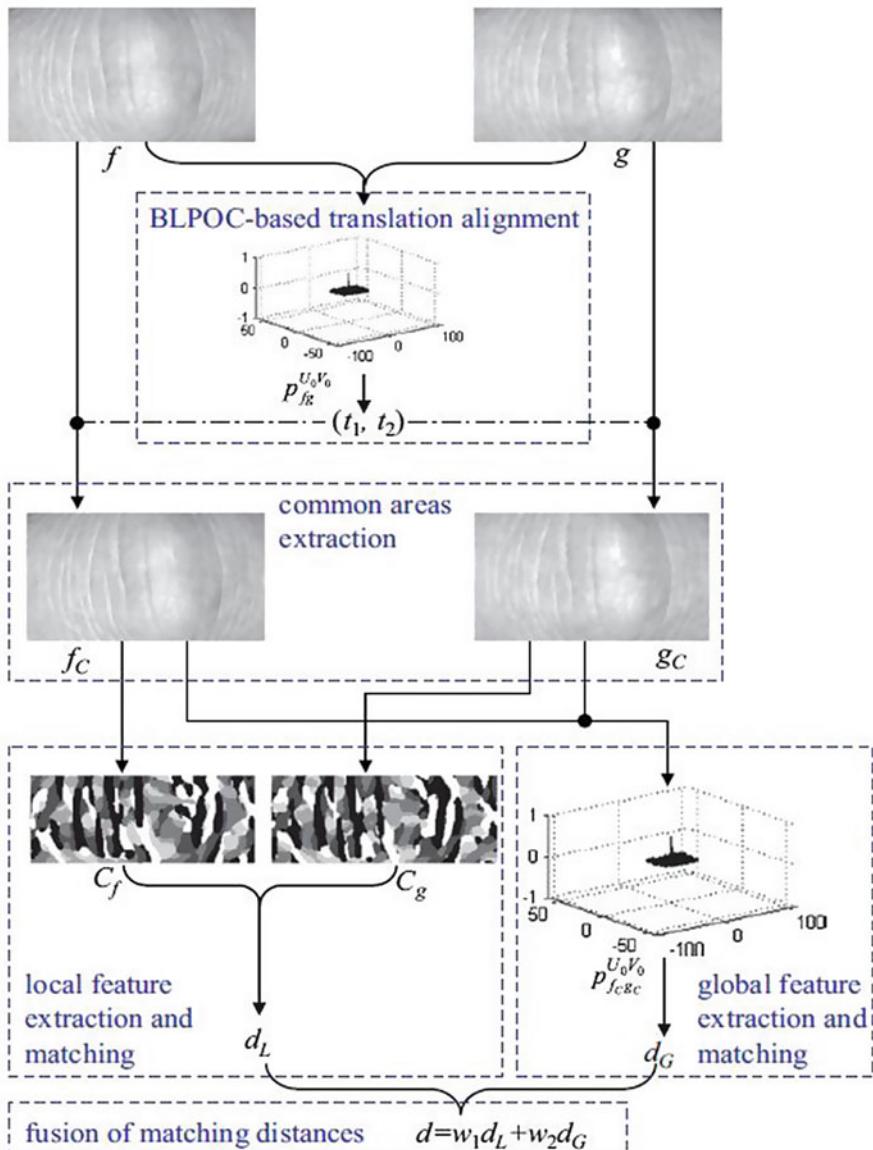


Fig. 7.4 Illustration for the matching distance computation between a pair of FKP ROI images with LGIC

Step 4: Fusion of Matching Distances

Until now, two matching distances d_L and d_G have been obtained. These two distances can be fused together to get the final matching distance. There are a couple of rules for the fusion of matching distances, such as the Simple-Sum

(SS) rule, the MIn-Score (MIS) rule, the MAX-Score(MAS) rule, and the Matcher-Weighting (MW) rule (Snelick et al. 2005). In our case, d_L and d_G can be considered to be obtained from two different matchers, matcher 1 (local feature based matcher) and matcher 2 (global feature based matcher), and we adopt the MW rule. With the MW fusion rule, weights are assigned according to the equal error rate (EER) obtained on a training data set by different matchers. Denote by e_k the EER of the matcher k , $k = 1, 2$. Then, the weight w_k associated with matcher k can be calculated as

$$w_k = \left(1 / \sum_{k=1}^2 \frac{1}{e_k} \right) / e_k \quad (7.15)$$

where $0 \leq w_k \leq 1$ and $\sum_{k=1}^2 w_k = 1$. It is obvious that the weights are inversely proportional to the corresponding EERs. Then, the final matching distance is calculated as

$$d = w_1 d_L + w_2 d_G \quad (7.16)$$

7.5 Experimental Results

7.5.1 FKP Database and Test Protocol

In our previous work (Zhang et al. 2009a, b, 2010), an FKP database was established using the developed FKP image acquisition device. This database is intended to be a benchmark to evaluate the performance of various FKP recognition methods, and it is available at (PolyU Finger-Knuckle-Print Database 2010). In this database, FKP images were collected from 165 volunteers, including 125 males and 40 females. Among them, 143 subjects were 20–30 years old and the others were 30–50 years old. We collected samples in two separate sessions. In each session, the subject was asked to provide 6 images for each of the left index finger, the left middle finger, the right index finger and the right middle finger. Therefore, 48 images from 4 fingers were collected from each subject. In total, the database contains 7920 images from 660 different fingers. The average time interval between the first and the second sessions was about 25 days. The maximum and minimum time intervals were 96 days and 14 days, respectively. In all of the following experiments, we took images collected at the first session as the gallery set and images collected at the second session as the probe set. To obtain statistical results, each image in the probe set was matched with all the images in the gallery set. If the two images were from the same finger, the matching between them was counted as a genuine matching; otherwise it was counted as an imposter matching.

The EER, which is the point where the false accept rate (FAR) is equal to the false reject rate (FRR), is used to evaluate the verification accuracy. The decidability index d' (Daugman 2003) is used to measure how well the genuine and the imposter distributions are separated. d_0 is defined as

$$d' = \frac{|\mu_1 - \mu_2|}{\sqrt{(\sigma_1^2 + \sigma_2^2)/2}} \quad (7.17)$$

where μ_1 (μ_2) is the mean of the genuine (impostor) matching distances and σ_1 (σ_2) is the standard deviation of the genuine (impostor) matching distances.

7.5.2 Determination of Parameters

In real implementation, parameters need to be determined for LGIC. To this end, we tuned the parameters based on a sub-dataset, which contained the first 300 FKP images. Parameters for the local feature and the global feature were retuned separately. Three parameters σ_x , σ_y and f need to be tuned for the local feature while two parameters U_0/M_0 and V_0/N_0 need to be tuned for the global feature. The tuning criterion was that parameter values that could lead to a lower EER would be chosen. As a result, the parameters used in this chapter were set as: $\sigma_x = 5.0$, $\sigma_y = 9.0$, $f = 0.0435$, $U_0/M_0 = 0.25$ and $V_0/N_0 = 0.2$. Moreover, two fusion weights w_1 and w_2 can be calculated using Eq. (7.15).

7.5.3 FKP Verification Results

Verification aims to answer the question of “whether the person is the one he/she claims to be”. In this experiment, all the classes of FKPs were involved. Therefore, there were 660(165×4) classes and 3960(660×6) images in the gallery set and the probe set each. Each image in the probe set was matched against all the images in the gallery set. Thus, the numbers of genuine matchings and imposter matchings were 23,760 and 15,657,840, respectively. In order to show its superiority, the proposed LGIC was compared with the other three state-of-the-art FKP verification methods, CompCode (Zhang et al. 2009a, b), BLPOC (Zhang et al. 2009a, b) and ImCompCode&MagCode (Zhang et al. 2010). Some optimizations have been made on ROI extraction and matching, so the experimental results for Comp- Code, ImCompCode&MagCode and BLPOC are better than the previous publications. The results in terms of the EER and d_0 are summarized in Table 7.1. In addition, the FRRs for each algorithm obtained with a fixed $\text{FAR} = 1.1 \times 10^{-3}$ are also presented

Table 7.1 Performance comparison of different FKP verification schemes

	EER(%)	d'	FRR(%) (when FAR = 1.1×10^{-3})
CompCode (Zhang et al. 2009a, b)	1.658	4.2989	3.4848
BLPOC (Zhang et al. 2009a, b)	1.676	2.4745	8.5939
ImCompCode&MagCode (Zhang et al. 2010)	1.475	4.3224	3.0818
LGIC	0.402	4.5356	0.9680

in Table 7.1 for comparison .Furthermore, by adjusting the matching threshold, a detection error trade off (DET) curve (Martin et al. 1997), which is a plot of false reject rate (FRR) against false accept rate (FAR) for all possible thresholds, can be created. The DET curve can reflect the overall verification accuracy of a biometric system. Figure 7.5a shows the DET curves generated by the four different FKP verification schemes. Distance distributions of genuine matching sand imposter matchings obtained by the proposed LGIC scheme are plotted in Fig. 7.5b. Table 7.2 lists five typical operating states obtained by using the LGIC scheme.

From the results listed in Table 7.1 and the DET curves shown in Fig. 7.5a, we can see that the proposed LGIC scheme performs significantly better in terms of the verification accuracy than the other state-of-the-art FKP verification methods evaluated, including CompCode, BLPOC and ImCompCode&MagCode. As stated, the local orientation information or the local magnitude information is used in CompCode and ImCompCode&MagCode, so they can be classified as local-based methods. By contrast, in BLPOC, the Fourier transform of the whole image is taken as the feature so it is actually a global-based method. Therefore, the experimental results also corroborate the claim that methods fusing local and global information together can outperform the methods depending on only a specific kind of features, local or global.

7.5.4 Speed

The FKP recognition software is implemented using Visual C#.Net 2005 on a Dell Inspiron 530s PC embedded Intel E6550 processor and 2 GB of RAM. Computation time for the key processes is listed in Table 7.3. The execution time for data preprocessing and ROI extraction is 198 ms. The time for BLPOC-based translation alignment is about 1.4 ms. The time for competitive coding is 60 ms. The time for calculating d_L and d_G is 0.3 and 2.1 ms, respectively. Thus, the total execution time for one verification operation is less than 0.5 s in our prototype system, which is fast enough for real-time applications. We believe that with the optimization of the implementation, the system's efficiency could be much further improved.

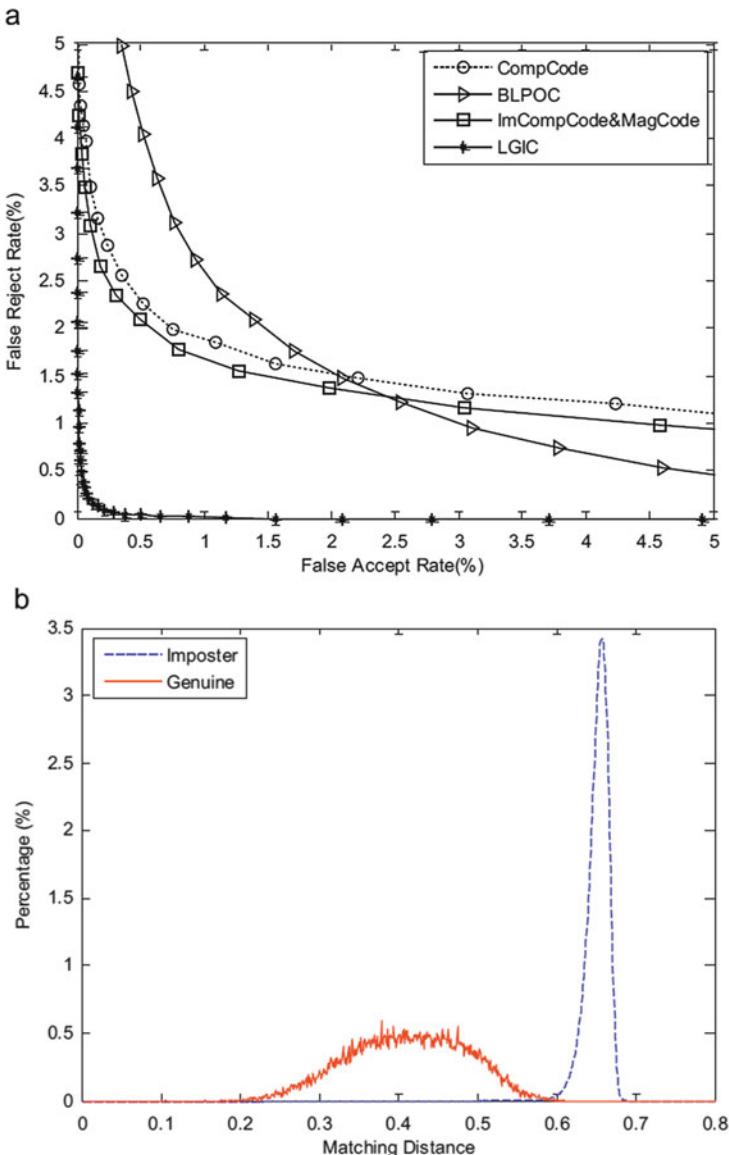


Fig. 7.5 (a) DET curves obtained by the four FKP recognition methods; (b) distance distributions of genuine matching sand imposter matchings with the proposed scheme LGIC

Table 7.2 Typical operating states using LGIC

	FAR (%)	FRR (%)
0.0515		1.5236
0.1068		0.9680
0.2148		0.6439
0.3396		0.4461
0.5982		0.2819

Table 7.3 Computation time for key processes

Operations	Time (ms)
ROI extraction	198
Translation alignment	1.4
Competitive coding	60
Calculation of d_L	0.3
Calculation of d_G	2.1

7.6 Summary

In this chapter, a novel local–global information combination (LGIC) based FKP recognition method was proposed. It is based on the fact that both local and global features are crucial for the image recognition and perception and they play different and complementary roles in such a process. In LGIC, the local orientation extracted by the Gabor filters based competitive coding scheme was taken as the local feature. From the perspective of time–frequency analysis, when the scale of the Gabor transform goes to infinity, it degenerates to the Fourier transform. Thus, the Fourier transform was naturally taken as the global feature in our work. LGIC exploits both local and global features for FKP verification, where the global features were also used to refine the alignment of FKP images in matching. Extensive experimental results conducted on our FKP database indicate that the proposed scheme could achieve much better performance in terms of EER and the decidability index than the other state-of-the-art competitors. Specifically, the EER of LGIC is 0.402% and it can operate at a low FRR of 1.5236% with a low FAR of 0.0515% on our FKP database.

References

- Aguilar J, Chen Y, Garcia J, Jain A (2006) Incorporating image quality in multi-algorithm fingerprint verification. Proc ICB 3832:213–220
- Bracewell R (1965) The fourier transform and its applications. McGraw-Hill, New York
- Daugman J (2003) The importance of being random: statistical principles of iris recognition. Pattern Recogn 36(2):279–291
- Fang Y, Tan T, Wang Y (2002) Fusion of global and local features for face verification. Proc ICPR 2:382–385
- Gabor D (1946) Theory of communication. J Inst Electr Eng 93(1946):429–457
- Gu J, Zhou J, Yang C (2006) Fingerprint recognition by combining global structure and local cues. IEEE Trans Image Process 15(7):1952–1964
- Ito K, Nakajima H, Kobayashi K, Aoki T, Higuchi T (2004) A fingerprint matching algorithm using phase-only correlation. IEICE Trans Fundam E87-A(3):682–691
- Ito K, Aoki T, Nakajima H, Kobayashi K, Higuchi T (2008) A palmprint recognition algorithm using phase-only correlation. IEICE Trans Fundam E91-A(4):1023–1030
- Jain A, Ross A, Pankanti S (1999) A prototype hand geometry-based verification system. In: Proceedings of the AVBPA 1999, pp 166–171
- Jain A, Flynn P, Ross A (2007) Handbook of biometrics. Springer, New York

- Jia W, Huang D, Zhang D (2008) Palmprint verification based on robust line orientation code. *Pattern Recogn* 41(5):1504–1513
- Kong A (2008) An evaluation of Gabor orientation as a feature for face recognition. In: Proceedings of the ICPR 2008, pp 1–4
- Kong W, Zhang D (2004) Competitive coding scheme for palmprint verification. *Proc ICPR* 1:520–523
- Kuglin C, Hines D (1975) The phase correlation image alignment method. In: Proceedings of the international conference on cybernetics and society'75, pp 163–165
- Kumar A, Ravikanth C (2009) Personal authentication using finger knuckle surface. *IEEE Trans Inf Forensics Secur* 4(1):98–109
- Kumar A, Zhou Y (2009) Personal identification using finger knuckle orientation features. *Electron Lett* 45(20):1023–1025
- Lee T (1996) Image representation using 2D Gabor wavelet. *IEEE Trans Pattern Anal Mach Intell* 18(10):957–971
- Li Q, Qiu Z, Sun D, Wu J (2004) Personal identification using knuckleprint. In: Proceedings of SinoBiometrics 2004, pp 680–689
- Li F, Leung K, Yu X (2007) A two-level matching scheme for speedy and accurate palmprint identification. In: Proceedings of the MMM 2007, pp 323–332
- Maltoni D, Maio D, Jain A, Prabhakar S (2003) Handbook of fingerprint recognition. Springer, New York
- Martin A, Doddington G, Kamm T, Ordowski M, Przybocki M (1997) The DET curve in assessment of detection task performance. In: Proceedings of the Eurospeech 1997, pp 1895–1898
- Miyazawa K, Ito K, Aoki T, Kobayashi K, Nakajima H (2008) An effective approach for iris recognition using phase-based image matching. *IEEE Trans Pattern Anal Mach Intell* 20(10):1741–1756
- Nanni L, Lumini A (2009) A multi-matcher system based on knuckle-based features. *Neural Comput Appl* 18(1):87–91
- Nikam S, Goel P, Tapadar R, Agarwal S (2007) Combining Gabor local texture pattern and wavelet global features for fingerprint matching. In: Proceedings of the ICCIMA 2007, pp 409–416
- Pan X, Ruan Q, Wang Y (2007) Palmprint recognition using fusion of local and global features. In: Proceedings of the ISPACS 2007, pp 642–645
- Pinto N, Cox D, Corda B, Doukhan D (2008) Why is real-world object recognition hard? Establishing honest benchmarks and baselines for object recognition. *Proceedings of the COSYNE 2008*
- Pinto N, DiCarlo J, Cox D (2009), How far can you get with a modern face recognition test set using only simple features? In: Proceedings of the CVPR 2009, pp 2591–2598
- PolyU Finger-Knuckle-Print Database (2010). <http://www.comp.polyu.edu.hk/~biometrics/FKP.htm>
- Ratha N, Bolle R (2004) Automatic fingerprint recognition systems. Springer, New York
- Reddy B, Chatterji B (1996) An FFT-based technique for translation, rotation, and scale-invariant image registration. *IEEE Trans Image Process* 5(8):1266–1271
- Sanchez-Reillo R, Sanchez-Avila C, Gonzalez-Marcos A (2000) Biometric identification through hand geometry measurements. *IEEE Trans Pattern Anal Mach Intell* 22(10):1168–1171
- Snelick R, Uludag U, Mink A, Indovina M, Jain A (2005) Large-scale evaluation of multimodal biometric authentication using state-of-the-art systems. *IEEE Trans Pattern Anal Mach Intell* 27(3):450–455
- Su Y, Shan S, Chen X, Gao W (2009) Hierarchical ensemble of global and local classifiers for face recognition. *IEEE Trans Image Process* 18(8):1885–1896
- Sun Z, Wang Y, Tan T, Cui J (2004) Cascading statistical and structural classifiers for iris recognition. In: Proceedings of the ICPR 2004, pp 1261–1262
- Sun Z, Wang Y, Tan T, Cui J (2005) Improving iris recognition accuracy via cascaded classifiers. *IEEE Trans Syst Man Cybern Part C* 35(3):435–441

- Wang J, Yau W, Suwandy A, Sung E (2008) Personal recognition by fusing palmprint and palm vein images based on “Laplacianpalm” representation. *Pattern Recogn* 41(5):1531–1544
- Woodard D, Flynn P (2005) Finger surface as a biometric identifier. *Comput Vis Image Underst* 100(3):357–384
- Zayed A (1996) Handbook of function and generalized function transformations. CRC Press, Boca Raton
- Zhang D, Kong W, You J, Wong M (2003) Online palmpoint identification. *IEEE Trans Pattern Anal Mach Intell* 25(9):1041–1050
- Zhang L, Zhang L, Zhang D (2009a) Finger-knuckle-print: a new biometric identifier. In: Proceedings of the ICIP 2009, pp 1981–1984
- Zhang L, Zhang L, Zhang D (2009b) Finger-knuckle-print verification based on band-limited phase-only correlation. In: Proceedings of the CAIP 2009, pp 141–148
- Zhang L, Zhang L, Zhang D, Zhu H (2010) Online finger-knuckle-print verification for personal authentication. *Pattern Recogn* 43(7):2560–2571

Chapter 8

Finger-Knuckle-Print Verification with Score Level Adaptive Binary Fusion

Abstract Recently, a new biometrics identifier, namely finger knuckle print (FKP), has been proposed for personal authentication with very interesting results. One of the advantages of FKP verification lies in its user friendliness in data collection. However, the user flexibility in positioning fingers also leads to a certain degree of pose variations in the collected query FKP images. The widely used Gabor filtering based competitive coding scheme is sensitive to such variations, resulting in many false rejections. We propose to alleviate this problem by reconstructing the query sample with a dictionary learned from the template samples in the gallery set. The reconstructed FKP image can reduce much the enlarged matching distance caused by finger pose variations; however, both the intra-class and inter-class distances will be reduced. We then propose a score level adaptive binary fusion rule to adaptively fuse the matching distances before and after reconstruction, aiming to reduce the false rejections without increasing much the false acceptances. Experimental results on the benchmark PolyU FKP database show that the proposed method significantly improves the FKP verification accuracy.

Keywords Biometrics • Finger-knuckle-print • Reconstruction • Score level fusion

8.1 Introduction

Refer to Chap. 5, although a triangular block is used to control the finger freedom in FKP image acquisition, there is still much flexibility for the users to position their fingers. This is good to increase the user friendliness but also allows much variation of the finger pose in query sample collection process. Figure 8.1 shows some examples. We can see obvious deformations between the two FKP samples due to the finger pose variations. Unfortunately, the CompCode and LFI based FKP recognition methods are sensitive to such variations, resulting in false rejections and degrading the FKP verification performance.

From the above discussions, we can see that the main difficulty in FKP recognition is the false rejections caused by finger-pose-variation in the query samples.

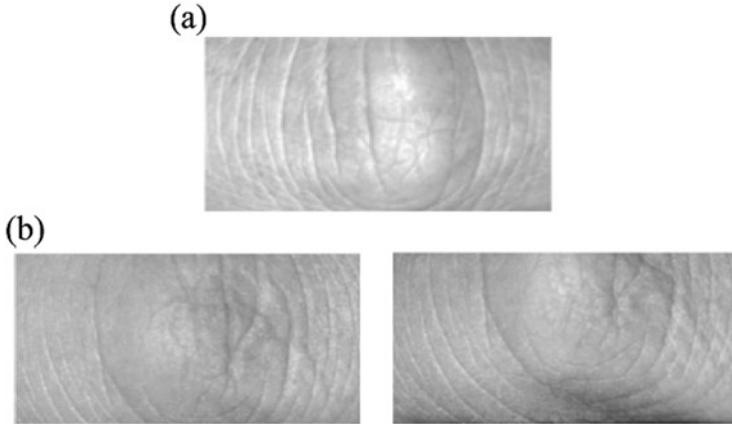


Fig. 8.1 (a) A template FKP sample and (b) two testing FKP samples from the same finger but collected at different sessions. Obvious pose variations can be observed from the samples in (b)

One strategy to solve this problem is to correct the pose deformations by affine transformations. However, estimating the affine transformation parameters is itself a very difficult problem, particularly for FKP images where very few distinctive key points can be extracted. Since our ultimate goal is FKP verification but not pose deformation correction, another strategy is to enhance the FKP matching process without pose deformation correction. Considering that the finger pose variation caused FKP image deformation enlarges the matching distance between two FKP images from the same person and hence results in false rejections, we propose a reconstruction based matching scheme to reduce the enlarged matching distance.

First, a dictionary is learned from the template FKP images, and this dictionary defines the subspace of the gallery FKP dataset. For a given query sample which may have pose variation, we represent it as the linear combination of the atoms in the learned dictionary. This process actually projects the query sample onto the subspace spanned by the gallery FKP images. The CompCode scheme can then be applied to the reconstructed image for feature extraction and matching. Nonetheless, the reconstruction of the query sample will not only reduce the intra-class distance, but also reduce the interclass distance. In other words, it can reduce false rejections but may also increase the false acceptances. To effectively exploit the discriminative information of the query sample before and after reconstruction, a simple yet powerful score level adaptive binary fusion (ABF) rule is proposed to make the final decision by fusing the matching scores before and after reconstruction. The ABF ensures a good reduction of false rejections without increasing much the false acceptances, leading to much lower equal error rates than state-of-the-art methods.

8.2 Competitive Coding Based FKP Verification

Gabor filtering has been widely used as an effective feature extraction technique in face, iris, fingerprint, palmprint, as well as FKP recognition systems. A 2D Gabor filter can be mathematically expressed as

$$G(n, m) = \exp\left(-\frac{1}{2}\left(\frac{n_\theta^2}{\sigma_n^2} + \frac{m_\theta^2}{\sigma_m^2}\right)\right) \cdot \exp(i2\pi f n_\theta) \quad (8.1)$$

where $n_\theta = n \cdot \cos \theta + m \cdot \sin \theta$, $m_\theta = -n \cdot \sin \theta + m \cdot \cos \theta$, f is the frequency of the sinusoid factor, θ is the orientation of the normal to the parallel stripes of the Gabor function, σ_n and σ_m are the standard deviations of the 2D Gaussian envelop.

Based on the observation that the FKP images contain abundant line-like structures, the orientation features can be extracted for FKP image recognition. Let's denote by GR the real part of a Gabor filter, and by I_{ROI} an FKP ROI (region of interest) image. With a bank of Gabor filters, at each pixel $I_{ROI}(n, m)$, the CompCode scheme (Kong and Zhang 2004; Zhang et al. 2009a, b, c; Zhang et al. 2010) extracts and codes the dominant orientation feature as follows:

$$\text{CompCode}(n, m) = \arg \min_j \{I_{ROI}(n, m) * G_R(n, m, \theta_j)\} \quad (8.2)$$

where symbol “*” denotes the convolution operation, and $\theta_j = j\pi/6$, $j = \{0, \dots, 5\}$. $\text{CompCode}(n, m)$ is assigned the orientation along which the smallest response is obtained.

In order for real-time recognition, CompCode uses three bits to represent each orientation (Kong and Zhang 2004). For matching two CompCode maps P and Q , the normalized Hamming distance based angular distance is commonly adopted (Kong and Zhang 2004):

$$s_h = \frac{\sum_{n=1}^{\text{Rows}} \sum_{m=1}^{\text{Cols}} \sum_{i=0}^2 (P_i(n, m) \otimes Q_i(n, m))}{3S} \quad (8.3)$$

where S is the area of the code map, $P_i(Q_i)$ is the i th bit plane of $P(Q)$, and \otimes represent the bitwise “exclusive OR” operation. In practice, multiple matches are performed by translating one of the two feature maps vertically and horizontally, and the minimum matching distance is regarded as the final angular distance. Nonetheless, the CompCode scheme is sensitive to the variation of FKP image and then resulting in false rejection. Even a small rotation and misalignment can lead to an incorrect matching.

8.3 Recognition via Reconstruction

8.3.1 Motivation

The CompCode scheme is simple and fast, and it leads to acceptable accuracy in FKP verification (Zhang et al. 2009a, b, c, 2010). If the query FKP image is well aligned after ROI extraction, the CompCode scheme can work very well. As we discussed in the introduction section, however, there can be certain degree of variations of the finger pose in the data collection process (refer to Fig. 8.1 please), which lead to deformations in the FKP images and consequently result in false rejections because CompCode is sensitive to image deformations.

Figure 8.2 shows the genuine and imposter matching distance distributions of a typical FKP verification system. Based on the curves, it is obvious that we can divide the matching distance into two parts: an uncertain interval, which is between thresholds t_1 and t_2 , and a confidence interval, which covers the remaining part. Generally speaking, if the matching distance of two FKP feature maps falls into the confidence interval, it can be easily decided if the query sample is a genuine or an imposter, while most of the false acceptance sand false rejections occur when the matching distance falls into the uncertain interval $[t_1, t_2]$. On the other hand, the finger pose variation is the main cause that increases the intra-class distance (see the long tail of the blue genuine curve in Fig. 8.2), making the false rejections happen.

The FKP verification accuracy can be improved if we could correct the finger-pose-variation caused deformations of query samples via affine transformation. Unfortunately, it is a particularly difficult problem to estimate the affine transformation parameters for FKP images because very few distinctive key points can be

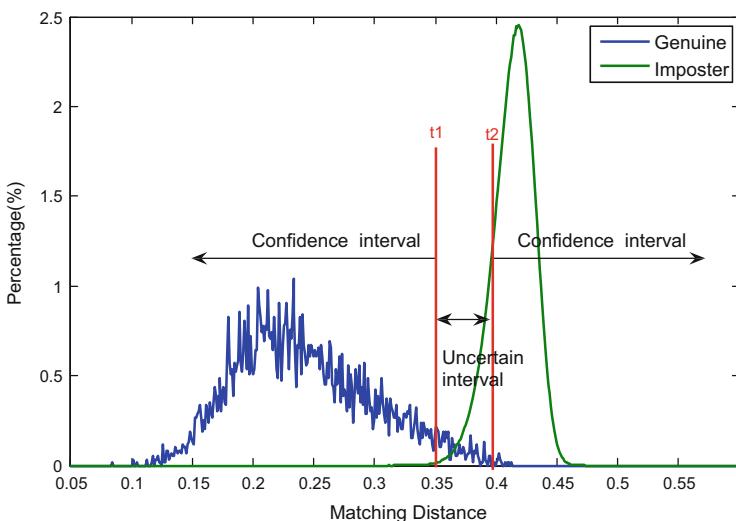


Fig. 8.2 The distribution of matching distance of a typical FKP verification system

extracted from them. Therefore, this solution is impractical. In this chapter, we propose to reduce the pose variation caused false rejections by enhancing the matching process without pose deformation correction.

Denote by $X = [x_1, x_2, \dots, x_k]$ the set of gallery FKP samples, where x_i is a vectorized FKP sample. We can use the hull of X , denoted by $H(X) = \{X \cdot w\}$, where $w = [w_1, w_2, \dots, w_k]^T$ is the vector of weights, to characterize the subspace of gallery FKP images. If the gallery set is big enough, we can reasonably assume that each regular FKP image will fall into the hull $H(X)$, i.e., it can be well represented as the linear combination of the template samples in X . For a query image y , we can project it into the hull $H(X)$, and rewrite it as $\hat{y} = X \cdot \hat{w}$, where weight \hat{w} is determined by minimizing the distance between y and its projection \hat{w} (usually some regularization will be imposed on \hat{w}), and e is the projection residual.

For a regular query image, there is no much finger pose variation caused deformation in it, and the projection residual e will be very small (i.e., $\|e\|_2$ is very small). In such case, the query sample y can be accurately recognized by using the efficient CompCode scheme. However, if the finger pose varies much in the data acquisition process, the query sample y can have much deformation and fall outside the hull $H(X)$, i.e., $\|e\|_2$ becomes much bigger. In this case, the CompCode scheme can fail since it is sensitive to image deformations, which can be reflected by the big values of $\|e\|_2$.

Let's plot the distribution of $\|e\|_2$ by using the PolyU FKP database (PolyU Palmprint Database 2006). We use the first six samples of each of the 660 classes in the PolyU FKP database to construct the gallery dataset X , and take the remaining six samples of each class as the query samples. In projecting each query sample into the hull $H(X)$, we use the l_2 -norm regularized least square to compute the weight: $\hat{w} = \operatorname{argmin}_w \|y - X \cdot w\|_2^2 + \lambda \|w\|_2^2$. Then e is computed as $e = y - X \cdot \hat{w}$. In Fig. 8.3 we plot the distribution of $\|e\|_2$ by using all the $660 \times 6 = 3960$ query samples. We can see that the distribution has a long tail, which is mainly caused by those samples with large deformations. Since it is mainly the big residual e that makes the CompCode feature of query sample y deviates much from the CompCode features of gallery samples in X , one intuitive idea is that we can compute the CompCode of $\hat{y} = X \cdot \hat{w}$ for verification.

Note that we use all the gallery samples, not only the samples from the class that y claims, to compute \hat{y} . There are two reasons for such a configuration. First, the number of samples of each class is usually small (e.g., six samples per class in the gallery set of PolyU FKP database), and thus reconstructing y using only the samples from one class is not accurate. Second, the query sample y can be an imposter, i.e., it may come from a class out of the gallery set, or from a class that is different from the class it claims. Reconstructing y using only the samples of one class will make the distance from \hat{y} to this class too small so that false acceptance will happen. Therefore, all the classes in the gallery set should be involved in reconstructing y .

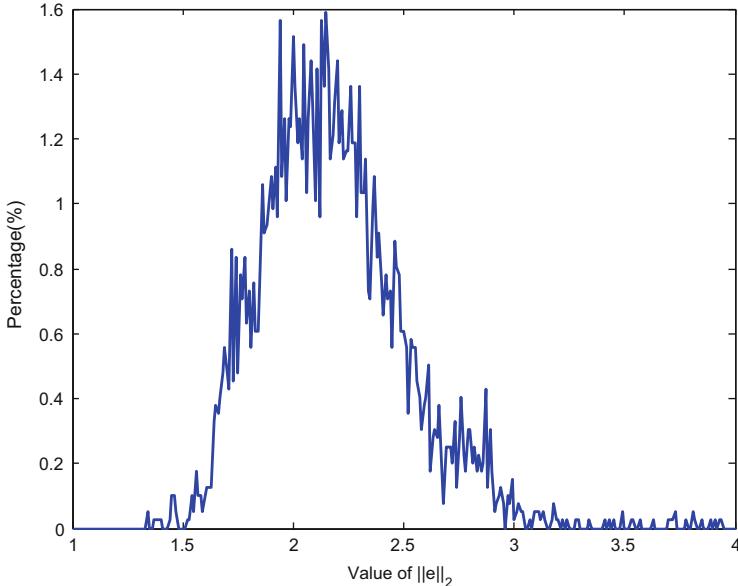


Fig. 8.3 The distribution of $\|e\|_2$

8.3.2 Reconstruction with l_1 -Norm Sparse Regularization

By approximating y with $X \cdot w$, one solution to w is the least square solution: $\hat{w} = \operatorname{argmin}_w \|y - X \cdot w\|_2^2$. It is easy to see that $\hat{w} = (X^T X)^{-1} X^T y$. Though the least square solution is simple to compute and it ensures the minimal l_2 -norm reconstruction residual of \hat{y} , it is not the best choice for the verification purpose. The least square solution aims to minimize the reconstruction residual, and the weights w tend to be densely distributed, hence many classes in X will contribute in reconstructing y . Finally, some discriminative features in y may be smoothed out in \hat{y} .

In order to preserve the discriminative features of y in \hat{y} , some regularization term could be imposed on w . Intuitively, we hope that only a small portion of the weights in w are significant so that only several classes are dominantly involved to reconstruct y . The l_1 -norm based sparse representation (or sparse coding) is a very good choice to this end. In recent years sparse coding has been successfully used in various image reconstruction and pattern classification applications (Candès and Romberg 2005; Kim et al. 2007; Aharon et al. 2006; Rubinstein et al. 2010; Yang et al. 2010; Wright et al. 2009). It represents a given signal as a sparse linear combination over a dictionary of atoms. By imposing the l_1 -norm constraint on w , we have

$$\hat{w} = \arg \min_w \|y - X \cdot w\|_2^2 + \lambda \|w\|_1 \quad (8.4)$$

where λ is a positive scalar balancing the reconstruction residual and the sparsity of w . Equation (8.4) can be solved by many convex optimization algorithms such as l_1 -magic (Candès and Romberg 2005), l_1-l_s (Kim et al. 2007), etc.

The sparse coding in Eq. (8.4) still has two problems. First, it is known that the commonly used l_1 -minimization solvers such as l_1-l_s have an empirical complexity of $O(z^2k^{1.3})$ (Kim et al. 2007), where z is the dimension of y and k is the number of samples in X . In practice, k can be very big so that the sparse coding complexity is high. Second, the atoms in X are the original gallery FKP images, which may contain noise and some trivial structures that can be negative to the representation of y .

To solve the above problems and considering the fact that the FKP images in X have much redundancy across samples, we can learn a more compact dictionary D from X , and then use D to code the input FKP image y . Dictionary learning has been widely used in image processing and pattern recognition (Aharon et al. 2006; Rubinstein et al. 2010; Yang et al. 2010, 2011; Wang et al. 2012). In this chapter, we simply adopt the method in (Yang et al. 2010) to learn the dictionary D . Denote by $D = [d_1, d_2, \dots, d_p]$, where $p \leq k$ and each d_j is a unit column vector. The dictionary learning can be formulated as the following minimization problem:

$$J_{D,W} = \arg \min_{D,W} \left\{ \|X - DW\|_F^2 + \lambda \|W\|_1 \right\} \quad (8.5)$$

Equation (8.5) is a joint optimization of dictionary D and the coefficient matrix W , and it can be solved by optimizing D and W alternatively (Yang et al. 2010). Once the dictionary D , which has less number of atoms than X , is computed, we use it to code the input FKP image y as follows:

$$\hat{w} = \arg \min_w \|y - D \cdot w\|_2^2 + \lambda \|w\|_1 \quad (8.6)$$

Finally, the image is reconstructed as:

$$\hat{y} = D \cdot \hat{w} \quad (8.7)$$

Let's show an example of the reconstruction results. We take 1980 images from the first 330 classes in the gallery set of PolyU FKP dataset as X , and learn from it a dictionary D with 1386 atoms. D is then used to reconstruct the input image y . The parameter λ is set as 0.1 and 2 in Eqs. (8.5) and (8.6), respectively. The left column of Fig. 8.5 shows an input FKP image, and the second column shows the reconstructed image using D . One can see that the reconstructed image is smoother than the original one because some details as well as variations in the query sample y , which cannot be represented by the learned dictionary D , are suppressed.

8.3.3 Reconstruction with l_2 -Norm Regularization

In Eq. (8.6), the l_1 -norm sparsity constraint is imposed on the coding coefficients to enforce that only a small portion of the atoms are dominantly used to reconstruct y . However, l_1 -minimization is time consuming. Though many fast l_1 -minimization solvers such as FISTA (Beck and Teboulle 2009), ALM (Yang et al. 2011) and Homotopy (Malioutov et al. 2005) have been developed, they may not be fast and accurate enough for practical use in the application of FKP verification, where real time implementation is expected. One intuitive solution is to relax the strong l_1 -regularization to the weaker l_2 -regularization in Eq. (8.6). The l_2 -regularization offers a closed form solution to w , which can be very efficiently computed. Though the resolved coefficient w is not sparse any more, the l_2 -regularization can still make w have a small energy. As we will see in the experimental results, the FKP verification accuracy by l_2 -regularized reconstruction is only a little lower than that by l_1 -regularized reconstruction, but the computational complexity is greatly reduced.

By using l_2 -norm to regularize w , the coding becomes a regularized least square problem:

$$\hat{w} = \underset{w}{\operatorname{argmin}} \|y - D \cdot w\|_2^2 + \lambda \|w\|_2^2 \quad (8.8)$$

and a closed form solution can be readily obtained as:

$$\hat{w} = (D^T D + \lambda \cdot 1)^{-1} D^T y \quad (8.9)$$

Let $P = (D^T D + \lambda \cdot 1)^{-1} D^T$. Clearly, P can be pre-calculated so that the coding vector can be obtained by projecting y over P : $\hat{w} = Py$. This process is very fast.

In the third column of Fig. 8.4, we show the reconstructed FKP image by l_2 -regularization in Eq. (8.8). The parameter λ is set as 4. (Note that the parameter λ in l_2 -regularized coding is usually set bigger than that in l_1 -regularized coding because l_2 -regularization is much weaker than l_1 -regularization.) For a better illustration, we also show in the right column of Fig. 8.4 the reconstructed FKP image without regularization [i.e., set $\lambda = 0$ in Eq. (8.8)]. It can be observed that due to the non-sparse regularization, more classes are involved in the reconstruction of , and thus the reconstructed image is smoother than that by l_1 -sparse regularization.

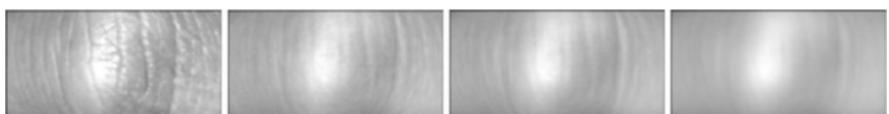


Fig. 8.4 From left column to right column: original FKP image; reconstructed FKP image by l_1 -regularization (Eq. 8.6), reconstructed FKP image by l_2 -regularization (Eq. 8.8), and reconstructed FKP image without regularization (i.e., let $\lambda = 0$)

This will lead to some loss of the distinctive features in the original FKP image. Nonetheless, this is the price we should pay for the great reduction in time complexity. In the section of experimental results, we will see that the reconstructed image by l_2 -regularization can still lead to quite competitive verification accuracy. It can also be observed from Fig. 8.4 that the reconstructed image without regularization is the smoothest and more distinctive features are lost. Our experiments also show that the verification accuracy by the reconstructed image without regularization is lower than that with, l_2 -regularization which validates that l_2 -regularization is very helpful for verification. Note that the complexity of reconstruction with l_2 -regularization is the same as that without l_2 -regularization.

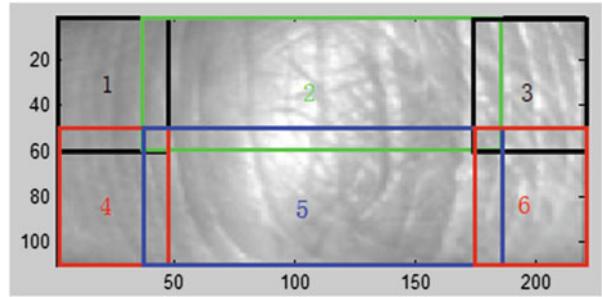
8.3.4 Patch Based Reconstruction

In Sects. 8.3.2 and 8.3.3, we stretch the whole FKP image as a vector y for coding and reconstruction. In coding y over D by Eq. (8.6) or Eq. (8.8), we actually enforce that all the elements (i.e., all the pixels) in , denoted by y_i , $i = 1, 2, \dots, n$, share the same coding vector w over their corresponding sub dictionary (i.e., the i th row of D). The good side of such a global coding strategy is that the solution is very stable because only one global coding vector needs to be solved. The bad side of such a coding strategy, however, lies in its less flexibility because it does not allow the different parts of the FKP image to have different coding vectors. Considering the fact that different portions of a query FKP sample y may have different variations, it is reasonable to allow them to have different coding vectors so that the reconstruction can be spatially adaptive. Therefore, we can partition the FKP image into several patches, reconstruct separately each patch, and then combine them to obtain the whole reconstructed image.

If we partition the image into too many patches, the size of each patch will become small and the sub-dictionary corresponding to each patch will tend to be under-determined (i.e., the sub-dictionary will tend to be a fat matrix). This will reduce the stability of the coding process, no matter l_1 -regularization or l_2 -regularization is used. Based on our experimental experience and considering the special pattern of FKP images, we partition the FKP image (size: 110×220) into 6 overlapped patches, as illustrated in Fig. 8.5. Four patches of size 60×45 lie in the four corners of the FKP images, and two patches of size 60×155 lie in the middle of the image. The reason that we set two fat rectangle patches in the middle is based on the observation that the pose variation along vertical direction has bigger effect than that along horizontal direction on the FKP recognition accuracy. Therefore, pay more attention to the pose variation along vertical direction in the partition.

Let's denote by y_j , $j = 1, 2, \dots, 6$, the six patches of an FKP image y . For each patch, we can learn a dictionary D_j from the training samples by using the same method described in Sect. 8.3.2. Then for an input query sample y , each patch y_j of it

Fig. 8.5 The patch partition of FKP images



can be reconstructed by D_j . Similarly, both l_1 -regularization and l_2 -regularization can be employed in the reconstruction of y_j via

$$\hat{w} = \underset{w_j}{\operatorname{argmin}} \ \|y_j - D_j \cdot w_j\|_2^2 + \lambda \|w_j\|_1 \quad (8.10)$$

and

$$\hat{w} = \underset{w_j}{\operatorname{argmin}} \ \|y_j - D_j \cdot w_j\|_2^2 + \lambda \|w_j\|_2^2 \quad (8.11)$$

respectively. After each patch is reconstructed, the whole reconstructed image can be obtained by combining them. For the overlapped area of neighboring patches, we simply average the results. Similar to our discussions in Sects. 8.3.2 and 8.3.3, the l_1 -regularization in Eq. (8.10) may preserve more discrimination information, but the l_2 -regularization in Eq. (8.11) is much faster.

At last, let's use an example to illustrate the performance of patched based reconstruction. Figure 8.6 shows a query sample y and a template image which is from the same class of y . We reconstruct y with l_2 -regularization in this example. Image \hat{y}_g is the output of global reconstruction by Eq. (8.8) (we set $\lambda = 1$), while \hat{y}_p is the output of patch based reconstruction by Eq. (8.11) (we set $\lambda = 0.1$). One can see that the patch based reconstruction can preserve more distinctive features than the global reconstruction due to the higher flexibility in coding coefficients of each patch. After extract the CompCode feature maps (refer to Sect. 8.2 please), the matching distances between the template image and y , \hat{y}_g and \hat{y}_p are 0.3783, 0.3761 and 0.3631, respectively. One can see that the global reconstruction can reduce the genuine matching distance, while the patched based reconstruction can further reduce the matching distance. Though imposter matching distance will also be reduced after reconstruction, we will see that with an adaptive binary fusion strategy proposed in next section, more accurate verification results can be obtained by adaptively fusing the matching distances before and after reconstruction. In the experimental results in Sect. 8.6, we will also see that patch based reconstruction can lead to better verification performance.

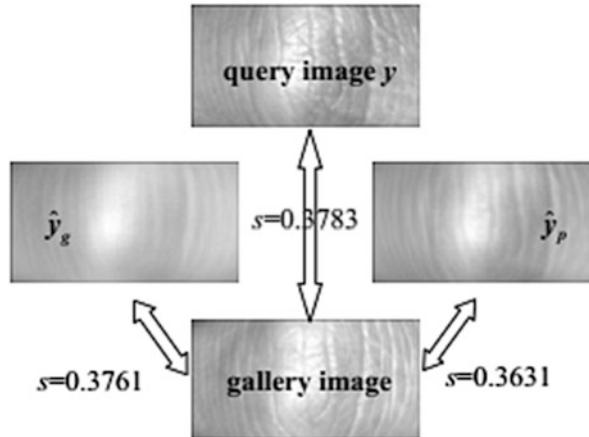


Fig. 8.6 A query sample y and the reconstructed images of it with l_2 -regularization. \hat{y}_g is the output of global reconstruction while \hat{y}_p is the output of patch based reconstruction. The matching distances between a gallery image (which is from the same class as y) and y , \hat{y}_g and \hat{y}_p are 0.3783, 0.3761 and 0.3631, respectively. One can see that the global reconstruction can reduce the genuine matching distance, while the patched based reconstruction can further reduce the matching distance

8.4 Verification by Binary Score Level Fusion

As can be observed in Figs. 8.4 and 8.6, the reconstruction can reduce the deformation caused matching distance so that the intra-class matching distance can be reduced. At the same time, however, the inter-class matching distance may also be reduced. Figure 8.7 shows some examples. One can see that the imposter inter-class matching distance is also reduced after the query sample is reconstructed (global reconstruction with l_2 -regularization is used). What we expected is that the intra class genuine matching distance can be reduced more than the inter-class imposter matching distance, but there is no guarantee for such an ideal situation. If we directly apply the CompCode scheme (Kong and Zhang 2004) to the reconstructed image for verification, incorrect decision can be made. To make this clear, we apply CompCode to the original images and the reconstructed images, respectively, for FKP verification. The equal error rates (EER) are shown in Table 8.1, from which we can see that using only the reconstructed images for verification leads to worse performance because some useful texture features are smoothed out in the reconstructed images. For a more robust and accurate verification, the matching scores (or distances) of both the original image y and the reconstructed image \hat{y} should be considered for decision making.

Denote by s_1 and s_2 the matching distances of y and \hat{y} to a gallery image, respectively. We propose to fuse the two distance scores for final decision making. In the following, we first briefly review the existing popular score level fusion

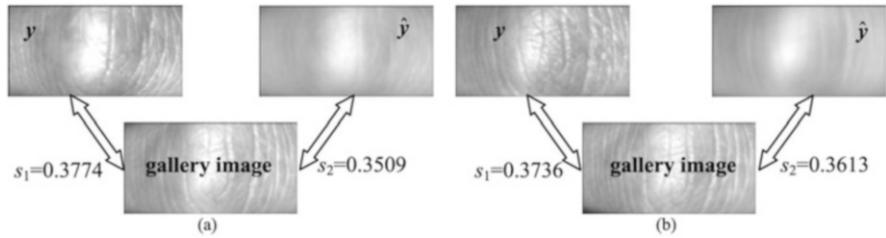


Fig. 8.7 Examples of matching distances before and after reconstruction. (a) inter-class; (b) intra-class

Table 8.1 FKP verification by using only the original image and only the reconstruction image, respectively

Images used	Original	Reconstructed
EER	1.65%	2.09%

methods, and then propose a simple but very effective adaptive fusion method, namely the adaptive binary fusion rule.

8.4.1 Popular Score Level Fusion Methods

The score level fusion is a kind of combination-based approach, where the matching scores of individual matchers are integrated to generate a single scalar score for final decision making. Denote by s the fusion result of s_1 and s_2 . Three commonly used score level fusion methods (Snelick et al. 2005) are the simple-sum (SS):

$$s = s_1 + s_2 \quad (8.12)$$

the min-score (MIN):

$$s = \min\{s_1, s_2\} \quad (8.13)$$

and the max-score (MAX):

$$s = \max\{s_1, s_2\} \quad (8.14)$$

The above three fusion rules do not use any additional information apart from the matching scores s_1 and s_2 . Another popular method is the matcher weighting (MW) scheme (Snelick et al. 2005). The fused score is the weighted average of the two scores:

$$s = \omega_1 \cdot s_1 + \omega_2 \cdot s_2 \quad (8.15)$$

where $0 \leq \omega_1, \omega_2 \leq 1$ and $\omega_1 + \omega_2 = 1$. To determine ω_1 and ω_2 , some prior knowledge needs to be known. Often the EERs of the two matchers (i.e., matching by y and matching by \hat{y}) are used. Using some training dataset, the two EERs, denoted by e_1 and e_2 , can be obtained, and the weight can then be calculated as

$$\omega_j = \left(1 / \left(\sum_{j=1}^2 (1/e_j) \right) \right) / e_j \quad (8.16)$$

It is obvious that the weight is inversely proportional to the corresponding EER. A higher weight will be assigned to a more reliable (i.e., lower EER) matcher, and vice versa.

There are also some other score level fusion formulas. For example, in (Kumar et al. 2010) Kumar et al. proposed the exponential sum rule $s = \sum_{j=1}^n \exp(s_j) \omega_j$ and the tan-hyperbolic sum rule $s = \sum_{j=1}^n \tanh(s_j) \omega_j$. The Particle Swarm Optimization (PSO) (Eberhart and Kennedy 2010) is employed to dynamically select the weights ω_j . Such dynamic rules may work better than the SS, MIN, MAX and MW rules when there are multiple biometric identifiers for fusion, but they need to optimize the fusion rules, weights, and decision thresholds. Overall, the PSO based dynamic score level fusion is complex and has high computational complexity (Kumar et al. 2010).

8.4.2 Adaptive Binary Fusion

In general, the MW fusion rule works better than SS, MIN and MAX rules. However, the MW rule has two drawbacks. First, it needs a preset training dataset to train the weights. Second, once the weights are learned, they are applied to all query images y and the reconstructed images \hat{y} . In other words, the MW rule is not adaptive to the input query image, limiting its performance. This phenomenon can be seen in Fig. 8.7. In Fig. 8.7b, the reconstructed image is more reliable for verification, so the weight ω_2 assigned to this matcher is higher ($\omega_1 < \omega_2$). However, the situation is opposite in Fig. 8.7a. Because the query sample is an imposter, it is hoped that the weight ω_1 assigned to the original image is higher ($\omega_1 > \omega_2$) so that the final matching distance can be larger. The MW rule cannot meet this requirement because ω_1 and ω_2 are fixed.

In this chapter, we propose a new fusion rule, which is adaptive to the query image and does not need a preset training dataset. The weights ω_1 and ω_2 are adaptively determined online based on the input image pair and \hat{y} . The idea is as follows. For the query image y which is claimed to belong to class c , we can calculate its within-class and between class matching distances. Using CompCode, those matching distances can be computed very fast by Eq. (8.3). So does for the

reconstructed query image \hat{y} . Then the higher weight is assigned to the matcher whose within-class and between-class matching distances are better separated.

Denote by $\mu_{1,w}$ and $\mu_{1,b}$ the mean values of the within class and between-class matching distances of y , respectively, and by $\sigma_{1,w}$ and $\sigma_{1,b}$ the standard deviations of the within class and between-class matching distances, respectively. For both genuine matching and imposter matching, the within class distances are calculated between the query image and the class which it claims to belong to, and the between-class distances are calculated between the query image and other classes. Because of the fast speed of CompCode, this process can be implemented in less than one second on the PolyU FKP database. For very large scale databases, computing the between-class distance using the whole dataset can be costly. To save cost, we can randomly select an enough number of samples from the classes other than the claimed class to compute a good approximation of the between-class distance.

The decidability index (Daugman 2003) can be used to measure the separability of the distributions of within-class and between class matching distances. The decidability index for y is calculated as follows:

$$d_1 = \frac{|\mu_{1,w} - \mu_{1,b}|}{\sqrt{(\sigma_{1,w}^2 + \sigma_{1,b}^2)/2}} \quad (8.17)$$

Similarly, we can calculate the decidability index of \hat{y} and denote it as d_2 . A bigger decidability index means that the within-class and between-class matching scores can be better separated, and hence the matcher is more accurate, vice versa. Therefore, d_1 and d_2 can be used to adaptively determine the weights ω_1 and ω_2 that are assigned to s_1 and s_2 . The higher the decidability index, the higher the weight. However, designing an optimal function to map d_1 and d_2 to ω_1 and ω_2 is not a trivial work. In this chapter, we propose to use the simple binary logic operation for the weight determination. The so-called adaptive binary fusion (ABF) rule is defined as follows:

Adaptive Binary Fusion (ABF): If $d_1 \geq d_2$, $\omega_1 = 1$ and $\omega_2 = 0$; otherwise, $\omega_1 = 0$ and $\omega_2 = 1$.

The above proposed ABF rule is a kind of “winner-take all” strategy and is similar to the notion of using cohort scores for multi-biometric fusion. However, there are clear differences between them. In cohort score based multi biometric fusion, the fusion weight is fixed for each biometric identifier once the weights are learned offline. All the users of one biometric identifier share the same pre-learned fusion weight. Instead, in the proposed ABF the fusion weight is adaptively determined for each user online. The ABF rule is simple but fits our application very well. The reconstruction of y can only lead to two situations: the reconstructed image \hat{y} is either better or worse than y for verification. Hence, it is reasonable and effective to adaptively choose one of them for the final decision making. Our

experimental results in Sect. 8.6 validate the effectiveness of the proposed ABF rule.

8.5 Summary of the Verification Algorithm

In practice, it is not necessary to reconstruct every input query image for verification. As shown in Fig. 8.2, if the matching distance of a query image y falls into the confident interval, we can directly make the decision; only when the matching distance falls into the uncertain interval $[t_1, t_2]$, there construction is needed, and the ABF rule is applied for the verification. The proposed algorithm of reconstruction based FKP verification with ABF is summarized in Table 8.2. The output is the final matching distance s . The final decision (accept or reject) is then made by applying a threshold to s , as in existing FKP verification systems (Zhang et al. 2009a, b, c, 2010, 2011, 2012).

8.6 Experimental Results

The PolyU FKP database, which can be freely downloaded at: <http://www.comp.polyu.edu.hk/~biometrics/FKP.htm>, was used in our experiments. This database contains the cropped FKP region of interest (ROI) images of four fingers (the left index finger, the left middle finger, the right index finger and the right middle finger) from 165 persons. Each finger knuckle was acquired 12 samples in two separated sessions with 6 samples per session, giving a total of $165 \times 4 \times 12 = 7920$ samples from 660 (i.e., 165×4) fingers.

In the following experiments, the gallery set is always extracted from the first session while the probe set is extracted from the second session. As in (Zhang et al. 2009a, b, c, 2010, 2011, 2012), each image in the probe set was matched with all the

Table 8.2 Summary of the proposed FKP verification algorithm

Input: The gallery dataset X and the trained dictionary D ;

The query image y ;

The interval $[t_1, t_2]$.

Output: Matching distance s .

1. Calculate the minimal matching distance, denoted by s_1 , from y to the class it claims by using the CompCode scheme.
 2. If s_1 is outside $[t_1, t_2]$, output $s = s_1$ and end the matching process.
 3. Otherwise, reconstruct \hat{y} from y by using the dictionary D .
 4. Calculate the minimal matching distance, denoted by s_2 , from \hat{y} to the class y claims by using the CompCode scheme.
 5. Calculate the decidability indices d_1 and d_2 , and then fuse s_1 and s_2 using the ABF rule.
 6. Output the fused matching distance s and end the matching process.
-

images in the gallery set. If the two images were from the same finger, a genuine matching was counted; otherwise, an imposter matching was counted.

8.6.1 Comparison Between Different Fusion Rules

In this section, we verify that the proposed ABF rule is more effective than the commonly used SS, MIN, MAX and MW rules. In the experiment, the gallery set is composed of the first 165 fingers in the PolyU FKP gallery set, and the probe set is composed of the first 330 fingers in the PolyU FKP probe set. (Other settings of the gallery and probe sets lead to similar conclusions.) That is, there are 165 classes out of the gallery set. The gallery set is used to train the dictionary D to code the samples in the probe set.

In this experiment the l_1 -regularization is used in the reconstruction of query samples. In the dictionary learning (Eq. 8.5) and sparse coding (Eq. 8.6), the parameter λ is set as 0.1 and 0.5. The number of atoms in the learned dictionary is 0.7 times the number of samples in the gallery set. The parameters for uncertain interval setting is $t_1 = 0.35$ and $t_2 = 0.39$ based on our experimental experience. If the class label of the query sample is out of the gallery set, only the imposter matching distance will be counted when calculating EER.

The algorithm described in Table 8.2 is used to perform the FKP verification experiments with different fusion rules in step 5. For the MW rule, the whole dataset is used to train the weights. The EER results by using the different fusion rules are listed in Table 8.3. We can clearly see that the lowest EER is obtained by the proposed ABF rule, which works much better than other rules. Even that the MW rule uses more information with a training dataset, it is only slightly better than the MIN, SS and MAX rules. In the following experiments, we only report the results by using the ABF rule.

8.6.2 Experiment Settings and Parameter Selection

We compare the proposed reconstruction based FKP verification method with ABF, denoted by “R-ABF”, with state of-the-art FKP verification methods, including the improved CompCode (ImCompCode) (Zhang et al. 2010), BLPOC (Zhang et al. 2009a, b, c), LGIC (Zhang et al. 2011) and the local feature integration (LFI) method (Zhang et al. 2012). Considering that the LGIC scheme is a combination of the CompCode (which employs the image local orientation features) and the BLPOC (which employs the image global Fourier transform features) methods,

Table 8.3 EER (%) values by different fusion rules

Fusion rules	SS	MIN	MAX	MW	ABF
EER	2.22	2.07	2.33	1.99	1.10

for fair comparison we will first compare R-ABF with ImCompCode, BLPOC and LFI, and then couple R-ABF with BLPOC and compare it with the LGIC method.

In the proposed method, the query image reconstruction can be regularized by either l_1 -norm or l_2 -norm, and can be done either globally or patch-by-patch. Therefore, there are four variants of the proposed R-ABF scheme, denoted by R-ABF-g- l_1 , R-ABF-g- l_2 , R-ABF-p- l_1 and R-ABF-p- l_2 , respectively, where “g” is for “global” and “p” is for “patch”. In order to evaluate the proposed R-ABF method more comprehensively, we conduct 3 experiments with different sizes of the gallery set (165, 330 and 660 fingers, respectively). In all the experiments, the gallery set is extracted from the FKP images collected in the first session, while all the FKP samples collected in the second session are used as the probe set (all the 660 fingers).

There are some parameters to set in our algorithm. For the parameters in CompCode, we adopt the settings in the original chapter (Zhang et al. 2009a, b, c). The uncertain interval is set by letting $t_1 = 0.35$ and $t_2 = 0.39$. The thresholds t_1 and t_2 can also be automatically determined based on the training set. By using the “leave-one-out” strategy, each image of one subject is matched with all the other images in the training set to obtain the distribution of matching distances. The point where the false acceptance rate equals to the false rejection rate is taken as the decision threshold t , and the width of the uncertain interval (refer to Fig. 8.2) can be set as w_0 . Then t_1 and t_2 can be determined as: $t_1 = t - w_0/2$ and $t_2 = t + w_0/2$. By using this strategy, the computed t_1 and t_2 are 0.34 and 0.38, respectively, which are very close to the manually set thresholds (i.e., 0.35 and 0.39).

In the global reconstruction scheme, we set the parameter λ in dictionary learning (i.e., in Eq. (8.5)) as 0.1 by experience. The selection of parameter λ in dictionary learning has small influence on the final verification result in terms of EER. The parameter λ for l_1 -regularized coding in Eq. (8.6) and l_2 -regularized coding in Eq. (8.8) is related to the number of atoms in dictionary D . We determine it by the following criterion: $\lambda = 0.5 \times (n/990)^2$ in l_1 -regularized coding and $\lambda = (n/990)^2$ in l_2 -regularized coding, where n is the number of atoms in the dictionary. In patch based reconstruction, the parameter λ in dictionary learning and coding is set as 1/10 of that in global reconstruction. The partition of patches is discussed in Sect. 8.6.5.

Our method needs to train a dictionary D with atom number $p \leq k$, where k is the number of samples in gallery set X . If p is too small, much information contained in the gallery sample set X can be lost and thus the test sample may not be well represented by the learned dictionary D . On the other hand, the commonly used l_1 -minimization solvers such as $l_1 - l_S$ have an empirical complexity of $O(z^2 p^{1.3})$ (Kim et al. 2007) (z is the dimension of samples). So if p is big, the computational cost can be high. To balance the computational cost and the representation capability, we set the ratio of p to k as 0.7 in our method, and this configuration leads to satisfying experimental results.

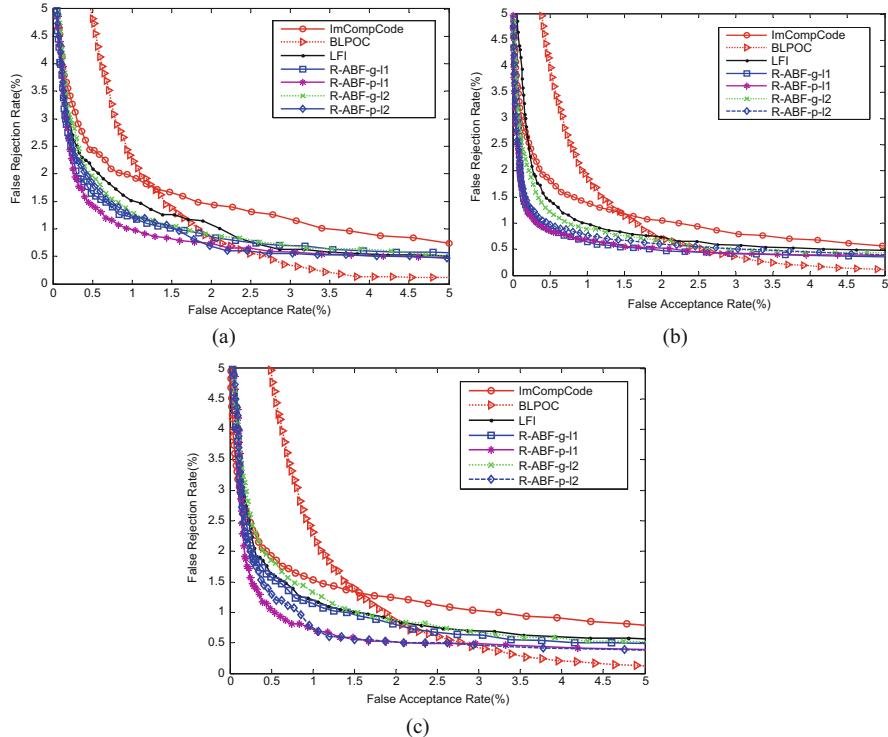


Fig. 8.8 DET curves by different methods in (a) experiment 1; (b) experiment 2; and (c) experiment 3

8.6.3 FKP Verification Results

1. Experiment 1: In this experiment, only the FKP images of the first 165 (out of the 660) fingers (i.e., $165 \times 6 = 990$ samples) are used as the gallery set. Hence, there are 5940 genuine matchings and 3,914,460 imposter matchings, respectively. Figure 8.8a plots the DET (Detection Error Tradeoff) curves, which are the plots of false rejection rates (FRR) against false acceptance rates (FAR) for all possible thresholds. Table 8.4 lists the results of competing methods in terms of EER and decidability index.

It can be seen that the R-ABF methods outperform much ImCompCode, BLPOC and LFI. This validates that the reconstruction of query sample y can reduce much the image deformation induced intra-class distance, and the ABF rule can prevent the less discriminative reconstruction of y from being adopted for final decision making. Among the four variants of R-ABF, the l_1 -regularized ones

Table 8.4 EER(%) values and decidability indices by the competing methods

Method	Experiment 1		Experiment 2		Experiment 3	
	EER (%)	Decidability index	EER (%)	Decidability index	EER (%)	Decidability index
ImCompCode (Zhang et al. 2010)	1.62	4.1478	1.28	4.4518	1.39	4.4302
BLPOC (Zhang et al. 2009a, b, c)	1.44	4.1925	1.35	3.4311	1.44	3.4218
LFI (Zhang et al. 2012)	1.32	4.1951	1.03	4.4821	1.13	4.4625
R-ABF-g- l_1	1.24	4.2291	0.84	4.5699	0.94	4.5305
R-ABF-p- l_1	1.12	4.2309	0.76	4.6385	0.82	4.2572
R-ABF-g- l_2	1.27	4.1948	0.91	4.5014	1.19	4.1882
R-ABF-p- l_2	1.28	4.2170	0.82	4.6188	0.92	4.2461

have higher accuracy than the l_2 -regularized ones, while the patch based ones have higher accuracy than the global based ones. Specifically, R-ABF-p- l_1 achieves the lowest EER. This is consistent with our discussions in Sect. 8.3.

Table 8.5 lists the percentage of matchings in which the score from \hat{y} or the reconstruction is selected as the final matching score. We can see that \hat{y} for genuine matchings, about 40% ~ 45% of the matching scores from \hat{y} are selected by the ABF rule.

For imposter matchings, about 28% ~ 35% of the matching scores from \hat{y} are adopted.

2. Experiment 2: In the 2nd experiment, 330 classes are involved in the gallery set. Therefore, the numbers of genuine and imposter matchings are 11,880 and 7,828,920, respectively. Figure 8.8b shows the DET curves by the different verification schemes, while the EER values and decidability indices are listed in Table 8.4. The percentage of matchings in which the score from the reconstruction \hat{y} is selected is listed in Table 8.5. Again, the proposed R-ABF methods get much better results than ImCompCode, BLPOC and LFI.
3. *Experiment 3:* At last, all the classes (i.e., all fingers) are involved in the gallery set, and the numbers of genuine and imposter matchings are 23,760 and 15,657,840, respectively. The DET curves by different verification schemes are illustrated in Fig. 8.8c. Table 8.4 lists the EER values and decidability indices, and Table 8.5 lists the percentage of matchings in which the score from the reconstruction \hat{y} is selected. Similar conclusions to the previous two experiments can be made.

It can be seen that the EER decreases from Experiment 1 to Experiment 2, and increases from Experiment 2 to Experiment 3. The reason can be as follows. From Experiment 1 to Experiment 3, the number of gallery classes is increasing. The increased number of gallery samples makes the dictionary D more capable to

Table 8.5 The percentage (%) of matchings in which the score from y or \hat{y} is selected

Matching type	Experiment 1		Experiment 2		Experiment 3	
	$\omega_1 = 1$	$\omega_2 = 1$	$\omega_1 = 1$	$\omega_2 = 1$	$\omega_1 = 1$	$\omega_2 = 1$
Genuine	60	40	55	45	57	43
Imposter	71.59	28.41	65.36	34.61	68.74	31.26

reconstruct the query sample, but it also makes the verification tasks more challenging. There are only 165 gallery classes in Experiment 1, so the learned dictionary D may not be representative enough to reconstruct the query sample. The number of gallery classes is increased to 330 in Experiment 2, and the representativeness of D is much improved so that the FAR and FRR are decreased simultaneously. As a consequence, the overall EER in Experiment 2 is reduced. With 330 gallery classes, the representativeness of learned dictionary D is already good. Thus, the benefit of using 660 gallery classes in Experiment 3 is not big in term of learning dictionary D ; however, the FAR and FRR are increased simultaneously due to the increased number of gallery classes, resulting in a bigger EER than Experiment 2.

8.6.4 Integrating with Global Features

The LGIC scheme (Zhang et al. 2011) combines CompCode, which employs the image local orientation features, and BLPOC, which employs the global Fourier transform features. In this section, for fair comparison, we also combine RABF, which basically employs the image local orientation features, with the BLPOC method in the same way as that in (Zhang et al. 2011). According to the reconstruction strategy, we denote the fused methods as R-ABF-g- l_1 + BLPOC, R-ABF-g- l_2 + BLPOC, R-ABF-p- l_1 + BLPOC and R-ABF-p- l_1 + BLPOC, respectively. We compare the performance of LGIC with the proposed methods under the same experiment settings as described in Sect. 8.6.3. The combination of LFI and BLPOC is also used for more comprehensive comparison. The DET curves of the competing methods are shown in Fig. 8.9, and the EER values and decidability indices are summarized in Table 8.6.

From the experimental results in both Sects. 8.6.3 and 8.6.4, we can see that the proposed R-ABF scheme leads to state-of-the-art verification accuracy, no matter using only the local orientation feature or using both local and global features. Specifically, the R-ABF-p- l_1 and R-ABF-p- l_1 + BLPOC methods achieve the best accuracy, respectively. By relaxing the l_1 -regularized sparsity constraint in the reconstruction, the l_2 -regularized reconstruction can also lead to very competitive verification results but with much less complexity, which is a good solution in practical FKP recognition systems.

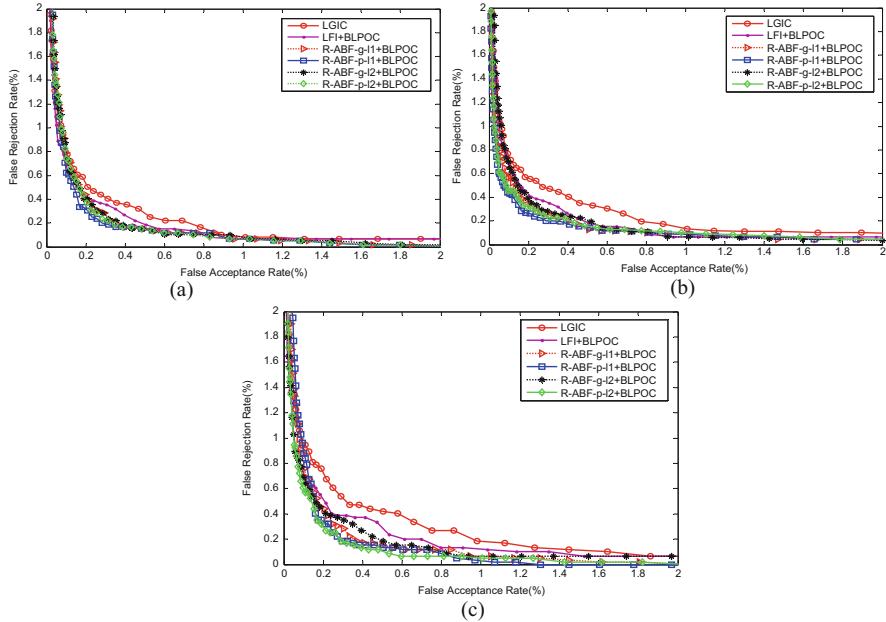


Fig. 8.9 DET curves by the fused methods in (a) experiment 1; (b) experiment 2; and (c) experiment 3

8.6.5 Discussions

In the proposed method, we learn a dictionary D from the gallery set to reconstruct a query sample. When a new subject is enrolled, we can update D by solving Eq. (8.5). However, if the dataset has a large scale, this can be very costly. Fortunately, it is not necessary to update D for a new enrollment in large-scale dataset.

The dictionary D in our algorithm is just used for reconstruction, and it is not used in the classification stage. This is very different from the works in (Yang et al. 2010, 2011; Wright et al. 2009), where the atoms in the dictionary have class labels and they will be used to calculate the class-specific distances for classification. The role of dictionary D in our work is similar to the dictionaries in image restoration such as K-SVD (Aharon et al. 2006). There are no class labels of the atoms in D , and D is a universal dictionary shared by all classes. Once enough gallery classes are involved in learning the dictionary D , this D will be able to well represent any FKP image. Therefore, when there are some new enrollments, we actually do not need to update the dictionary. Kindly note that the online learning algorithms in (Masip et al. 2009) and (Singh et al. 2010) aim for updating classifiers when new individuals are enrolled. The problem is very different from ours.

Table 8.6 EER (%) values and decidability indices by different methods

Method	Experiment 1		Experiment 2		Experiment 3	
	EER (%)	Decidability index	EER (%)	Decidability index	EER (%)	Decidability index
LGIC (Zhang et al.)	0.38	4.3882	0.39	4.6182	0.41	4.4302
LFI+BLPOC	0.33	4.3946	0.35	4.6378	0.35	4.6635
R-ABF-g- l_1 + BLPOC	0.29	4.4197	0.25	4.6886	0.29	4.6940
R-ABF-p- l_1 + BLPOC	0.25	4.4150	0.21	4.7479	0.22	4.4823
R-ABF-g- l_2 + BLPOC	0.29	4.4048	0.29	4.5937	0.33	4.4641
R-ABF-p- l_2 + BLPOC	0.27	4.4076	0.26	4.7354	0.26	4.4764

Let us use two experiments to validate the above statement. We use the first 600 classes out of the 660 classes in the PolyU FKP database as the gallery set (6 samples per class in the first session) to learn a dictionary, denoted by D_1 . We then use all the 660 classes to learn another dictionary, denoted by D_2 . In the first experiment, we take D_1 as the dictionary and take the 60 new classes (6 samples per class) as the query set, and the EER is 0.8% by the proposed R-ABF-g- l_2 method. If we take as the dictionary D_2 and take the 60 same classes as the query set, the EER is 0.76% by R-ABF-g- l_2 . In the second experiment, with D_1 and using all the 660 classes as the query set (6 samples per class in the second session), the EER is 1.21% by R-ABF-g- l_2 , while with D_2 and using all the 660 classes as the query set, the EER is 1.19% by R-ABFg- l_2 . Clearly, in both the two experiments, the EER values by D_1 and D_2 are very close, implying that there is no necessary to further update the dictionary since D_1 is already good in reconstruction.

If we do want to update the dictionary D when new enrollments come, there are two strategies to save cost. First, we can update D once a batch of new enrollments is available. Second, we can let the new dictionary be $D_{new} = [D_o D_a]$, where D_o is the old dictionary and D_a includes the new atoms to be added. In this way, we only need to learn the several new atoms by using dictionary learning algorithms such as (Yang et al. 2010).

8.7 Summary

This chapter presented a novel reconstruction based finger knuckle-print (FKP) verification method to reduce the false rejections caused by finger pose variations in data collection process. For an input query image whose matching distance falls into the uncertain interval, we reconstructed a new version of it by using a dictionary learned from the gallery set. Then a new matching distance can be

obtained. An adaptive binary fusion (ABF) rule was then proposed to fuse the two matching distances for the final decision making. The proposed reconstruction based FKP verification with ABF, denoted by R-ABF, can effectively reduce the false rejections without increasing much the false acceptances. Our extensive experimental results demonstrated that the R-ABF can result in much lower equal error rate than existing state-of-the-art methods.

References

- Aharon M, Elad M, Bruckstein A (2006) K-SVD: an algorithm for designing over complete dictionaries for sparse representation. *IEEE Trans Signal Process* 54(11):4311–4322
- Beck A, Teboulle M (2009) A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J Imag Sci* 2(1):183–202
- Candès E, Romberg J (2005) 1-Magic: recovery of sparse signals via convex programming [Online]. <http://www.acm.caltech.edu/1magic/>
- Daugman J (2003) The importance of being random: statistical principles of iris recognition. *Pattern Recogn* 36(2):279–291
- Eberhart R, Kennedy J (2010) Swarm intelligence. Morgan Kaufmann, San Diego, CA
- Kim S, Koh K, Lustig M, Boyd S, Gorinevsky D (2007) An interior point method for large-scale ℓ_1 -regularized least squares. *IEEE J Sel Topics Signal Process* 1(4):606–617
- Kong W, Zhang D (2004) Competitive coding scheme for palmprint verification. In: Proceedings of IEEE international conference on pattern recognition, vol 1, pp 520–523
- Kumar A, Kanhangad V, Zhang D (2010) A new framework for adaptive multimodal biometrics management. *IEEE Trans Inf Forensics Secur* 5(1):92–102
- Malioutov D, Cetin M, Willsky A (2005) Homotopy continuation for sparse signal representation. In: Proceedings of IEEE international conference on acoustics, speech signal processing, vol 5, pp 733–736
- Masip D, Lapedriza A, Vitria J (2009) Boosted online learning for face recognition. *IEEE Trans Syst Man Cybernet B Cybernet* 39(2):530–538
- Rubinstein R, Bruckstein A, Elad M (2010) Dictionaries for sparse representation modeling. *Proc IEEE* 98(6):1045–1057
- Singh R, Vatsa M, Ross A, Noore A (2010) Biometric classifier update using online learning: a case study in near infrared face verification. *Image Vis Comput* 28(7):1098–1105
- Snelick R, Uludag U, Mink A, Indovina M, Jain A (2005) Large-scale evaluation of multimodal biometric authentication using state-of-the-art systems. *IEEE Trans Pattern Anal Mach Intell* 27(3):450–455
- Wang S, Zhang L, Liang Y, Pan Q (2012) Semi-coupled dictionary learning with applications to image super-resolution and photo-sketch image synthesis. *Proc IEEE Conf Comput Vis Pattern Recogn*:2216–2223
- Wright J, Yang A, Ganesh A, Sastry S, Ma Y (2009) Robust face recognition via sparse representation. *IEEE Trans Pattern Anal Mach Intell* 31(2):210–227
- Yang J, Zhang Y (2011) Alternating direction algorithms for ℓ_1 -problems in compressive sensing. *SIAM J Sci Comput* 33(1):250–278
- Yang M, Zhang L, Yang J, Zhang D (2010) Metaface learning for sparse representation based face recognition. In: Proceedings of IEEE international conference on image processing, pp 1601–1604
- Yang M, Zhang L, Feng X, Zhang D (2011) Fisher discrimination dictionary learning for sparse representation. In: Proceedings of IEEE international conference computer vision, pp 543–550

- Zhang D, Lu G, Li W, Zhang L, Luo N (2009a) Palmprint recognition using 3-D information. *IEEE Trans Syst Man Cybern C Appl Rev* 39(5):505–519
- Zhang L, Zhang L, Zhang D (2009b) Finger-knuckle-print: a new biometric identifier. In: Proceedings of 16th IEEE international conference on image processing, pp 1981–1984
- Zhang L, Zhang L, Zhang D (2009c) Finger-knuckle-print verification based on band-limited phase-only correlation. In: Proceedings of international conference on computer analysis of images patterns, pp 141–148
- Zhang L, Zhang L, Zhang D, Zhu H (2010) Online finger-knuckle print verification for personal authentication. *Pattern Recogn* 43(7):2560–2571
- Zhang L, Zhang L, Zhang D, Zhu H (2011) Ensemble of local and global information for finger-knuckle-print recognition. *Pattern Recogn* 44(9):1990–1998
- Zhang L, Zhang L, Zhang D, Guo Z (2012) Phase congruency induced local features for finger-knuckle-print recognition. *Pattern Recogn* 45(7):2522–2531

Part III

Other Hand-Based Biometrics

Chapter 9

3D Fingerprint Reconstruction and Recognition

Abstract The chapter studies a 3D fingerprint reconstruction technique based on multi-view touchless fingerprint images. This technique offers a solution for 3D fingerprint image generation and application when only multi-view 2D images are available. However, the difficulties and stresses of 3D fingerprint reconstruction are the establishment of feature correspondences based on 2D touchless fingerprint images and the estimation of the finger shape model. In this chapter, several popular used features, such as scale invariant feature transformation (SIFT) feature, ridge feature and minutiae, are employed for correspondences establishment. To extract these fingerprint features accurately, an improved fingerprint enhancement method has been proposed by polishing orientation and ridge frequency maps according to the characteristics of 2D touchless fingerprint images. Therefore, correspondences can be established by adopting hierarchical fingerprint matching approaches. Through an analysis of 440 3D point cloud finger data (220 fingers, 2 pictures each) collected by a 3D scanning technique, i.e., the structured light illumination (SLI) method, the finger shape model is estimated. It is found that the binary quadratic function is more suitable for the finger shape model than the other mixed model tested in this chapter. In our experiments, the reconstruction accuracy is illustrated by constructing a cylinder. Furthermore, results obtained from different fingerprint feature correspondences are analyzed and compared to show which features are more suitable for 3D fingerprint images generation.

Keywords 3D fingerprint reconstruction • Finger shape model • Finger-print features correspondences • Orientation map • Frequency map • Touch-less multi-view imaging

9.1 Introduction

As one of the most widely used biometrics, fingerprints have been investigated for more than a century (Maltoni et al. 2009). Advanced Automated Fingerprint Recognition Systems (AFRSs) are available in the market everywhere and most of them capture fingerprint images by using the touch-based technique, since it is easy to obtain images with high ridge-valley contrast. However, the touch-based

imaging technique introduces distortions and inconsistencies to the images due to the contact of finger skin with device surface. In addition, the curved 3D finger surface flattens into 2D plane during image acquisition, destroying the 3D nature of fingers. To deal with these problems, 3D fingerprint imaging techniques start to be considered (Parziale and Diaz-Santana 2006; Hartley 2000; Hernandez et al. 2008; Blais et al. 1988; Wang et al. 2010; Stockman et al. 1988; Hu and Stockman 1989). Usually, these techniques capture fingerprint images at a distance and provide the 3D finger shape feature simultaneously. The advent of these techniques brings new challenges and opportunities to existing AFRSs.

Currently, there are three kinds of popular 3D imaging techniques: multi-view reconstruction (Parziale and Diaz-Santana 2006; Hartley 2000) and (Hernandez et al. 2008), laser scanning (Blais et al. 1988; Rusinkiewicz et al. 2002) and (Bradley et al. 2002), and structured light scanning (Wang et al. 2010; Stockman et al. 1988; Hu and Stockman 1989). Among them, the multi-view reconstruction technique has the advantage of low cost but the disadvantage of low accuracy. Laser scanning normally achieves high resolution 3D images but costs too much and the collecting time is long (Blais et al. 1988; Rusinkiewicz et al. 2002; Bradley et al. 2002). As mentioned in (Bradley et al. 2002), the currently available commercial 3D scanning systems cost from \$2500 to \$240,000USD. The time of scanning a turtle figurine (18 cm long) is from 4 to 30 min for different scanners (Rusinkiewicz et al. 2002). The status (wet or dry) of objects also affects the accuracy of 3D images due to surface reflection. The wetter the surface is, the lower the accuracy will be (Blais et al. 1988). Different from the multi-view reconstruction and laser scanning, structured light imaging has high accuracy as well as a moderate cost. However, it also takes much time to collect 3D data and suffers from the instability problem such that one needs to keep still when it projects some structured light patterns to the human finger (Wang et al. 2010; Stockman et al. 1988; Hu and Stockman 1989). Thus, it is necessary and important to study the reconstruction technique based on multi-view 2D fingerprint images when considering the cost, friendliness, as well as the complexity of device design. It is well known that the 3D spatial coordinates of an object are available from its two different plane pictures captured at one time according to binocular stereo vision theory, if some camera parameters and the corresponding matched pairs are provided (Hartley 2000). In (Parziale and Diaz-Santana 2006), the authors briefly introduce the 3D reconstruction method since it is the same as those methods used to reconstruct any other type of 3D objects. There are several drawbacks with adopting general methods for 3D fingerprint reconstruction. For instance, it is time-consuming for the reason that the coordinate of each pixel needs to be calculated. Only the 3D coordinates of correspondences which represent the same portion of the skin between a pair of neighbor images can be calculated. 3D visualization of finger is unavailable, if correspondences cannot be found between two neighbor images.

To overcome the disadvantages mentioned above, a new 3D fingerprint reconstruction system using feature correspondences and the prior estimated finger model is proposed in this chapter. Comparative little research has been carried out into touchless fingerprint matching due to the characteristics of touchless fingerprint imaging, and hardly any work can be found for finger shape model

analyses. This chapter for the first time analyzes touchless fingerprint features for correspondences establishment and studies the model of human finger shape. 3D fingerprints are then reconstructed based on the images captured by a touchless multi-view fingerprint imaging device designed by us (Liu et al. 2013). Figure 9.1 shows the schematic diagram of our designed acquisition device and an example of 2D fingerprint images. Finally, 3D fingerprint reconstruction results based on different feature correspondences are given and compared with those based on manually labeled correspondences. It is concluded that such reconstruction results are helpful to 3D fingerprint recognition.

The chapter is organized as follows. In Sect. 9.2, the imaging device and the procedure of the proposed 3D fingerprint reconstruction system are briefly introduced. Sect. 9.3 is devoted to presenting the methods proposed to establish finger-print feature correspondences. The approach to estimating the finger shape model is described in Sect. 9.4. The feature extraction and matching method are introduced in Sect. 9.5. Experimental results and the reconstructing error analysis are given in Sect. 9.4. Section 9.4 concludes the chapter and indicates the future work.

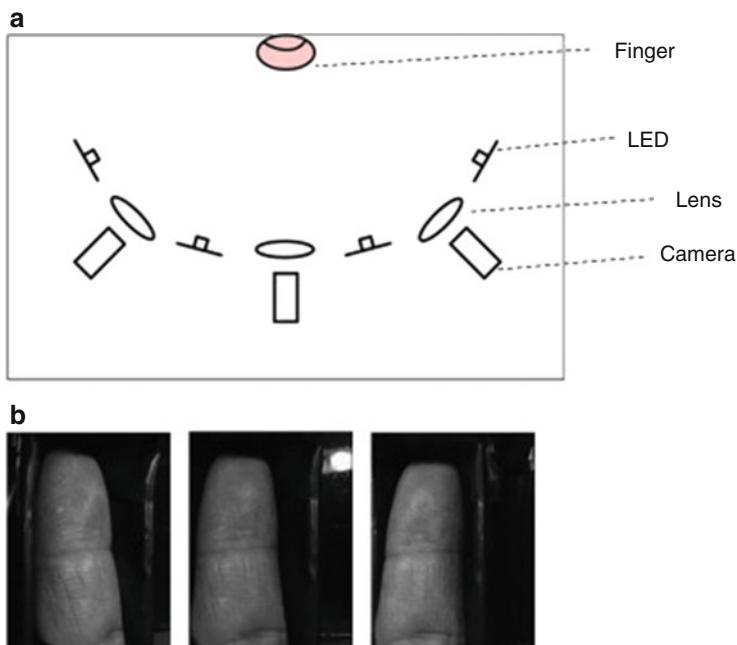


Fig. 9.1 Device and captured touchless multi-view fingerprint images. (a) Schematic diagram of our designed touchless multi-view fingerprint acquisition device, (b) images of a finger captured by the device (*left, frontal, right*)

9.2 3D Fingerprint Reconstruction System

Before reconstruction, multi-view fingerprint images need to be provided. The images used in this chapter are captured by the touchless multi-view fingerprint acquisition device designed by us. The schematic diagram of the acquisition device is shown in Fig. 9.1a. One central camera and two side cameras are focused on the finger. Four blue LEDs are used to light the finger and arranged to give uniform brightness. A hole is designed to place the finger in a fixed position. All of the three cameras are JAI CV-A50. The lens focal length is 12 mm and the object-to-lens distance is set to 91 mm due to the consideration of image quality and device size. The angle between the central camera and the side cameras is roughly 30° . The image size of each channel is restricted to 576×768 pixels and the resolution of the image is ~ 400 dpi. The three view images of a finger captured by the device are shown in Fig. 9.1b.

According to the theory of binocular stereo vision in computer vision domain (Hartley 2000), the 3D information of an object can be obtained from its two different plane pictures captured at one time. As shown in Fig. 9.2a, given two images and C_r captured at one time, the 3D coordinate of A can be C_l calculated if some camera parameters (e.g., focal length of the left camera f_l , focal length of the right camera f_r , principal point of the left camera O_l , principal point of the right camera O_r) and the matched pair $(a_l(u_l, v_l)) \leftrightarrow (a_r(u_r, v_r))$, where $a_*(*)$ represents a 2D point in the given images C_r or C_l ; u_* is the column-axis of the 2D image, and v_* is the row-axis of the 2D image) are provided. Once the shape model and several calculated 3D coordinates of the 3D object are known, the shape of the 3D object can be obtained after interpolation. As can be seen in Fig. 9.2b, the triangle in 3D space is obtained after computing 3D coordinates of three vertices and interpolating according to a triangle model. Therefore, the reconstruction method is divided into five steps, including the camera parameters calculation, correspondences

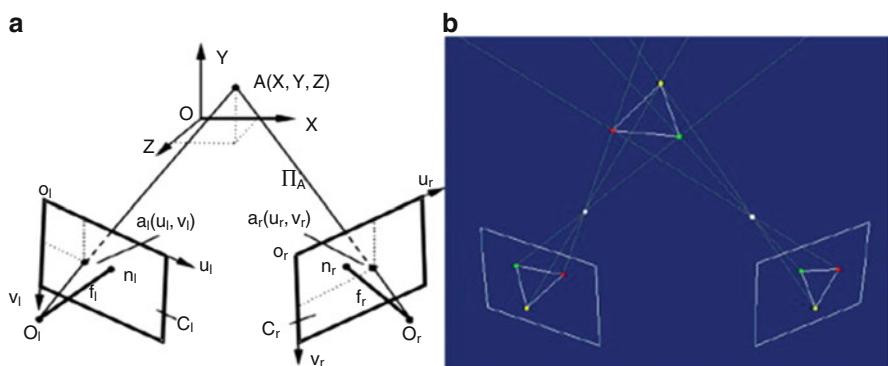


Fig. 9.2 An illustration of constructing a 3D triangle based on binocular stereo vision. (a) 3D coordinates calculation on 3D space, (b) 3D triangle reconstruction

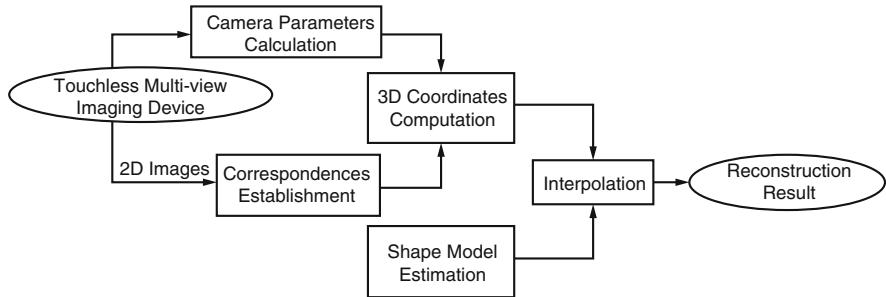


Fig. 9.3 The flow chart of our reconstruction system

establishment, 3D coordinates computation, shape model estimation, and interpolation. The flow chart of the reconstruction system in this chapter is shown in Fig. 9.3.

Camera calibration is the first step for 3D reconstruction. It provides the intrinsic parameters (Focal Length, Principal Point, Skew, and Distortion) of each camera and extrinsic parameters (Rotation, Translation) between cameras necessary for reconstruction. It is usually implemented off-line. In this chapter, the methodology proposed in (Zhang 2000) and the improved algorithms coded by Bouguet (Bouguet online) are employed. The free codes can be obtained from the website (Bouguet online). It can be noted that there are three cameras used in our fingerprint capturing device. The position of the middle camera is chosen as the reference system, because the central part of the fingerprint is more likely to be captured by this camera, where the core and the delta are usually located. The frontal image captured by the middle camera is also selected as the texture image when the final 3D fingerprint image is generated. To ensure that the frontal view of finger is captured by the middle camera of the device, a simple guide is given for users to correctly use the device.

Correspondences establishment is of great importance to the 3D reconstruction accuracy. It will be introduced in detail in Sect. 9.3.

Once camera parameters and matched pairs between fingerprint images of different views are both obtained, the 3D coordinate of each correspondence can be calculated by using the stereo triangulation method (Bouguet online).

Since it is very hard to identify all of the correspondences which represent the same portion of the skin between two neighboring fingerprint image pairs, it is very important to calculate the 3D finger shape for 3D fingerprint visualization. This chapter for the first time analyzes finger shape models. The details will be presented in Sect. 9.4.

Based on the calculated 3D coordinates of limited feature correspondences and the estimated shape model, a 3D finger shape can be finally reconstructed by interpolation. Here, the classical approach, namely, multiple linear regression using least squares (Chatterjee and Hadi 1986; Draper and Smith 1981), is adopted for interpolation because of its simplicity and effectiveness.

9.3 Fingerprint Feature Correspondences Establishment

Fingerprints are distinguished by their features. Different fingerprint features can be observed from different resolution fingerprint images. There are three frequently-used features for low resolution fingerprint images, namely Scale Invariant Feature Transformation (SIFT) feature, ridge map and minutiae (Choi et al. 2010; Zhang et al. 2011; Kumar and Zhou 2011; Park et al. 2008; Feng 2008; Malathi and Meena 2010; Jain and Ross 2002; Shah et al. 2005; Choi et al. 2007). This chapter thus tries to extract such features and establish correspondences between different views of fingerprint images.

9.3.1 Correspondences Establishment Based on SIFT

SIFT (Lowe 2004) is popular in object recognition and image retrieval, since it is robust to low quality image. For touchless fingerprint images, they have the characteristic of low ridge–valley contrast. This feature makes true correspondences possible to be established when minutiae and ridge features cannot be correctly extracted. Moreover, it is robust to deformation variation and rich in quantity (Park et al. 2008; Malathi and Meena 2010). Figure 9.4b, d illustrate the extracted 1911 and 1524 SIFT features, respectively. 108 pairs are matched by using the point wise matching method to Fig. 9.4a, c, as shown in Fig. 9.4e. From Fig. 9.4e, we can see that there exist false correspondences and hence refined algorithms are needed to be employed to select true ones. To this end, the classical RANSAC algorithm, which is insensitive to initial alignment and outliers (Fishler and Bolles 1981) is utilized. It should be noted that the TPS model which is popularly used in fingerprint domains (Choi et al. 2010; Shah et al. 2005; Ross et al. 2005) is adopted in the RANSAC algorithm due to the curved surface of finger and distortions introduced by cameras. Figure 9.4f gives the final selected true correspondences when RANSAC with the TPS model acts on the initial correspondences of Fig. 9.4e.

9.3.2 Correspondences Establishment Based on Ridge Map

Before establishing correspondences between ridge maps, ridges must be extracted and recorded. In general, ridge map refers to the thinning image where ridges are one-pixel-width, and ridge pixels have value 1 and background pixels have value 0. Figure 9.5 shows the flowchart of steps for ridge map extraction. However, touchless fingerprint images have low ridge–valley contrast and their ridge frequency increases from center to side, as shown in Fig. 9.4a, c. These make it difficult to extract the ridge map accurately due to the difficulty of fingerprint enhancement. Currently, there are a number of fingerprint enhancement approaches, such as Gabor filter-based, STFT-based, DCT-based and Diffusion filter-based methods (Hong et al. 1998;

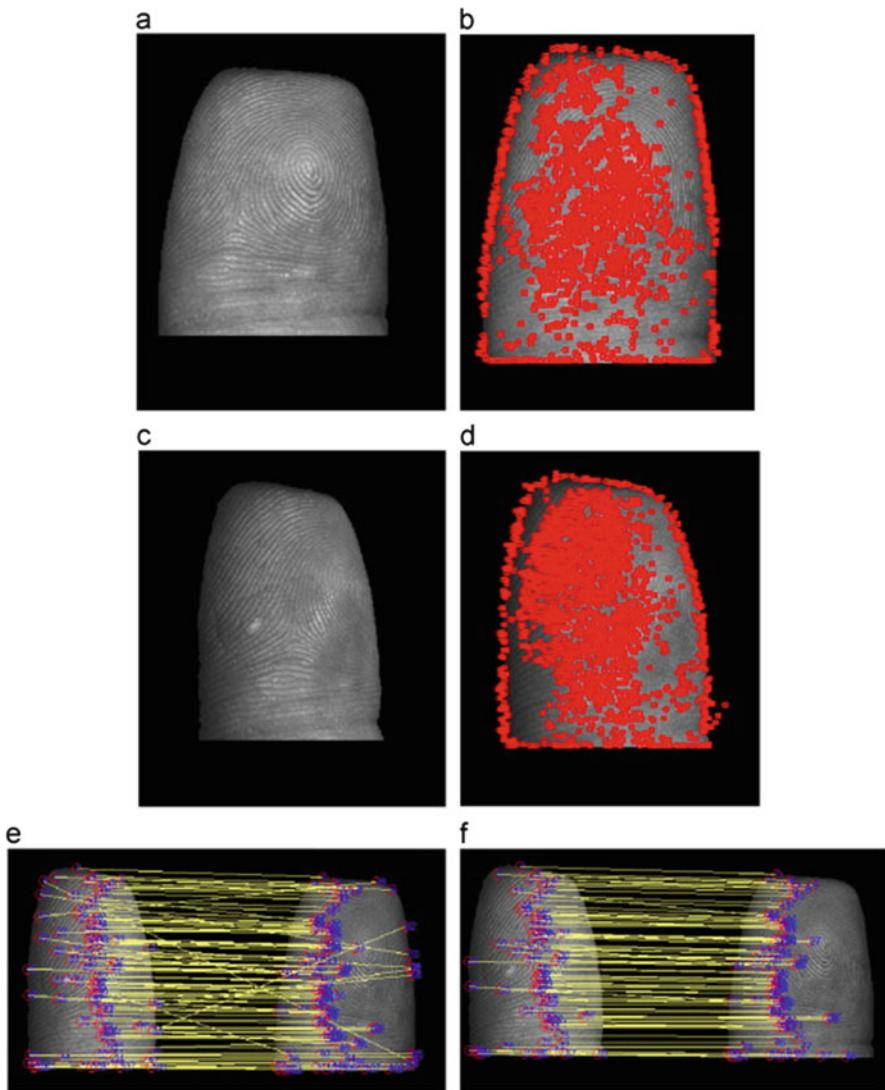


Fig. 9.4 Example of correspondences establishment based on SIFT features. (a) Original frontal image, (b) extracted SIFT feature from (a), (c) original left-side image, (d) extracted SIFT feature from (c), (e) initial correspondences established by point wise matching, (f) final correspondences after refining by the RANSAC method

Chikkerur et al. 2007; Jirachaweng and Areekul 2007; Weichert 1999; Chen and Dong 2006; Hao and Yuan 2004; Hastings 2007; Almansa and Lindeberg 2000; Xie and Wang 2004). Among them, the Gabor filter based method is the simplest and the most traditional one. It is finally adopted in this chapter. Fingerprint images are enhanced by a bank of Gabor filters generated from given fingerprint orientation and frequency.

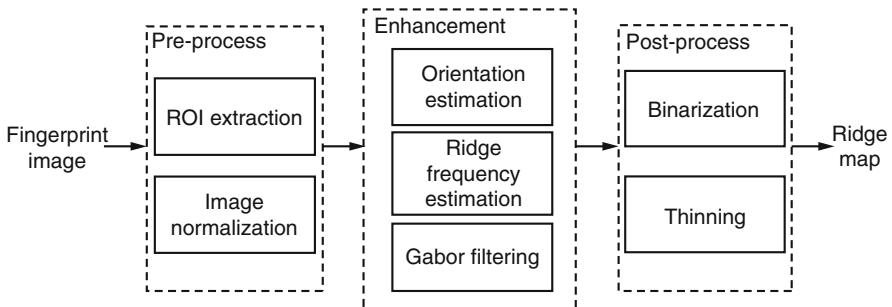


Fig. 9.5 Flowchart of ridge map extraction

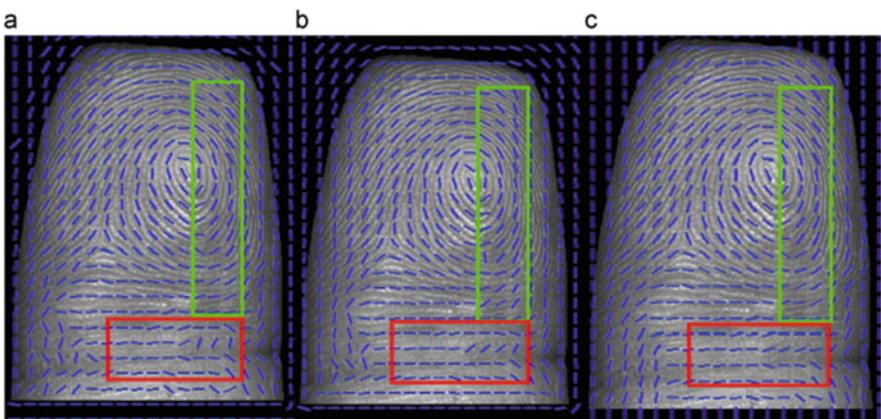


Fig. 9.6 Fingerprint ridge orientation maps. (a) Original orientation map, (b) smoothed orientation map of (a), (c) improved orientation map by our proposed method

Orientation and frequency maps play an important role in the enhancement approach. This chapter thus tries to improve the orientation map and frequency map so as to acquire better enhanced results.

As introduced in (Maltoni et al. 2009), the gradient-based ridge orientation estimation method is the simplest and most intuitive one. It is efficient and popularly used in fingerprint recognition studies. However, it also has some drawbacks, such as sensitivity to noise when orientation is estimated at too fine a scale, low accuracy when smooth factors are used to the orientation map, as shown in Fig. 9.6a (lower rectangle) and Fig. 9.6b (right rectangle). To keep the estimation accuracy of a good quality area and correct the orientation where noises exist, a method is proposed to act on original orientation map to improve the orientation map. The main steps include: (i) part the original orientation map into eight uniform regions. Small blocks in the uniform regions represent the wrong estimated orientation results (see Fig. 9.7a, in the red circles); (ii) sort uniform regions with the same color in a descending manner; such regions whose size is smaller than the mean size of all regions with the same color are

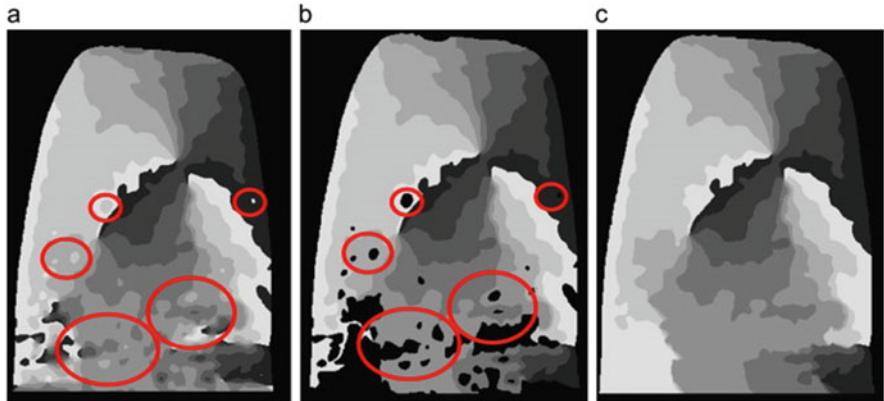


Fig. 9.7 Partition results according to orientation map. (a) Partition result according to original orientation map, (b) partition result according to our improved orientation map

set to zero (see Fig. 9.7b, the dark regions in ROI); (iii) assign values to the points with zero value set by step (ii) according to the nearest neighbor method. The improved orientation map is obtained by following these three steps. Figure 9.6c shows the improved orientation map based on Fig. 9.6a, and Fig. 9.7c gives the partition map according to Fig. 9.6c. The results show that the estimation accuracy of a good quality area is kept and the wrong orientation area is corrected (Fig. 9.6c, rectangle).

Frequency maps record the number of ridges per unit length along a hypothetical segment and orthogonal to the local ridge orientation. The simplest and most popular ridge frequency estimation method is the x-signatures based method (Maltoni et al. 2009). However, this kind of method does not work with blurry or noisy fingerprint areas. In this situation, interpolation and filtering is used to post-process the original estimated frequency map. For touchless fingerprint images, frequency maps are harder to estimate than touch-based fingerprint images due to the low ridge–valley contrast of touchless fingerprint images, and simple interpolation or filtering is invalid when the frequency is wrongly estimated in neighborhoods. By observing the ridges on touchless fingerprint images, we find their frequency increases from the central part to the side part for horizontal section and decreases from the fingertip to the distal interphalangeal crease for the vertical section, as shown in Fig. 9.8 (ridge frequency is calculated with blocks of 32×32 pixels). This phenomenon can be explained from the touchless capturing technique and the observation of the human finger. As shown in Eq. (9.1), M is the optical magnification. p and q are the lens-to-object and lens-to-image distances, respectively. For a fix q , a large p will lead to a small magnification M . Figure 9.9 illustrates three different values of p . It can be seen that the distance from the side parts to the lens (i.e., D_2 or D_3) is larger than the distance from the central part to the lens (i.e., D_1), which leads to smaller M on the side parts than on the central part. The smaller the magnification M is, the larger the ridge frequency will be. Thus, it is larger in the central part of the ridge period than side-view ones for the horizontal section. The vertical distribution of ridge period increases from the fingertip to the distal interphalangeal

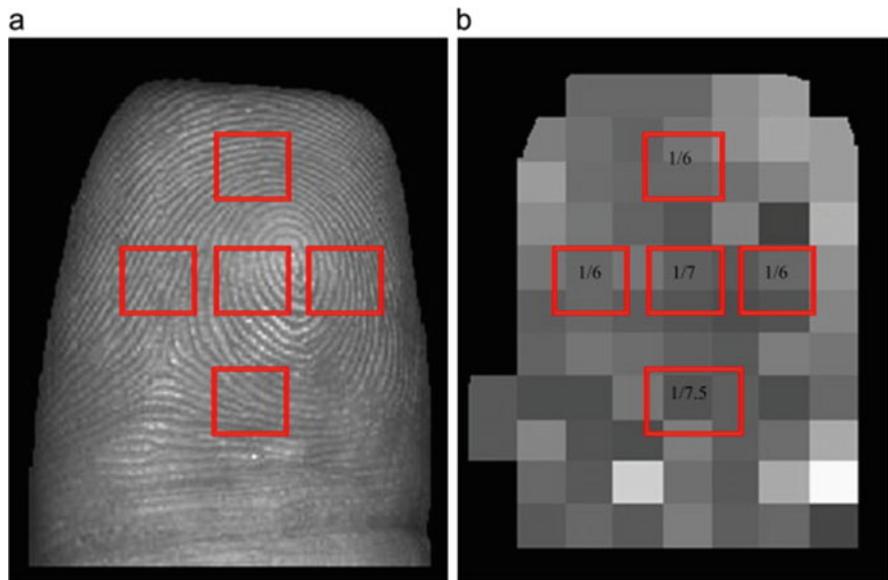


Fig. 9.8 Frequency variation of touchless fingerprint images. (a) Original touchless fingerprint image and (b) corresponding frequency map

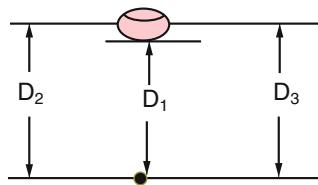


Fig. 9.9 Distance between lens and different parts of the finger

crease, because p increases from the tip to the center part of the finger and the ridges are wider near the distal interphalangeal crease than the other parts by observation.

$$M = \frac{q}{p} \quad (9.1)$$

According to the distribution of ridge frequency of touchless fingerprint images, this chapter proposes to use monotone increasing function (logarithmic function) to fit the ridge period (1/ridge frequency) map along the vertical direction and quadratic curve along the horizontal direction. The improved ridge period map is finally achieved by fitting original ridge period map with a mixed model of logarithmic function and quadratic curve.

Once the orientation and ridge frequency maps are calculated, a series of Gabor filter can be generated based on them. The enhanced fingerprint image was then obtained, as shown in Fig. 9.10. After binarizing the enhanced fingerprint image by simple threshold and morphology approaches, the final ridge map is acquired.

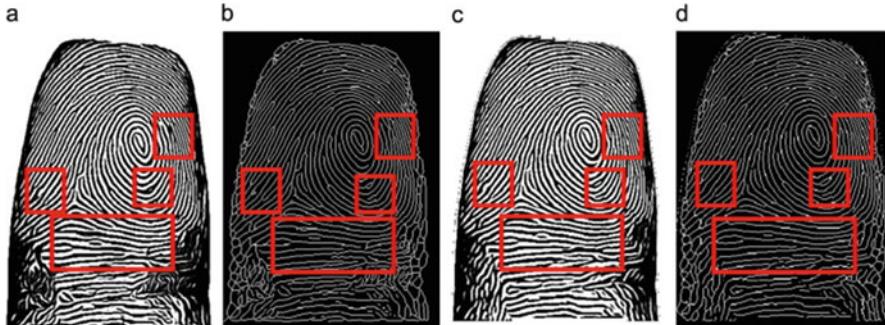


Fig. 9.10 Ridge maps. (a) Ridge map of Fig. 9.4a enhanced by using original orientation and ridge frequency maps, (b) thinned ridge map of (a), (c) ridge map of Fig. 9.4a enhanced by using improved orientation and ridge frequency maps, (d) thinned ridge map of (c)

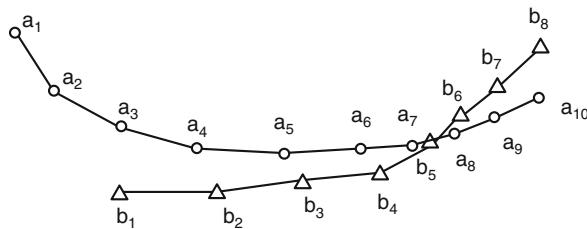


Fig. 9.11 Correspondences establishment between two ridges

Figure 9.10a, b shows the ridge maps of Fig. 9.4a enhanced by using the original orientation map and the original ridge frequency map interpolated by mean value of the frequency map. Figure 9.10c, d shows the enhanced ridge maps of Fig. 9.4a using the improved orientation and ridge frequency maps. Better results are achieved when comparing Fig. 9.10c, d with a, b (labeled in rectangles). It should be noticed that the pre-process steps of ROI extraction and normalization are the same as those proposed in (Liu et al. 2013).

Before correspondences establishment, ridges are recorded at tracing starting from minutiae where ridges are disconnected. Due to the existence of noise, a ridge image often has some spurs and breaks. In some cases of insignificant noise, the ridge structure can be correctly recovered by removing short ridges or connecting broken ridges. However, in other cases of strong noise, it is difficult to recover the correct ridge structure by removing short ridges or connecting broken ridges. In such cases, we remove all related ridges. Finally, the down sampled ridge point coordinates of each ridge are recorded in a list.

Coarse alignment of two ridge maps is done by using the global transform model calculated in Sect. 9.3.1 when SIFT features matched. Ridges in ridge maps are then matched by adopting the Dynamic Programming (DP) method. As shown in Fig. 9.11 and Table 9.1, $\{a_1, a_2, \dots, a_{10}\}$ represents a ridge line in the template ridge map and $\{b_1, b_2, \dots, b_8\}$ denotes a ridge line in the test ridge map. For any ridge in template and test ridge maps, the Euclidian distance between each pair of compared ridge lines is

Table 9.1 Record of status among ridge points in Fig. 9.11

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
b_1	0	0	0	0	0	0	0	0	0	0
b_2	0	0	0	0	0	0	0	0	0	0
b_3	0	0	0	0	1	0	0	0	0	0
b_4	0	0	0	0	0	1	1	0	0	0
b_5	0	0	0	0	0	0	1	1	0	0
b_6	0	0	0	0	0	0	0	1	0	0
b_7	0	0	0	0	0	0	0	0	1	0
b_8	0	0	0	0	0	0	0	0	0	0

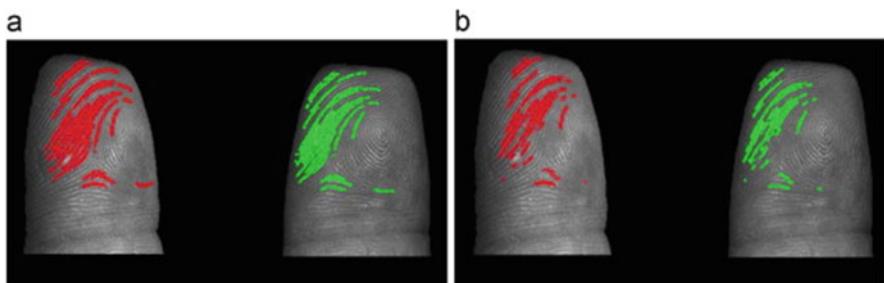


Fig. 9.12 Ridge correspondences establishment. (a) Initial correspondences and (b) final correspondences after RANSAC

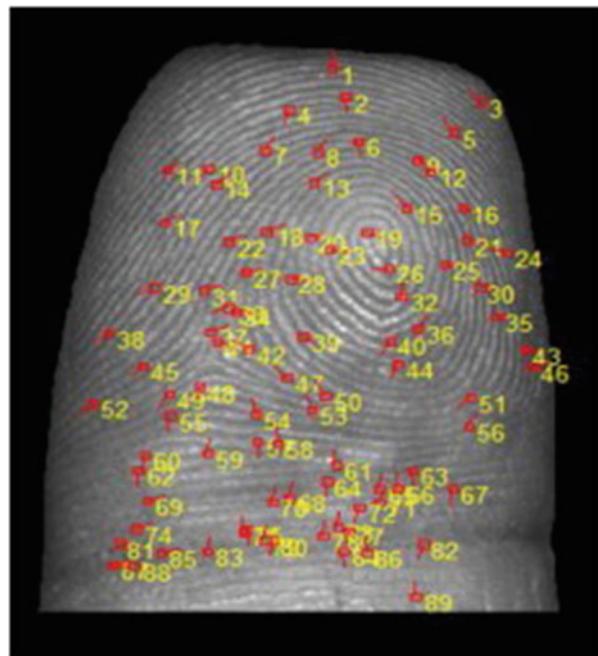
calculated. The status will be 1 if the distance of a pair of ridge points is smaller than a threshold (it is set to five points in this chapter), otherwise, the status will be 0. The DP method is adopted to find matched ridge pairs with the largest number. Coarse ridge correspondences are then established after DP. RANSAC algorithm introduced in Sect. 9.3.1 is then adopted to select true ones from the coarse set. Figure 9.12 shows the results of the established ridge correspondences.

9.3.3 Correspondences Establishment Based on Minutiae

Due to their distinctive ability, minutiae are widely used for fingerprint recognition and also considered in the chapter. They are extracted from the ridge map calculated in Sect. 9.3.2. An example of extracted minutiae using the method introduced in (Jain et al. 1997) is shown in Fig. 9.13.

Since the transformation model is obtained when SIFT correspondences are established, minutiae sets can be coarsely aligned by the calculated transformation model. Then, initial minutiae correspondences are established by the nearest neighbor method, and the final result is achieved by the RANSAC algorithm with a TPS model. This kind of minutiae correspondences establishment is demonstrated in Fig. 9.14.

Fig. 9.13 Example of minutiae extraction result



9.4 Finger Shape Model Estimation

To reconstruct the finger shape, it is necessary to know the shape model after certain 3D points of the finger are calculated. Unfortunately, exact model for human's finger shape is not directly available, and hence, it should be estimated. To this end, we propose to estimate the finger shape model by analyzing 440 3D point cloud data collected from human fingers (220 fingers, 2 pictures each) in this chapter. The 3D point cloud data are defined as the depth information of each point on the finger. They are collected by a camera together with a projector using the Structured Light Illumination (SLI) method (Wang et al. 2010; Zhang et al. 2010). The structure diagram of the collection device is shown in Fig. 9.15. 13 structured light stripes generated by a computer are projected onto the finger surface by using the Liquid Crystal Display (LCD) projector. The camera then captures the fingerprint images formed with projected stripes on it. 3D point cloud data, which consists of depth information of each point on the finger, can be calculated using transition and phase expansion techniques (Saldner and Huntley 1997). Since this technique is well studied and proved to acquire 3D depth information of each point on the finger with high accuracy (Wang et al. 2010; Stockman et al. 1988; Hu and Stockman 1989; Zhang et al. 2009, 2010; Saldner and Huntley 1997), 3D point cloud data obtained using this technique are taken as the ground truth of the human finger to build the database for finger shape model estimation.

Fig. 9.14 Minutiae correspondences establishment. (a) Initial correspondences and (b) final correspondences after RANSAC

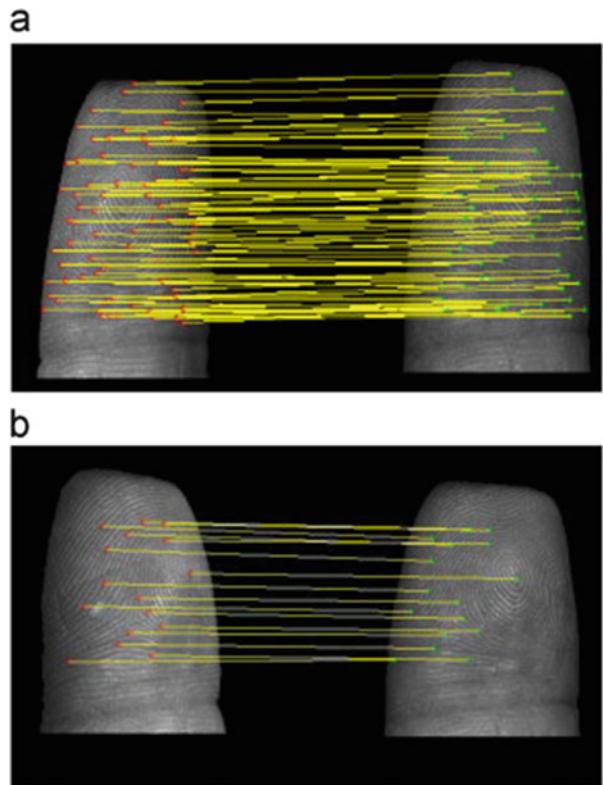


Fig. 9.15 Structure diagram of device used to capture 3D point cloud data of human finger (Wang et al. 2010)

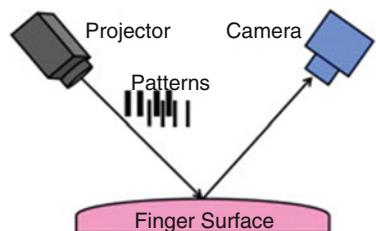


Figure 9.16a displays an example of 3D point cloud data we collected from a thumb. We randomly selected and drew the horizontal profile and the vertical profile of the 3D point cloud data, as shown in Fig. 9.17 (thick rugged line). The horizontal profile is in a parabola-like shape, as shown in Fig. 9.17a, while the vertical profile can be represented by a quadratic curve or a logarithmic function (see Fig. 9.17b). Thus, both of the binary quadratic function

$$f_1(x, y) = Ax^2 + By^2 + Cxy + Dx + Ey + F \quad (9.2)$$

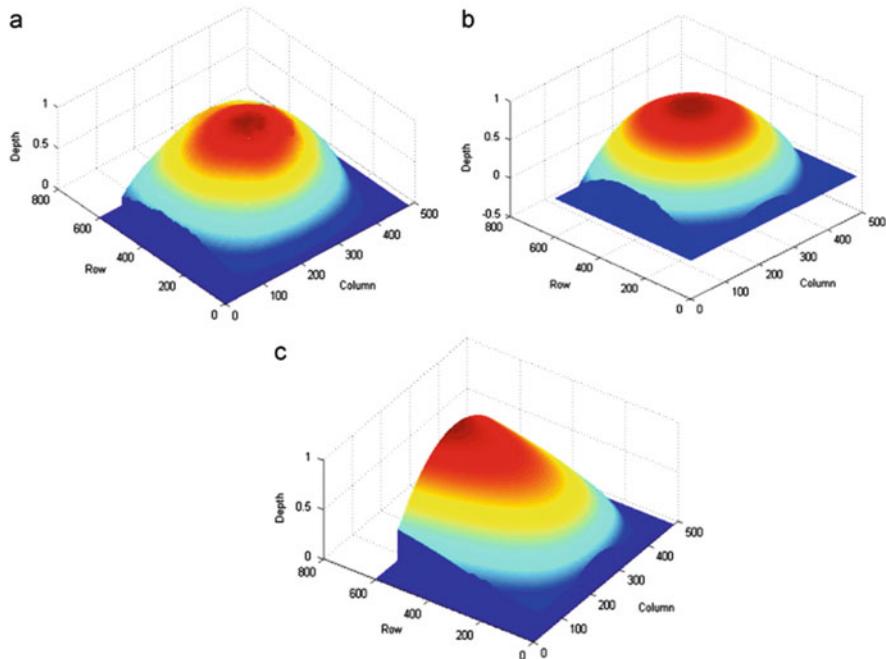


Fig. 9.16 Example 3D finger point cloud data and its fitting results by different models. (a) 3D point cloud data of a thumb, (b) fitting result of (a) by binary quadratic function, (c) fitting result of (a) by a mixed model with parabola and logarithmic function

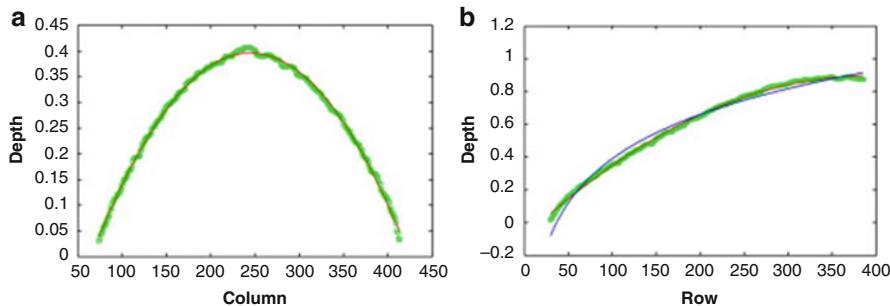


Fig. 9.17 Randomly selected profiles of Fig. 9.16a. (a) Horizontal profile, *thick rugged line* depicts real data, *thin smooth line* is fitting by Parabola, (b) vertical profile, *thick rugged line* depicts real data, *thin smooth lines* are fitting by Quadratic Curve (closer to real data) and logarithmic Function, respectively

and the mixed model with parabola and logarithmic function

$$f_2(x, y) = Ax^2 + By + C \ln(y) + D \quad (9.3)$$

are chosen to fit all of our collected 440 3D point cloud finger data by the regression method (Chatterjee and Hadi 1986) and (Draper and Smith 1981). Note that, in (9.2) and (9.3), A, B, C, D, E, and F represent the coefficients of the functions, x is the variable of column-coordinate of the image, and y is the variable of row-coordinate of the image. Figure 9.16b gives the fitting result of Fig. 9.16a (denoted by V) by the binary quadratic function (denoted by $\bar{V}_{Eq.2}$), while Fig. 9.16c gives the fitting result of Fig. 9.16a by the mixed model (represented by $\bar{V}_{Eq.3}$). It can be seen that binary quadratic function is closer to the finger shape model. Therefore, binary quadratic function in Eq. (9.2) is finally adopted in this chapter.

9.5 Curvature Features Extraction and Matching

Since our 3D fingerprint image is reconstructed from multi-view fingerprint images, there is a one-to-one correspondence between the 3D points and the 2D fingerprint image pixels. Preprocessing such as ROI extraction and pose correction can be done in 2D fingerprint images, and implement into 3D situation. The ROI extraction steps are shown in Fig. 9.18. Since the uncontrolling of fingerprint image collection (tilted of finger position, see Fig. 9.18a), pose correction is necessary. We accomplished it by simple rotating the original image as follows: (i) Scan each line of ROI horizontally and find the center point (green line in Fig. 9.18b); (ii) Fit such center points by a line (red line in Fig. 9.18b); (iii) Calculate the angle between the fitted center line and vertical axis (θ shown in Fig. 9.18b); (iv) Rotate the image anticlockwise by θ . Figure 9.18c shows the final correct 2D fingerprint image and Fig. 9.18e shows the correct 3D finger shape of original 3D shape of Fig. 9.18d.

Given a corrected 3D fingerprint image, stable and unique features are expected to be extracted for the following pattern matching and recognition. 3D depth information reflects the overall structure of human finger. However, there are many invalid points in the whole 3D finger shape due to the structure of human finger. Wrinkles and scars in finger also affect local structure of finger shape. Thus we proposed to extract curve-skeleton of finger shape. As shown in Fig. 9.19, Different 3D objects almost fully represented by their curve-skeletons.

Since 3D finger shape model is close to binary quadratic function, profile of horizontal section can be fitted by parabola and reflects the changes of finger width, while vertical profile depicts variation tendency of depth from finger tip to distal interphalangeal crease. The curve-skeleton of 3D fingerprint image consists of representative vertical and horizontal lines. Since each horizontal profile is parabola-like shape, we extracted the extreme value of each fitted parabola line to form the representative vertical line (blue line in Fig. 9.20a). Three representative horizontal lines are selected at a certain step length (100). The distal

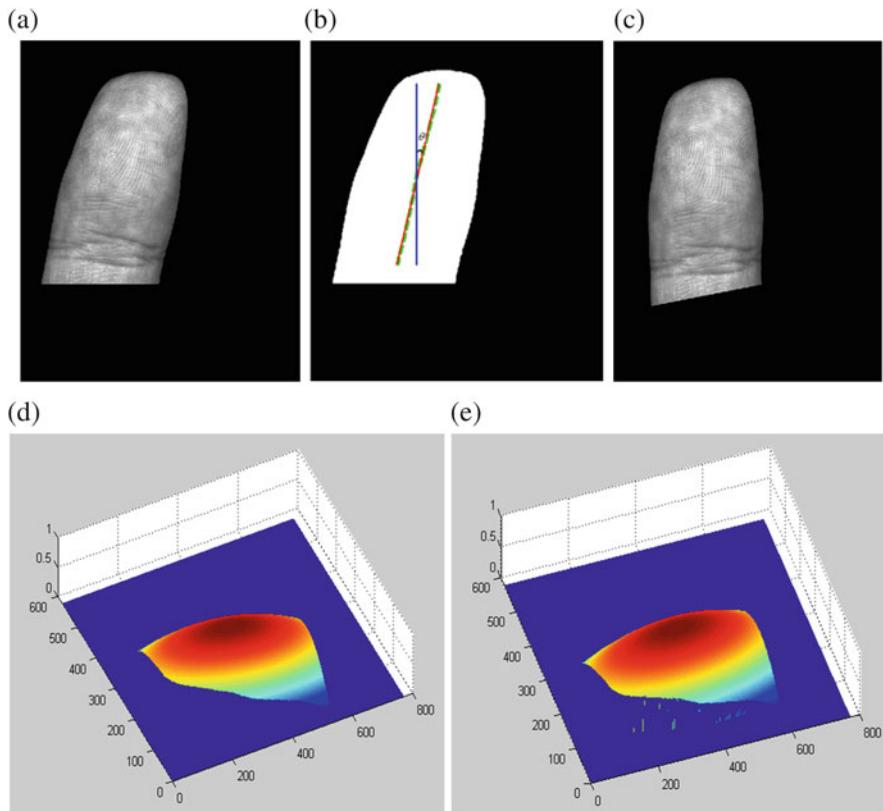


Fig. 9.18 Position Correction: (a) Original tilted fingerprint image, (b) ROI of (a), (c) Fingerprint image after pose correction, (d) Original 3D finger shape, (e) Corrected 3D finger shape

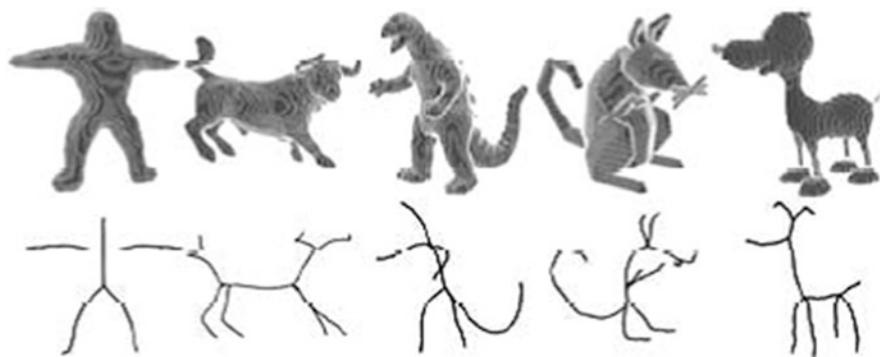


Fig. 9.19 Examples of curve-skeletons of different 3D objects

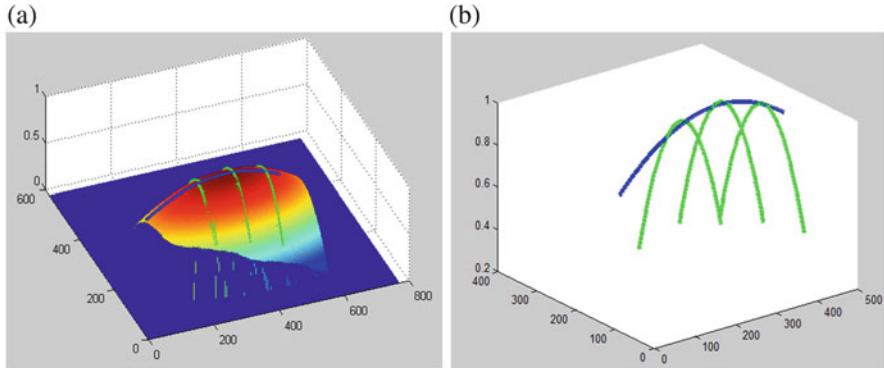


Fig. 9.20 Examples of curve-skeleton for 3D finger: (a) 3D finger shape, (b) Extracted curve-skeleton

interphalangeal crease is chosen as the base line (green line in Fig. 9.20a). Figure 9.20b then shows the curve-skeleton we extracted from 3D finger depth map. For overall curvatures, they can be easily calculated since our 3D finger shape is reconstructed by model fitting. The coefficients of the binary quadratic function control the maximal horizontal and vertical curvatures of 3D finger, namely the parameters of A and B in Eq. (9.4). Thus, these two coefficients of the binary quadratic function are maintained to represent the maximal horizontal and vertical curvatures, namely the defined overall curvatures.

$$f(x, y) = Ax^2 + By^2 + Cxy + Dx + Ey + F \quad (9.4)$$

From Fig. 9.20b, we can see that curve-skeleton consists of several 3D lines. Intuitively, the iterative closest point (ICP) algorithm is suitable to solve such matching problem. ICP method is widely used in many 3D object recognition systems for matching. In this chapter, we slightly modified the ICP method to measure the distances between two sets of points. The algorithm is given below and Fig. 9.21 shows an example of matching two curve skeletons by our modified ICP method.

ICP algorithm:

1. Input: Medel point set: D_1 ; Test point set D_2 ;
2. Parameters initialization: stop criterion for distance $T_d = 0.1$; initial rotation matrix $R_0 = I$; initial translation vector $T_0 = [0 \ 0 \ 0]^T$;
3. While [new correspondences set found between D_1 and D_2]


```
{[corr, D_i] = dsearchn(D_1, D_2);
K_i = D_i > T_d;
Discard corr(K_i);
Update R_i, T_i;
D_2 = R_i*D_2 + T_i;}
```
4. Output: distance vector D, registered D2, rigid transform parameters: R and T.

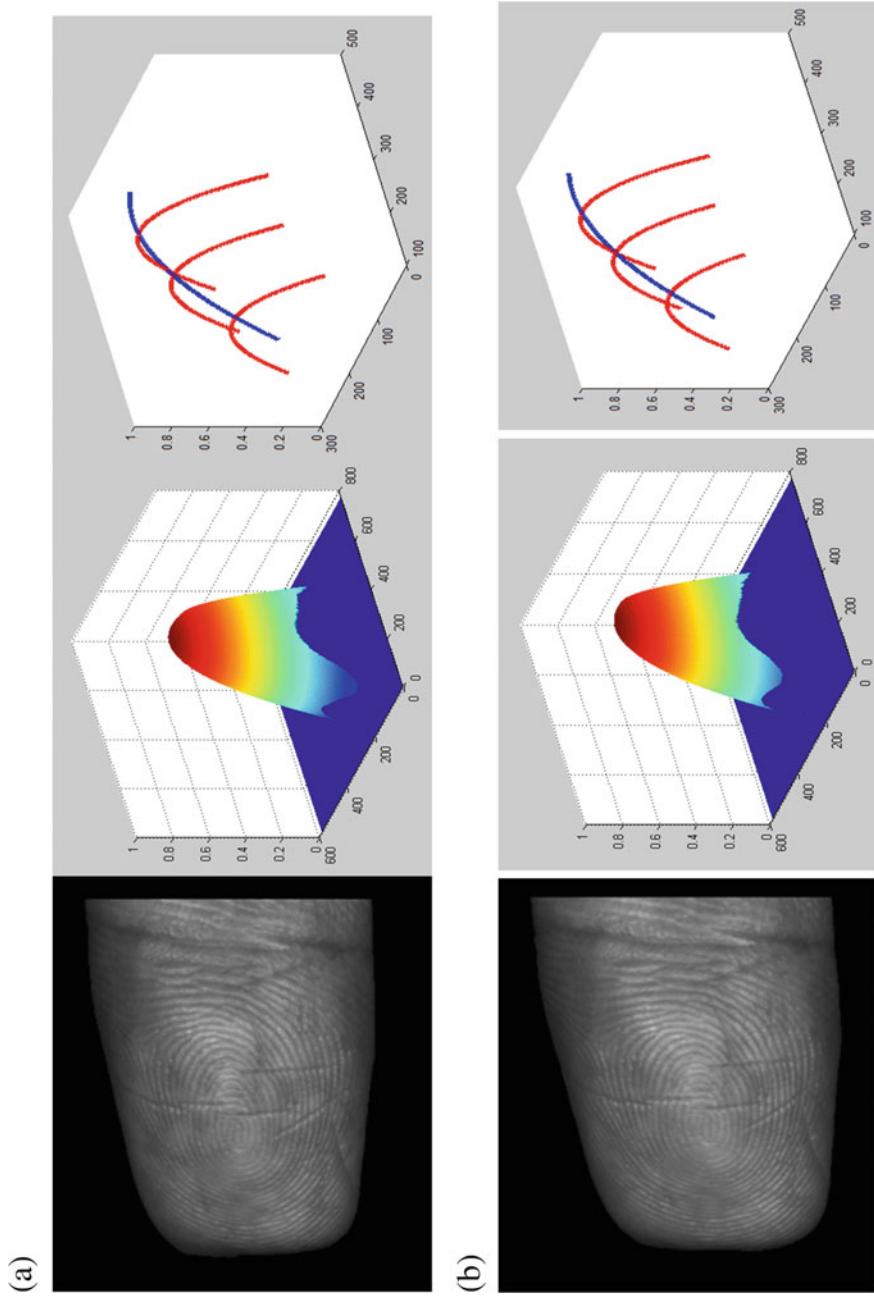


Fig. 9.21 Example of curve-skeleton matching by icp method: (a) The model 2D fingerprint image, 3D finger shape, and extracted curve-skeleton feature, (b) Matching result by icp method The test 2D fingerprint image, 3D finger shape, and extracted curve-skeleton feature, (c) Matching result by icp method

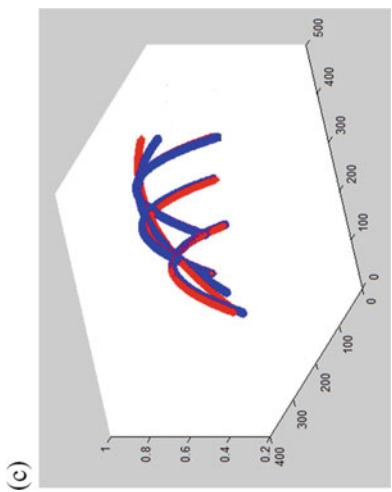


Fig. 9.21 (continued)

9.6 Experimental Results and Analysis

Reconstruction and system errors are inevitable. To acquire these errors, the reconstruction of an object with the standard cylinder shape and of radius 10 mm is given. The example object is shown in Fig. 9.22a. The surface of the object is wrapped by a grid chapter to facilitate feature extraction. Three 2D pictures (left-side, frontal, and right-side) of the cylinder are captured by the touchless multi-view imaging device we designed. Figure 9.22b, c shows two grouped images (left-side and frontal, right-side and frontal). As mentioned in Sect. 9.2, there are five main steps in our reconstruction technique. Camera parameters are first calculated off-line. The corner features of the wrapped grid chapter are then labeled and their correspondences between grouped images are established manually, as shown in Fig. 9.22b, c. Figure 9.22d, e illustrates the calculated 3D coordinates corresponding to the matched pairs shown in Fig. 9.22b, c based on the given camera parameters and feature correspondences. Shape model estimation is unnecessary since the cylinder model is known as a prior knowledge. By using the calculated 3D coordinates and the known shape model of cylinder, the cylindrical surface is finally generated by interpolation based on the multiple linear regressions using the least squares method (Zhang et al. 2009) and (Chatterjee and Hadi 1986). Figure 9.22f, g are the reconstructed cylinders shown by a 3D display software called Image ware 12.1. This software is used for 3D point cloud data display and analysis. The error maps shown in Fig. 9.22h, i are also obtained by this software. From Fig. 9.22f, g, we can see that the radius of reconstructed cylinders from 40 3D points of Fig. 9.22d, e are ~ 9.91 and ~ 9.84 mm compared with the real radius 10 mm. Figure 9.22h, i gives the error maps of 3D points corresponding to Fig. 9.22d, e when fitted by cylinder shape with radius of 10 mm. The error ranges are $[-0.07\text{--}0.06 \text{ mm}]$ and $[-0.1\text{--}0.06 \text{ mm}]$. The results demonstrate that the reconstruction error of our device is within ~ 0.2 mm.

9.6.2 Comparison and Analysis of Reconstruction Results

By following the five steps introduced in Sect. 9.2, reconstructed 3D fingerprint images can be obtained. Since there are three fingerprint images captured at one time and the central camera is selected as the reference system, the proposed reconstruction system consists of two parts (left-side camera and central camera, right-side camera and central camera) according to binocular stereo vision theory. In this chapter, these two parts are combined before the fourth steps by normalizing the calculated depth value of correspondences into $[0, 1]$. Here, the Min-Max strategy of normalization is used. This combination is adopted for two reasons. One is that there are parts of overlapping region between two adjacent fingerprint

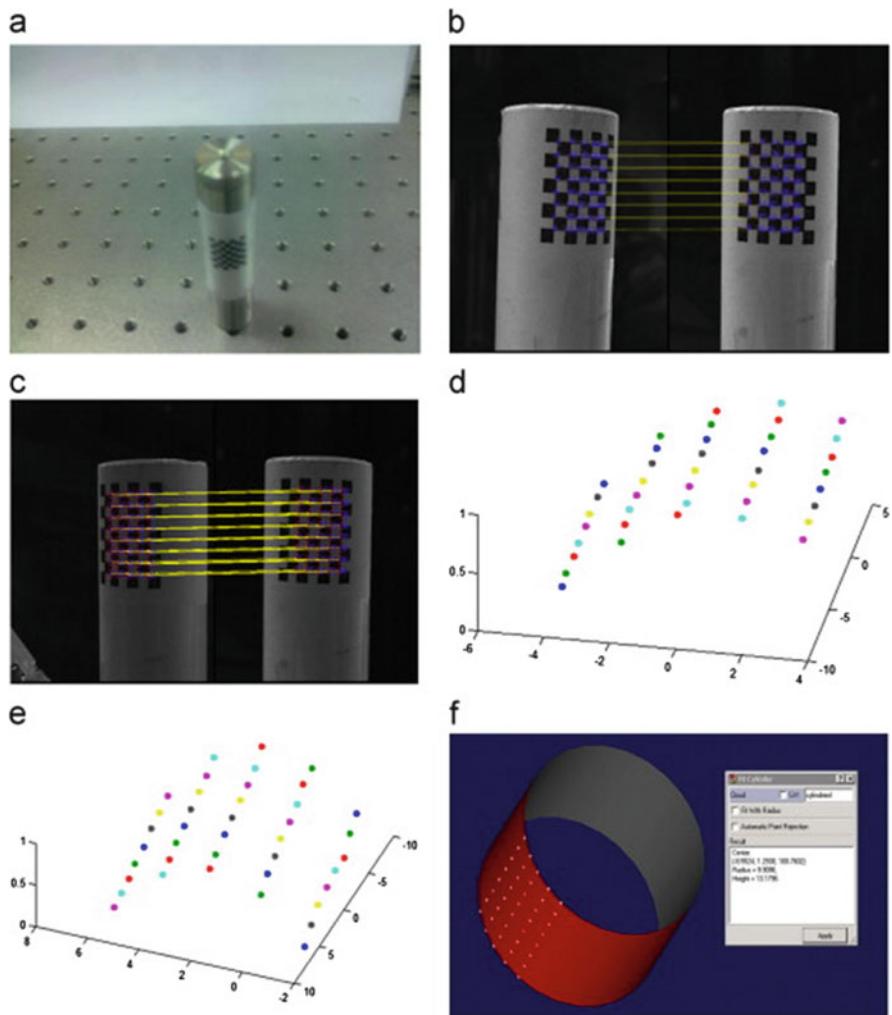


Fig. 9.22 Reconstruction accuracy analysis of cylinder shape object. (a) Original cylinder shape object wrapped with grid chapter, (b) correspondences established between *left-side* and *frontal* images captured by our device, (c) correspondences established between *right-side* and *frontal* images captured by our device, (d) 3D space points corresponding to (b), (e) 3D space points corresponding to (c), (f) fitting result corresponding to (d), (g) fitting result corresponding to (e), (h) error map corresponding to (d) when fitting by cylinder shape with radius of 10 mm, (i) error map corresponding to (e) when fitting by cylinder shape with radius of 10 mm

images, and the distribution of correspondences may focus on a small part of fingerprint images. Larger areas of fingerprint image can be covered by discrete correspondences through combining two parts of the system. The other is that very simple to accomplish and the system error of combining two parts before model

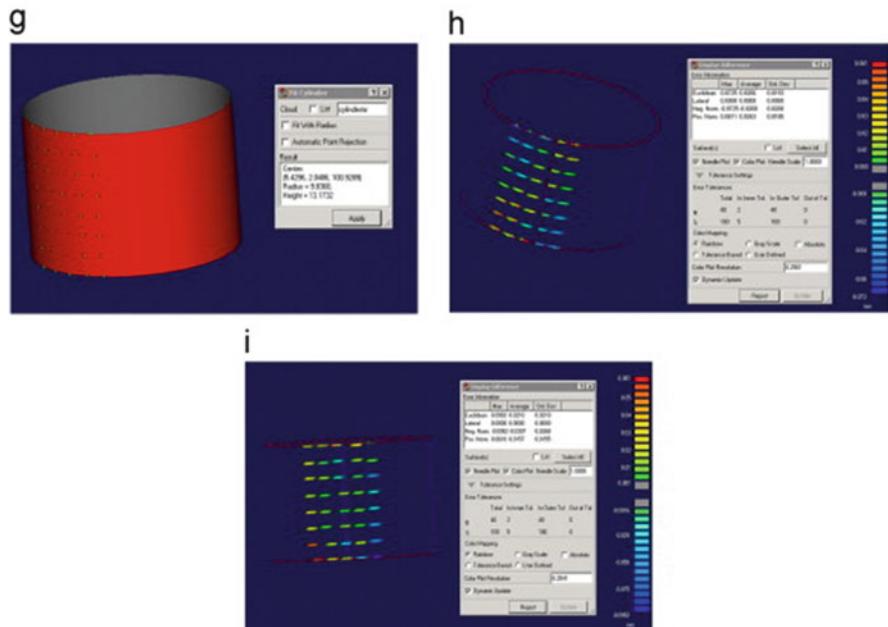


Fig. 9.22 (continued)

fitting is alleviated. Table 9.2 shows the reconstruction results based on three different fingerprint feature correspondences using the example images shown in Fig. 9.23. The results are different corresponding to different feature matched pairs due to different numbers and distribution of established fingerprint feature correspondences and the existence of false correspondences.

To investigate which features are more suitable for 3D fingerprint reconstruction, we also manually labeled fingerprint correspondences, as shown in Fig. 9.24. The histogram of error map between reconstructed results in Table 9.3 and Fig. 9.24 is shown in Fig. 9.25. The results show that for single feature used, a reconstruction result based on SIFT features achieves the best result, while the ridge feature-based is the worst one. When combining with other features, best reconstruction results can be generated if all three features of correspondences are used. However, comparable results are obtained by using SIFT and minutiae. Considering the computational complexity, it is recommended to simply use SIFT and minutiae.

9.6.3 Validation of Estimated Finger Shape Model

The effectiveness of the proposed finger shape model is validated by analyzing the fitting errors. Table 9.3 presents the errors measured by the mean distance and the standard variation between the estimated finger shape and the original 3D point cloud

Table 9.2 Reconstruction results from different feature correspondences of Fig. 9.23

Results Used feature	Established correspondences	Reconstructed 3D fingerprint image
SIFT feature		
Minutiae		
Ridge feature		
Feature Combination	Reconstructed 3D fingerprint image	
SIFT feature and minutiae		
SIFT and ridge feature		
Minutiae and ridge feature		
SIFT feature, minutiae and ridge feature		

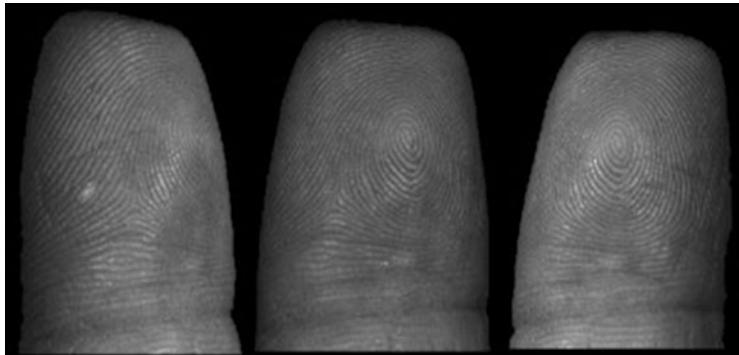


Fig. 9.23 Example fingerprint images captured by our device (*left, middle, right*)

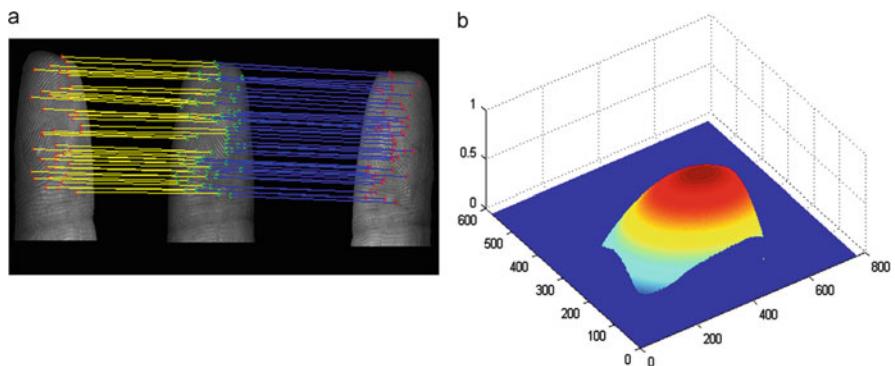


Fig. 9.24 Reconstruction of 3D finger shape of Fig. 9.23. **(a)** Manually labeled correspondences between fingerprint images, **(b)** reconstructed 3D finger shape based on **(a)**

Table 9.3 Mean distance and standard variation of error map between estimated finger shape and real finger shape of example images in Fig. 9.16

Index factor fitting mode function	Mean distance mean $(V - \tilde{V})$	Standard variation $std(V - \tilde{V})$
$f_1(x, y)$	0.024	0.019
$f_2(x, y)$	0.082	0.057

data in Fig. 9.16a. It can be seen that the error between V and $\tilde{V}_{Eq.9.2}$ is smaller than the one between V and $\tilde{V}_{Eq.9.3}$. Next, the errors between the 3D point cloud data and their corresponding fitting results of all 440 fingers we collected are computed. It can be seen from Fig. 9.26 that the binary quadratic function is more suitable for the finger shape model since smaller errors are obtained between the original 3D point cloud data

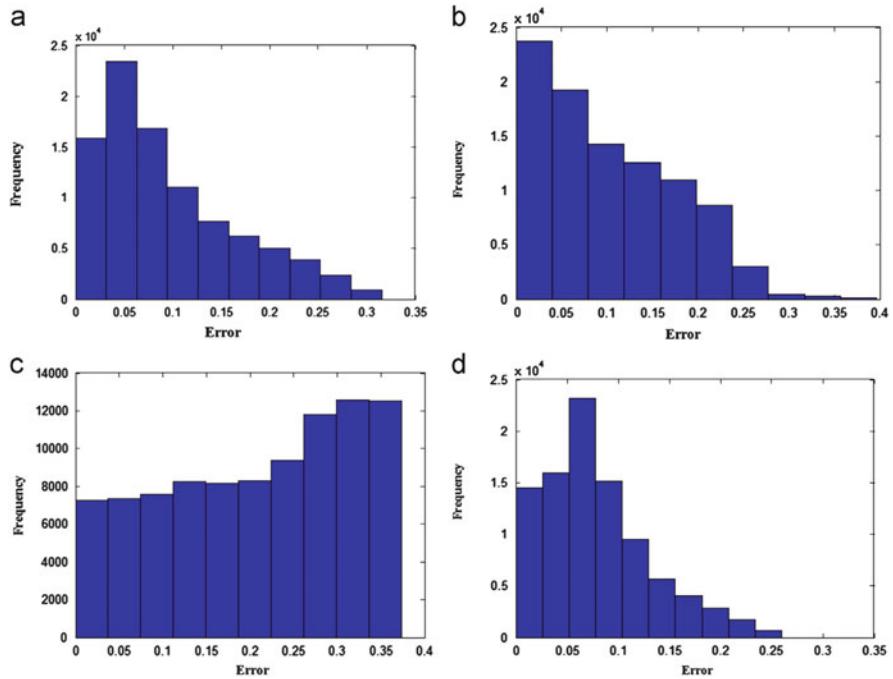


Fig. 9.25 Histogram of error maps between reconstructed results in Table 9.2 and Fig. 9.24b. (a) Histogram of err map between Fig. 9.24b and reconstruction result by using SIFT feature only, (b) histogram of err map between Fig. 9.24b and reconstruction result by using minutiae only, (c) histogram of err map between Fig. 9.24b and reconstruction result by using ridge feature only, (d) histogram of err map between Fig. 9.24b and reconstruction result by using both SIFT feature and minutiae, (e) histogram of err map between Fig. 9.24b and reconstruction result by using both SIFT feature and ridge feature, (f) histogram of err map between Fig. 9.24b and reconstruction result by using both minutiae and ridge feature, (g) histogram of err map between Fig. 9.24b and reconstruction result by using SIFT feature, minutiae and ridge feature

and their corresponding fitting results by the binary quadratic function. For this reason, the binary quadratic function is chosen as the finger shape model in this chapter.

Since the final 3D finger shape is obtained after interpolation according to the prior estimated finger shape model, we compared the reconstruction result with the 3D point cloud data of the same finger to verify the effectiveness of the model. From the results shown in Fig. 9.27, it can be seen that the profile of the finger shape reconstructed from multi-cameras is similar to the 3D point cloud data even though it is not that accurate. The real distance between the upper left core point and the lower left delta point is also calculated and shown in Fig. 9.27a, c, the values are 0.357 and 0.386 respectively. As a result, it is concluded that the estimated finger shape model is effective even though there is an error between the reconstruction result and the 3D point cloud data.

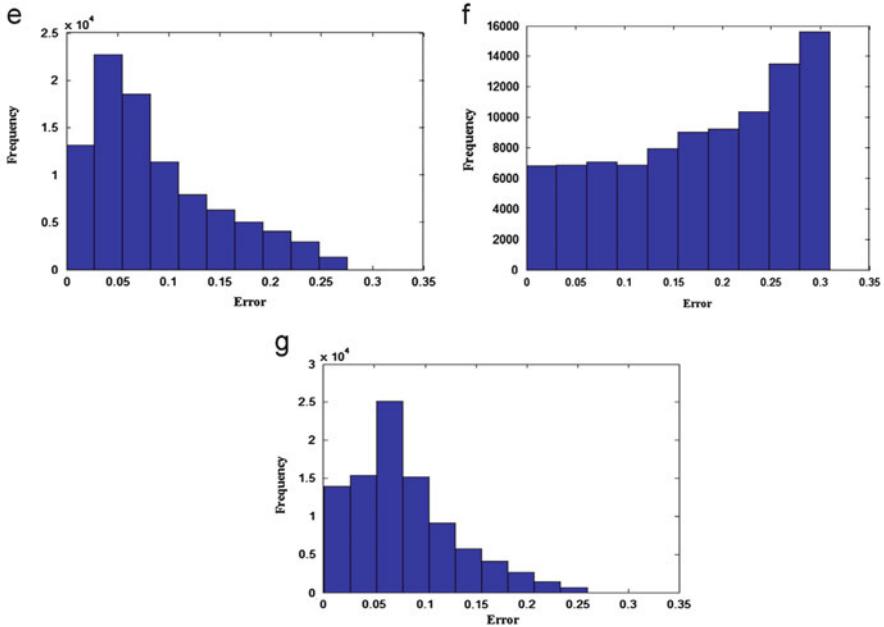


Fig. 9.25 (continued)

9.6.4 Reconstruction System Computation Time Analysis

There are six main parts included in our reconstruction system from image acquisition to result generation, as the block diagram shows in Fig. 9.3. The reconstruction method is implemented by Matlab on Fujitsu notebook embedded with Intel Core 2 Duo CPU, T9600 (2.80 GHz) processor. For image acquisition, it consumes no more than 100 ms to capture three views of fingerprint images since the frame rate of each camera is 30 frames/s. Because both of the camera parameters calculation and shape model estimation are done off-line, they do not occupy any time in the whole system. The correspondences establishment step consists of feature extraction and matching, which consumes considerable time. This time is variable for different images. The average time statistically calculated in our database is then used to measure. They are ~ 60.3 and ~ 24.32 s. It takes ~ 0.31 s to compute the 3D coordinates of feature correspondences. For interpolation, the code included in the matlab toolbox is employed and the consumption time is ~ 1.21 s. To summarize, it takes ~ 1.5 min to generate a 3D image by using the proposed system. It is believed, however, this time will be largely reduced once the code is compiled by C/C++ language and the multithread processing technique is used.

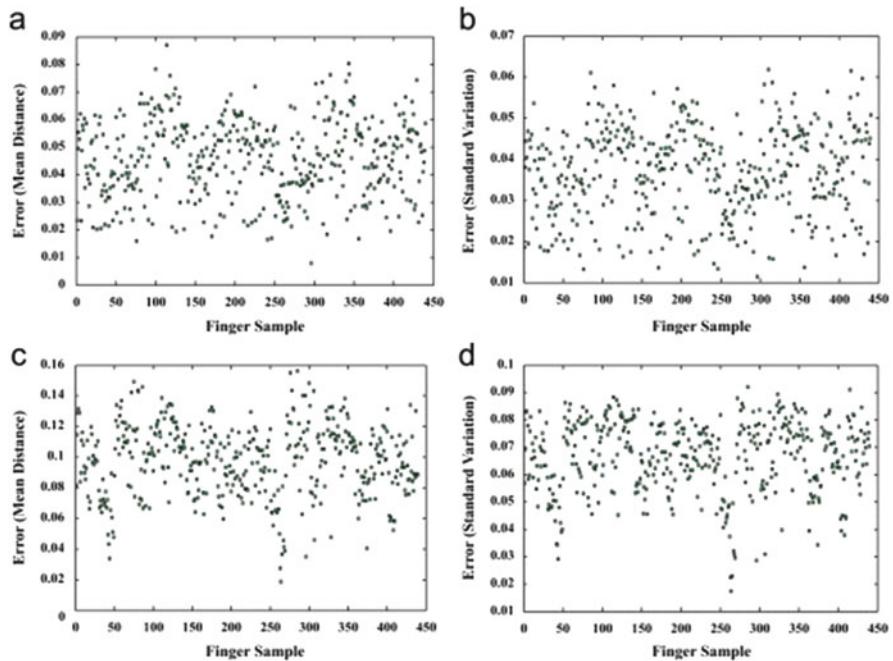


Fig. 9.26 Errors between the original 3D point cloud data of all 440 fingers we collected and their corresponding fitting results by different models. (a) Errors represented by the mean distance between the original 3D point cloud data and their corresponding fitting result by binary quadratic function, (b) errors represented by the standard variation between the original 3D point cloud data and their corresponding fitting result by binary quadratic function, (c) errors represented by the mean distance between the original 3D point cloud data and their corresponding fitting result by the mixed model, (d) errors represented by the standard variation between the original 3D point cloud data and their corresponding fitting result by the mixed model

9.6.5 3D Fingerprint Recognition

To study the distinctiveness of curve-skeleton features of human fingers, we show examples of matching results of different gender and different fingers. As shown in Table 9.4, examples of curve-skeletons from a female and a male with thumb, index finger and little finger captured at different sessions are given. We then matched them by ICP method. The percentage of matched points (Pm) and the mean distance between matched pairs (Mdist) are taken as the match score.

We firstly matched the curve-skeletons from the same finger but captured at different time, as listed in Table 9.5. Results show that the mean distance between matched pairs are smaller than one and the percentage of matched points are larger than 70%. Figure 9.28 also shows the matching results of different gender and finger types, the match scores are listed in Table 9.6. The results show that big

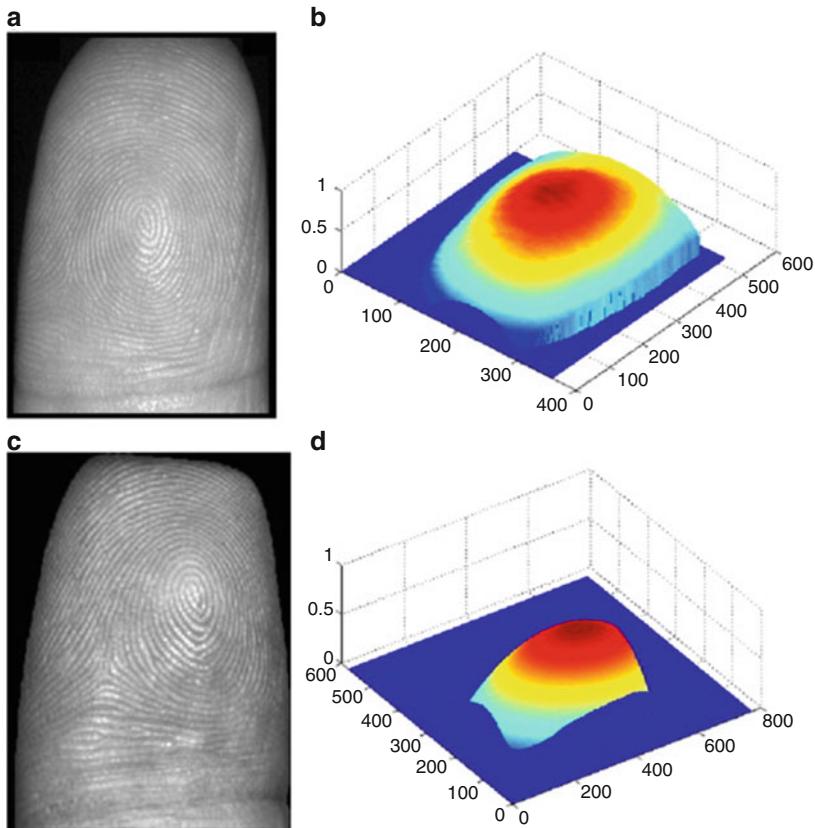


Fig. 9.27 Comparison of 3D fingerprint images from the same finger but different acquisition technique. (a) Original fingerprint image captured by the camera when collecting 3D point cloud, (b) 3D point cloud collected by one camera and a projector using the SLI method, (c) original fingerprint image captured by our device, (d) reconstructed 3D fingerprint image with labeled correspondences

difference existed between different fingers and different genders in curve-skeleton, since such feature reflects the ridge width feature of human finger and curvatures are different for human finger from finger tip to the distal interphalangeal crease.

Fingerprint identification experiment based on curve-skeletons is then implemented on our established database. Figure 9.29 shows the ROCs of different match score indexes. The EERs were obtained from 541 genuine scores and 292,140 imposter scores (generated from 541 fingers, 2 pictures of each finger). From the results, we can see that an EER of around 15% can be obtained when matching 3D fingerprint curve-skeleton features by simple ICP algorithm. The index of mean distance between matched pairs is better than the percentage of matched points. Curve-skeleton feature of 3D fingerprint image can be used to distinguish different fingers even though it is not as accurate as other higher level fingerprint features.

Table 9.4 Examples of extracted curve-skeletons from different gender and different fingers

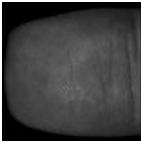
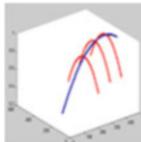
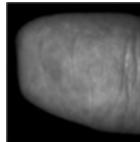
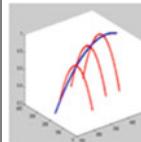
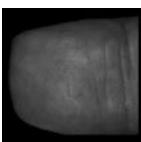
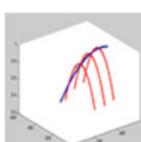
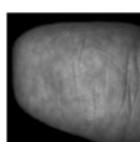
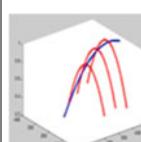
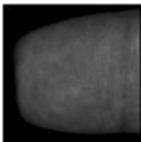
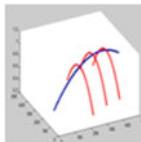
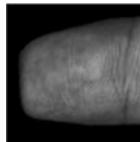
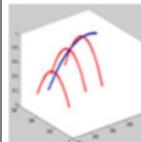
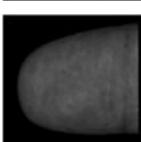
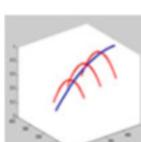
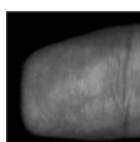
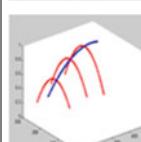
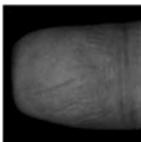
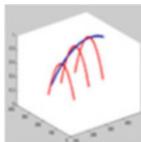
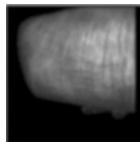
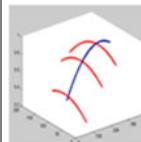
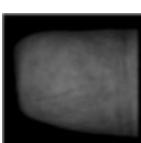
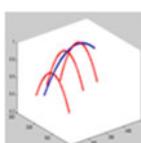
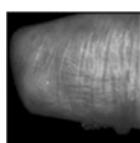
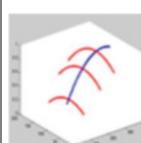
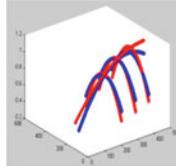
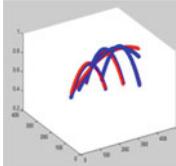
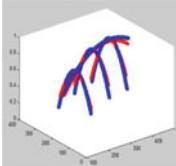
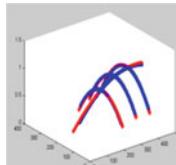
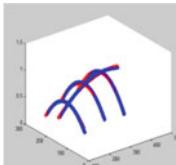
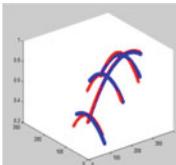
Finger type		Gender			
		Male		Female	
		Original 2D image	Curve-skeleton	Original 2D image	Curve-skeleton
Thumb	Session 1 (a1)				
	Session 2 (a2)				
Index finger	Session 1 (b1)				
	Session 2 (b2)				
Little finger	Session 1 (c1)				
	Session 2 (c2)				

Table 9.5 Matching result of curve-skeletons from the same finger but different session

Gender	Finger type		
	Thumb (a1)–(a2)	Index finger (b1)–(b2)	Little finger (c1)–(c2)
Male			
	Pm = 74%; Mdist = 0.20	Pm = 93%; Mdist = 0.39	Pm = 79%; Mdist = 0.25
Female			
	Pm = 94%; Mdist = 0.72	Pm = 97%; Mdist = 0.09	Pm = 90%; Mdist = 0.32

9.7 Summary

This chapter investigates a 3D reconstruction technique based on limited feature correspondences in 2D fingerprint images captured by the designed multi-view touchless fingerprint imaging device. Specific to the characteristic of low ridge-valley contrast of touchless fingerprint images, an improved fingerprint enhancement method is proposed, so as to extract more robust fingerprint features. Then, three frequently used features, i.e., SIFT feature, ridge feature and minutiae, having different numbers and various distributions, are considered for correspondences establishment. Correspondences are finally established by adopting the hierarchical fingerprint matching approaches. The finger shape model in this chapter is estimated by analyzing 3D point cloud finger data collected by one camera and a projector using the SLI method. Results show that the binary quadratic function is more suitable for the finger shape model compared with another mixed model proposed in the chapter. By reconstructing a standard cylinder object, it is shown that it is reasonable and feasible for us to adopt the methodology of the reconstruction technique, as well as the capturing device. The comparison and analysis of 3D fingerprint reconstruction results based on different fingerprint feature correspondences illustrates that best reconstruction results can be generated if all three features of correspondences are used. However, it is recommended to simply use SIFT and minutiae since comparable results are achieved by using them. The effectiveness of the estimated finger shape model is verified by comparing the reconstructed 3D finger shape with the corresponding 3D point cloud finger data.

This chapter further studied the recognition of the reconstructed 3D fingerprint image. Some fingerprint features which are coarser than Level 1 features-Curvature

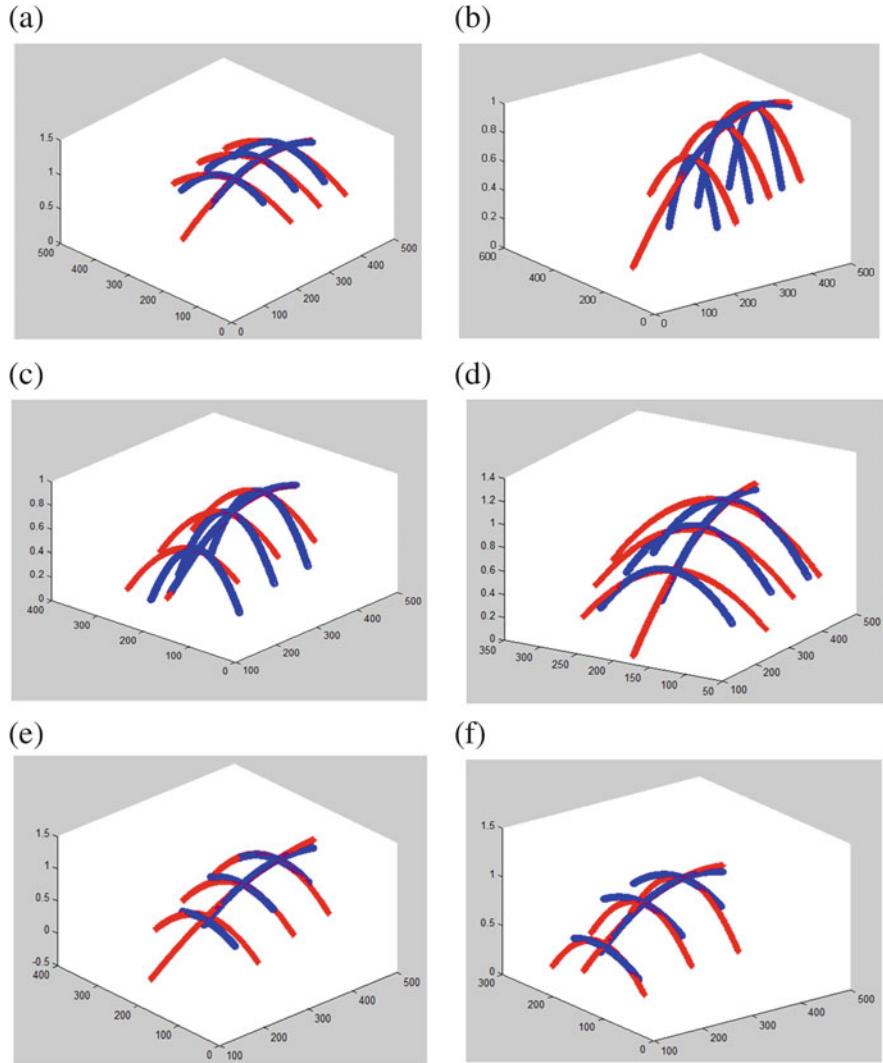


Fig. 9.28 Example of matching results of curve-skeletons from different gender and finger types in Table 9.4: (a) Matching result of [(male, thumb)—(male, index finger)], (b) Matching result of [(male, thumb)—(male, little finger)], (c) Matching result of [(male, index finger)—(male, little finger)], (d) Matching result of [(female, thumb)—(female, index finger)], (e) Matching result of [(female, thumb)—(female, little finger)], (f) Matching result of [(female, index finger)—(female, little finger)], (g) Matching result of [(male, thumb)—(female, thumb)], (h) Matching result of [(male, index finger)—(female, index finger)], (i) Matching result of [(male, little finger)—(female, little finger)]

Fingerprint Features, are firstly defined. These features are then used for assisting fingerprint recognition. Experimental results show that an EER of $\sim 15\%$ can be achieved when using 3D curve-skeleton for recognition.

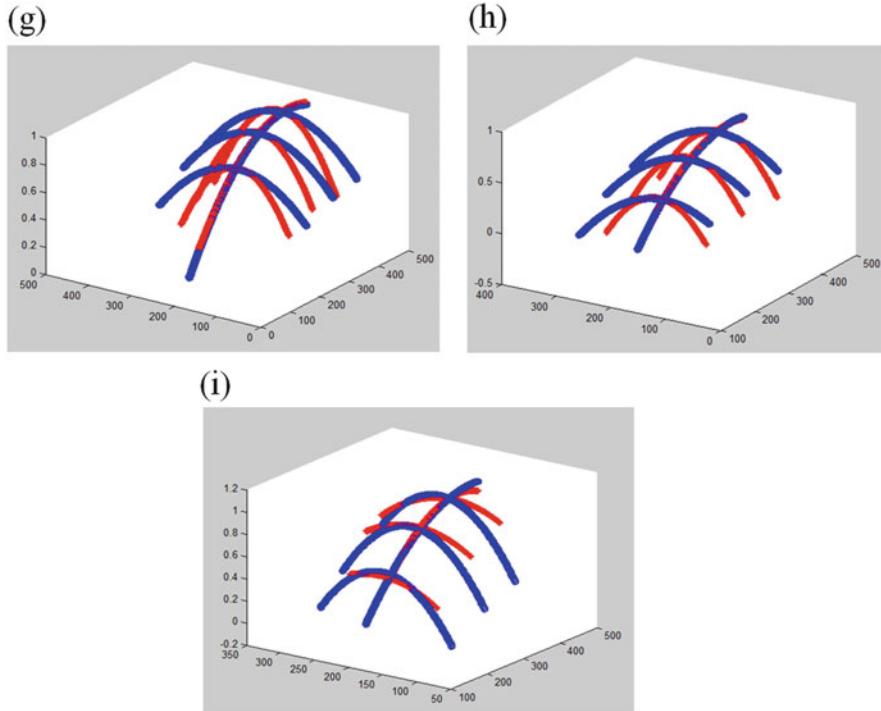


Fig. 9.28 (continued)

Table 9.6 Match scores corresponding to Fig. 9.28

Match score index	Corresponding labels in Fig. 9.28								
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)
Pm (%)	57	38	53	55	45	62	50	53	57
Mdist	8.3	13.7	2.9	6.8	14.8	3.1	4.0	4.1	4.6

Currently, researchers find that 3D fingerprint images provide more attributes for fingerprint features than 2D fingerprint images. For instance, a minutia feature in 2D fingerprint image is usually represented by its location $\{x, y\}$ and orientation θ . While in 3D case, it may be noted by $\{x, y, z, \theta, \phi\}$, where x, y and z are the spatial coordinates. Two angles of orientation of the ridge in 3D space θ and ϕ are available. Thus, fingerprint recognition with higher security can be achieved by matching features in 3D space [e.g. 3D minutia matching (Parziale and Niel 2004)]. By observing fingerprint in 3D images, we find that the center part of the finger is higher than the side parts, and the core point on fingerprints is located at almost the highest part of the finger. These characteristics of 3D fingerprint images benefit alignment when two fingerprint images are compared. Thus, our future work will investigate the application of such 3D information for fingerprint recognition.

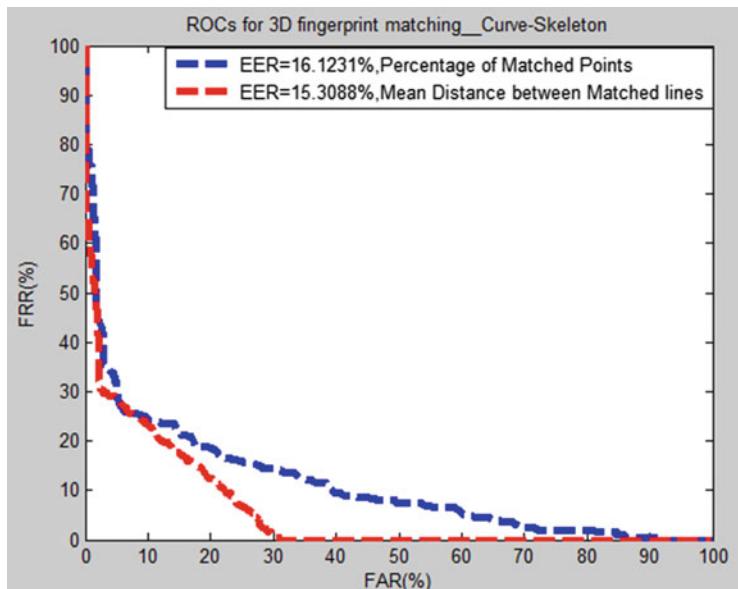


Fig. 9.29 ROCs for 3D fingerprint matching by ICP with Curve-Skeleton feature

References

- Almansa A, Lindeberg T (2000) Fingerprint enhancement by shape adaptation of scale-space operators with automatic scale selection. *IEEE Trans Image Process* 9(12):2027–2042
- Blais F, Rious M, Beraldin J (1988) Practical considerations for a design of a high precision 3-Daserscanner system. *Proc SPIE*:225–246
- Bouguet J. Camera calibration toolbox for Matlab. http://www.visioncaltech.edu/bouguetj/calib_doc/download/index.html
- Bradley B, Chan A, Hayes M (2002) A simple, low cost, 3D scanning system using the laser light-sectioning method. In: Proceedings of the IEEE international instrumentation and measurement technology conference, Victoria, Vancouver Island, Canada, pp 299–304
- Chatterjee S, Hadi A (1986) Influential observations, high leverage points, and outliers in linear regression. *Stat Sci* 1(3):379–416
- Chen H, Dong G (2006) Fingerprint image enhancement by diffusion processes. In: Proceedings of the 13th international conference on image processing, pp 297–300
- Chikkerur S, Cartwright A, Govindaraju V (2007) Fingerprint enhancement using STFT analysis. *Pattern Recogn* 40(1):198–211
- Choi K, Choi H, Lee S, Kim J (2007) Fingerprint image mosaicking by recursive ridge mapping. *IEEE Trans Syst Man Cybern B* 37(5):1191–1203
- Choi H, Choi K, Kim J (2010) Mosaicing touchless and mirror-reflected fingerprint images. *IEEE Trans Inf Forensics Secur* 5(1):52–61
- Draper N, Smith H (1981) Applied regression analysis, 2nd edn. Wiley, New York
- Feng J (2008) Combining minutiae descriptors for fingerprint matching. *Pattern Recogn* 41 (1):342–352
- Fishler M, Bolles R (1981) Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun ACM* 24(6):381–395

- Hao Y, Yuan C (2004) Fingerprint image enhancement based on nonlinear anisotropic reversed if fusion equations. In: Proceedings of the 26th annual international conference of the IEEE engineering in medicine and biology society, pp 1601–1604
- Hartley R (2000) Multiple view geometry in computer vision. Cambridge University Press, Cambridge
- Hastings R (2007) Ridge enhancement in fingerprint images using oriented diffusion. In: DICTA'07 Proceedings of the 9th biennial conference of the Australian Pattern Recognition Society on digital image computing techniques and applications pp 245–252
- Hernandez C, Vogiatzis G, Cipolla R (2008) Multiview photometric stereo. *IEEE Trans Pattern Anal Mach Intell* 30(3):548–554
- Hong L, Wan Y, Jain A (1998) Fingerprint image enhancement: algorithm and performance evaluation. *IEEE Trans Pattern Anal Mach Intell* 20(8):777–789
- Hu G, Stockman G (1989) 3-D surface solution using structured light and constraint propagation. *IEEE Trans Pattern Anal Mach Intell* 11(4):390–402
- Jain A, Ross A (2002) Fingerprint mosaicking. In: Proceedings of the IEEE international conference on acoustics, speech, and signal processing (ICASSP), Orlando, Florida, 4: pp 4064–4067
- Jain A, Hong L, Bolle R (1997) On-line fingerprint verification. *IEEE Trans Pattern Anal Mach Intell* 19(4):302–314
- Jirachaweng S, Areekul V (2007) Fingerprint enhancement based on discrete cosine transform. In: Proceedings of the international conference on biometrics, LNCS, pp 96–105
- Kumar A, Zhou Y (2011) Contactless fingerprint identification using level zero features. In: Proceedings of CVPR 2011, pp 121–126
- Liu F, Zhang D, Lu G, Song C (2013) Touchless multi-view fingerprint acquisition and mosaicing. *IEEE Trans Instrum Meas* 62:2492–2502. doi:[10.1109/TIM.2013.2258248](https://doi.org/10.1109/TIM.2013.2258248)
- Lowe D (2004) Distinctive image features from scale-invariant key points. *Int J Comput Vis* 60 (2):91–110
- Malathi S, Meena C (2010) Partial fingerprint matching based on SIFT features. *Int J Comput Sci Eng* 4(2):1411–1414
- Maltoni D, Maio D, Jain A, Prabhakar S (2009) Handbook of fingerprint recognition. Springer, New York
- Park U, Pankanti S, Jain A (2008) Fingerprint verification using SIFT features. In: Proceedings of SPIE 6944, 69440 K-69440K-9
- Parziale G, Diaz-Santana E (2006) The surround imager: a multi-camera touchless device to acquire 3D rolled-equivalent fingerprints. In: Proceedings of international conference on biometrics (ICB), Hong Kong, China, pp 244–250
- Parziale G, Niel A (2004) A fingerprint matching using minutiae triangulation. In: Proceedings of the international conference on biometric authentication (ICBA), LNCS, 3072, pp 241–248
- Ross A, Dass S, Jain A (2005) A deformable model for fingerprint matching. *Pattern Recogn* 38 (1):95–103
- Rusinkiewicz S, Holt O, Levoy M (2002) Real-time 3D model acquisition. In: Proceedings of the 29th annual conference on computer graphics and interactive techniques, pp 438–446
- Saldner H, Huntley J (1997) Temporal phase unwrapping: application to surface profiling of discontinuous objects. *Appl Opt* 36(13):2770–2775
- Shah S, Ross A, Shah J, Crihalmeau S (2005) Fingerprint mosaicking using thin plate splines. In: Proceedings of the biometric consortium conference
- Stockman G, Chen S, Hu G, Shrikhande N (1988) Sensing and recognition of rigid objects using structured light. *IEEE Control Syst Mag* 8(3):14–22
- Wang Y, Hassebrook L, Lau D (2010) Data acquisition and processing of 3-D fingerprints. *IEEE Trans Inf Forensics Secur* 5(4):750–760
- Weichert J (1999) Coherence-enhancing diffusion filtering. *Int J Comput Vis* 31(2–3):111–127
- Xie M, Wang Z (2004) Fingerprint enhancement based on edge-directed diffusion. In: Proceedings of the 11th international conference on image processing, pp 274–277

- Zhang Z (2000) A flexible new technique for camera calibration. *IEEE Trans Pattern Anal Mach Intell* 24(11):1330–1334
- Zhang D, Lu G, Li W (2009) Palmprint recognition using 3-D information. *IEEE Trans Syst Man Cybern Part C Appl Rev* 39(5):505–519
- Zhang D, Kanhangad V, Luo N, Kumar A (2010) Robust palmprint verification using 2D and 3D features. *Pattern Recogn* 43(1):358–368
- Zhang D, Liu F, Zhao Q, Lu G, Luo N (2011) Selecting a reference high resolution for fingerprint recognition using minutiae and pores. *IEEE Trans Instrum Meas* 60(3):863–871

Chapter 10

Hand Back Skin Texture for Personal Identification

Abstract Human hand back skin texture (HBST) is often consistent for a person and distinctive from person to person. In this chapter, we study the HBST pattern recognition problem with applications to personal identification and gender classification. A specially designed system is developed to capture HBST images, and an HBST image database was established, which consists of 1920 images from 80 persons (160 hands). An efficient text on learning based method is then presented to classify the HBST patterns. First, textons are learned in the space of filter bank responses from a set of training images using the l_1 -minimization based sparse representation (SR) technique. Then, under the SR framework, we represent the feature vector at each pixel over the learned dictionary to construct a representation coefficient histogram. Finally, the coefficient histogram is used as skin texture feature for classification. Experiments on personal identification and gender classification are performed by using the established HBST database. The results show that HBST can be used to assist human identification and gender classification.

Keywords Biometrics • Hand back skin texture • Texton learning • Orientation map • Sparse representation

10.1 Introduction

Skin, as the outermost part of the human body, is known to provide much useful information, such as health status (Skin disease atlas 2012; Cula et al. 2004) and human identity information (Cula et al. 2005). Skin appearance can be viewed as a kind of texture surface, and skin texture analysis can be used in various applications. For example, in (Cula et al. 2004), skin texture analysis is applied to computer-aided diagnosis in dermatology, where the dermatologist can use the computational texture representation to make an initial diagnosis for the patient. Meanwhile, biomedical evaluation based on skin texture can provide some tests for topical skin treatments, which can be used to judge whether these treatments are effective or not in the early stages. In addition, skin texture analysis can be used to estimate human skin age (Tanaka et al. 2008; Kim et al. 2009).

With the rapid development of computer techniques, researchers have investigated the use of various biometric traits, including fingerprints (Jain et al. 2007; Ratha and Bolle 2004), face (Delac and Grgic 2007; Wechsler 2006), iris (Daugman 1993, 2004), retina (Hill 1999; Borgen et al. 2008), palmprints (Guo et al. 2009; Sun et al. 2005) and finger-knuckle-prints (Zhang et al. 2010), etc., for the purposes of personal authentication. Moreover, face (Bruce et al. 1993; Moghaddam and Yang 2002) and gait (Li et al. 2008) have been used for gender classification. In (Moghaddam and Yang 2002), the authors demonstrated that the SVM classifier is able to learn and classify gender from a set of hairless low resolution face images with high classification accuracy. For gait-based gender recognition, a number of combinations of gait components (Li et al. 2008) are extracted to classify gender with the SVM classifier. Skin texture, as a potential biometric identifier to assist existing biometric traits, has also received certain attention in the past years. Based on the locally consistent property of the fingerprint skin tissue, Rowe (2007) extracted texture features of the fingerprint skin for human identification while reducing the size of the fingerprint sensing area. Cula (Cula et al. 2005, 2004; Cula and Dana 2001) used the bidirectional texture function, which is analogous to the bidirectional reflectance distribution function, to model skin texture to assist face recognition. For each skin texture surface, the bidirectional texture function is sampled in multiple camera views and illumination directions. However, obtaining accurate bidirectional image measurements of skin texture surface is hard, because the skin surface is non-planar, non-rigid and can be stretched.

It can be observed that the human hand back skin has a clear and consistent texture pattern which is uniformly distributed over a large portion of hand back. Based on our daily life experience, we know that the hand based skin texture (HBST) pattern is not permanent and it will change over time. For example, young people will have finer (i.e., smoother and smaller size of micro-cells) HBST than old people, while females will have much finer HBST than male. Nonetheless, over a relatively long period, the HBST of a person is stable. Based on (Tatsumi et al. 1999), the changes in skin associated with age can be visualized by gloss and wrinkles, and thus some measurements of wrinkles, gloss and density of microgrooves of skin can be used for age estimation. In (Tatsumi et al. 1999), the number of pixels in the binary image of the epidermal cross-section is used to estimate the age. From the curve of measured peripheral length vs. age in (Tatsumi et al. 1999), one can see that the peripheral length changes little in 1–2 years, which means that skin texture can remain stable for a relatively long period. These motivated us to investigate the possibility that the HBST pattern can be used to aid personal identification and gender classification. Many biometric identifiers such as fingerprints, faces, iris and palmprints, etc., have been proposed for human identification, and our goal is not to compete with those biometric identifiers, but to validate whether HBST has a sufficient level of accuracy so that it can be helpful to assist the existing biometric authentication techniques. Moreover, apart from biometric applications, as a specific kind of texture patterns, the established HBST dataset can also be used to evaluate the texture feature extraction

and classification algorithms in the community of computer vision and pattern recognition.

In this chapter, we study the use of HBST for personal identification and gender classification. To this end, an HBST imaging device was designed to capture HBST images. Since HBST is a type of fine scale feature, a high resolution (about 450 dpi) is set to capture the detailed texture patterns in hand back images. Different from the method in (Cula et al. 2005), where skin texture is modeled as a 3D texture and the bidirectional texture function is used to describe the skin appearance, we model HBST as a kind of 2D appearance texture because the hand back can be approximately viewed as a 2D plane. Therefore, we directly capture the HBST image using a CCD camera with the fixed position under the fixed illumination direction. Such a design makes the HBST image acquisition very efficient and feasible for the purpose of personal identification and gender classification. In the 3D model (Cula and Dana 2001, 2004), multiple cameras and multi-illuminations are needed to collect samples, which makes the imaging system very complex. Compared with the 3D model, modeling the hand back skin surface with the 2D model makes our imaging system much easier to design and more convenient to collect samples. In addition, our goal is to analyze the texture pattern in hand back skin so that 2D modeling is more suitable.

By using the designed HBST imaging device (please refer to Sect. 10.2 for more details), an HBST image database is established, which consists of 1920 images from 80 volunteers (160 hands). A texton learning based method is then proposed for HBST pattern classification. The HBST images are passed through a bank of filters, and a set of textons are learned from the filter responses with the sparse representation (SR) technique. Then, under the SR framework, the representation coefficient histograms of HBST images are computed and used for classification. The performance of the proposed method is evaluated by using the established HBST database in comparison with state-of-the-art texture classification schemes, including the multi-fractal spectrum (Xu et al. 2006), original LBP (Ojala et al. 2004), dominant LBP (Liao et al. 2009), completed LBP (Guo et al. 2010) and k -means based texton learning method (Varma and Zisserman 2002, 2005). Experimental results demonstrated that HBST could achieve interesting personal identification and gender classification accuracy, which implies that HBST can be used to aid existing biometric authentication techniques and improve the overall performance.

In summary, the major contributions of this work lie in that we developed the HBST imaging device, established an HBST dataset and proposed a sparse texton learning based HBST texture classification method. We validated that HBST pattern has potential to do personal identification and especially gender classification. In addition, as a special type of texture images, the established HBST dataset is very challenging, and it provides a good platform to evaluate and develop high performance texture feature extraction and classification algorithms.

10.2 Hand Back Skin Texture Imaging System

A schematic diagram of the major components of the developed hand back skin texture (HBST) imaging system is shown in Fig. 10.1. It is composed of a ring of LED light source, a lens, an associated CCD camera, and a data acquisition card. When it works to collect data, the LED light source will illuminate the hand back skin, and then the CCD camera will capture the HBST image and pass it to the data acquisition card. The data acquisition card will then transmit the image to the data processing unit (e.g., the CPU in a PC).

Figure 10.2a illustrates the inner structure of the HBST imaging device and Fig. 10.2b shows its exterior. A critical issue in HBST data acquisition is how to

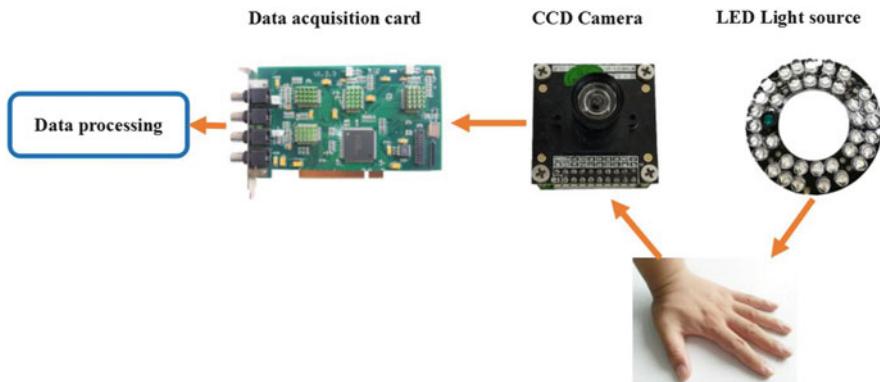


Fig. 10.1 The schematic diagram of the developed hand back skin texture imaging system

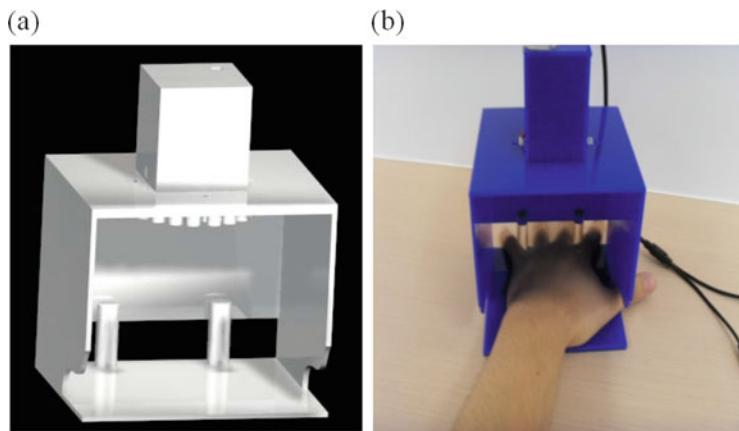


Fig. 10.2 (a) The inner structure of the developed hand back skin texture imaging system; (b) The outside view of the imaging system

make the data collection environment as stable and consistent as possible so that undesired disturbances (e.g., the background and environmental illumination disturbances) can be reduced. Meanwhile, a stable data collection environment can effectively reduce the complexity of feature extraction and improve the classification accuracy. Specifically speaking, in our imaging system how to keep the illumination uniform and constant and how to fix the position of the hand are of the most importance. To this end, a ring of LED light source (in visible spectrum, 390–780 nm) and a CCD camera are enclosed in a box to keep the illumination nearly constant. The LEDs are arranged in a circle around the camera to make the illumination uniform. Referring to Fig. 10.2a, in order to capture the central part image of the hand back skin texture, two pegs are used to fix the hand, which can guide the position of index and little fingers with a user friendly interface. This can also reduce largely the pose variations of the hand in different capturing sessions. In addition, our design could make the skin texture surface as flat as possible so that we can model the skin surface as a 2D planar texture image. Note that there are some differences between our device and the palmprint device (Zhang et al. 2003). First, in order to capture the micro-structures of HBST, the resolution of the chosen camera in our device is higher than that in the palmprint device. Second, the light source is different from that in the palmprint device. In our device, the ring LED is used while the halogen light source is used in the palmprint device. Finally, the architecture of the device is different. In our HBST imaging system, we employ the micro-industrial CCD camera board, LED light source and USB data acquisition card to collect data. However, in the palmprint device, the commonly used industrial CCD camera, halogen light source and PCI data acquisition card are used to collect data. Compared to the palmprint device, the size of the HBST imaging device is much smaller due to the use of micro-industrial CCD camera board and LED light source. In addition, the cost of HBST device is also lower.

The texture pattern of human hand back skin can only be clearly observed in a relatively fine scale. In order to capture the HBST image in a high enough resolution while avoiding the HBST image size to be too big, the focal of the lens should be carefully designed. In our imaging system, due to the limited distance between the camera and the hand back, we chose to use a 12 mm focal length lens to capture the HBST images. Further reduction in the focal length will distort the captured image. The size of the CCD output image is 576×768 (the raw image is saved in the 24-bitmap format and we convert it into 8-bit gray level image), and finally the HBST image is captured under a resolution of about 450 dpi. In designing our imaging device, we tested different resolution settings of the HBST image, and found the resolution of about 450 dpi can satisfy our requirements. If the resolution of the image is too low, the micro-structures such as wrinkles in the image cannot be captured clearly. If the resolution of the image is too high, the cost of the camera will be high and the computational cost will also increase. A resolution of 450 dpi is good enough to capture clear HBST images at a low cost.

In our HBST imaging system, since we use two pegs to fix the hand position, the top and bottom boundary of the captured skin texture image can be roughly fixed. Although the hand back skin can be viewed nearly as a 2D plane in the central part,

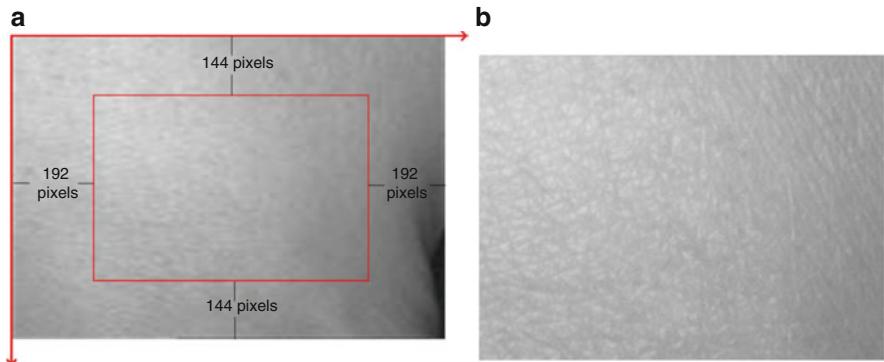


Fig. 10.3 (a) is the raw image (size 576×768) captured by our device and (b) is the sub-image (size 288×384) cropped from the central part of (a)

the boundary part of the hand back can be quite distorted in the captured HBST image. In order to reduce the effect of the hand back boundary area on the later feature extraction and recognition procedures, we crop a sub-image from the captured raw image by removing the four boundary areas. Referring to Fig. 10.3, we simply set the top left corner of the HBST image as the origin point, and based on our experimental experience we crop the central part of size 288×384 from the original image of size 576×768 . Such a sub-image cropping process cannot only make the feature extraction more stable and accurate, but also reduce computational cost.

Figure 10.4 shows some example cropped HBST images captured in two different sessions with a time interval of about 30 days. Figure 10.4a, b are the left-hand HBST images from one person in the two sessions, while Fig. 10.4c, d are the right-hand HBST images from the same person. Figure 10.4e, h are the left and right-hand HBST images from another person. Figure 10.5 shows the HBST images from one male subject and one female subject. From these HBST example images, we can have the following observations. (1) First, the left-hand and right-hand HBST patterns of a person are similar. (2) Second, the HBST patterns captured in different sessions from the same person are similar. (3) Third, the HBST patterns from different persons are different, which implies its potential for human identification. (4) At last, the HBST patterns of male and female subjects are different, which makes HBST pattern a good feature for gender classification.

10.3 HBST Feature Extraction and Classification

Texture classification is a classical topic in computer vision and pattern recognition. Although some well-known texture classification methods (Xu et al. 2006; Ojala et al. 2004; Lazebnik et al. 2005) can obtain good performance on some benchmark

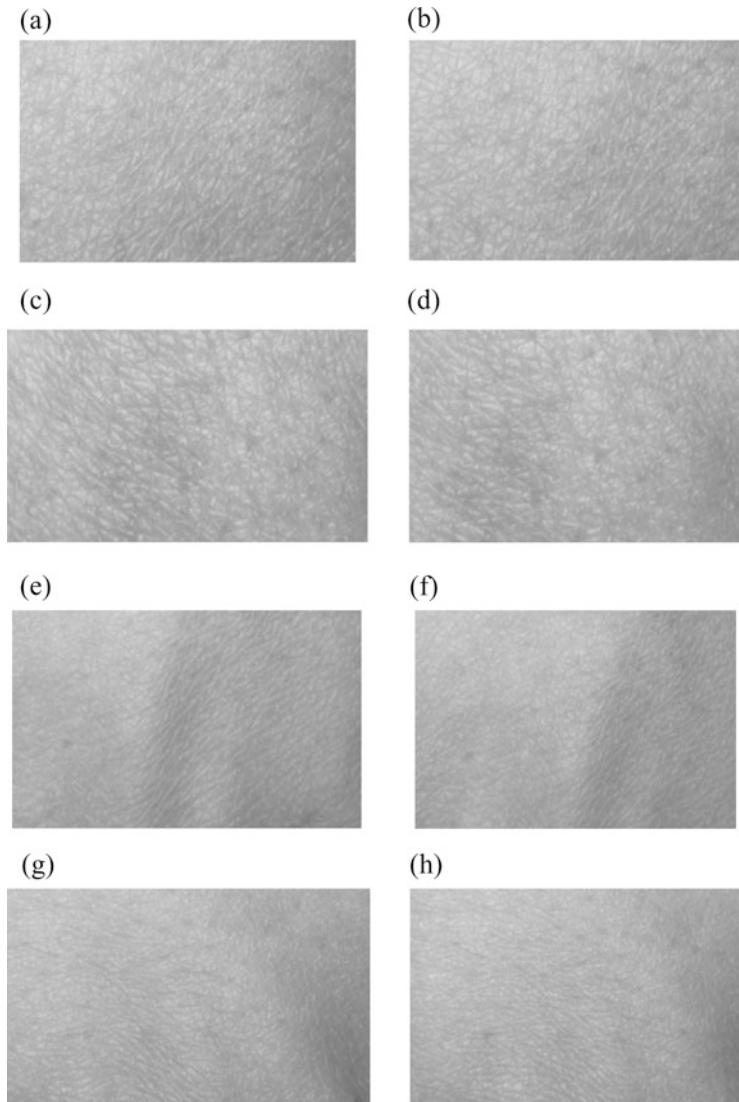


Fig. 10.4 (a) and (b) are the cropped left-hand HBST images of a person collected in two different sessions, while (c) and (d) are the right-hand HBST images from the same person. (e) and (f) are the cropped left-hand HBST images from another person, while (g) and (h) are the right-hand HBST images from this person

databases such as the UIUC (Lazebnik et al. 2005) and CUReT (Dana et al. 1999) texture databases, they may not be suitable for HBST patterns due to the special micro-structure of the hand back skin. The multi-fractals spectrum method (Xu et al. 2006) and the LBP method in (Ojala et al. 2004) cannot obtain good results because the features generated by them cannot characterize the appearance

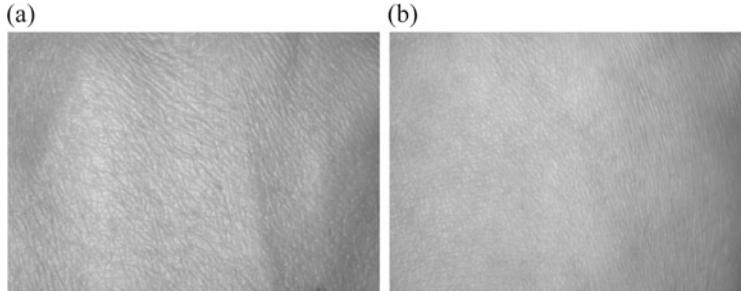


Fig. 10.5 (a) and (b) are the HBST images from one male and one female, respectively

of the skin texture well. Since there are no obvious interest points or interest regions in HBST images, the method in (Lazebnik et al. 2005) cannot detect accurately the affine invariant regions for a robust statistical description of the skin texture, and thus it will fail to classify the HBST patterns.

With a more careful look of the HBST images, we can observe some properties of the HBST patterns. First, there are no clear edges and corner points in the HBST images. Second, the HBST patterns are made up of some micro-cellular structures. Third, those micro-structures are generally distributed uniformly across the whole HBST image. Based on these observations, we choose to learn the micro-structures (i.e., textons) from the training HBST images, and then use them to describe the query HBST image for classification. Our experimental results in Sect. 10.4 also verify that the texton learning based method performs well for HBST pattern recognition.

As in (Varma and Zisserman 2002, 2005), the texton learning is performed in the space of MR8 filter bank responses. Different from (Varma and Zisserman 2002, 2005), which use the k-means method to learn the textons, in this chapter we employ the technique of sparse representation (SR) (Donoho 2006; Candes et al. 2006) to learn an over-complete dictionary of textons via the l_1 -norm minimization. And under the SR framework, we extract the SR coefficient histogram as the HBST feature for recognition. By filtering the training images with the MR8 filter bank, for each class of HBST images we can construct a training dataset $X = [x_1, x_2, \dots, x_n]$, where $x_i = 1, 2, \dots, n$, is an 8-dimensional MR8 filtering response vector at a pixel of the training sample images of this class. A dictionary of textons, denoted by $D = [d_1, d_2, \dots, d_l]$, will be trained from the constructed training dataset X , where d_j $j = 1, 2, \dots, l$, is a texton. The number of textons is generally much smaller than that of the elements in the training dataset, i.e., $l \ll n$. In the following sub-sections, we present in detail the method for HBST feature extraction and classification.

10.3.1 MR8 Filter Bank

The MR8 filter bank (Varma and Zisserman 2002, 2005) is a nonlinear filter bank with 38 filters but only eight filter responses. It contains 36 bar and edge filters,

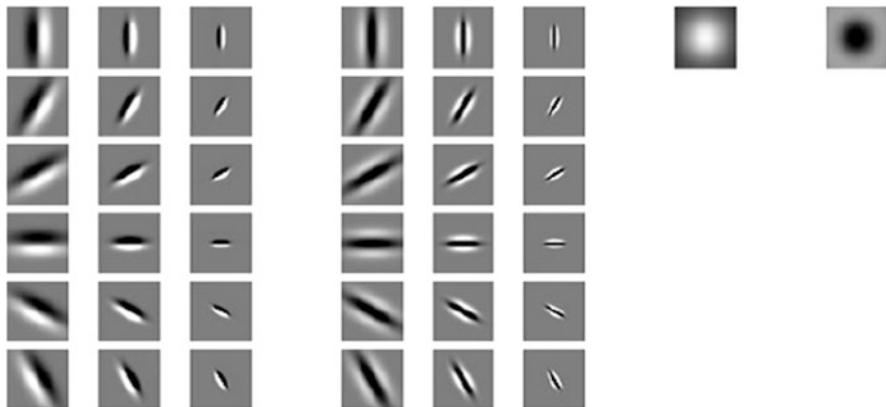


Fig. 10.6 The MR8 filter bank

which are along six orientations and across three scales, as well as a Gaussian filter and a Laplacian of Gaussian filter on a single scale. In order to obtain rotation invariance, for the edge and bar filters, the maximum filtering response along six orientations is selected for each scale. Moreover, using only the maximum orientation response can reduce the number of responses from 38 to 8. Figure 10.6 illustrates the MR8 filter bank. The motivation for using the MR8 filter bank is to extract rotation invariant skin texture features since there will be some rotation variations in the collected HBST images from the same subject. The MR8 filter bank responses are rotation invariant while preserving the distinctive features of the texture images. In (Varma and Zisserman 2002, 2005), the classical k -means clustering method is used to learn textons for texture image feature extraction and classification. Specifically, the textons are determined by solving the following problem:

When computing the MR8 filter bank responses, there are some pre-processing steps to follow in order to reduce some effects on feature extraction. Before convolving the original HBST image with the MR8 filter bank, all HBST images are normalized to have zero mean and unit standard deviation. This normalization can reduce the variations caused by illumination changes. After computing the MR8 filter responses, the filter response x_i at pixel i is normalized using the Weber's law (Varma and Zisserman 2002): $x_i \log(1 + L/0.03)/L$, where $L = \|x_i\|_2$ is the magnitude of the filter response vector x_i .

10.3.2 Texton Learning Based on SR

SR reveals the fact that if the input signal is intrinsically sparse in some domain, it can be sparsely represented over the dictionary which can define the sparse domain. For a given signal $x \in R^m$, we say that x has a sparse approximation over a

dictionary $D = [d_1, d_2, \dots, d_l] \in R^{m \times l}$, if we can find a linear combination of only “a few” atoms from D that is “close” to the signal x . Under this assumption, the sparsest representation of x over D is the solution of the following minimization problem:

$$\arg \min_{\alpha} \|\alpha\|_0 \text{ s.t. } \|x - D\alpha\|_2^2 \leq \varepsilon \quad (10.1)$$

where $\alpha \in R^l$ is the sparse representation coefficient vector by coding signal x over dictionary D such that $x \approx D\alpha$ and most of the elements in α are close to zero. The l_0 -norm counts the number of non-zero elements in the representation vector α . Because the l_0 -norm minimization is an NP hard problem, an alternative way is to solve the l_1 -norm minimization problem:

$$\arg \min_{\alpha} \|\alpha\|_1 \text{ s.t. } \|x - D\alpha\|_2^2 \leq \varepsilon \quad (10.2)$$

In the application of HBST analysis, the signal x is an 8-dimensional feature vector of MR8 filtering response at a pixel of the HBST image. In order to better represent the query image for the classification propose, the dictionary D needs to be learned from the training HBST images. For each class of HBST images, we filter the training images with the MR8 filter bank and construct a training data set $X = [x_1, x_2, \dots, x_n]$, where $x_i = 1, 2, \dots, n$, is the MR8 filtering response vector at pixel i of the training images of this class. The dictionary D associated with this class can be learned by optimizing the following objective function:

$$\arg \min_{D, \Lambda} \|\Lambda\|_1 \text{ s.t. } \|x - D\Lambda\|_F^2 \leq \varepsilon \quad (10.3)$$

where $\Lambda = [\alpha_1, \alpha_2, \dots, \alpha_n]$ and $\|\cdot\|_F$ is the Frobenius matrix norm. We can rewrite Eq. (10.3) into an unconstrained optimization problem with a penalty term:

$$\arg \min_{D, \Lambda} \|x - D\Lambda\|_F^2 + \lambda \|\Lambda\|_1 \quad (10.4)$$

The optimization problem in Eq. (10.4) is non-convex. Usually we can have a local minimum by alternatively optimizing D and Λ ; that is, from some initialization of D , we can solve Λ , and then by fixing Λ , we can update D . Such a procedure iterates until convergence. In this chapter, we adopt the alternating direction method in (Yang and Zhang 2011) to solve Λ (when D is fixed) and the Lagrange dual method (Lee et al. 2006) to update D (when Λ is fixed). After we learn the dictionary of textons for each class of HBST images, we combine these textons into one big over-complete dictionary, and use it to extract skin texture features.

In (Varma and Zisserman 2002, 2005), the classical k -means clustering method is used to learn textons for texture image feature extraction and classification. Specifically, the textons are determined by solving the following problem:

$$\arg \min_{d_j} \sum_{j=1}^l \sum_{x_i \in \Omega_j} \|x_i - d_j\|_2^2 \quad (10.5)$$

The k -means clustering will partition the training set $X = [x_1, x_2, \dots, x_n]$ into l groups $\Omega_1, \Omega_2, \dots, \Omega_l$, and the texton d_j is defined as the mean vector of all the vectors within α_i . The k-means clustering based texton learning method can be viewed as a special case of the SR based dictionary learning. If we require that α_i has only one non-zero element, which is 1, then the problem in Eq. (10.5) will be basically the same as the problem in Eq. (10.1). In this case, we use only one texton to represent the feature vector x_i and assign the label of x_i to that texton. For an input vector x_i which may lie in the boundary of two or more clusters, the k-means clustering will randomly assign it to one of the classes. However, such a representation may not be accurate enough in practice. In contrast, by using SR, x_i will be coded as a linear combination of more than one texton, which can achieve a much lower reconstruction error due to the less restrictive constraint. In the experiments in Sect. 10.4, we will see that by using the SR technique to learn the textons and the associated feature description method in Sect. 10.3.3, the HBST recognition accuracy can be much improved.

10.3.3 Feature Extraction and Classification

Denote by D_k the texton dictionary for the k th class of HBST, the dictionary for all c classes of HBST images can be formed by amalgamating the c dictionaries, $D = [D_1, D_2, \dots, D_c]$. With this dictionary D , each training HBST image can generate a model by mapping it to the texton dictionary.

In the method of (Varma and Zisserman 2002, 2005), each pixel of a texture image, is labeled with the element in the dictionary D that is closest to the feature vector at this pixel. A histogram of texton labels of this image is then formed for classification. Different from this method, under the SR framework, we can construct a histogram of the SR coefficients of a texture image for classification. The representation coefficient vector can be obtained by coding the feature vector x_i over D with the SR technique. However, the computational cost of solving the l_1 -norm minimization problem to obtain the SR coefficient is very heavy because of the highly over-complete dictionary D . To reduce the cost of sparse coding, we can use only a subset of D to represent x_i . Specifically, we use the closest t textons ($t \ll z$, z is the total number of textons learned from the c classes of HBST images) to x_i in D to form a sub-dictionary for x_i . Denote by $d_1^i, d_2^i, \dots, d_t^i$ the t closest textons to x_i , and the sub-dictionary for x_i is then $D_i = [d_1^i, d_2^i, \dots, d_t^i]$. The representation vector of x_i over D_i , denoted by $\alpha_i = [\alpha_1^i, \alpha_2^i, \dots, \alpha_t^i]$, can then be computed by solving the following l_1 -norm minimization problem:

$$\arg \min_{\alpha_i} \|x_i - D_i \alpha_i\|_F^2 + \lambda \|\alpha_i\|_1 \quad (10.6)$$

The alternating direction method in (Yang and Zhang 2011) can be used to solve Eq. (10.6). Since D_i is a subset of D , once we have α_i , we can easily construct another representation vector h_i over D such that:

$$D_i \alpha_i = D h_i \quad (10.7)$$

Obviously, most of the entries in h_i will be 0, and only the entries corresponding to the same textons as those in D_i will have non-zeros values, which are the same as those in α_i .

Finally, for each pixel at position i , we have a representation vector h_i . Hence, we can form a presentation coefficient histogram, denoted by H_f , for this HBST image by summing all the vectors of $|h_i|$:

$$H_f = \sum_{i=1}^N |h_i| \quad (10.8)$$

where N is the number of pixels in the HBST image. The H_f can be taken as the final feature descriptor of the HBST image for the classification purpose. Figure 10.7 shows the coefficient histograms of some HBST images from different persons.

We denote by H_j , $j = 1, 2, \dots, J$, the histogram of a training texture image. Similarly, for an input test image Y , we can construct a representation coefficient histogram for it, denoted by H_Y . The similarity between H_j and H_Y can be computed as:

$$\chi^2(H_j, H_Y) = \frac{1}{2} \sum \frac{(H_j - H_Y)^2}{H_j + H_Y} \quad (10.9)$$

The test HBST image Y can then be classified with the nearest neighbor classifier. That is, it is classified to the class whose training sample has the shortest χ^2 distance to it.

10.4 Experimental Results

10.4.1 Database Establishment

In order to evaluate the proposed HBST analysis method for personal identification and gender classification, we established an HBST image database using the developed HBST imaging device. Those HBST sample images were collected from 80 volunteers (160 hands), including 61 males and 19 females whose ages ranged from 20 to 50 years old.

The samples were collected in two different sessions. In each session, each person was asked to provide six left-hand and six right-hand HBST images,

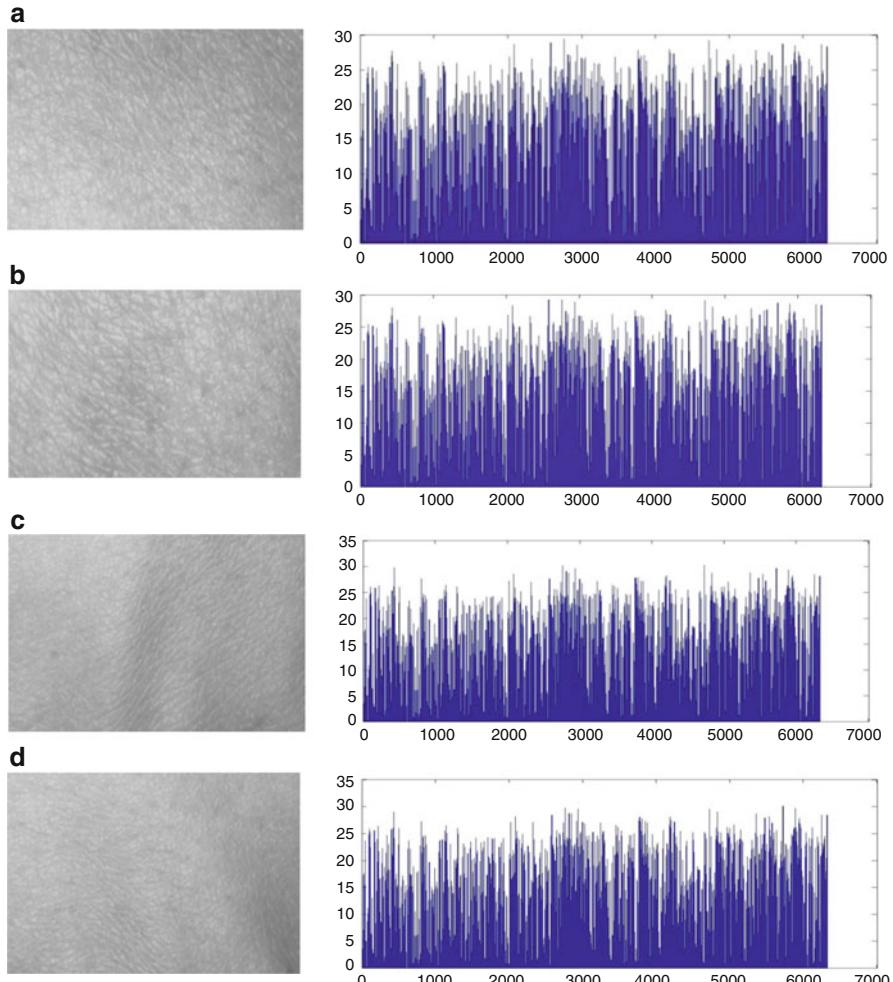


Fig. 10.7 The coefficient histograms of HBST images from different persons. (a) and (b) are the histograms of the *left-hand* HBST images from the same person while (c) and (d) are the histograms of the *left-hand* HBST images from another person

respectively. Therefore, 12 samples from one person were collected in each session. In total, the database contains 1920 samples from 160 hand backs. The average interval between the first and second session is about 30 days, and the maximum and minimum intervals are 40 days and 14 days, respectively. In the following experiments, without specific instructions, we use the samples collected in the first session as the training set and the samples in the second session as the test set.

Due to the various difficulties in data collection (e.g., the funding support, the recruitment of volunteers, etc.), our established database may not be large and comprehensive enough to support very strong conclusions. Nonetheless, we believe

that its size is reasonably large to illustrate if HBST patterns can be used to assist personal identification and gender classification. We are planning to collect more samples from more subjects in the following years, making our database more comprehensive and more balanced in terms of male and female subjects.

10.4.2 Personal Identification

In this section we aim to answer the question that whether HBST can be used as a kind of biometric trait to aid personal identification. To this end, we conducted five experiments using the proposed texton learning method with SR (TL_SR), and we compare the proposed TL_SR method to some representative texture classification methods such as the multi-fractal spectrum method (Xu et al. 2006), original LBP (Ojala et al. 2004), dominant LBP (DLBP) (Liao et al. 2009), completed LBP (CLBP) (Guo et al. 2010) and the texton learning method using the k -means clustering (TL_KM) (Varma and Zisserman 2002). For the multi-fractal spectrum method, the dimension of the multi-fractal spectrum vector is set as 26. In the original LBP, dominant LBP and completed LBP method, the radius of the neighborhood is set to 2 and the number of sampled points in the neighborhood is set to 8. For TL_KM, 40 textons are learned for each class of HBST images. In the proposed TL_SR method, 40 textons are also learned per class. Moreover, in the stage of feature description, for each descriptor x_i , t is set as 100, which means that 100 closest textons to x_i in D are chosen to form a sub-dictionary to obtain the SR coefficient. In the following experiments, we use the classification accuracy to evaluate these HBST classification methods. The classification accuracy is computed as $r = n_c/n$, where n_c is the number of correctly classified test samples and n is the number of all test samples.

Experiment 1

In the first experiment, all classes of HBST images are involved. The left and right hand HBST images from the same person are taken as from different classes. Therefore, in this experiment, there are 160 classes and each class has six training and six test samples. Since the multi-fractal spectrum vector and the histogram generated by the original LBP method cannot characterize well the appearance (e.g., cell-like micro-structures) of skin texture, they lead to poor experimental results in our task. The multi-fractal spectrum and original LBP methods can only achieve the classification accuracy of 35.65% and 46.52%, respectively. Hence, in the following experiments, we only compare TL_SR with DLBP, CLBP and TL_KM.

Table 10.1 shows the classification accuracies by the competing methods. We can see that the TL_SR method that uses the SR coefficient histogram as feature is superior to the TL_KM method that uses the texton label histogram for HBST classification. Also, the proposed method is better than the CLBP method, which combines the central pixel, magnitude and sign information of the neighborhood to completely model the LBP operator.

Table 10.1 Classification accuracies by competing methods

Method	DLBP (Liao et al. 2009)	CLBP (Guo et al. 2010)	TL_KM (Varma and Zisserman 2002)	TL_SR
Accuracy	75.56%	84.51%	84.40%	86.81%

For one person, the left hand and right hand HBST images are viewed as from two different classes. Thus there are 160 classes in this experiment

Table 10.2 Classification accuracies by competing methods

Method	DLBP (Liao et al. 2009)	CLBP (Guo et al. 2010)	TL_KM (Varma and Zisserman 2002)	TL_SR
Accuracy	78.59%	86.29%	88.40%	90.17%

For one person, the left hand and right hand HBST images are viewed as from the same class. Thus there are 80 classes in this experiment

Table 10.3 Classification accuracies on the left-hand HBST images

Method	DLBP (Liao et al. 2009)	CLBP (Guo et al. 2010)	TL_KM (Varma and Zisserman 2002)	TL_SR
Accuracy	80.38%	85.51%	84.54%	88.60%

Table 10.4 Classification accuracies on the right-hand HBST images

Method	DLBP (Liao et al. 2009)	CLBP (Guo et al. 2010)	TL_KM (Varma and Zisserman 2002)	TL_SR
Accuracy	82.91%	86.44%	85.24%	89.71%

The interesting HBST image classification accuracies validate that the proposed HBST identification system can well capture the characteristics of skin textures, allowing good discrimination between different classes. These results also suggest that human identification can be aided by HBST analysis.

Experiment 2

In the second experiment, all HBST images are involved. Different from Experiment 1, here the left and right hand HBST images from the same person are viewed as from the same class. Therefore, in this experiment there are 80 classes and each class has 12 training and 12 test samples. The experimental results using the DLBP, CLBP, TL_KM and TL_SR method are compared in Table 10.2. We can see that for all methods the classification accuracy is increased. This is mainly because the total number of classes is smaller than that in Experiment 1, and the left hand and right hand HBST images of one person are similar.

Experiment 3

The aim of this experiment is to evaluate the performance on the left and right hand HBST separately. For either left hand or right hand HBST images, there are 80 classes and 480 images in the training and test sets, respectively. The classification accuracies by different methods are listed in Tables 10.3 and 10.4. From the experimental results, one can see that the classification accuracy on the right-hand

HBST images is slightly higher than that on the left-hand HBST. This is probably because most people who provided their HBST samples to our database are right handed so that they feel more convenient to use our imaging device with the right hand. Therefore, compared to the left-hand HBST samples, the right-hand HBST samples collected in our database have less deformation, which results in a slightly higher classification accuracy for personal identification.

Experiment 4

In this experiment, we fuse the left-hand and right-hand HBST for identification. That is, both the left-hand and right-hand HBST samples of a person will be collected to identify his/her identity. Therefore, there are 480 pairs of left-hand and right-hand samples in the training set, which are from 80 subjects. In the test set there are also 480 pairs of HBST samples. For the left-hand and right-hand test samples, we calculate two distances χ_l^2 and χ_r^2 , where χ_l^2 is the distance between the left-hand test sample and left-hand training sample, and χ_r^2 is the distance between the right-hand test sample and right-hand training sample from the same pair. Then the two distances can be used by the simple weighted average method. The final distance for classification is $\chi_f^2 = w \times \chi_l^2 + (1 - w) \times \chi_r^2$, where the weight w can be trained from the training dataset using the “leave-one-out” strategy. For the four competing classification methods in our experiment, the weights are 0.4, 0.5, 0.45 and 0.4, respectively. The classification accuracies by fusing the left-hand and right-hand HBST with different methods are listed in Table 10.5. Compared with the results in Experiments 1–3 (please refer to Tables 10.1, 10.2, 10.3 and 10.4), one can see that the classification accuracy by fusing the left-hand and right-hand HBST images is much increased, showing that the left-hand and right-hand HBST patterns have complementary information.

Experiment 5

As we mentioned in the Introduction section, one goal of this work is to investigate whether hand back skin texture patterns can be used to aid other biometrics identifiers to improve personal identification accuracy. Therefore, in this experiment we fuse palmprint and HBST for personal identification. Since there are 160 hand backs (80 left hands and 80 right hands) in our HBST dataset, we randomly extract from the PolyU palmprint database (Zhang et al. 2003) 1920 palmprint images, which belong to 160 palms (80 left hands and 80 right hands). Each palm has 12 samples collected from two separated sessions, 6 samples per session. We then assume that each hand has six palmprint images and six HBST images in each session, and use the data from the first session for training, and use the data from the second session for testing.

Table 10.5 Classification accuracies by fusing the left hand and right hand HBST

Method	DLBP (Liao et al. 2009)	CLBP (Guo et al. 2010)	TL_KM (Varma and Zisserman 2002)	TL_SR
Accuracy	85.23%	87.24%	89.03%	92.51%

We use the competitive code scheme (Kong and Zhang 2004) to extract the palmprint feature, and use the Hamming distance to measure the similarity between palmprint features.

We fuse the palmprint and HBST matching distances by the weighted average method. The final distance for classification is $d_f = w \times d_p + (1 - w) \times d_h$, where d_p is the distance between palmprint sample sand d_h is the distance between HBST samples. In our experiment, the weight w is set to 0.8 by experience. The classification accuracies of palmprint, HBST and the fusion of palmprint and HBST are listed in Table 10.6. Compared with the identification rate by either palmprint or HBST individually, one can see that the accuracy is improved by fusing palmprint and HBST matching distances. This validates that HBST can be used to aid the existing biometric traits for personal identification.

10.4.3 Gender Classification

As can be seen in Fig. 10.5, the hand back skin appearance differs much from male to female. In most cases, the HBST surface from female is much smoother than that from male, and size of micro-cells in female HBST samples is smaller than those for males. Therefore, it is very interesting to verify that if the HBST patterns are distinctive enough to distinguish males from females. In this section, we conduct such experiments for gender classification.

In our HBST database, there are 61 males and 19 females. In gender classification, there are only two classes: male and female. The samples from all the 61 male subjects are taken as the samples of the male class, and the samples from all the 19 females are taken as those of the female class. The 960 samples collected from the first session are used as the training set, and the other 960 samples from the second session are taken as test samples. Table 10.7 shows the results by the DLBP, CLBP, TL_KM and TL_SR method. One can see that the gender classification accuracy can be higher than 98%, which implies that HBST can be aided to distinguish males from females.

Moreover, in Table 10.8 we present the numbers of falsely classified male and female samples by the proposed TL_SR method. As illustrated in Table 10.8, among the 732 male test samples, nine samples are incorrectly classified. Among

Table 10.6 Classification accuracies by palmprint, HBST and the fusion of them

Feature	Palmprint	HBST	Fusion
Accuracy	98.65%	86.81%	99.58%

Table 10.7 Gender classification accuracies by different methods

Method	DLBP (Liao et al. 2009)	CLBP (Guo et al. 2010)	TL_KM (Varma and Zisserman 2002)	TL_SR
Accuracy	95.46%	97.63%	98.60%	98.65%

Table 10.8 Numbers and rates of falsely classified male and female samples by the proposed TL_SR method

	Male	Female
Number	9	4
Rate	1.23%	1.75%

the 228 female samples, four samples are falsely classified. The classification error rates of male and female samples are 1.23% and 1.75%, respectively. Although the numbers of male and female subjects in our database are not balanced, the classification error rate on female samples is only slightly higher than that on male samples. Certainly, we need to collect more samples and make the dataset more balanced to further validate this conclusion.

10.4.4 Discussion

Currently, compared with the biometric traits such as fingerprints, iris scans and palmprints, etc., the personal identification accuracy of HBST is much lower than for those. However, each biometric trait has its pros and cons, and no one can supersede another one. In practice using two or more biometric traits together will provide a more robust solution. In this work, our goal is to investigate whether hand back skin texture patterns can be used to aid personal identification and gender classification. Considering that HBST images can be collected when capturing fingerprint or palmprint images, fusing fingerprint/palmprint and HBST can be a good way for multi-modal biometrics, as we demonstrated in Sect. 10.4.2.

Furthermore, as a specific type of texture images, the established HBST dataset can be used to test texture classification algorithms in the community of computer vision and pattern recognition. Different from the commonly used texture datasets such as UIUC (Lazebnik et al. 2005), CUReT (Dana et al. 1999) and KTH_TIPS (Hayman et al. 2004), which are challenging in terms of scaling, viewpoint and illumination variations, the established HBST dataset is also challenging but in a very different aspect: the high inter-class similarity. In CUReT, KTH_TIPS and UIUC, different materials are viewed as different classes. However, in our HBST dataset, samples are from different persons but they are all from the same material: skin texture. Although there are no significant scale, viewpoint and illumination changes in the HBST dataset, the high inter-class similarity makes it challenging to achieve a high classification rate. Some classical texture classification methods such as LBP and multi-fractal spectrum, which work well on the CUReT, KTH_TIPS and UIUC datasets, do not work well on the HBST dataset. This motivates us to develop more advanced texture classification methods.

It should be noted that although HBST analysis can assist personal identification and gender classification, there are some factors, such as hairs on skin and humidity of skin, to affect the performance of personal identification and gender

classification. In our established HBST database, most of samples are collected from oriental people so that there are relatively few hairs on the hand back skin. In our future work, we will collect more samples from more subjects and investigate the influences of these factors on skin texture analysis. In addition, modeling skin texture over a long period is a challenging problem since there are large variations between skin textures in different ages. Hence, in the future we will study how to model skin texture over a long period more effectively to improve the performance of biometric tasks with skin texture analysis.

10.5 Summary

This chapter studied the problem of using hand back skin texture (HBST) for assisting personal identification and gender classification. An effective skin texture imaging system was developed for capturing HBST images. Moreover, we employed the sparse representation (SR) technique to learn the dictionary of textons to model the HBST pattern. Then, based on the learned textons of HBST images, we extracted the SR coefficient histogram as feature for classification. To evaluate the performance of the proposed system, an HBST database was established, consisting of 1920 images from 160 hands of 80 persons. Extensive experiments were conducted and the experimental results showed that human identification and gender classification can be aided by HBST analysis with good performance. In the future, more HBST samples need to be collected to verify the different aspects of HBST analysis and algorithm development. Meanwhile, some factors (hairs on the skin, humidity, etc.) will be investigated for HBST analysis.

References

- Borgen H, Bours P, Wolthusen S (2008) Visible-spectrum biometric retina recognition. In: Proceedings of the international conference on intelligent information hiding and multimedia signal processing, Harbin, China, pp 1056–1062
- Bruce V, Burton A, Dench N, Hanna E, Healey P, Mason O, Coombes A, Fright R, Linney A (1993) Sex discrimination: how do we tell the difference between male and female faces? Perception 22:131–152
- Candes E, Romberg J, Tao T (2006) Stable signal recovery from incomplete and in accurate measurements. Commun Pure Appl Math 8:1207–1223
- Cula O, Dana K (2001) Compact representation of bidirectional texture functions. In: Proceedings of computer vision and pattern recognition, Saint Petersburg, Russia, pp 1041–1047
- Cula O, Dana K (2004) 3D texture recognition using bidirectional feature histograms. Int J Comput Vis 1:33–60
- Cula O, Dana K, Murphy F, Rao B (2004) Bidirectional imaging and modeling of skintexture. IEEE Trans Biomed Eng 12:2148–2159
- Cula O, Dana K, Murphy F, Rao B (2005) Skin texture modeling. Int J Comput Vis 1:97–119

- Dana K, Ginneken B, Nayar S, Koenderink J (1999) Refelctance and texture of real world surfaces. *ACM Trans Graph* 2:1–34
- Daugman J (1993) High confidence visual recognition of persons by a test of statistical independence. *IEEE Trans Pattern Anal Mach Intell* 11:1148–1161
- Daugman J (2004) How iris recognition works. *IEEE Trans Circuits Syst Video Technol* 1:21–30
- Delac K, Grgic M (2007) Face recognition. I-Tech Education, Vienna
- Donoho D (2006) For most large underdetermined systems of linear equations the minimal ℓ_1 -normsolution is also the sparsest solution. *Commun Pure Appl Math* 6:797–829
- Guo Z, Zhang D, Zhang L, Zuo W (2009) Palmprint verification using binary orientation co-occurrence vector. *Pattern Recogn Lett* 30:1219–1227
- Guo Z, Zhang L, Zhang D (2010) Completed modeling of local binary pattern operator for texture classification. *IEEE Trans Image Process* 6:1657–1663
- Hayman E, Caputo B, Fritz M, Eklundh O (2004) On the significance of real-world conditions for material classification, vol 3024. ECCV 2004, Prague, Czech Republic, pp 253–266
- Hill R (1999) Retinal identification, in biometrics: personal identification in networked society. Kluwer Academic, Dordrecht
- Jain A, Flynn P, Ross A (2007) Handbook of biometrics. Springer, Berlin
- Kim K, Choi Y, Hwang E (2009) Wrinkle feature-based skin age estimation scheme. In: Proceedings of the international conference on multimedia and expo, pp 1222–1225
- Kong A, Zhang D (2004) Competitive coding scheme for palmprint verification, vol 1. In: International conference on pattern recognition, Cambridge, UK, pp 520–523
- Lazebnik S, Schmid C, Ponce J (2005) A sparse texture representation using local affine regions. *IEEE Trans Pattern Anal Mach Intell* 2:1265–1278
- Lee H, Battle A, Raina R, Ng A (2006), Efficient sparse coding algorithms. In: Proceedings of NIPS, Vancouver, BC, Canada, pp 801–808
- Li X, Maybank S, Yan S, Tao D, Xu D (2008) Gait components and their application to gender classification. *IEEE Trans Syst Man Cybern Part C Appl Rev* 38:145–155
- Liao S, Law M, Chung A (2009) Dominant local binary patterns for texture classification. *IEEE Trans Image Process* 5:1107–1118
- Moghaddam B, Yang M (2002) Learning gender with support faces. *IEEE Trans Pattern Anal Mach Intell* 24:707–711
- Ojala T, Pietikainen M, Maenpaa T (2004) Multi-resolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans Pattern Anal Mach Intell* 7:971–987
- Ratha N, Bolle R (2004) Automatic fingerprint recognition systems. Springer, Berlin
- Rowe R (2007) Biometrics based on multispectral skin texture. In: Proceedings of ICB, Seoul, Korea, pp 1144–1153
- Skin disease atlas (2012) <http://www.dermnet.com>. Accessed 20 Feb 2012
- Sun Z, Tan T, Wang Y, Li S (2005) Ordinal palmprint representation for personal identification. In: Proceedings of CVPR, San Diego, CA, USA, pp 279–284
- Tanaka H, Nakagami G, Sanada H (2008) Quantitative evaluation of elderly skin based on digital image analysis. *Skin Res Technol* 2:192–200
- Tatsumi S, Noda H, Sugiyama S (1999) Estimation of age by epidermal image processing. *Leg Med* 1:266–230
- Varma M, Zisserman A (2002), Classifying images of materials: achieving viewpoint and illumination independence. In: Proceedings of European conference on computer vision, Copenhagen, Denmark, pp 255–271
- Varma M, Zisserman A (2005) A statistical approach to texture classification from single images. *Int J Comput Vis* 1:61–81
- Wechsler H (2006) Reliable face recognition methods-system design, implementation and evaluation. Springer, Berlin
- Xu Y, Ji H, Fermuller C (2006) A projective invariant for textures. In: Proceedings of computer vision and pattern recognition, New York, USA, pp 1932–1939

- Yang J, Zhang Y (2011) Alternating direction algorithms for ℓ_1 -problems in compressive sensing. SIAM Sci Comput 1:250–278
- Zhang D, Kong W, You J, Wong M (2003) Online palmprint identification. IEEE Trans Pattern Anal Mach Intell 9:1041–1050
- Zhang L, Zhang L, Zhang D, Zhu H (2010) Online finger-knuckle-print verification for personal authentication. Pattern Recogn 7:2560–2571

Chapter 11

Line Scan Palmpoint Recognition System

Abstract Biometric recognition systems have been widely used globally. However, one effective and highly accurate biometric authentication method, palmpoint recognition, has not been popularly applied as it should have been, which could be due to the lack of small, flexible and user-friendly acquisition systems. To expand the use of palmpoint biometrics, we propose a novel palmpoint acquisition system based on the line-scan image sensor. The proposed system consists of a customized and highly integrated line-scan sensor, a self-adaptive synchronizing unit, and a field-programmable gate array controller with a cross-platform interface. The volume of the proposed system is over 94% smaller than the volume of existing palmpoint systems, without compromising its verification performance. The verification performance of the proposed system was tested on a database of 8000 samples collected from 250 people, and the equal error rate is 0.048%, which is comparable to the best area camera-based systems.

Keywords Identification of persons • Pattern recognition equipment • Image processing • Image sensors • Optical imaging

11.1 Introduction

BIOMETRICS, led by the fast development of imaging technologies and pattern recognition algorithms, has been utilized in complex fields, from physical/logical access control to justice/law enforcement; from time and attendance to healthcare biometrics (Biometric Applications 2011; Jain and Feng 2009; Laadjel et al. 2009a, b; Dai and Zhou 2011). Now, biometrics has been required to be real-time, online, user-friendly, and flexible for these complicated cross-disciplinary applications other than traditional high performance and robust requirements. For instance, hand-held fingerprint capturing devices and iris capturing devices, fingerprint sensors integrated with laptops, fingerprint sensors embedded in locks, iris sensors integrated with the steel safe. However, one of the best biometric technology, the palmpoint recognition, has been weak in these new features. Current palmpoint recognition systems have been limiting the palmpoint recognition applications.

The first palmprint recognition system was invented in 1998 (Shu and Zhang 1998). Since then, numerous systems have been created, such as flatbed scanners (Goh et al. 2003, 2006; Connie et al. 2005; Han et al. 2003; Lin et al. 2005; Struc and Pavesic 2008; Badrinath and Gupta 2007; Zheng et al. 2007), web camera based systems (Chaudhary and Nath 2009; Han et al. 2007a, b; Kumar et al. 2003; Goh et al. 2008, 2010; Zhu and Zhang 2010), and palmprint systems with the pegged flat platen surface (Kong et al. 2006, 2009, 2003; Kong and Zhang 2002; Li and Zhang 2009; Wong et al. 2005; Hao et al. 2008; Wang et al. 2008; Zhang et al. 2010a, b, 2009). These systems achieved good performance. However, limited by the imaging structure, the dimensions of the devices of these systems cannot be reduced further. In addition, the user interaction and applications were limited to current situation.

Line scan technique would be an ideal solution for an online palmprint recognition system. Using line scan technique, palmprint-capturing devices could save much space for a comfortable user interface and a flexible structure without sacrifice in image quality and system performance. In a line scan sensor (also called the linear image sensor, the one dimensional image sensor), pixels array in one line, which is different from area sensors. Line scan sensors are of high dynamic range and high resolution with the same cost compared with area image sensors.

In this study, a novel line scan sensor, a Contact Image Sensor (CIS) module in specific, based palmprint recognition system was proposed. This system was simpler and smaller than area sensor based solutions in fundamental imaging structure. It overcame the dimension problem and the noise sensitive optical path problem, which were caused by area imaging structure. To solve the motion feedback problem in the proposed CIS based palmprint system, we designed a novel synchronizing unit. With this unit, linear CIS module could capture a image when a hand rolls the rollers. In this chapter, the proposed system including image-capturing device, extraction of region of interest (ROI), feature extraction is presented. The verification performance of the proposed system was comparable with current area based ones, which was supported by the experiment results in a database of 8000 images.

Some operational definitions used in this chapter are listed below.

Sensor: A sensor means a photoelectric sensor chip, for example, a CCD image sensor chip or a CMOS sensor chip.

Device: A device refers to the full set of image capturing hardware which consists of lightings, lens, a camera, a frame grabber and the supportive structure of the device.

System: A system means all the hardware and software, especially the methods and algorithms in all palmprint recognition procedures.

This chapter is organized as follows. Section 11.2 reviews existing systems. In Sect. 11.3, the details of the structure and each component of our line scan palmprint recognition system are presented. In Sect. 11.4, the system performance is tested by a verification experiment. Comparisons between our system and current area-based systems are also presented here. Section 11.5 concludes the entire chapter.

11.2 Existing System

Palmprint systems could be divided into three categories: Flatbed scanners, Palmprint systems with a digital camera or a web camera and Palmprint systems with a pegged flat platen surface.

11.2.1 Flatbed Scanner

A flatbed scanner was an off-line palmprint capturing device, which collects off-line palmprint images in high resolution. It was taken as a palmprint scanner by many researchers (Chaudhary and Nath 2009; Han et al. 2007a, b; Kumar et al. 2003; Goh et al. 2008, 2010; Zhu and Zhang 2010). The scanning time of a flatbed scanner was typically 10–20 s for a 300 dpi A4 size scan.

11.2.2 Web Camera Based Systems

Web cameras are fast, flexible, and very compact. They are suitable for real-time video surveillance applications.

The palmprint authentication systems based on web cameras were examined by plenty of researchers, mainly Han (Han et al. 2007a, b), Connie's group (Goh et al. 2008) and Zhu's group (Zhu and Zhang 2010). They made great contributions in real-time palmprint tracking and the optimization of lightings, as shown in Fig. 11.1.

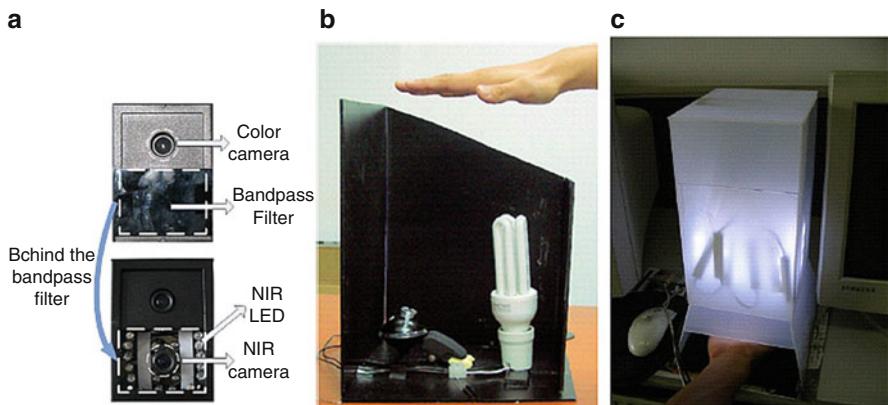


Fig. 11.1 Web camera based palmprint capturing devices. (a) Han's double camera design (Han et al. 2007a, b), (b) Connie's environment retrained design (Goh et al. 2008). (c) Zhu's close design (Zhu and Zhang 2010)

Han et al. (2007a, b) proposed a system based on two web cameras in order to track a palm by comparing the visible spectrum images and near-infrared images taken by the two cameras respectively.

Connie's group (Goh et al. 2008) proposed a web camera based system, in which a simple case was built to protect the system against environmental light. They used an effective palm tracking method in this system. This system used a web camera with a 640*480 resolution to capture hand images. After the hand being tracked in the images, the hand images were further down-sampled using two-dimensional (2D) wavelet transform. The ROIs extracted from hand images were normalized to a 150*150 resolution.

Zhu's group (Zhu and Zhang 2010) proposed another web camera based system in 2010, in which the light uniformity was carefully examined. Under the evenly distributed light, the hand image quality was improved, and the system equal error rate (EER) was 0.17%.

In web camera based systems, the problem of the calibration of hand pose variations in three-dimensional (3D) space is challenging (Goh et al. 2010).

11.2.3 *Palmprint Systems with Pegged Flat Platen Surface*

Palmprint systems with pegged flat platen surface (Wong et al. 2005; Zhang et al. 2003) have been popular in recent research. This design was good in hand stability, background noise, image quality, and recognition performance. The hand was held by the pegged flat platen surface in this kind of devices, as Fig. 11.2a, b show. The

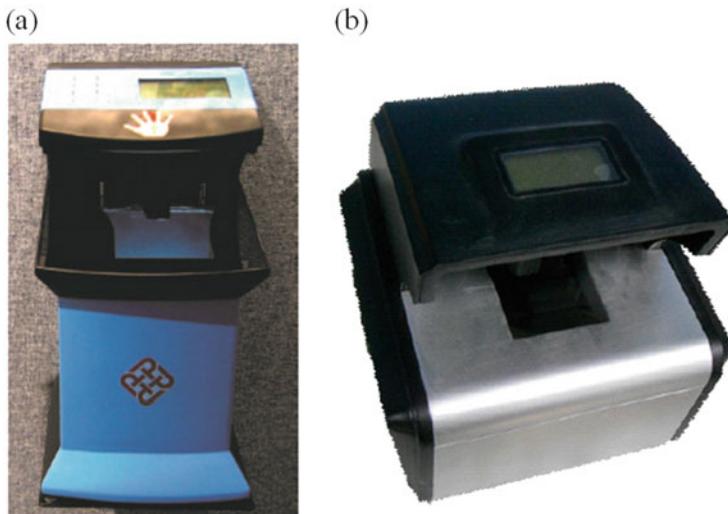


Fig. 11.2 (a) PC based area palmprint device, (b) Embedded area palmprint device. Typical palmprint devices with a pegged flat platen surface

user interface platen of Zhang's palmprint system covered most of the background area and the only left parts between fingers were covered by the upper cover. Currently the best results reported in research articles were mostly achieved in the database captured by Zhang's device, and the EERs was summarized to be distributed from 0.267% to 0.012% (Hao et al. 2008; Wang et al. 2008; Zhang et al. 2010a, b; Guo et al. 2009a, b; Laadjel et al. 2009a, b; Ito et al. 2006; Zuo et al. 2008), which would be generally better than other systems' EERs, which were outlined from 0.26% to over 4.73% (Goh et al. 2003, 2006, 2010; Connie et al. 2005; Han et al. 2003, 2007a; Lin et al. 2005; Struc and Pavesic 2008; Badrinath and Gupta 2007; Kumar et al. 2003).

This palmprint system used an area scan camera. The area scan camera is a camera with an area scan charge-coupled device (CCD) or complementary metal-oxide-semiconductor (CMOS) sensor. When an area scan camera captures images, all pixels of a frame are captured at the same time. Then electrons of pixels transfer line by line to the output channel and are amplified and sent to the output port pixel by pixel.

11.3 Line Scan Palmprint System Design

In order to take advantages of cutting-edge line scan image sensors, we proposed an online palmprint system based on the line scan image sensor. A synchronizing unit, which was composed of a roller module, an encoder, and an FPGA, was designed for our system. It synchronized the motion of hands with the capturing of images of the line scan image sensor. With this unit, the proposed system captured the palmprint adaptively to the motion of hands in real-time. Meanwhile, the resolution and quality of images were similar as current area systems. A control board was implemented based on an FPGA. The USB interface was utilized. With this universal interface, the proposed capturing device worked with either a desktop computer or an embedded platform, which was a major advantage for next generation real-time online applications. The rest of this section starts with the line scan imaging scheme, followed by the system framework. After the introduction of system framework, the three hardware parts: the line scan image sensor, the synchronizing unit, and the controller board are presented. Then the ROI extraction is reported. The system is proposed in the last.

11.3.1 Line Scan Imaging Scheme

In a line scan sensor, there is only one line of pixels. Pixels in a line scan sensor can be extended to enlarge the imaging area. With a larger imaging area, the dynamic range could be improved. This is different from the area sensors. In are a sensors, the size of pixels is a trade-off between resolution and dynamic range given a

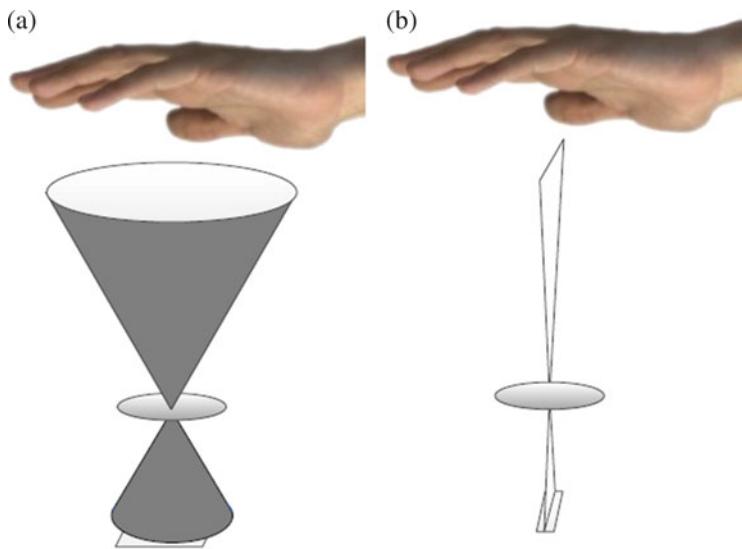


Fig. 11.3 (a) Area image sensor based palmprint capturing device imaging scheme. (b) Line scan image sensor based palmprint device imaging scheme. A comparison of optical path space between area scan image sensor based palmprint-capturing devices and line scan image sensor based devices

certain kind of pixels. In area cameras, a larger pixel could bring a higher dynamic range, but lower resolution. This problem does not exist in line scan sensors. Line scan image sensors could capture high-resolution images without compensating in dynamic range.

Furthermore, in an imaging system, line scan sensors require much less optical space than area image sensors, which is the most important limitation of current palmprint capturing devices. As Fig. 11.3b shows, a line scan sensor takes a line image. In a line scan sensor imaging system, the field of view is only one pixel length in the perpendicular direction. Comparing to the angle of view in the length direction, it is merely one of hundreds or one of thousands in the perpendicular direction. Therefore, in a line scan imaging system, the optical path, which is required by the traveling of rays from the object to the sensor, is like a thin plane in three-dimensional space. On the other hand, the optical path in a traditional area image sensor is two cone-shaped-space, as Fig. 11.3a shows. Furthermore, in an imaging system, this optical path is critical to the field of view, which is determined by the dimensions of objects. The field of view can be extended by increasing the angle of view. However, while the angle of view is increasing, the off-axis lights come into the optical system, and there will be bigger distortion aberration. Because of the reason above, the place taken by the line scan sensor optical path is much smaller than the space taken by the area image sensor, which cannot be further reduced in theory.

When building a system based on the line scan sensor, there is a challenge of motion synchronizing. A line scan image sensor captures images only with the help of a scanning unit who moves objects or the sensor along vertical direction. The benefit of this structure is that continuous images of objects could be captured and a long or large object could be captured by a thin or small sensor. However, an extra unit, which synchronizes the movement of the sensor and the object, is required.

This synchronization problem was solved in industrial machine vision systems. In industrial machine vision systems, line scan image sensors and cameras have been used for decades. A typical structure of an industrial line scan system consisted of a line scan camera, a conveyor belt and a motion controller (Kim et al. 2001; Baykal and Jullien 2004). Objects were moving smoothly on the conveyor belt when the line scan camera was taking line images. A motion controller monitored and adjusted the moving speed of the conveyor belt, and triggered the line scan camera to capture images synchronously.

However, this problem was different in building a lines can palmprint system, and it was still a challenge. First, in palmprint system, the imaging object is a part of human body. The palm cannot be moved as a physical subject. The synchronizing should be driven by the palm of people. Second, the motion of palm is not in a constant speed without any guide. Typically, the human palm moves in a way of various speeds and directions unless instructed or guided. Last, when building a palmprint system, the user interface should be user-friendly. The motion of palm should not be limited to a rigid way. In summary the motion synchronizing should be responsive, adaptive and user-friendly.

A line scan camera device in the fingerprint recognition was invented by Mil'shtein et al. (2008), which is different from the palmprint case. In the finger-print line scan device, the camera was rotating around the finger when capturing. The camera captured one vertical line of fingerprint image at a time, until 180° of the whole finger surface was scanned. This design captured a cylindrical fingerprint image. However, it was not suitable to be applied in palmprint. The skin surface of palm is a bumping plane, and the palmprint is not a cylinder. This structure could not be used in line scan palmprint systems.

Though a line scan sensor resembles a swipe fingerprint sensor, they are not the same. A swipe sensor captures a rectangle image, whose aspect ratio is as small as a bar is. The bar images are matched with each other to assembled the full fingerprint image. Although swipe sensors are very successful in fingerprint capturing, they are not suitable for palmprints. First, the palm is not completely flat. The center area of the palm is too far to be captured by the swipe sensor. The capacitive swipe sensor has to be touched and the ultrasonic wave based swipe sensor gets true skin layer fingerprint in less than 1 mm distance. When working with a longer working distance, the sensor chip generates tremendous heat. Second, the silicon chip is expensive. The palm is over four times larger as a fingerprint in width. The larger the chips' dimension, the lower conformance rate of chips. The swipe sensor customized for palmprint would not likely be economical.

11.3.2 System Framework

The framework of a general optical biometric system was composed of three parts: the image-capturing device, the computation platform, and the algorithms. The image-capturing device collects biometric samples. It is composed of a light source, lens, a camera or a sensor module, a frame grabber or an A/D converter. The computation platform is for the storage and data process purpose. A desktop computer is the most common platform. The algorithm is the way to extract ROIs and features and match encoded palmprints, which is used to identify or verify the subject.

Biometric applications deployed in hand held or mobile platforms are very popular (Shen et al. 2012; Han et al. 2007a, b; Jia et al. 2012). The embedded version of biometric systems is expected other than the desktop ones. A framework, in which both the desktop computation platforms and the embedded platforms were compatible with the same image-capturing device, was proposed for our lines can sensor based system. The framework is illustrated in Fig. 11.4. This device used an embedded controller, which controlled all the device parts and communicated with the computation platform through a universal interface. With this structure, the proposed palmprint capturing device could be used in both desktop applications and embedded applications with respective drivers.

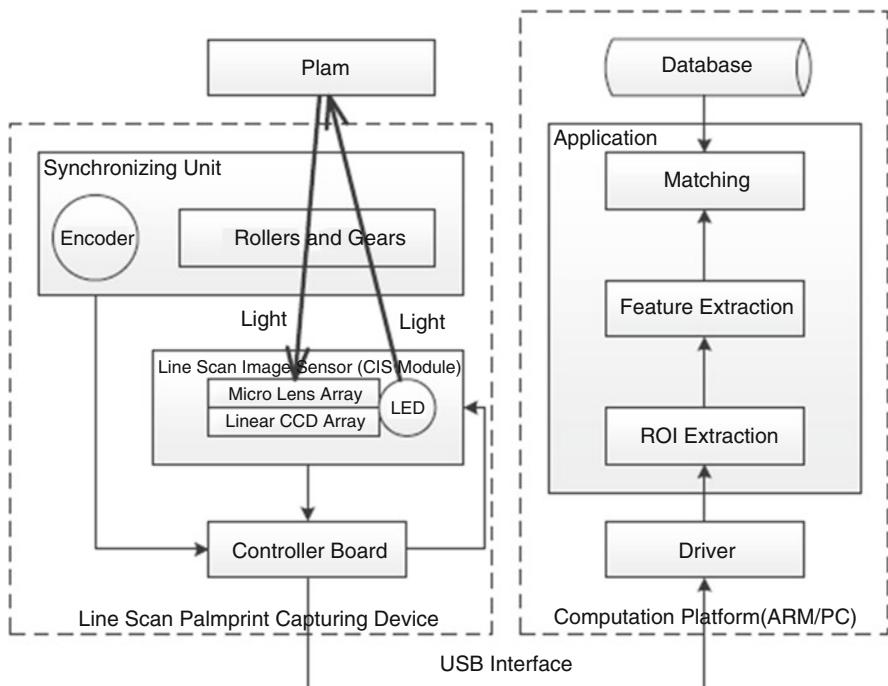


Fig. 11.4 The diagram of the system framework

In order to achieve the goal of the proposed system, to get high quality images for practical online applications, the image capturing speed, the image quality and the user interface design are the most important parts in palmprint recognition systems.

The data flow speed of the system is defined by the slowest part. Hence, each part should be optimized to maximum speed. First, the synchronizing unit should generate accurate synchronizing pulses according to the movement of the hand to minimize the lag. Second, the sensor should be fast enough for the real-time image capturing to minimize the data waiting time in buffer. The driver of the sensor should generate driver signals at a proper speed and accurate timing to minimize the lag. In addition, the speed of A/D converter should be faster than the output of the pixels. A buffer could help to cache the data during waiting for transferring through the interfaces.

To capture images with proper quality, the sensor, the optical system and the synchronizing unit should be selected with much care. The sensor resolution should be large enough to present the palmprint texture. Capturing images of proper resolution requires that the resolution of the image sensor and optical parts should be matched. The most challenging part is the synchronizing between the hand motion and the image capturing. The resolution of vertical and horizontal direction of the image should be matched accurately. The horizontal resolution of the system has to fulfill the palmprint feature requirements. According to Wong (Wong et al. 2005), the feature of a palmprint is composed of three parts, principal lines, wrinkles and ridges which lie from 75 dpi to 150 dpi. For online palmprint system, the resolution should be equal or over 75 dpi, which is analyzed by (Zhang et al. 2003) after the study of (Shu and Zhang 1998; Zhang and Shu 1999). In summary, 100 dpi was a good balance between image quality and speed for palmprint recognition system. The synchronizing unit is required to maintain a correct resolution aspect ratio between the horizontal resolution and the vertical resolution. In order to capture image lines in exactly same density as the horizontal resolution, which is a built-in parameter of lines can sensor, the synchronizing unit should generate accurate line image capturing signals. These signals keeps the same density of line images in vertical as the density of the pixels in the line scan sensor in the length direction, and it maintains the right aspect ration.

11.3.3 Line Scan Image Sensor-CIS Module

A line scan sensor is a kind of image sensor with a very simple structure (Chang et al. 2012; Fischer and Radil 2003; Luna et al. 2010; Marino et al. 2007; Watanabe and Hokari 2006). The photo diodes are simplest arranged in a linear array or a single line. Applications using line scan sensors are designed with either the camera or object moving in a direction perpendicular to the row of sensors. They are applied in the applications in which the object motion is under rigid control, for example, document scanning. They can be made by CCD or CMOS technology.

Table 11.1 Line scan image sensor (CIS module) characteristic

Items	Value
Scanning width	183 mm
Element density	100 dpi
Scanning speed	45 μ s/line
Light sources	Red: $\lambda_p = 630 \pm 15$ nm 60 mA Green: $\lambda_p = 520 \pm 15$ nm 60 mA Blue: $\lambda_p = 465 \pm 15$ nm 60 mA Infra red: $\lambda_p = 940 \pm 15$ nm 60 mA
Data output	Three analog parallel output channels

Instead of traditional line scan CCD or CMOS sensor chip, a new highly integrated CMOS line scan sensor module, which is called Contact Image Sensor (CIS), is designed for the proposed system. This CMOS based line scan sensor module was an integrated module including LED lights, micro lens and several CMOS line scan sensors. These parts could be integrated in one package, and all the control signals could be pinned out through one common connector. This module was customized by us and then was produced in a professional factory. The factory assured the performance by calibrating the lens, the sensor chip, and the LED lights in according to the specifications.

The CIS sensor was customized for high-speed line scan purpose. The basic scanning resolution was 200 dpi, and could be configured as 100 dpi or 50 dpi resolution. The major features are listed in Table 11.1.

Table 11.1 shows that, this sensor was 183 mm long, which is designed to capture the whole hand. The mediocre hand size is 189 mm for male and 172 mm for female (Average Hand Size 2012). The sensor size was set 4 mm larger than the average masculine hand, because the palm was the focus of the proposed system. In addition, the average hand width, which is also the width of the palm, is 84 mm for male and 74 mm for female. 183 mm was sufficient for the capturing of human palm both for male and female. It also covered a little more area of fingers, which was taken as a stable background in image processing later. This black background was necessary for the ROI extraction.

The resolution of this CIS sensor could be configured to 200 dpi, 100 dpi, and 50 dpi. The palmprint features are mostly in from 100 dpi to 150 dpi resolution according to (Wong et al. 2005). Considering the trade-off of the resolution and the computation cost, which affected the computation speed greatly, 100 dpi would be an optimal resolution for the proposed device. This 100 dpi preserved the majority of the palmprint features and saved lots of storing space and computation cost.

As is shown in Fig. 11.5, this CIS sensor scans at a speed of 45 μ s per line in 100 dpi mode. It captures 720 lines of image at a time to get the full hand image including the wrist part. The image capturing takes 33 dpi. This capturing speed ensures that the processing time was not delayed by the sensor. However, the entire processing time largely depended on the movement of hands.

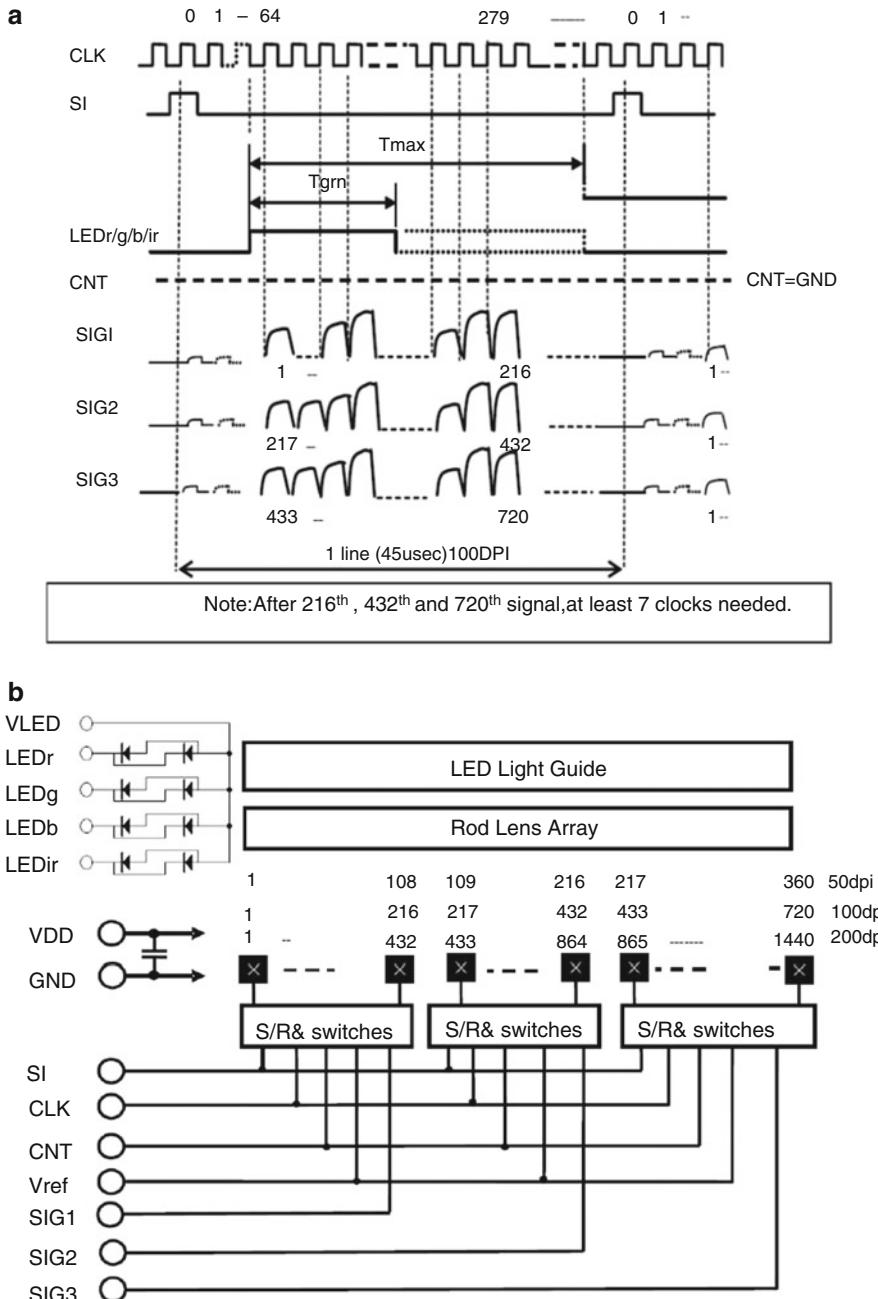


Fig. 11.5 (a) The timing graph of the CIS module. (CLK—clock timing, SI—starting pulse, LEDr/g/b/ir—LED integration switches, SIG1—output signal 1, SIG2—output signal 2, SIG3—output signal 3). (b) The controlling block diagram with three parallel output channels: SIG1, SIG2 and SIG3. CIS module timing diagram and block diagram

After capturing the image, this CIS sensor outputs the data through three parallel channels. The output structure of this CIS module is shown in Fig.11.5b. The CIS module outputs the first 1–216 pixels in SIG1 pin; then outputs 217–432 pixels in SIG2 pin; outputs the rest pixels in SIG3 pin. This structure maintained the fastest output speed, and it accelerated the output in a most convenient way. A typical A/D converter was working in three parallel channels, because the most typical application—video application requires three channels for BGR (Blue, Green and Red) signals. AD9822 is chosen as the A/D converter of the proposed device. This A/D converter was designed to work with BGR CCD sensor. There are three parallel input pins for three BGR channels. It converts three pixels in parallel within a period at 67 ns, which is over twice faster as the customized CIS module. The AD9822 A/D converter converts a line, which contains 720 pixels in three channels, at the speed of $45 \mu s$ per line, which is the speed of the CIS module. The delay of this A/D converting could be neglected compared with the entire reading time.

This CIS module equipped with a Rod Lens Array (RLA). The RLA is an array of cylindrical rod lens, which are highly polished small diameter rods made of optically transparent homogeneous materials. They are working as cylindrical lens for the linear array CMOS sensor pixel by pixel. The typical working distance of this kind of lens is about several millimeters. In the designed module, it was 3 mm ($-2/+4$ mm), which was set according to our synchronizing unit to ensure the identical image resolution along vertical and horizontal directions.

As shows in Table 11.1, there were four types of LEDs deployed in our CIS module, 630 nm for red color, 520 nm for green color, 465 nm for blue color and 940 nm for infrared spectrum. Typical palmprint features lay in the visible spectrum, from 380 to 780 nm. They make a hybrid of white color with the color temperature 6500 K. Recent research (Zhu and Zhang 2010) has discovered that multi-spectral palmprint images are better than traditional palmprint recognition using the visible spectrum only. In the proposed CIS module, near infrared LEDs were deployed with the same intensity as the white light. The combination of visible and infrared illumination was good for anti-spoofing purpose.

11.3.4 Synchronizing Unit

A synchronizing unit is designed to synchronize the motion of the hand with the CIS module capturing. There are three purposes of this unit. First, it sends synchronizing pulse signals to the CIS sensor. These signals should be distributed in 100 dpi density, which is the same as the horizontal resolution of the CIS module. Then, it holds the hand flattened and stretched. If the palm touches the surface of the CIS sensor, it squeezes and the lines and texture are distorted. At last, it keeps the hand skin 2–4 mm above the surface. The CIS module only captures objects in a distance of 1–7 mm. The depth of view of it is very limited.

The synchronizing unit consists of a pair of rollers, a gear set, and a photoelectric encoder, as is shown in Fig. 11.6. A pair of rollers holds the hand three millimeters above the surface of CIS module; meanwhile it stretches the palm skin to avoid the squeeze. First, this distance is optimal working distance of the CIS module. Second, when a hand is pressing on the rollers, the skin inside the field of view between two contact rollers is actually 1–2 mm lower than the top of rollers. Because it requires pressure to push the roller, the hand is pressing on the rollers with a moderate force. Though the force is meant to move the roller by the user, it stretched the skin of the palmprint, and kept it flattened. Third, the distance between the two rollers is minimum to avoid too much variation of the hand skin. It is the minimum distance to keep the CIS module under the rollers top surface. In addition, the two rollers are made of the same size to maintain a stable motion. They are also synchronized by the gear set. It means they are rolling at the same angular velocity. The hand motion on these rollers is both stable and synchronized.

The material of the rollers is copper. Three kinds of metal were tested in the making of these rollers. Either aluminum alloy or steel was not strong enough for them. The earlier samples were abraded by sweaty hands within a month. Eventually the copper rollers have worked for over half a year, and it was still shining brightly. Tiny ridges and valleys were carved in the cylindrical surface of rollers. They ensured that the frictional force was strong enough. In addition the rolling of the rollers was accurately driven by the hand motion.

Then the motion on the rollers is digitalized by the photo electric encoder to send out synchronizing pulses through the gear set, as illustrated in Fig. 11.7. The image resolution on the motion direction is defined by the rollers and the gear ratio. The photoelectric encoder here is industry standard 500 pulses per round. The resolution S_L is defined as Eq. (11.1).



Fig. 11.6 CIS module and the synchronizing unit composed of a pair of rollers, a set of gears and an optical encoder

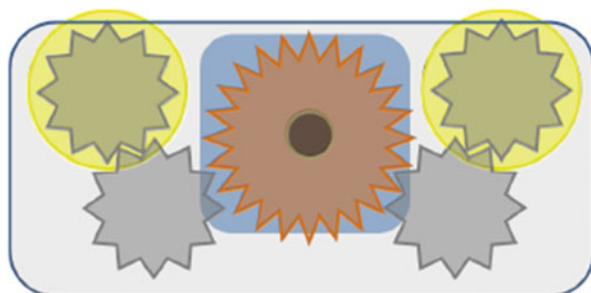


Fig. 11.7 The gear set keeps the two rollers synchronized and sends rotations to the encoder

$$S_L = \frac{25.4 \times 500}{\pi \cdot D_R \cdot R_g \cdot R_f} \quad (11.1)$$

Here D_R is the diameter of both rollers; R_g is the gear ratio; R_f is the filter ratio, which is set in the controller. In our prototype device, the rollers' diameter is 10 mm, the gear ratio is 2:1, and R_f is 2:1. Under this condition, the vertical (the rolling direction) resolution is $100 \text{ dpi} \pm 1:5 \text{ dpi}$, which is almost the same with the resolution along the width direction of the CIS module setting. The gears have to be large enough to transmit power through to the encoder. The gear ratio cannot be 4:1, in which case the output gear is too small to maintain a stable transmission. A pulse filter is set in the controller. It sends out one pulse when receiving two pulses (2: 1). With this method, the sampling rate is set to 100 dpi as expected.

11.3.5 Controller Board

An FPGA board was built as the controller. It comprised a CIS driver, an A/D controller, a data buffer and a USB interface. The block diagram of the controller is shown in Fig. 11.8a.

The FPGA is the core of this board, and it controls the CIS driver, the A/D controller, and the data buffer. It reads in the motion pulses, and sends out driving timing pulses to the CIS module. When the CIS module is ready for one line, the FPGA lets the AD9822 A/D converter start to read in analog pixel signals. Then the digitalized 8 bit image signals are read and stored in the SDRAM buffer. Finally, the data are sent out through the USB interface.

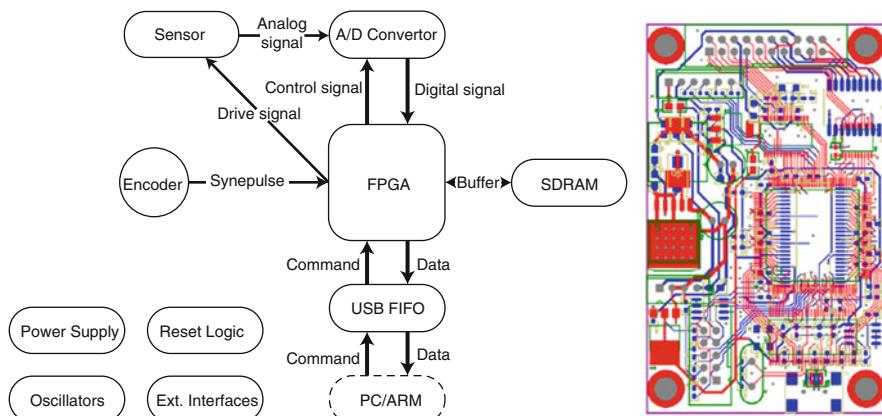


Fig. 11.8 (a) Scheme block diagram, **(b)** PCB layout, The controller board of line scan palmprint capturing device

11.3.6 The Device

Figure 11.9 shows the image-capturing device, which is composed of the line scan sensor, the synchronizing unit and the controller board. The proposed device could capture lines can palmprint images with either a desktop computer or an embedded ARM platform. The only two interfaces connected to it are a 12 V power supply and a USB mini B plug.

11.3.7 ROI Extraction

The images captured by the proposed device are shown in Fig. 11.10. The position of hands varies in the images indifferent captures. An automatic ROI extraction method was implemented. This ROI extraction method was an optimized method based on Guo's work (Guo et al. 2009a, b). Because the background of the line scan palmprint image was pure black, the line scan image was free from noise and environment light, which were commonly appeared in images captured by area sensor based devices. The optimization focused on the reduction of the noise removal part, which affected every step of the ROI extraction.

The extraction of ROI contains five steps: preprocessing, binarization, contour extraction, finding tangent points and computing ROI location.

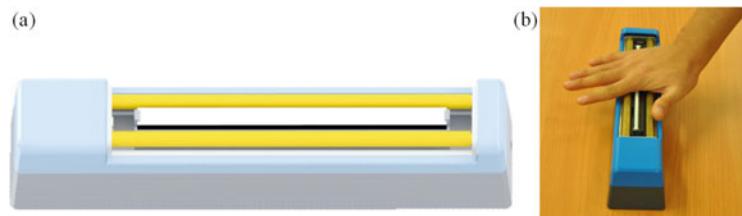


Fig. 11.9 (a) The 3D design model of the capturing device, (b) A right hand was testing on the device, The proposed image-capturing device

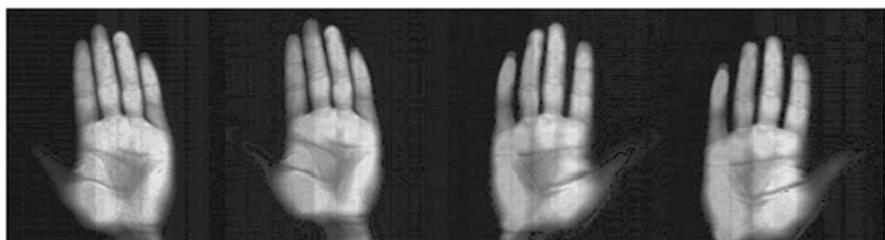


Fig. 11.10 The captured palmprint images

Preprocessing: In the preprocessing step, the raw image was normalized to remove the noise of the sensor chip. The grayscale value of the raw line scan image pixel was defined by both the characteristics of the time of the exposure and the pixel sensor. The noise of the same position of the same object between lines was caused by the time of the exposure. In order to remove the exposure time noises, standard white and dark images were captured as reference to normalize the line scan sensor. An exposure vector was extracted from 50 vertical lines to remove the exposure time noise. The images were normalized using this exposure vector by dividing each pixel of the image by it. The noise between pixels in one horizontal line was caused by the differences of the characteristics of the pixels, which was caused by dark current. Then dark current noises of the pixels were extracted from ten images of the standard black objects. As Fig. 11.11a shows, after subtracting the dark current noise from the image, images were normalized to 0–255 gray-scale. The preprocessing was implemented in the FPGA. The normalizing was implemented after capturing each line image.

Binarization: In the second step, fast Otsu method was used to binarize images with a dynamic threshold. Since the depth of view of proposed system is optimized to palm skin, there is no environmental interference in the background. As Fig. 11.11b shows, the background always appears as a black area, even though there is no back cover above the sensor surface. Comparing with area based images, extra morphological operations and blob detection can be saved.

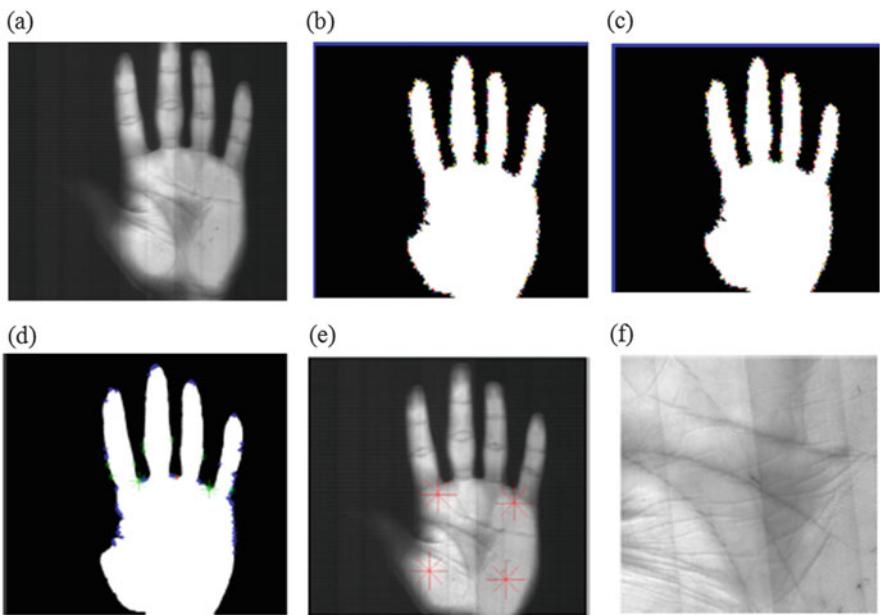


Fig. 11.11 (a) Palmprint sample example, (b) Binarization result, (c) Contour extracted, (d) Tangents points found, (e) ROI located, (f) Extracted ROI. Five-step ROI extraction

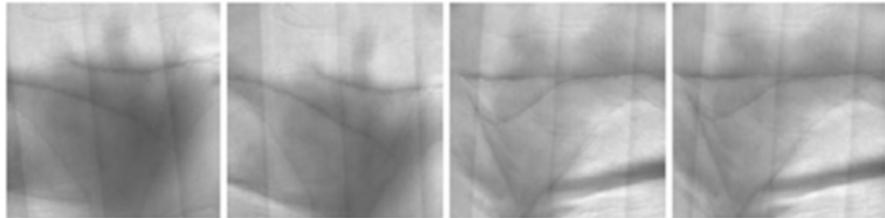


Fig. 11.12 Extracted palmprint ROIs

Finding Tangent Points: Here the tangent points are defined as the point of tangency where lowest tangent line meets the first and third valleys. The first valley is between the index finger, and the third valley is between ring finger and little finger. To find these two valleys, distances from points on the contour to the hand center is computed first. Local minima, which represent the valley points between fingers, are found as valley candidates. Second both clockwise and counter clockwise directions of hand contour are found to remove local minima of minor noise on hand shape and repeat recursively to find three valley points between fingers. Then the contour curve segments near the first and third valleys are extracted, which is depicted in Fig. 11.11d. The tangent line and two tangent points are found on these curve segments.

Computing ROI region: To locate the palmprint ROI region in the hand image, a coordinate system was set up. The line connecting the two tangent points were taken as y-axis. The middle point between them was taken as original point. Then in this coordinate system, the ROI region was obtained in a 128*128 pixel area between coordinates (50, 64), (178, 64), (50, -64), (178, -64), which is shown in Fig. 11.11e. The ROI was cropped, which is illustrated in Fig. 11.11f. There are some examples of extracted ROIs shown as Fig. 11.12.

11.4 Experiment and Comparison

In this section, the experiment of the proposed system on a large database is presented.

11.4.1 Line Scan Palmprint Database

A line scan palmprint sample database was built including 8000 line scan palmprint samples from 250 people, in order to evaluate the performance of proposed system. The subjects were volunteers from universities and neighboring communities. In the database, 189 people are male and the age distribution is from 20 to 63 years

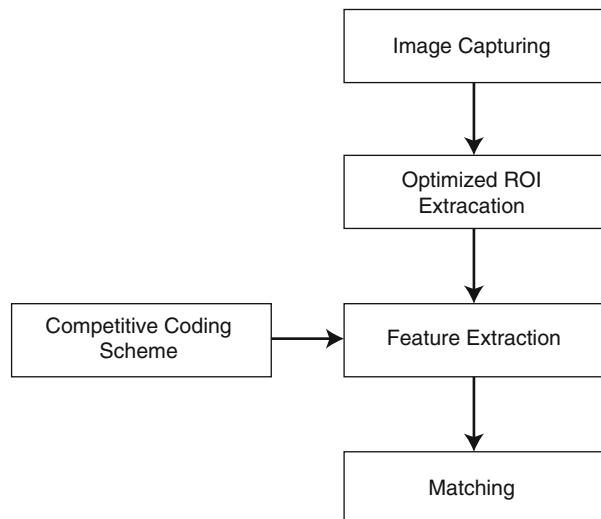
old. The samples from both left and right hand were collected in two separate sessions. Eight samples from each hand were collected in each session. The interval between the two sessions was 24.5 days in average. When collecting samples, the subject was asked to roll the rollers of the line scan palmprint-capturing device. A subject was trained for less than 30 s before the sampling. In this short-time training, a helper demonstrated the sampling process for two to three times, and then each subject tried for three to eight times on our line scan palmprint device, before the sampling in each session. Our database contains 8000 samples from 500 different palms. The resolution of the samples is 720*720 (100 dpi).

11.4.2 Verification Experiment

To compare with current real-time online area based palmprint systems (Wong et al. 2005; Zhu and Zhang 2010; Zhang et al. 2003), a verification experiment was built under the environment similar to area systems.

Figure 11.13 presents the design of the experiment. The experiment was designed following the area palmprint systems' convention. This experiment consisted of four parts: image capturing, ROI extraction, feature extraction, and matching. Image capturing and the ROI extraction are discussed above. In feature extraction part, the competitive coding scheme was implemented to extract feature and to encode the feature of palmprint ROIs. The competitive coding scheme (CompCode) (Kong and Zhang 2004) was a very effective 2D Gabor based palmprint feature extraction method. It extracted the orientation information of

Fig. 11.13 Line scan palmprint system verification experiment flowchart



palm lines. This method was widely used in area based palmprint systems. This method was the most promising method in real-time online palmprint feature extraction study.

In the matching part, the Humming distance was adopted following the best practice of area-based convention (Zhang et al. 2010a, b), which was proved to be efficient and effective in palmprint feature matching. With this design of experiment, the line scan palmprint system was tuned to work in the identical setting as area sensor-based ones. The verification accuracy is computed in the following tests, each palmprint sample is matched with all the other palmprint samples in the database. A match is counted as a genuine if the two samples are from the same palm; otherwise, it counted as an impostor. The total number of matchings is 31,996,000 and the number of genuine matchings is 60,000. The EER, which is the point when false accept rate (FAR) is equal to false reject rate (FRR), is used to evaluate the accuracy.

As is shown in Table 11.2, the EER is 0.048% when the FAR equals with FRR. The proposed system achieved a verification performance comparable with area sensor-based designs without compromising in other aspects. Though the performance (EER) was not better than the best result in all area counterparts, it was still comparable. This frame work and structure showed its capability as a promising palmprint application.

11.4.3 Comparisons with Current Palmprint Systems

A detailed comparison with area sensor-based palmprint system is shown in Table 11.2. The outer shape of line scan palmprint system is only $22 \times 5 \times 5 \text{ cm}^3$. Compared with current area image sensor based palmprint devices, the proposed system was much smaller. The size of the proposed system was less than 6% in size as area based ones. The size of area sensor-based systems could not be reduced further in area-based design. The area-based design could not reduce the outer shape smaller than $16 \times 16 \times 20 \text{ cm}^3$ in theory considering the space taken by the optical path and the back cover. In contrast, the outer shape of line scan palmprint systems could be further reduced greatly after moderate improvement.

In addition, the line scan system improved the speed performance of the palmprint sampling process. The line scan palmprint system works faster than area ones when capturing images. The more important is that the line scan palmprint system captures samples according the user's movement. The speed of capturing was self-adaptive to the user.

With all above improvement, the line scan palmprint system didn't compromise in verification performance. The line scan palmprint system achieved a verification performance comparable with most advanced area sensor-based counterparts.

Table 11.2 Comparison between the line scan palmprint system and area scan systems

Items	Line scan palmprint system	System (Wong et al. 2005)	System (Zhang et al. 2003)	System (Zhang et al. 2010a, b)
Type	Line scan	Pegged platen	Pegged platen	Pegged platen
		Surface	Surface	Surface
Hand pose	Stretched and roll	Stretched	Stretched	Stretched
Sensor	CIS module	CCD camera	CCD camera	CCD camera
Image resolution	50 dpi/100 dpi/ 200 dpi	75 dpi/100 dpi/ 125 dpi/150 dpi	75 dpi	<100 dpi
Illumination	Multispectrum LED (Fused)	White Fluorescent	White Fluorescent	Multispectrum LED (Alternative)
Lenses	Rod lens array (Integrated)	Industrial Std C mount	Industrial Std C mount	Industrial Std C mount
A/D converter	Three parallel Channels 8 bit	PCI 8bit	PCI 8bit	PCI 8bit
Cold start time	40 ms	90 ms	90 ms	90 ms
Image-capturing time	20–500 ms	40 ms	40 ms	40 ms
	Self-adaptive			
ROI extraction time	90 ms	–	538 ms	138 ms
Feature extraction time	36 ms	–	84 ms	36 ms
Matching time	0:06 ms	–	1:7 ms	0:06 ms
Batch matching time (s/samples)	1:5/1000	1:5/400	1:1/100	1:5/1000
Database (Population)	250	235	193	250
Database (Number of samples)	8000	9400	7752	6000
Verification Performance (EER)	0:048%	–	0:6%	0:021 ~ 0:052%
Dimensions (cm × cm × cm)	22 × 5 × 5	32 × 16 × 19	32 × 16 × 19	34 × 28 × 26
Cost without PC and case	\$160	–	–	~\$220
Platforms supported	Desktop and embedded	Desktop	Desktop	Desktop
	Windows and Linux	Windows	Windows	Windows

11.5 Summary

A novel line scan sensor based palmprint recognition system is proposed to improve the palmprint biometrics. An online line scan image sensor based palmprint system is made. This line scan palmprint is featured of a customized highly integrated line scan sensor, a self-adaptive motion feedback, and a cross-platform control board. The size of the proposed system is less than 6% of the size of current palmprint systems, without compromising in verification performance. The verification performance is tested in a database of 8000 samples from 250 people and the EER is 0.048%. This system is proved to be suitable for online palmprint biometric applications. The future research of line scan palmprint system could be in three directions. First, using the same structure provided by the proposed system, an even smaller system could be made with further study on optimization of the mechanical structure of the current design. Second, the materials and the structure of rollers should be studied. Self-cleaning materials should be used to keep the surface clean. More durable materials of the roller can contribute in reducing the size of the rollers. Third, the motion feedback can be integrated into the line scan module. With moderate improvements on the synchronizing unit, the ideal size of the proposed system in mass production should be less than to $15 \times 2 \times 2 \text{ cm}^3$ in theory. We hope that, using the proposed system, the line scan palmprint biometrics could be used in many space critical and portable situations, where previous area sensor-based palmprint systems cannot.

References

- Average Hand Size (2012) Your resource for hand size, palm size and finger length averages [Online]. Available http://www.theaveragebody.com/averagehand_size.php
- Badrinath G, Gupta P (2007) An efficient multi-algorithmic fusion system based on palmprint for personnel identification. In: Proceedings of 15th international conference advanced computing and communications, Washington, DC, USA, pp 759–764
- Baykal I, Jullien G (2004) Self-synchronization of time delay and integration cameras. *J Electron Imaging* 13(4):680–687
- Biometric Applications (2011) [Online]. Available <http://www.findbiometrics.com/>
- Chang J, Cheng K, Hsieh C, Chang W, Tsai H, Chiu C (2012) Linear CMOS image sensor with time-delay integration and interlaced super-resolution pixel. *IEEE Sensors J*:1–4. doi:[10.1109/ICSENS.2012.6411292](https://doi.org/10.1109/ICSENS.2012.6411292)
- Chaudhary S, Nath R (2009) A multimodal biometric recognition system based on fusion of palmprint, fingerprint and face. In: Proceedings of 2009 international conference advances in recent technologies in communication and computing, Washington, DC, USA, pp 596–600
- Connie T, Teoh A, Goh M, Ngo D (2005) An automated palmprint recognition system. *Image Vis Comput* 23(5):501–515
- Dai J, Zhou J (2011) Multifeature-based high-resolution palmprint recognition. *IEEE Trans Pattern Anal Mach Intell* 33(5):945–957
- Fischer J, Radil T (2003) DSP based measuring line scan CCD camera. In: Proceedings of 2nd IEEE international workshop intelligent data acquisition and advanced computing systems: technology and applications, pp 345–348

- Goh M, Connie T, Teoh A, Ngo D (2003) A single sensor hand geometry and palmprint verification system. Proceedings of 2003 ACM SIGMM workshop biometrics methods and applications (WBMA), pp 100–106
- Goh M, Connie T, Teoh A, Ngo D (2006) A fast palm print verification system. In: Proceedings of international conference computer graphics, imaging and visualisation (CGIV), Washington, DC, USA, pp 168–172
- Goh M, Connie T, Teoh A (2008) Touch-less palm print biometrics: novel design and implementation. *Image Vis Comput* 26(12):1551–1560
- Goh M, Connie T, Teoh A (2010) Design and implementation of a contactless palm print and palm vein sensor. In: Proceedings of 11th international conference control automation robotics vision (ICARCV), pp 1268–1273
- Guo Z, Zhang D, Zhang L, Zuo W (2009a) Palmprint verification using binary orientation co-occurrence vector. *Pattern Recogn Lett* 30(13):1219–1227
- Guo Z, Zuo W, Zhang L, Zhang D (2009b) Palmprint verification using consistent orientation coding. In: Proceedings of 16th IEEE international conference image process (ICIP), pp 1965–1968
- Han C, Cheng H, Lin C, Fan K (2003) Personal authentication using palm-print features. *Pattern Recogn* 36(2):371–381
- Han Y, Sun Z, Wang F, Tan T (2007a) Palmprint recognition under unconstrained scenes. In: Proceedings 8th Asian conference on computer vision, Berlin, Springer, pp 1–11
- Han Y, Tan T, Sun Z, Hao Y (2007b) Embedded palmprint recognition system on mobile devices, in advances in biometrics. *Lect Notes Comput Sci* 4642:1184–1193
- Hao Y, Sun Z, Tan T, and Ren C (2008) Multispectral palm image fusion for accurate contact-free palmprint recognition. In: 15th IEEE international conference on image processing, ICIP 2008, pp 281–284
- Ito K, Aoki T, Nakajima H, Kobayashi K, Higuchi T (2006) A palmprint recognition algorithm using phase-based image matching. In: Proceedings of IEEE International conference image processing (ICIP), pp 2669–2672
- Jain A, Feng J (2009) Latent palmprint matching. *IEEE Trans Pattern Anal Mach Intell* 31 (6):1032–1047
- Jia W, Hu R, Gui J, Zhao Y, Ren X (2012) Palmprint recognition across different devices. *Sensors* 12(6):7938–7964
- Kim J, Ahn S, Jeon J, and Byun J (2001) A high-speed high resolution vision system for the inspection of TFT LCD. In: Proceedings of IEEE international symposium on industrial electronics, vol 1, pp 101–105
- Kong A, Zhang D (2002) Palmprint texture analysis based on low resolution images for personal authentication. In: Proceedings of 16th international conference on pattern recognition (ICPR), pp 807–810
- Kong A, Zhang D (2004) Competitive coding scheme for palmprint verification. In: Proceedings of 17th international conference on pattern recognition (ICPR), vol 1, pp 520–523
- Kong W, Zhang D, Li W (2003) Palmprint feature extraction using 2-D Gabor filters. *Pattern Recogn* 36(10):2339–2347
- Kong A, Zhang D, Kamel M (2006) Analysis of brute-force break-ins of a palmprint authentication system. *IEEE Trans Syst Man Cybern B Cybern* 36(5):1201–1205
- Kong A, Zhang D, Kamel M (2009) A survey of palmprint recognition. *Pattern Recogn* 42(7):1408–1418
- Kumar A, Wong D, Shen H, Jain A (2003) Personal verification using palmprint and hand geometry biometric. In: Proceedings of 4th international conference audio- and video-based biometric person authentication (AVBPA), Berlin, Springer, pp 668–678
- Laadhel M, Kurugollu F, Bouridane A, Boussakta S (2009a) Degrade dpartial palmprint recognition for forensic investigations. In: Proceedings of 16th IEEE international conference image processing, Piscataway, NJ, USA, pp 1497–1500

- Laadjel M, Kurugollu F, Bouridane A, Yan W (2009b) Palmprint recognition based on subspace analysis of Gabor filter bank. In: Proceedings of 10th Pacific Rim conference on multimedia: advances in multimedia information processing (PCM). Berlin, pp 719–730
- Li W, Zhang D (2009) Three dimensional palmprint recognition. In: Proceedings of IEEE international conference on systems man and cybernetics, pp 4847–4852
- Lin C, Chuang T, Fan K (2005) Palmprint verification using hierarchical decomposition. *Pattern Recogn* 38(12):2639–2652
- Luna C, Mazo M, Lazaro J, Vazquez J (2010) Calibration of line scan cameras. *IEEE Trans Instrum Meas* 59(8):2185–2190
- Marino F, Distante A, Mazzeo P, Stella E (2007) A real-time visual inspection system for railway maintenance: automatic hexagonal-headed bolts detection. *IEEE Trans Syst Man Cybern C Appl Rev* 37(3):418–428
- Mil'shtein S, Palma J, Liessner C, Baier M, Pillai A, Shendye A (2008) Line scanner for biometric applications. In: Proceedings of IEEE international conference on technologies for homeland security (HST), pp 205–208
- Shen L, Wu S, Zheng S, Ji Z (2012) Embedded palmprint recognition system using OMAP 3530. *Sensors* 12(2):1482–1493
- Shu W, Zhang D (1998) Automated personal identification by palmprint. *Opt Eng* 37(8):2359–2362
- Struc V, Pavescic N (2008) A palmprint verification system based on phase congruency features, in biometrics and identity management. *Lect Notes Comput Sci* 5372:110–119
- Wang J, Yau W, Suwandy A, Sung E (2008) Person recognition by fusing palmprint and palm vein images based on Laplacianpalm representation. *Pattern Recogn* 41(5):1514–1527
- Watanabe K, Hokari M (2006) Measurement of 3-D loci and attitudes of the golf driver head while swinging. *IEEE Trans Syst Man Cybern A Syst Humans* 36(6):1161–1169
- Wong M, Zhang D, Kong W, Lu G (2005) Real-time palmprint acquisition system design. *IEEE Proc Vis Image Signal Process* 152(5):527–534
- Zhang D, Shu W (1999) Two novel characteristics in palmprint verification: datum point invariance and line feature matching. *Pattern Recogn* 32(4):691–702
- Zhang D, Kong W, You J, Wong M (2003) Online palmprint identification. *IEEE Trans Pattern Anal Mach Intell* 25(9):1041–1050
- Zhang D, Lu G, Li W, Zhang L, Luo N (2009) Palmprint recognition using 3-D information. *IEEE Trans Syst Man Cybern C Appl Rev* 39(5):505–519
- Zhang D, Guo Z, Lu G, Zuo W (2010a) An online system of multispectral palmprint verification. *IEEE Trans Instrum Meas* 59(2):480–490
- Zhang D, Kanhangad V, Luo N, Kumar A (2010b) Robust palmprint verification using 2D and 3D features. *Pattern Recogn* 43(1):358–368
- Zheng Y, Shi G, Zhang L, Wang Q, Zhao Y (2007) Research on offline palmprint image enhancement. In: Proceedings of IEEE international conference image processing (ICIP), pp I-541–I-544
- Zhu L, Zhang S (2010) Multimodal biometric identification system based on finger geometry, knuckle print and palm print. *Pattern Recogn Lett* 31(12):1641–1649
- Zuo W, Yue F, Wang K, and Zhang D (2008) Multiscale competitive code for efficient palmprint recognition. In: Proceedings of 19th international conference on pattern recognition (ICPR), pp 1–4

Chapter 12

Door Knob Hand Recognition System

Abstract Biometric applications have been used globally in everyday life. However, conventional biometrics is created and optimized for high security scenarios. Being used in daily life by ordinary untrained people is a new challenge. Facing this challenge, designing a biometric system with prior constraints of ergonomics, we propose ergonomic biometrics design model, which attains the physiological factors, the psychological factors, and the conventional security characteristics. With this model, a novel hand based biometric system, door knob hand recognition system, is proposed. Door knob hand recognition system has the identical appearance of a conventional door knob, which is an optimum solution in both physiological factors and psychological factors. In this system, a hand image is captured by door knob imaging scheme, which is a tailored omni-vision imaging structure and is optimized for this predetermined door knob appearance. Then features are extracted by local Gabor binary pattern histogram sequence method and classified by projective dictionary pair learning. In the experiment on a large data set including 12,000 images from 200 people, the proposed system achieves competitive recognition performance comparing with conventional biometrics like face and fingerprint recognition systems, with an equal error rate of 0.091%. This study shows that a biometric system could be built with a reliable recognition performance under the ergonomic constraints.

Keywords Biometrics • Ergonomics • Feature extraction • Image processing • Machine learning • Optical imaging • Pattern recognition • User centered design

12.1 Introduction

In the last decade, biometrics has expended dramatically and globally. Biometrics came under the spotlight after the counter-terrorism war began. It has become a major solution for identity recognition and authentication. Since 2006, America has been requiring biometric passports for travelers entering the United States under the visa waiver program ([Embassy 2009](#)). Not only in America, biometric technologies

have been applied in identity documents in 15 countries (Eidan 2013; Jeng and Chen 2009; Levush 2014).

With the wide spread of biometrics, it has been used not only in border control, forensics and law enforcement agencies (Jain and Kumar 2012), but also in everyday life, such as, in smart cars (Padmapriya and Kalajames 2012) and smart homes (Carvalho and Rosa 2010). Applications like these have educated the general public about the convenience and high-security level of biometric systems. People have accepted biometric systems as a common, convenient, and secure access control solution. Therefore, though biometrics has been designed to be used in high-security applications, there has been a great demand of biometrics in everyday life.

However, biometrics has not been designed in a user-friendly way. In current biometrics design, ergonomics (human factors) has not been a priority target. In the biometric system design (Dunstone and Yager 2009), ergonomic study was limited. When designing a biometric system, only seven elements have been addressed: universality, distinctiveness, permanence, collectability, performance, acceptability, and circumvention (Jain and Kumar 2012; Jain et al. 2004). When testing a biometric system, only recognition related performance has been considered (Mansfield and Wayman 2002). Ergonomics, to some extent, has been neglected.

In this chapter, we present ergonomic biometrics design (EBD) model that considers ergonomics in all aspects of the design in Sect. 12.2. We propose a door knob hand recognition system (DKHRS) in Sect. 12.3, which is shaped like a standard door knob, but incorporates a customized imaging device, a robust feature extraction, and a discriminative classification method. When addressing the imaging problem of this new device in this space-limited and shape-confined case, we propose a simplified catadioptric imaging structure-door knob imaging (DKI) scheme. The DKI scheme captures the surrounding hand skin surface in one omni-vision image in a cost-efficient structure. In the proposed system, we employ a local Gabor binary pattern histogram sequence (LGBPHS) method, which extracts robust histograms of dense local feature from DKHRS images. The features are classified using the dictionaries learned by projective dictionary pair learning (DPL). Combining DKI scheme, LGBPHS method, and DPL method, we make the proposed DKHRS effective and efficient under the ergonomic constraints. The DKHRS has been used to collect an experimental data set of a significant scale. The experiment result on this data set is promising in Sect. 12.4. The conclusion and future work is summarized in Sect. 12.5.

12.2 Ergonomic Biometrics Design

12.2.1 *Development of Biometric Systems*

Biometric systems are identity authentication systems utilizing various biological and/or behavioral traits, including fingerprint (Vatsa et al. 2009), face (Medioni et al. 2009), hand/finger geometry, iris (Kalka et al. 2010; Burge and Bowyer 2013;

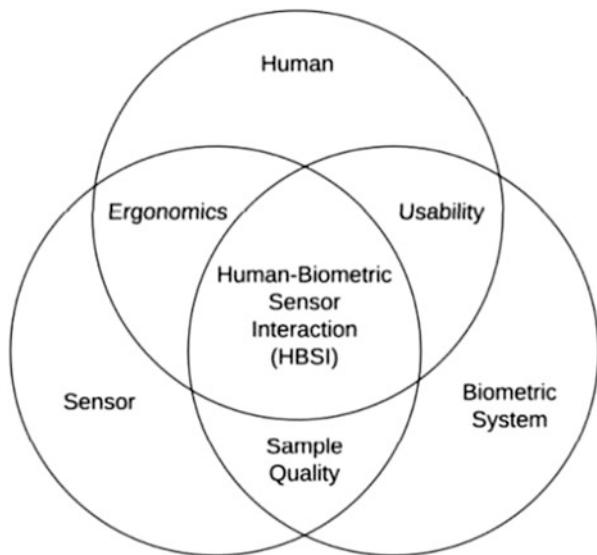
Gong et al. 2013), sclera (Zhou et al. 2012; Alkassar et al. 2015), signature, gait, palmprint (Zhang et al. 2003; Wu et al. 2006), voice pattern, ear (Bustard and Nixon 2010), hand vein, odor, and the DNA information of a person. A biometric system includes a sample collecting module (device), a feature extraction module, a database module, and a classification module (Jain et al. 2011). A biometric system provides verification and/or identification functions (Jain et al. 2011). The design cycle of a biometric system includes understanding the nature of the application, choosing biometrics trait, collecting biometric data, choosing features and classification algorithm and evaluating the system. During the development of a biometric system, seven factors: universality, distinctiveness, permanence, collectability, performance, acceptability, and circumvention are considered important (Jain and Kumar 2012; Jain et al. 2004, 2011). Another influential study summarizes that a biometric system should be assessed by attributes including distinctiveness, stability, scalability, usability, inclusiveness, insensitivity, vulnerability, privacy, maintenance, health, quality, integration and cost (Dunstone and Yager 2009).

12.2.2 Ergonomics Studies in Biometrics

From the above studies, it is noticed that ergonomics takes a significant role in developing biometric systems. Ergonomics (Human Factors) is to design products and systems in considering the interaction with people (Schlick 2009). In biometrics, physical ergonomics and cognitive ergonomics are two critical factors. Physical ergonomics focuses on physical motion related traits including human anatomical, anthropometric, physiological and biomechanical characteristics. Cognitive ergonomics focuses on human-system interaction related mental activities including perception, memory, reasoning, and motor response. In biometric systems, ergonomics presents various significant factors. The Collectability/Health might be categorized to physical ergonomics. The Acceptability/Usability might fall into the area of cognitive ergonomics. In biometric systems, the ideal ergonomic solution would make the user barely notice the authentication process (Dunstone and Yager 2009). Also, a biometric system with poor ergonomics would jeopardize the quality of collected biometric samples (Alonso-Fernandez et al. 2012). For example, iris systems require fixed height of eyes. When using these systems, tall or short people encounter frustrations (Dunstone and Yager 2009). Another example is that small fingerprint sensors without guides capture fingerprints of poor quality (Dunstone and Yager 2009).

Early ergonomics studies in development of biometric systems aimed at user acceptance (Albrecht 2001; Eschenburg et al. 2005; Elliott et al. 2007), latent fingerprint examination (Wertheim 2010; Expert 2012), and collectability (Jain and Kumar 2012; Dunstone and Yager 2009; Jain et al. 2011; Faddis et al. 2011; Mordini and Tzovaras 2012). According to Albrecht (Albrecht 2001), users accept natural and everyday motions the most readily, which conforms with ergonomic principles both in physical and cognitive. The analysis of latent fingerprint is widely

Fig. 12.1 HBSI model shows the interactions between human, biometric sensors, and biometric systems



depended on the human judgments because the analysis is semi-automatic (Wertheim 2010; Expert 2012). In collectability, the physical ergonomics is the challenge from an engineering perspective to next generation biometrics (Jain and Kumar 2012; Mordini and Tzovaras 2012). Then, addressing ergonomic issues in a systematic view, human-biometric sensor interaction (HBSI) model is proposed (Kukula et al. 2006, 2007).

HBSI model, which is illustrated in Fig. 12.1 focuses on the interactions between target subjects and the biometric sensors (Kukula et al. 2006, 2007; Kukula 2008). HBSI model utilizes the metrics from both biometrics and ergonomics to assess the functionality and performance of biometric systems. Human-sensor intersection focuses on the physical ergonomics. Human-biometric system intersection represents the interactions between users and systems, which include sensors, software and implementations of systems. The aim of this intersection is comprised of three factors: effectiveness, efficiency, and satisfaction. Sensor-biometric system intersection addresses the image/sample quality issue.

HBSI model has been used to examine the ergonomics of swipe fingerprint sensors (Kukula et al. 2006, 2010; Kukula 2008), and hand geometry machine (Kukula et al. 2007; Elliott et al. 2010). HBSI model provides an adequate and thorough evaluation framework for biometric systems. Under the assessment of HBSI model, the functionality and performance of a biometric system can be characterized. However, HBSI has several restrictions. First, HBSI model, to a large extent, is an evaluation model. It contributes significantly in assessing a variety of biometric systems, but not in crafting new biometrics. Second, HBSI model evaluates interactions in the sensor level. In HBSI evaluations, either fingerprint sensors or hand geometry machines are taken as the elementary subject.

The appearance and structure of the sensor constrained the enhancement of ergonomics in biometric systems. Third, in HBSI model, the ergonomics should be addressed at feature level. It is the biometric feature which defines the interactions among human, biometric sensors and biometric systems. The biometric feature determines the structure of sensors. For example, fingerprints are captured by semiconductor swipe fingerprint sensors or optical fingerprint sensors; hand geometries are obtained by hand geometry image capturing devices. In addition, the biometric feature defines a large part of interactions. For example, the iris system requires users to stand straight in front of the camera; the swipe fingerprint sensors require users to swipe with fingers. In order to enhance the functionality and the performance of biometric systems, ergonomics should be addressed at feature level.

12.2.3 *Ergonomic Biometrics Design Model*

EBD model is proposed to address the above limitations. EBD model implements four concepts.

1. Considering ergonomics in the first stage—selecting biological and behavioral characteristics.
2. Considering ergonomics in all developing stages—selecting biological and behavioral characteristics, designing the sample-collecting device and designing the feature extraction and classification method.
3. Considering both physical and cognitive ergonomics in each stage.
4. Collaborating the recognition performance with ergonomics.

EBD model is illustrated in Fig. 12.2. This model is to provide a guideline for creating a new ergonomic biometric system. In design theory, Rubin (Rubin and Chisnell 2008) argue that there are five reasons why a machine or a system is difficult to use. The fundamental issue is that the focus has been on the machine or

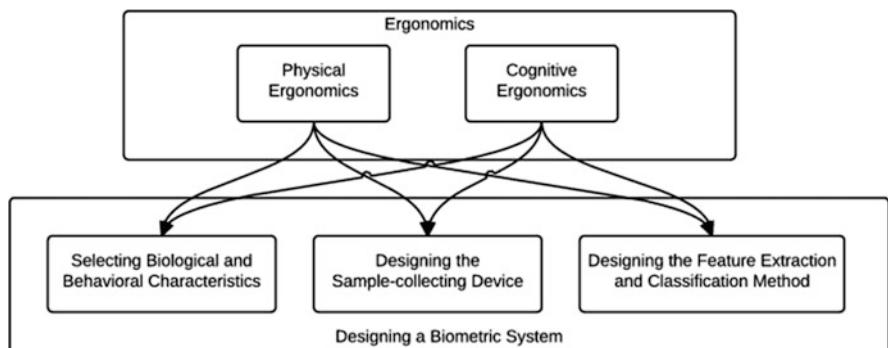


Fig. 12.2 EBD model considers both physical ergonomics and cognitive ergonomics in all three stages of biometric system development

the system and not on users during the development. In a biometric system, the emphasis should be users. The consideration of ergonomics should be as early as in the first design stage and also in the full design process. Conventionally, the ergonomics analysis is occurred in the implementation stage, which is after the production of sensors. When the feature and the sensor is determined, the room for ergonomics is limited. In EBD model, we insist that ergonomics should be included in the full biometrics system design process. In all three stages: selecting biological and behavioral characteristics, designing the sample-collecting device and designing the feature extraction and classification method, ergonomics should be considered. Furthermore, both categories of ergonomics should be considered including the physical ergonomics and the cognitive ergonomics. Though ergonomics is an essential element in EBD model, recognition performance should not be ignored. The ergonomics and recognition performance should collaborate with each other during the design process.

12.3 Door Knob Hand Recognition System

In this section, EBD model applies to the design of DKHRS.

12.3.1 Concept and Framework

The concept of DKHRS comes from the applications of biometric systems. The biometric systems have been widely employed in the access control scenario (Jain and Kumar 2012; Dunstone and Yager 2009; Burge and Bowyer 2013; Jain et al. 2008; Jain 2009). However, in the access control scenario, the most accepted device is the door knob, which can be traced back to 1893 (Russell 1893). Since then, the door knob has become a standard attachment to the door. Therefore, it occurs to us that how to remake the door knob by designing a biometric system with the appearance of the door knob. The appearance of the door knob has been optimized for human hands for over a hundred years (Russell 1893), which would be an ideal physical ergonomic design of a biometric device, because it is natural and it has become a daily motion (Albrecht 2001). With the appearance of the door knob, users can use DKHRS installed in the door without extra burden of thinking, which could achieve a better performance in cognitive ergonomics. The function and the form unite in DKHRS.

The goal of DKHRS is to be user-friendly. The design concept of DKHRS can be depicted in one sentence, open the door just like it is not locked. To fulfill this requirement, EBD model has been applied in the design process.

In the development of a biometric system using EBD model, both the physical and cognitive ergonomics are considered in the three designing stages: selecting biological and behavioral characteristics, designing the sample-collecting device

and designing the feature extraction and classification method, as illustrated in Fig. 12.2.

First, the unique and robust biometric characteristics of DKHRS are the skin texture of the hand. DKHRS imitates the door knob. When using a traditional door knob, people hold the door knob with their hands. The hand encloses the door knob. The hand skin has multiple layers of texture, which have been used in various biometric systems (Zhang et al. 2003; Kong and Zhang 2002, 2004; Wong et al. 2005). The texture of a hand could be a promising biometric characteristic.

Second, the image capturing device of DKHRS is to capture hand images with the appearance of a traditional door knob. The appearance of a traditional door knob is inherited from the historical design experience, which has been physiologically optimized for the human hand. Because of the long history of the door knob being used, this appearance is unconsciously perceived as an access control device. With the standard door knob appearance, DKHRS is identical to the traditional door knob in appearance and function. However, DKHRS is a biometric system and the recognition performance is a major factor. A good recognition performance requires good image quality and rich features. Thus, a novel DKI scheme is proposed. This imaging scheme is designed to capture quality images under the constraints of the appearance.

Third, the feature extraction and classification method of DKHRS should be fast and robust to dislocations. The recognition speed of a biometric system is critical to cognitive user experience. The dislocations of the hand are inescapable in physical device design. The dislocation problem can also be a challenging issue in cognitive reception if a user has to try multiple times for successful recognition. In feature extraction part, LGBPHS method (Zhang et al. 2005; Yan et al. 2011) is adopted to extract texture feature from the image of the hand. LGBPHS method extracts local binary patterns from patches in the Gabor surface of the original image and concatenates all histograms of these patterns to form a feature vector. LGBPHS is robust to illumination variations and minor dislocations (Zhang et al. 2005; Yan et al. 2011). However LGBPHS feature is of a very high dimension. Thus, an efficient and discriminative representation of this feature is required. In classification aspect, DPL (Gu et al. 2014) is adopted. DPL learns a synthesis dictionary and an analysis dictionary to represent the feature. This method converges fast in both training and testing stages and shows promising performance in face recognition, object recognition, and action recognition (Gu et al. 2014).

12.3.2 Imaging

The imaging structure is challenging to DKHRS. Considering the physical ergonomic factor, the dimensions of the imaging structure is constrained; the appearance of the imaging structure is predefined as a standard door knob; and the usage of the imaging structure is to be held with a hand. The imaging structure should be

small and compact, and the hand images around it should be captured. Therefore, DKHRS cannot use the most popular conventional imaging method.

The most popular used imaging method is the conventional imaging (Kong and Zhang 2002, 2004; Zhang et al. 2003; Wong et al. 2005; Jain et al. 2008; Dunstone and Yager 2009; Jain 2009; Jain and Kumar 2012; Burge and Bowyer 2013). This imaging scheme places a light source and a camera in front of the object. This structure is simple and stable. It is widely used in most of biometric systems (Kong and Zhang 2004; Wong et al. 2005; Jain et al. 2008; Dunstone and Yager 2009; Jain 2009; Jain and Kumar 2012; Burge and Bowyer 2013).

However, there are three limitations of the conventional imaging scheme when this scheme is being applied in DKHRS. First, conventional imaging requires a large space, which is illustrated in Fig. 12.3. The conventional imaging needs two clear cone-shaped spaces: one is between the camera and the lens, and the other is between the lens and the object. Second, in conventional imaging, the camera only captures objects in front of the lens. The field of view captured by this scheme is proportional to the distance between the lens and the object. Third, a large space

Fig. 12.3 Conventional imaging scheme requires a large space and only captures objects in front of the lens



and the open light path of the conventional imaging are very sensitive to the environment light. To address the above limitations and the requirements of the imaging, the catadioptric scheme is adopted.

The catadioptric scheme is an imaging structure, in which the ray is reflected once or several times to increase the distance and to enlarge the field of view (Jones 2009; Zhang and Li 2012 catadioptric). It is widely used in omnidirectional cameras. In omnidirectional cameras, the catadioptric scheme is used to get a round image of the environment (Zhang and Li 2012). The omnidirectional cameras being used in machine vision can be grouped into three categories: cameras that use special lens, cameras that use a convex mirror and a set of lens, and cameras that use two mirrors and a set of lens (Jones 2009). Most directional mirrors use these four types of shapes: conical mirrors, hemispherical mirrors, hyperboloidal mirrors, and paraboloidal mirrors (Jones 2009; Zhang and Li 2012). The mirrors and the minimum working distance of the systems restrict the minimum size of omnidirectional cameras (Zhang and Li 2012).

When the catadioptric scheme is being used in the proposed DKHRS, the advantage of this scheme is that the hand image, which is around the camera and the door knob device of DKHRS, can be captured in one image. However, there are still several limitations. First, the catadioptric scheme is designed to capture images far away from the camera. In the proposed DKHRS, the distance from the hand to the camera is constrained by the dimension of the door knob. Second, the directional mirrors of the catadioptric scheme, which are a crucial part of this imaging scheme, are expensive. The directional mirrors are of a complex curved shape and tend to be of expensive material. To overcome these limitations, we proposed the door knob imaging (DKI) scheme.

DKI is a simplified catadioptric scheme, which is customized for DKHRS. The simplifications are as follows. First, the imaging structure is simplified to capture the surrounding area near the case of the device only. DKI structure is different from the catadioptric scheme, in which the object or the environment is at a relatively long distance. In the proposed device, the hand would be holding the device and touching the case of the device. Second, the imaging structure is simplified to reduce the cost of the device, and to reduce the complexity. In the catadioptric scheme, the reflective mirror is a directional mirror, which reflects the surrounding environment in a linear proportion.

DKI scheme is shown in Fig. 12.4. This simplified scheme is composed of a door knob case, a flat reflective mirror, a lens of a large angle of view and a camera. The door knob case is transparent. It holds the hand and enables the capturing of the hand image. The door knob case is a little larger in the far from the door end than the near door end, which is identical with a door knob. The flat mirror reflects the image of the hand. The camera and the lens capture images of the reflected image of the surrounding through the transparent case. With DKI scheme, the device can capture the image of the hand holding the device. Furthermore, to stabilize the position of the hand, there are two holding pegs fixed on the lower surface of the door knob case. The image-capturing device of DKHRS, which is made following DKI scheme, is shown in Fig. 12.5.

Fig. 12.4 DKI scheme is simple and cost-effective capturing the surrounding hand texture in one image

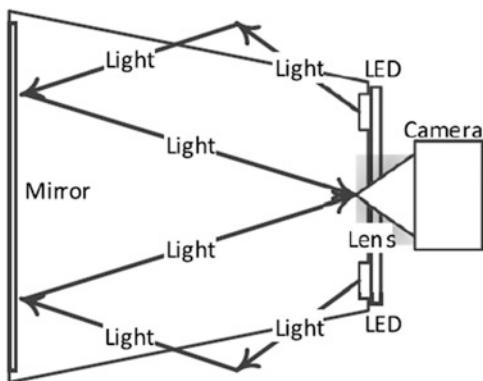
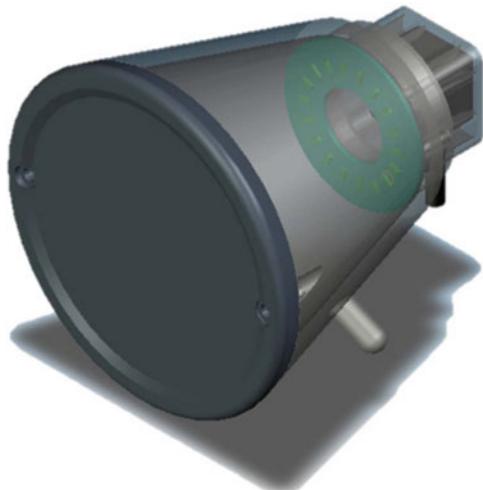


Fig. 12.5 Image-capturing device of DKHRS was made following the DKI scheme, which is small, compact and of a shape identical to a standard door knob



The device is made of these components listed below. The door knob is made of acrylic glass. The acrylic is a kind of highly transparent plastic transparent for over 90% energy and across a large spectrum including visible light and near infrared light and is also strong, lightweight and easy to process. The flat mirror reflects over 95% energy of light across the visible and infrared spectrum with a customized coating. The camera and the lens together capture the surrounding image in the mirror. The frame grabber digitizes analog images to digital images for further processing.

The components are as follows.

1. Acrylic transparent door knob.
2. Over 95% full spectrum reflective mirror.
3. White LED board.
4. 6 mm focal length pinhole lens.
5. 1/3 inch mini CCD camera.
6. A USB 2.0 frame grabber.

12.3.3 Feature Extraction and Classification

DKHRS captures hand images surrounding the door knob, as depicted in Fig. 12.6.

A biometric feature should be unique, stable, and persistent. DKHRS captures the hand image surrounding the door knob, as depicted in Fig. 12.6. The feature of a DKHRS image is the texture feature on the skin of the hand. Different from a palmprint image or a fingerprint image, a DKHRS image is an image of a hand in the holding gesture. Though this hand image is not taken with a flattened standard pose, the holding gesture of the hand still can be stable and reliable. Meanwhile, this hand image contains unique texture features. After preprocessing of the raw DKHRS image, the features of a DKHRS image can be extracted with LGBPHS method. After the extraction, the features are classified by DPL.

1. *Preprocessing:* In the preprocessing, the hand area of a DKHRS image, which is the region of interest (ROI), is cropped. The preprocessing is simple and straight forward after a manual calibration of the acquisition device. The calibration consists of two parts. First, the axes of the door knob and the camera are calibrated to the same, as shown in Fig. 12.7a. The calibration is performed at the pixel level with the help of images. Second, the parameters (center and radius) of the annular hand area are measured after the calibration. Because after the calibration, the round mirror and the camera image in the mirror shares the



Fig. 12.6 Images captured by DKHRS

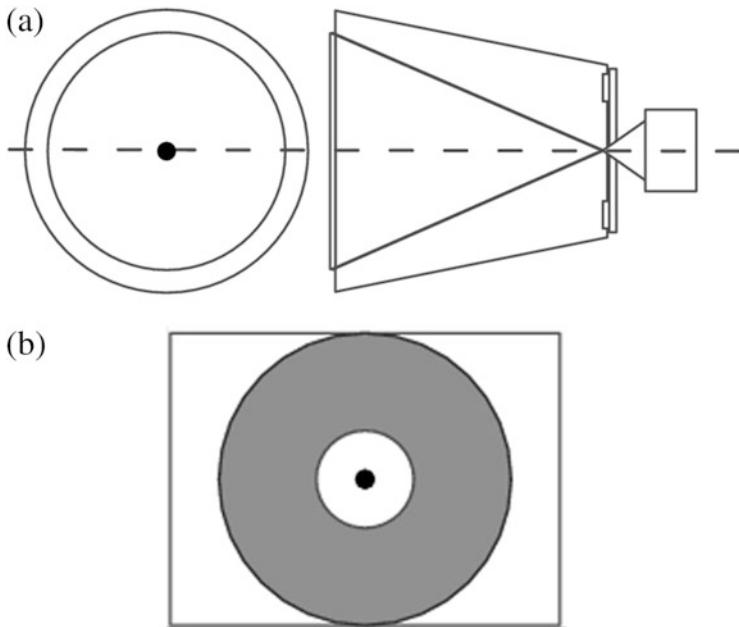


Fig. 12.7 DKHRS is calibrated before the extraction of the ROI, which is a wide ring area. **(a)** Optical axis of the LED light, the knob, the lens, the sensor and the center of the round mirror are calibrated to the same point. **(b)** ROI region of a DKHRS image is a ring area cropped by two circles with the same center, which is stable after the calibration

same center, the hand skin area is an annular area and can be simply cropped with two concentric circles, which is depicted in Fig. 12.7b.

2. Feature Extraction using Local Gabor Binary Pattern Histogram Sequence: In the proposed system, the features are extracted by LGBPHS (Zhang et al. 2005; Shan et al. 2006) method. LGBPHS method is a combination of local binary pattern (LBP) method and Gabor wavelet method. LBP was created for texture feature extraction (Ojala et al. 2002), then was applied in face feature extraction (Ahonen et al. 2004; Huang et al. 2004; Shen et al. 2008; Zhu et al. 2015; Guo et al. 2010a, b, c, d; Yan et al. 2011). Gabor wavelet method also achieved a promising performance in face recognition (Liu and Wechsler 2002; Shan et al. 2006). LGBPHS method combines these two methods. It is robust against rotation and translation, and against illumination interference.

In the proposed system, a patch-based LGBPHS method is employed, whose scheme is illustrated as Fig. 12.8b. First, 35 patches of the same size (64 by 64) are cropped from one DKHRS image, as shown in Fig. 12.8a. The center of these patches is located in the middle of the ring area, 130 pixels away from the center, to cover most hand skin. Second, each DKHRS patch is filtered by multi-orientation and multi-scale Gabor filters to get Gabor magnitude pictures (GMPs). The Gabor filters used are as follows:

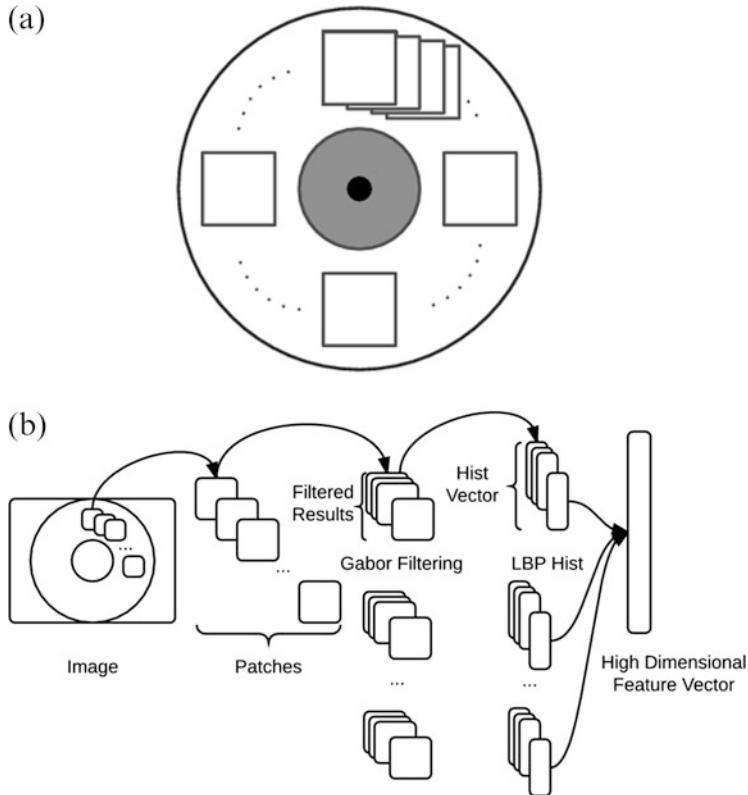


Fig. 12.8 Patch wise LGBPHS feature extraction method extracts a high dimensional vector from a DKHRS image. (a) Number of patches perpendicular to the radius direction are extracted for feature extraction. (b) LGBPHS feature extraction scheme comprises of four stages: localizing patches, filtering patches with Gabor filter bank, computing LBP histogram feature for each patch, concatenates all the histogram vector into one high-dimensional vector

$$\Psi_{\mu, v}(z) = \frac{\|k_{\mu, v}\|^2}{\sigma^2} e^{-\frac{\|k_{\mu, v}\|^2 \|z\|^2}{2\sigma^2}} \left[e^{ik_{\mu, v} z} - e^{-\frac{\sigma^2}{2}} \right] \quad (12.1)$$

where μ and v are the orientation and scale of the Gabor filters, $z(x, y)$, $\|\cdot\|$ denotes the norm operator, and the wave vector $k_{\mu, v} = k_v e^{i\phi_\mu}$, where $k_v = k_{max}/\lambda^v$ and $\phi_\mu = \pi\mu/8$. λ is the spacing factor between filters in the frequency domain. The patches images convolve with the Gabor filters. Given a patch $f(x, y)$, the convolution with a Gabor filter $\Psi_{\mu, v}(z)$ is defined as follows:

$$G_{\Psi f}(x, y, \mu, v) = f(x, y)^* \Psi_{\mu, v}(z) \quad (12.2)$$

where $*$ denotes the convolution operator. Five scales $v \in 0, \dots, 4$ and eight orientations $\mu \in 0, \dots, 7$ Gabor filters are used. Convolving each patch with

each of the 40 Gabor filters results in the Gabor feature. Here, one magnitude value is generated at each pixel position for each Gabor filter. After the filter, 40 GMPs are generated.

Third, for each patches, GMPs are encoded by LBP operator. LBP operator labels the pixels of an image by binarizing the 3×3 -neighborhood of each pixel $f_p (p = 0, 1, \dots, 7)$ with the center value f_c and considering the result as a binary number.

$$S(f_p - f_c) = \begin{cases} 1, & f_p \geq f_c \\ 0, & f_p < f_c \end{cases} \quad (12.3)$$

Then, by assigning a binomial factor 2^p for each $S(f_p - f_c)$, the LBP pattern at the pixel is achieved as

$$LBP = \sum_{p=0}^7 S(f_p - f_c) 2^p \quad (12.4)$$

Which characterizes the spatial structure of the local image texture. We denote the LBP transform result at position (x, y) of (μ, v) -GMP as $G_{lgbp}(x, y, \mu, v)$, which composes the (μ, v) -LGBP Map.

Fourth, the LBP codes are summarized to histograms. The local feature histogram summarizes the region property of the LGBP patterns.

At last, all the histograms from all LGBP maps are concatenated into a histogram sequence. The above process is formulated as follows: The histogram h of an image $f(x, y)$, with gray levels in the range $[0, L - 1]$ could be defined as

$$h_i = \sum_{x,y} I\{f(x, y) = i\}, \quad i = 0, 1, \dots, L - 1 \quad (12.5)$$

where i is the i th gray level, h_i is the number of pixels in the image with gray level i and

$$I\{A\} = \begin{cases} 1, & A \text{ is true} \\ 0, & A \text{ is false} \end{cases} \quad (12.6)$$

Assume each LGBP Map is divided into m regions R_0, R_1, \dots, R_{m-1} (in our case $m = 35$). The histogram of r th region of the specific LGBP Map (from (μ, v) -GMP) is computed by

$$H_{\mu, v, r} = (h_{\mu, v, r, 0}, h_{\mu, v, r, 1}, \dots, h_{\mu, v, r, L-1}) \quad (12.7)$$

where

$$H_{\mu, v, r, i} = \sum_{(x,y) \in R_r} I\{G_{lgbp}(x, y, \mu, v) = i\} \quad (12.8)$$

Finally, all the histogram pieces computed from the patches are concatenated to a high dimensional feature vector, which is the final feature vector to represent the given DKHSR image.

3. *Classification using Projective Dictionary Pair Learning:* In the proposed system, DPL (Gu et al. 2014) is used. DPL is dictionary learning method proposed in 2014, which is inspired by K-SVD (Aharon et al. 2006; Rubinstein et al. 2013) and LC-KSVD (Jiang et al. 2011, 2013). DPL learns both a synthesis dictionary and an analysis dictionary in the training stage. Both dictionaries are used in classification. DPL method is fast in both training and testing stages, and the classification accuracy is promising.

Denote by $X = [X_1, \dots, X_k, \dots, X_K]$ a set of p -dimensional training samples from K classes, where $X_k \in R^{p \times n}$ is the training sample set of class k , and n is the number of samples of each class.

DPL method learns a synthesis dictionary D and an analysis dictionary, denoted by $P \in R^{mK \times p}$ using the following objective function:

$$\begin{aligned} \{P, D\} = \operatorname{argmin}_{P, D} & \sum_{k=1}^K \|X_k - D_k P_k X_k\|_F^2 + \lambda \|P_k \bar{X}_k\|_F^2 \\ \text{s.t. } & \|d_i\|_2^2 \leq 1. \end{aligned} \quad (12.9)$$

It is different from conventional discriminative dictionary learning that the sparse code P , X is analytically obtained, instead of computing the time-consuming l_1 -norm sparse coding. With this feature, the DPL model is very efficient, and detailed description and reference can be found in (Gu et al. 2014).

In the optimization of the DPL model, a matrix A is introduced and the model is relaxed as following:

$$\begin{aligned} \{P^*, A^*, D^*\} = \operatorname{argmin}_{P, A, D} & \sum_{k=1}^K \left\{ \|X_k - D_k P_k X_k\|_F^2 \right. \\ & \left. + \tau \|P_k X_k - A_k\|_F^2 + \lambda \|P_k \bar{X}_k\|_F^2 \right\} \\ \text{s.t. } & \|d_i\|_2^2 \leq 1. \end{aligned} \quad (12.10)$$

where τ is a scalar constant. P and D are initialized as random matrices with unit Frobenius norm, then A , D , P are updated alternatively.

In the DPL model, the analysis sub-dictionary P_k^* is trained to produce small coefficients for samples from classes other than k , and it can only generate significant coding coefficients for samples from class k . Meanwhile, the synthesis sub-dictionary D_k^* is trained to reconstruct the samples of class k from their projective coefficients $P_k^* X_k$; that is, the residual $\|X_k - D_k^* P_k^* X_k\|_F^2$ will be small. On the other hand, since $P_k^* X_i, i \neq k$ will be small and D_k^* is not trained to

reconstruct X_i , the residual $\|X_i - D_k^* P_k^* X_i\|_F^2$ will be much larger than $\|X_i - D_k^* P_k^* X_k\|_F^2$.

Based on the property above, a matching scheme is designed as follows:

$$\text{matching}(y, i) = \begin{cases} 1, & \|D_i P_i y\|_2 \leq \text{threshold} \\ 0, & \|D_i P_i y\|_2 > \text{threshold} \end{cases} \quad (12.11)$$

The reconstruction residual is taken as a dissimilarity measure.

12.4 Experiments

The major concern of DKHRS, which is designed following EBD model, is the performance. In DKHRS, the hand is not in a standard gesture and the camera captures non-conventional and reflected image of the hand skin. Though LGBPHS and DPL methods are reported to be robust against distortions, it is still a myth whether the system could achieve a competitive performance. There is no doubt that the recognition performance is the critical criteria to a biometric system. Experiments have been devised to evaluate the performance of the proposed system. A large data set was built, and recognition experiments with different settings are presented.

12.4.1 Data Set

The DKHRS data set including 12,000 DKHRS images from 200 people was built, in order to evaluate the recognition performance of the proposed system. The subjects were volunteers from universities and neighboring communities. The images from both left and right hand were collected. There were two sessions of data collection. In each session, 15 images were collected for one hand. The interval between the two sessions was 1 week. When collecting images, the subject was asked to move away and hold the door knob again after taking each sample. Each subject was trained for less than 30 s before the collection. In this short-time training, a helper demonstrated the sampling process, then the subject tried once or twice with DKHRS device before the data collection. The resolution of the images is 640×480 . The prototype system used in the data collection is shown in Fig. 12.9. In the experiment, the data set is separated into a training set and a testing set, which contain equal number of images. For each hand, there are 15 randomly selected training images, and the other 15 images are used for testing. In the experiment, the classification is to evaluate the similarity between each sample and each class (hand). Thus, for one testing sample, there is one genuine match to

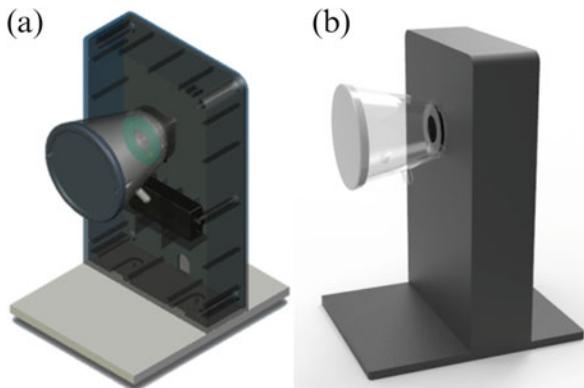


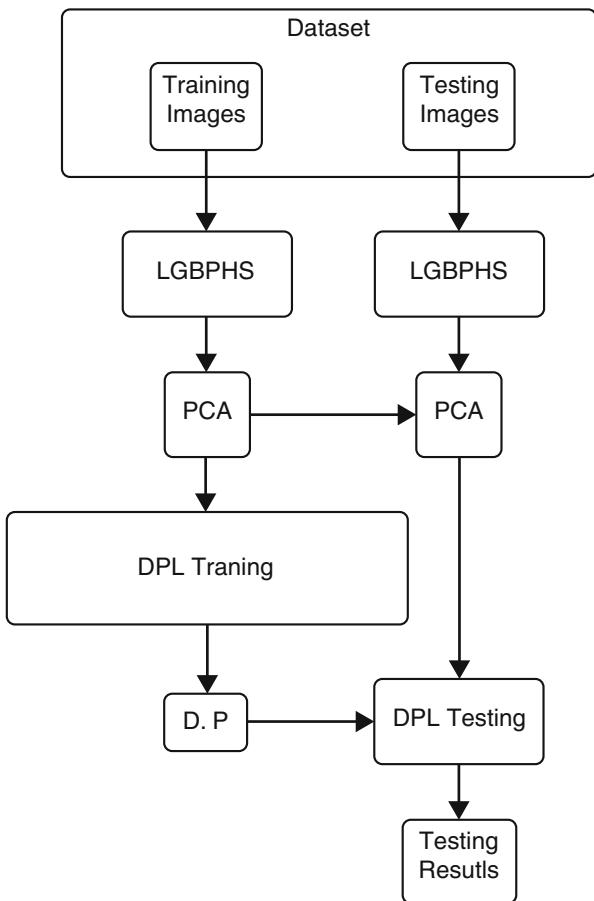
Fig. 12.9 Prototype system used in the data collection. The DKHRS is installed in a door simulator, which is a black acrylic box and is positioned on a balance board. This prototype system collected 12,000 images from 200 volunteers. (a) 3-D design model shows that the DKHRS is installed on the door. The little end is fastened and the power cord, and control wires are buried in the door. (b) Appearance of prototype system is a small black door with a transparent door knob. The prototype is made portable for the convenience of the experiment. The balance board is used to prevent accidental damage in experiments, because the system is light

the class it belongs to and 399 impostor matches to other classes. In total, there are 6000 genuine matches and 2,394,000 impostor matches.

12.4.2 Experimental Results

In the experiment, the recognition process comprises of four parts, as shown in Fig. 12.10. First, LGBPHS method is applied in all training images in the data set to extract the feature. Second, the high dimensional features are down-sampled with PCA. Third, the compressed features are fed to the DPL to learn the two dictionaries. Fourth, testing features, which are produced from the testing images by the same procedures, are tested using the learned dictionaries. The test result is evaluated with BioSecure Performance Evaluation Tool (Mansfield and Wayman 2002; Bolle et al. 2000, 2004). In this evaluation tool, three criteria are provided to evaluate the performance of a biometric system including recognition rate, receiver operating characteristic (ROC) curve, and equal error rate (EER). The recognition rate is the rate that right matches divided by all the matches, which is simple and direct, but is biased because of the imbalance of genuine matches and impostor matches. ROC curve depicts the overall performance of a biometric system. It exhibits the percentage of impostor attempts accepted, which is also called false acceptance rate (FAR), against the percentage of genuine attempts accepted (GAR), which equals to 1—false rejection rate (FRR). ROC curve does not depend on the threshold. When FAR curve crosses over FRR curve ($\text{FAR} = \text{FRR}$), the rate in the

Fig. 12.10 Experiment is following this flowchart. Both training and testing features are extracted using LGBPHS and down-sampled using PCA. Then the dictionaries learned from the training are used to test testing features



crossing point is EER. At EER point, the recognition rate (here it equals to $1 - \text{EER}$) is not affected by the imbalance of the genuine attempts and the impostor attempts, because FAR equals to FRR. EER is widely used as a brief criteria of a biometrics.

The testing result is shown in Fig. 12.11, including the recognition rate curve and the EER curve in Fig. 12.11a and the ROC curve in Fig. 12.11b. Figure 12.11a shows the recognition rates and EERs when the feature is reduced to different dimensions. When the feature dimension is higher than 200, the recognition rate is steadily above 99.5%. EER is lower than 0.15%, when the feature dimension is higher than 500. When the feature dimension is higher than 800, the EER becomes steady. When the feature dimension is 850, the proposed system achieves the best performance with a 0.091% EER. Figure 12.11b shows the ROC curve when the feature dimension is 850. Even when FAR is as lower as 10^3 , GAR is still above 98%.

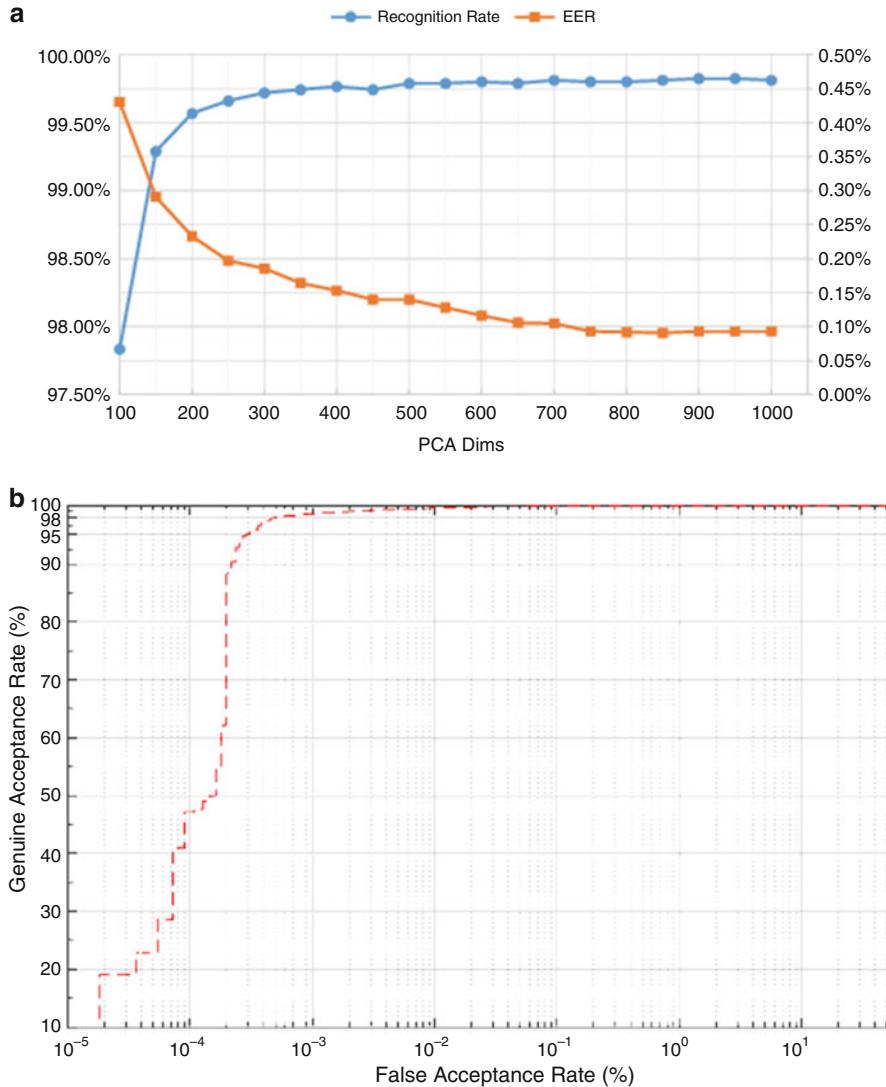


Fig. 12.11 Performance evaluating curves of DKHRS. (a) Recognition rates and EERs along different down-sampled dimensions. (b) ROC curve when the feature dimension is 850

12.4.3 Comparison with Conventional Biometrics

The recognition rate of DKHRS is over 99%, and its EER can be lower than 0.1%. Generally speaking, the recognition performance of DKHRS is much better than hand back skin texture (Xie et al. 2012), gait (Lai et al. 2014) and face recognition (Gu et al. 2014; LFW 2015); it is even surpass fingerprint recognition [about 1% EER on STFV-STD-1.0 dataset] and 3D fingerprint [3.4% EER (Liu et al. 2015)];

but it is still not as good as iris recognition [$<0.003\%$ EER (Daugman 2007)], and palmprint recognition [EER from 0.062% to 0.012% (Zuo et al. 2008; Guo et al. 2009a, b; Laadjel et al. 2009; Zhang et al. 2010a, b; Li et al. 2012; Qu et al. 2016)]. It should be noted that the data sets, the matching schemes and the experiment settings are different in these biometric systems. Thus, direct comparison of EERs may not be precise and fair. However, it still can be found that the performance of the proposed system is very promising. The reasons could be as follows. First, the door knob hand recognition is a contact recognition system. The imaging structure is well protected against environmental light. With this stable environment, the illumination is stable. DKHRS is free from illumination problem, which is a critical challenge for face recognition (Medioni et al. 2009). Second, the position of the hand is stable with the help of the pegs. Thus, the ROI area is stable. A stable ROI makes the DKHRS free from the ROI localization challenge which widely exists in many non-contact biometric systems. Third, DKHRS is a hand based biometrics, then the multiple layers of texture and the large area of skin provides DKHRS a rich and diverse features.

For the user experience aspect, DKHRS is better than conventional biometrics both in physical and cognitive aspects. In physical aspect, the appearance of DKHRS is identical to a common door knob. The dynamics of opening a door with DKHRS is also identical to opening a door with a common door knob. In the data collection, 30 s of training makes all volunteers familiar and experienced with the DKHRS device. In cognitive aspect, a DKHRS device is easy to be recognized as an access control device. Its appearance is identical to a door knob, which has been widely used in all occasions in everyday life, and has been accepted and used for over a century. The function and dynamics of DKHRS is identical to a door knob. This resemblance relieves the burden of understanding the function and training of the interactions with this new system. When collecting samples in the second session, one week after the first, 94% of volunteers remembered the operations of DKHRS.

Comparing with other modern biometrics, for example, keystroke (Liu et al. 2014), and mouse dynamics (Nakkabi et al. 2010). The advantages of DKHRS are that the performance of DKHRS is very competitive. The limitation of DKHRS is that the applications can only be with the door in an access control system. The keystroke and mouse dynamics requires no special devices, which can be used with a common keyboard or a common mouse. And both keystroke and mouse dynamics can be used as a continuous biometric system, which can verify the identity of the subject continuously. However, in access control systems, DKHRS simplifies the interaction of opening a door, which provides a unique advantage.

12.4.4 Discussion

The best EER that DKHRS achieved is 0.091% in the experiment, which is very competitive comparing with conventional biometrics. And the ergonomics are well considered both in physical and in psychological aspects. This is a major challenge

to a hidden assumption that it is only in the standard pose that a biometrics could achieve a good performance. In DKHRS, the hand is in a holding gesture. In a holding gesture, the hand is not stretched in a plane. But still, stable and robust features could be extracted. The reason of this could be two-folded. From the texture feature point of view, the texture feature of the skin in a curved surface can be extracted as a stable feature, as long as the curved surface is stable and well calibrated. From the behavior point of view, the pose, the gesture or even the action of a person could be persistent after thousands or even millions of repetitions along a long time period, which is also the foundation of many behavior biometrics. Only in DKHRS, these two aspects, biological feature and behavioral feature, are combined together, and realized in an ergonomics optimized form.

There are still four limitations of this system. First, the data set covers short term variances only. For a mature biometrics, a data set consists of samples across years variances is expected, which is a great challenge to a newly-born biometric system like DKHRS. Second, the data set in the experiment is lack of diversity of people. In this data set, all volunteers are Chinese. The color of skin may be a challenge to DKHRS, which is not tested in the experiment. Third, an auxiliary signal could be implemented for better cognition. The prototype is a silent system. A green light for a successful match and a red light for a failure could be very helpful. In addition, a sound notification, which reminds the subject that the door is unlocked could greatly enhance the user experience. Fourth, the feature extraction and classification could be improved. A DKHRS image covers the thumb, the index finger, the middle finger, and the upper half of the palm. Currently, DKHRS uses LGBPHS and DPL to extract high dimensional dense features and learn discriminative dictionaries to classify. The essence of the feature is not fully exploited. This should be studied in the future.

The proposed EBD model is only a concept model, though DKHRS system developed following this model is promising. In EBD model, the two aspects of ergonomics are addressed in three stages of designing a biometrics. And DKHRS shows a potential that designing a biometric system with the constraints of ergonomic concerns may not be definitely a compromise in security performance. However, unlike HBSI model, EBD model lacks detailed guides in designing and testing. This should be further studied.

EBD model exhibits an inspiring angle in designing a new biometric system. With the fast development of both electronics and algorithms and the increasingly use of biometrics in everyday life, the users of biometrics have been expanding from professionals to common people. The requirements of biometrics have also been shifted from solely high reliability to convenience, ergonomics, good user acceptance, and good user experience. Before, when biometric systems were only used in high security scenarios, a special job description could be posted for a suitable candidate to use a particular biometric system. However, nowadays biometric information has been embedded into ID cards. Biometrics has to be suitable for a large population. It is either to educate the public of biometrics or to design ergonomic biometrics. Thus, it is easy to predict that more ergonomic biometric systems will be invented, and many of them could benefit from EBD model.

12.5 Summary

In this chapter, we propose a novel door knob hand recognition system using ergonomic biometrics design model, which considers ergonomics in all three design stages. We invent door knob imaging scheme to capture an omni-vision hand image. Combining LGBPHS and DPL methods, the proposed system achieves promising recognition performance. Experimental result shows that the EER could be as lower as 0.091%. In addition, the proposed system shows that designing a biometric system with prior constraints of ergonomics does not definitely mean a worse performance. We expect that the proposed system and model could inspire more ergonomic biometric systems.

References

- Aharon M, Elad M, Bruckstein A (2006) K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans Signal Process* 54(11):4311–4322
- Ahonen T, Hadid A, Pietikäinen M (2004) Face recognition with local binary patterns. In: Pajdla T, Matas J (eds) Computer vision – ECCV 2004 (LNCS 3021). Springer, Heidelberg, pp 469–481
- Albrecht A (2001) Understanding the issues behind user acceptance. *Biom Technol Today* 9(1):7–8
- Alkassar S, Woo W, Dlay S, Chambers J (2015) Robust sclera recognition system with novel sclera segmentation and validation techniques. *IEEE Trans Syst Man Cybernet Syst* 47:474–486. doi:10.1109/TSMC.2015.2505649
- Alonso-Fernandez F, Fierrez J, Ortega-Garcia J (2012) Quality measures in biometric systems. *IEEE Secur Privacy* 10(6):52–62
- Biosecure, Biosecure Reference and Evaluation Framework [Online]. http://biosecure.it-sudparis.eu/AB/index.php?option=com_content&view=article&id=12&Itemid=15
- Bolle R, Pankanti S, Ratha N (2000) Evaluation techniques for biometrics-based authentication systems (FRR). In: Proceedings 15th international conference on pattern recognition, pp 831–837
- Bolle R, Ratha N, Pankanti S (2004) Error analysis of pattern recognition systems – the subsets bootstrap. *Comput Vis Image Underst* 93(1):1–33
- Burge M, Bowyer K (2013) In: Burge MJ, Bowyer KW (eds) *Handbook of iris recognition*. Springer, London
- Bustard J, Nixon M (2010) Toward unconstrained ear recognition from two-dimensional images. *IEEE Trans Syst Man Cybernet A Syst Humans* 40(3):486–494
- Carvalho R, Rosa P (2010) Identification system for smart homes using footstep sounds. In: Proceedings of IEEE international symposium industrial electronics (ISIE), Bari, Italy, pp 1639–1644
- Daugman J (2007) New methods in iris recognition. *IEEE Trans Syst Man Cybernet B Cybernet* 37(5):1167–1175
- Dunstone T, Yager N (2009) *Biometric system and data analysis – design, evaluation, and data mining*. Springer, New York
- Eidan R (2013) Hand biometrics: overview and user perception survey. In: Proceedings of 2nd international conference on informatics and applications (ICIA), Łódź, Poland, pp 252–257

- Elliott S, Massie S, Sutton M (2007) The perception of biometric technology: a survey. In: Proceedings of IEEE workshop on automatic identification advanced technologies, Alghero, Italy, pp 259–264
- Elliott S, Senjaya B, Kukula E, Werner J, Wade M (2010) An evaluation of the human biometric sensor interaction using hand geometry. In: Proceedings of IEEE international Carnahan conference on security technology (ICCST), San Jose, CA, USA, pp 259–265
- Embassy London (2009) Biometric passports and travel to the United States. Visa Services. [Online]. <http://www.usembassy.org.uk/visaservices/?p=420>
- Eschenburg F et al (2005) User acceptance: the BioSec approach. *Biom Technol Today* 13(7):80–10
- Expert (2012) Expert Working Group on Human Factors in Latent Print Analysis, Latent print examination and human factors: improving the practice through a systems approach, U.S. Department of Commerce, National Institute of Standards and Technology, Gaithersburg, MD, USA, Technical report, 2012 [Online]. http://www.nist.gov/manuscript-publicationsearch.cfm?pub_id=910745
- Faddis K, Howard J, Stracener J (2011) Enhancing the usability of human machine interface on the handheld interagency identification detection equipment (HIIDE). In: Proceedings of 21st international conference on systems engineering (ICSEng), Las Vegas, NV, USA, pp 305–310
- Gong Y, Zhang D, Shi P, Yan J (2013) Handheld system design for dual-eye multispectral iris capture with one camera. *IEEE Trans Syst Man Cybernet Syst* 43(5):1154–1166
- Gu S, Zhang L, Zuo W, Feng X (2014) Projective dictionary pair learning for pattern classification. In: Ghahramani Z, Welling M, Cortes C, Lawrence ND, Weinberger KQ (eds) Advances in neural information processing systems 27. Newry, UK, Curran Association, pp 793–801
- Guo Z, Zhang D, Zhang L, Zuo W (2009a) Palmprint verification using binary orientation co-occurrence vector. *Pattern Recogn Lett* 30(13):1219–1227
- Guo Z, Zuo W, Zhang L, Zhang D (2009b) Palmprint verification using consistent orientation coding. In: Proceedings of 16th IEEE international conference on image processing (ICIP), Cairo, Egypt, pp 1965–1968
- Guo Z, Zhang L, Zhang D, Mou X (2010a) Hierarchical multiscale LBP for face and palmprint recognition. In: Proceedings of 17th IEEE international conference on image processing (ICIP), Hong Kong, pp 4521–4524
- Guo Z, Zhang L, Zhang D, Zhang S (2010b) Rotation invariant texture classification using adaptive LBP with directional statistical features. In: Proceedings of 17th IEEE international conference on image processing (ICIP), Hong Kong, pp 285–288
- Guo Z, Zhang L, Zhang D (2010c) A completed modeling of local binary pattern operator for texture classification. *IEEE Trans Image Process* 19(6):1657–1663
- Guo Z, Zhang L, Zhang D (2010d) Rotation invariant texture classification using LBP variance (LBPV) with global matching. *Pattern Recogn* 43(3):706–719
- Huang X, Li S, Wang Y (2004) Shape localization based on statistical method using extended local binary pattern. In: Proceedings of IEEE 1st symposium on multi-agent security survivability, Piscataway, NJ, USA, pp 184–187
- Jain A (2009) Next generation biometrics, Department of Computer Science and Engineering, Michigan State University [Online]. http://www.cse.msu.edu/rgroups/biometrics/Presentations/Next_generation_biometrics_Korea_Dec2010.pdf
- Jain A, Kumar A (2012) Biometric recognition: an overview. In: Mordini E, Tzovaras D (eds) Second generation biometrics: the ethical, legal and social context (The International Library of Ethics, Law and Technology), vol 11. Springer, Dordrecht, pp 49–79
- Jain A, Ross A, Prabhakar S (2004) An introduction to biometric recognition. *IEEE Trans Circuits Syst Video Technol* 14(1):4–20
- Jain A, Flynn P, Ross A (2008) Handbook of biometrics. Springer, New York
- Jain A, Ross A, Nandakumar K (2011) Introduction to biometrics. Springer, New York

- Jeng A, Chen L (2009) How to enhance the security of e-passport. In: Proceedings of international conference on machine learning cybernetics, vol 5. Baoding, China, pp 2922–2926
- Jiang Z, Lin Z, Davis L (2011) Learning a discriminative dictionary for sparse coding via label consistent K-SVD. In: Proceedings of IEEE conference on computer vision and pattern recognition, Colorado Springs, CO, USA, pp 1697–1704
- Jiang Z, Lin Z, Davis L (2013) Label consistent K-SVD: Learning a discriminative dictionary for recognition. *IEEE Trans Pattern Anal Mach Intell* 35(11):2651–2664
- Jones L (2009) Reflective and catadioptric objectives. In: Bass M, DeCusatis C, Enoch JM (eds) *Handbook of optics*, vol 1: Geometrical and physical optics, polarized light, components and instruments, 3rd edn. McGraw-Hill, New York, pp 1–29
- Kalka N, Zuo J, Schmid N, Cukic B (2010) Estimating and fusing quality factors for iris biometric images. *IEEE Trans Syst Man Cybernet A Syst Humans* 40(3):509–524
- Kong W, Zhang D (2002) Palmprint texture analysis based on low-resolution images for personal authentication. In: Proceedings 16th international conference on pattern recognition (ICPR), vol 3. Quebec City, QC, Canada, pp 807–810
- Kong A, Zhang D (2004) Competitive coding scheme for palmprint verification. In: Proceedings of 17th international conference on pattern recognition (ICPR), Cambridge, UK, vol 1, pp 520–523
- Kukula E (2008) Design and evaluation of the human-biometric sensor interaction method, Ph.D. dissertation, The Center for Education and Research in Information Assurance Security, Purdue University, West Lafayette, IN, USA
- Kukula E, Elliott S, Tamer S, Senarith P (2006) Biometrics and manufacturing: a recommendation of working height to optimize performance of a hand geometry machine, Report BSPA/09-0001, Biometrics Standards, Performance, and Assurance Laboratory, Purdue University, West Lafayette, IN, USA
- Kukula E, Elliott S, Duffy V (2007) The effects of human interaction on biometric system performance. In: Duffy VG (ed) *Digital human modeling*. Springer, Heidelberg, pp 904–914
- Kukula E, Sutton M, Elliott S (2010) The human–biometricsensor interaction evaluation method: Biometric performance and usability measurements. *IEEE Trans Instrum Meas* 59(4):784–791
- Laadjel M, Kurugollu F, Bouridane A, Yan W (2009) Palmprint recognition based on subspace analysis of Gabor filter bank. In: Proceedings of 10th Pacific Rim conference multimedia advances in multimedia information processing (PCM), Bangkok, Thailand, pp 719–730
- Lai Z, Xu Y, Jin Z, Zhang D (2014) Human gait recognition via sparse discriminant projection learning. *IEEE Trans Circuits Syst Video Technol* 24(10):1651–1662
- Levush R (2014) Biometric data retention for passport applicants and holders. Global Legal Research Center for Law Library of Congress, Washington, DC, USA, Technical Report [Online]. <http://www.loc.gov/law/help/biometric-data-retention/>
- LFW (2015) Results [Online]. <http://vis-www.cs.umass.edu/lfw/results.html>
- Li W, Zhang D, Lu G, Luo N (2012) A novel 3-D palmprint acquisition system. *IEEE Trans Syst Man Cybernet A Syst Humans* 42(2):443–452
- Liu C, Wechsler H (2002) Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. *IEEE Trans Image Process* 11(4):467–476
- Liu J et al (2014) The Beihang keystroke dynamics systems, databases and baselines. *Neurocomputing* 144:271–281
- Liu F, Zhang D, Shen L (2015) Study on novel curvature features for 3D fingerprint recognition. *Neurocomputing* 168:599–608
- Mansfield A, Wayman J (2002) Best practices in testing and reporting performance of biometric devices, Technical report CMSC 14/02. Centre for Mathematical Scientific and Computing, National Physical Laboratory, Teddington, UK
- Medioni G, Choi J, Kuo C, Fidaleo D (2009) Identifying noncooperative subjects at a distance using face images and inferred three-dimensional face models. *IEEE Trans Syst Man Cybernet A Syst Humans* 39(1):12–24

- Mordini E, Tzovaras D (2012) In: Mordini E, Tzovaras D (eds) Second generation biometrics: the ethical, legal and social context (The International Library of Ethics, Law and Technology), vol 11. Springer, Dordrecht
- Nakkabi Y, Traore I, Ahmed A (2010) Improving mouse dynamics biometric performance using variance reduction via extractors with separate features. *IEEE Trans Syst Man Cybernet A Syst Humans* 40(6):1345–1353
- Ojala T, Pietikainen M, Maenpaa T (2002) Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans Pattern Anal Mach Intell* 24(7):971–987
- Padmapriya S, KalaJames E (2012) Real time smart car lock security system using face detection and recognition. In: Proceedings of international conference on computer communication and informatics (ICCCI), Coimbatore, India, pp 1–6
- Qu X, Zhang D, Lu G (2016) A novel line-scan palmprint acquisition system. *IEEE Trans Syst Man Cybernet Syst* 46(11):1481–1491
- Rubin J, Chisnell D (2008) Handbook of usability testing: how to plan, design, and conduct effective tests, 2nd edn. Wiley, Indianapolis, IN
- Rubinstein R, Peleg T, Elad M (2013) Analysis K-SVD: a dictionary learning algorithm for the analysis sparse model. *IEEE Trans Signal Process* 61(3):661–677
- Russell H (1893) Design for a door-knob. U.S. Patent D4 114 S
- Schllick C (2009) Industrial engineering and ergonomics. Springer, Heidelberg
- Shan S, Zhang W, Su Y, Chen X, Gao W (2006) Ensemble of piecewise FDA based on spatial histograms of local (Gabor) binary patterns for face recognition. In: Proceedings of 18th international conference on pattern recognition, Hong Kong, pp 606–609
- Shen L, Bai L, Auer D (2008) 3D Gabor wavelets for evaluating SPM normalization algorithm. *Med Image Anal* 12(3):375–383
- Vatsa M, Singh R, Noore A (2009) Unification of evidence-theoretic fusion algorithms: a case study in level-2 and level-3 fingerprint features. *IEEE Trans Syst Man Cybernet A Syst Humans* 39(1):47–56
- Wertheim K (2010) Human factors in large-scale biometric systems: a study of the human factors related to errors in semiautomatic fingerprint biometrics. *IEEE Syst J* 4(2):138–146
- Wong M, Zhang D, Kong W, Lu G (2005) Real-time palmprint acquisition system design. *IEE Proc Vis Image Signal Process* 152(5):527–534
- Wu X, Zhang D, Wang K (2006) Palm line extraction and matching for personal authentication. *IEEE Trans Syst Man Cybernet A Syst Humans* 36(5):978–987
- Xie J, Zhang L, You J, Zhang D, Qu X (2012) A study of hand back skin texture patterns for personal identification and gender classification. *Sensors (Basel)* 12(7):8691–8709
- Yan K, Chen Y, Zhang D (2011) Gabor surface feature for face recognition. In: Proceedings of 1st Asian conference on pattern recognition (ACPR), Beijing, China, pp 288–292
- Zhang B, Li Y (2012) Catadioptric vision system. In: Automatic calibration and reconstruction for active vision systems (Intelligent systems, control and automation: science and engineering), Ch 6, pp 117–150
- Zhang D, Kong W, You J, Wong M (2003) Online palmprint identification. *IEEE Trans Pattern Anal Mach Intell* 25(9):1041–1050
- Zhang W, Shan S, Gao W, Chen X, Zhang H (2005) Local Gabor binary pattern histogram sequence (LGBPHS): a novel non-statistical model for face representation and recognition. In: Proceedings of 10th IEEE international conference computer vision, Beijing, China, vol 1, pp 786–791
- Zhang D, Guo Z, Lu G, Zuo W (2010a) An online system of multispectral palmprint verification. *IEEE Trans Instrum Meas* 59(2):480–490
- Zhang D, Kanhangad V, Luo N, Kumar A (2010b) Robust palmprint verification using 2D and 3D features. *Pattern Recogn* 43(1):358–368
- Zhou Z, Du E, Thomas N, Delp E (2012) A new human identification method: sclera recognition. *IEEE Trans Syst Man Cybernet A Syst Humans* 42(3):571–583

- Zhu Z et al (2015) Three-dimensional Gabor feature extraction for hyperspectral imagery classification using a memetic framework. *Inf Sci* 298:274–287
- Zuo W, Yue F, Wang K, Zhang D (2008) Multiscale competitive code for efficient palmprint recognition. In: Proceedings 19th international conference on pattern recognition (ICPR), Tampa, FL, USA, pp 1–4

Part IV

Some New Head-Based Biometrics

Chapter 13

Dynamic Tongueprint Recognition

Abstract Biometrics, which use human physiological or behavioral features for personal identification, currently face the challenge of designing a secure biometric system that will accept only the legitimate presentation of the biometric identifiers without being fooled by the doctored or spoofed measurements that are input into the system. More biometric traits are required for improving the performance of authentication systems. In this chapter, we present a new number for the biometrics family, i.e. tongueprint, which uses particularly interesting properties of the human tongue to base a technology for noninvasive biometric assessment. The tongue is a unique organ which can be stuck out of the mouth for inspection, whose appearance is amenable to examination with the aid of a machine vision system. Yet it is otherwise well protected in the mouth and difficult to be forged. Furthermore, the involuntary squirm of the tongue is not only a convincing proof that the subject is alive, but also a feature for recognition. That is to say, the tongue can present both static features and dynamic features for authentication. However, little work has hitherto been done on the tongue as a biometric identifier. In this work, we make use of a database of tongue images obtained over a long period to examine the performance of the tongueprint as a biometric identifier. Our research shows that tongueprint is a promising candidate for biometric identification and worthy of further research.

Keywords Tongueprint • Verification • Identification

13.1 Introduction

There has been considerable research in biometrics (Zhang 2000; Jain and Healey 1998) over the last two decades. The list of physiological and behavioral biometric characteristics that have to date been developed and implemented is long and includes the face (Li and Juwei 1999; Abate et al. 2007), iris (Daugman 2004; Bowyer et al. 2007), fingerprint (Ratha et al. 1996), palmprint (Zhang et al. 2000), hand shape (Sanchez-Reillo et al. 2000), voice (Wan and Renals 2005), signature (Lee et al. 1996) and gait (Wang et al. 2003). Notwithstanding this great and

increasing variety of biometrics, no biometric has yet been developed that is perfectly reliable or secure. For example, fingerprints and palmprints are usually frayed; voice, signatures, hand shapes and iris images are easily forged; face recognition can be made difficult by occlusions or face-lifts; and biometrics such as fingerprints, iris and face recognition are susceptible to spoofing attacks (O’Gorman 2003), i.e. the biometric identifiers can be copied and used to create artifacts that can deceive many currently available biometric devices. The great challenge to biometrics is thus to improve recognition performance and be maximally resistant to deceptive practices (Jain et al. 2004). To this end, many researchers have sought to improve reliability and frustrate spoofers by developing biometrics that are highly individuating; yet at the same time, present a highly complex, hopefully insuperable challenge to those who wish to defeat them.

In this chapter, we propose a dynamic tongueprint, a biometric identifier that for the first time, makes use of both static features of the biometric as well as dynamic features. As a biometric identifier, tongueprint has the following properties. To begin with, the geometric features of the tongue, as well as the cracks and textures on its surface, are distinctive to each person. By way of illustration, Fig. 13.1a, b show some different tongue geometries and (c) shows some different types of cracks and (d) shows some different textures. Second, the tongue is the only internal organ that can easily be exposed for inspection and it is the exposed portion of the tongue that carries a great deal of the physiological information including geometric features, crack features and texture features, which are named as “tongueprint” synthetically. Third, according to our long time observations, the geometric features, crack features and texture features of an individual tongue are stable. Aside from the above three points, there are two unique features of the human tongue. First, contained in the mouth, the human tongue is protected from the external environment, unlike for example the fingers that could be damaged easily. Second, the involuntary squirm of the human tongue not only is a natural and convincing proof that a subject is alive, but can also be utilized for discriminating individuals.

The dynamic tongueprint makes use of a dynamic feature, which is the involuntary squirm of the human tongue when it is stuck out of the mouth. To date, this feature has been regarded as mostly a hindrance to the process of capturing biometric images, although it has been recognized that it does provide a natural and convincing proof that a subject is alive. However, it has not been used previously to discriminate individuals. Our authentication system, (Fig. 13.2) which makes use of both static and dynamic features, contains two modules: enrollment and recognition. The recognition module operates in two phases: in the first phase, liveness detection, dynamic information about the squirm of the tongue is used to detect whether the tongue is in fact living; the second phase, feature extraction, extracts both static physiological features and dynamic squirm features. The use of dynamic as well as static information not only means that our proposed system can provide liveness detection, but also has the potential to improve recognition (Tistarelli et al. 2009). Recent psychophysical and neural studies (OToole et al. 2002; Knight and Johnston 1997) have described the crucial

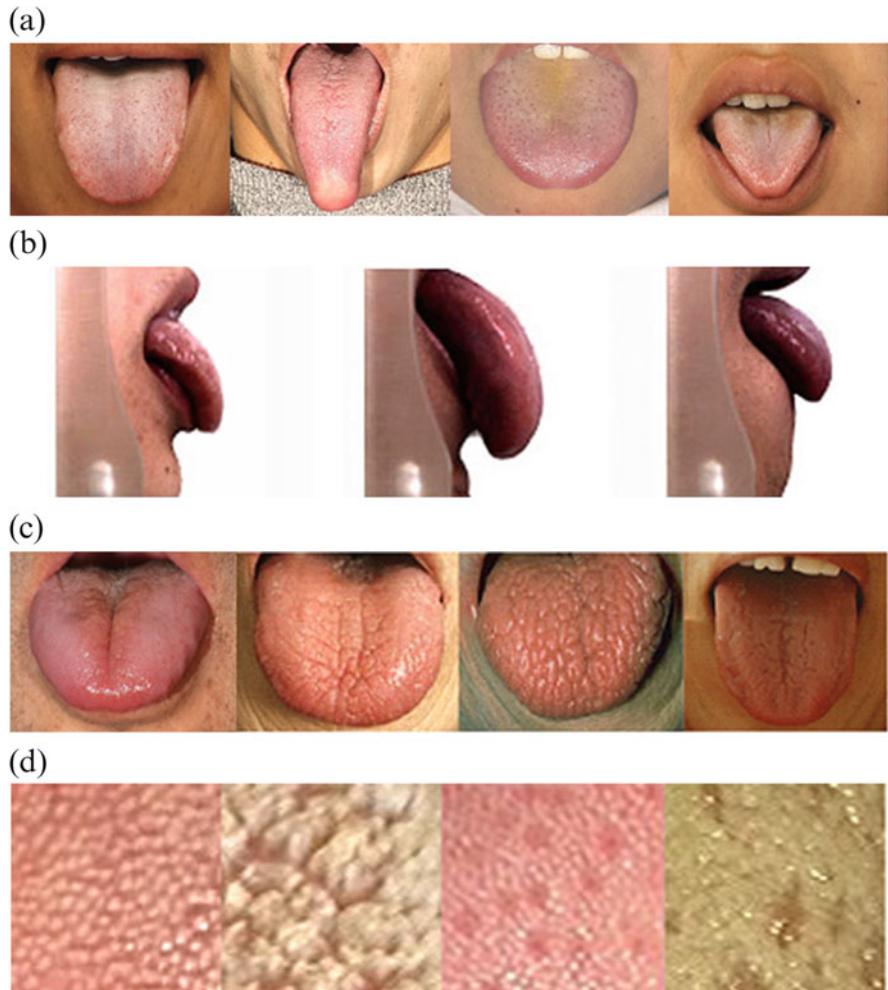


Fig. 13.1 The distinctiveness of the human tongue patterns. (a) Some different tongue geometries (front view). (b) Some different tongue geometries (profile view). (c) Some different types of tongue surface cracks. (d) Some different tongue surface textures

role of dynamic information in the human visual recognition process while work by Haxby et al. (2000) demonstrated that visual perception makes use of a double architecture of two connected neural activation patterns: one for processing static, unchanging and invariant features and the other for changing features (Tistarelli et al. 2009), an architecture reflected in this proposed approach.

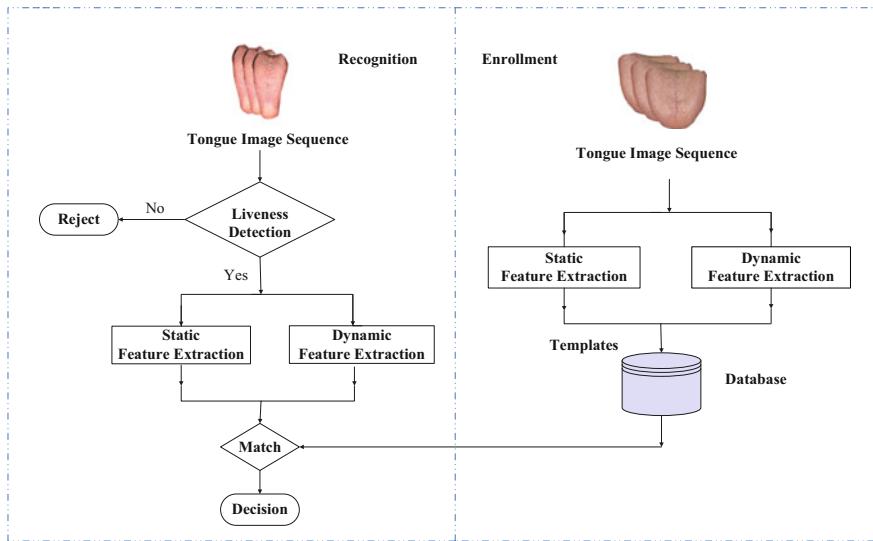


Fig 13.2 The flowchart of tongueprint recognition

13.2 Background

To touch upon the novel concept of dynamic tongueprint as a pattern for recognition, it is useful to introduce tongue anatomy firstly. According to the study of zoology, the invertebrate began to have the tongue structure, such as locusts. And the structure and functions of the tongue of vertebrates become much more complex, such as snakes or lizards. It is obvious that different animals have different tongues. In terms of the human being, the tongue normally resides inside the mouth and mainly composed of muscles (Brand and Isselhard 1998), covered with a mucous membrane. Two parts of the tongue are separated for easy description, which are the root and body. Note that this chapter only focuses on the visual information from the tongue body (see Fig. 13.3 for the anatomic structure of the tongue).

The tongue is composed of muscle tissues. It is the muscles that move the tongue and give it the dynamic property that can be used for liveness detection. In addition, the muscles make up most of its mass and various shapes.

Small nodules of tissue (papillae) cover the upper surface of the tongue. Between the papillae are the taste buds, which provide taste, heat, pain, and tactile information. Note that the various distributions and shape of fungiform papilla show the various textures on the tongue surface. In addition to taste, the tongue functions in moving food to aid chewing and swallowing, and it is important in speech. Figure 13.4 shows a micrograph of the surface of a human tongue featuring both fungiform and filiform papillae.

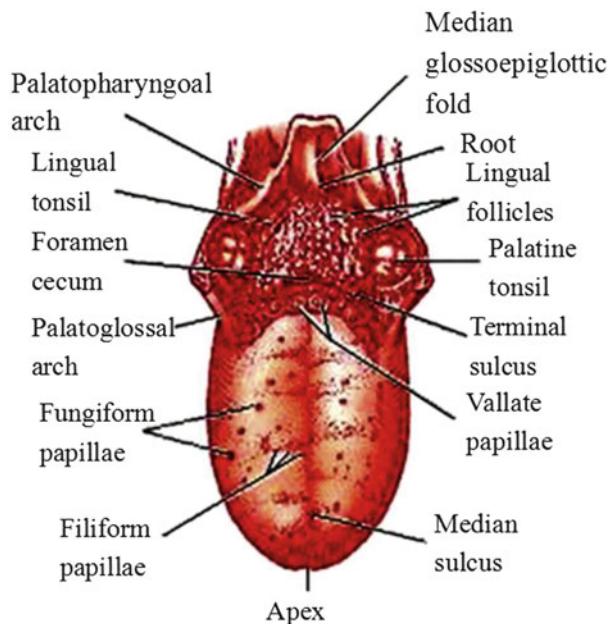


Fig. 13.3 Anatomy of the human tongue



Fig. 13.4 A micrograph of a fungiform papilla (the large round structure in the center of the image) surrounded by hairlike filiform papillae. In this image, the filiform papillae are “combed” down so that they are lying side by side

13.3 Tongue Squirm and Its Applications

When a person sticks his or her tongue out, the tongue squirms. Tongue squirm is a type of physical reaction when the muscles of the tongue are stretching. This squirming is both continuous and involuntary. From the point of view of biometrics, this squirm constitutes a very powerful “dynamic signature” which can augment other information obtained from the tongue and has the potential to dramatically improve the performance of the tongue as a biometric for individual identification. In the following, we describe how it can be applied to liveness detection and squirm feature extraction.

13.3.1 Liveness Detection

Notwithstanding its performance on other criteria, any biometric authentication system will be open to spoofing as long as it cannot distinguish between a photograph and biometric features of a living individual (Kollreider et al. 2009). Liveness detection (or vitality detection as it is sometimes known) seeks to ensure that input patterns are not from an inanimate object (Jain et al. 2004), determining whether the biometric data being captured is indeed from a legitimate live user who is physically present in front of the acquisition device. There has been research (Kollreider et al. 2005; Moon et al. 2005; Sandström 2004) that counter the threat of physical spoofing of biometric samples and various liveness detection methods have been implemented in some devices, for example, using perspiration in association with fingerprints and iris shrink with iris-based applications. However, it is not possible to claim success in this area as proposed and implemented approaches to date have either required complex hardware or hard to defeat. However, research continues in this area (Toth 2005).

Tongue squirm is an especially attractive liveness feature for three reasons. First, it is an involuntary action. It results from the muscles of the tongue stretching when the tongue is stuck out of the mouth and cannot be controlled. Moreover, unlike iris shrinking, it does not need to be externally stimulated (in the case of iris shrinking, it is by illumination) and so requires no additional hardware. Second, tongue squirm is random. It is not a performance that any spoofer can learn and imitate. Third, tongue squirm is a very slight and subtle movement and as such, it does not affect the global shape of the tongue during the capture process. Finally, tongue squirm is thus available to be captured in a sequence of continuous images reflecting live action using purely software-based methods rather than additional hardware.

An initial consideration in capturing tongue squirm features is the formidable task of learning the activity in the original image space. However, because the captured images of the squirm can be regarded as points in high-dimensional space and these may generally be expected to lie on a low-dimensional manifold embedded in the high-dimensional image space, we can reduce the size of the task by

analyzing tongue squirm in a low-dimensional subspace rather than the original image space. The goal thus becomes finding a compromise between preserving the global and local structure. It is well known that, of the current dimensionality reduction methods, the principal component analysis (PCA) (Turk and Pentland 1991) is good at preserving the global structure and locally linear embedding (LLE) (Roweis and Saul 2000) at preserving local structure. A compromise between these two goals can be found in Orthogonal Neighborhood Preserving Projections (ONPP) (Kokiopoulou and Saad 2007), a recently proposed method in image recognition for feature extraction and dimensionality reduction for visualization. ONPP is a linear dimensionality reduction technique and tends to preserve not only the local, but also the global geometry of high dimensional data samples. For this reason, we use ONPP to embed the tongue squirm into a lower dimensional feature space.

The tongue image sequence can be represented as a data set $I = \{I_1, I_2, \dots, I_N\} \in R^{m \times n}$, where N is the number of the images. This data set can be transformed into a matrix $X = [x_1, x_2, \dots, x_N] \in R^{h \times N}$, where x_i is the vector representation of $I_i \in I$ and $h = m \times n$. The key to ONPP is to construct the matrix $\bar{M} = X(I - W^T)(I - W)$ X^T and solve a generalized eigenvalue problem to create a transform matrix V for the dimensionality reduction. This mapping can be formulated as:

$$Y = V^T X \quad (i = 1, 2, \dots, N) \quad (13.1)$$

where W is the weight matrix, $Y \in R^{d \times N}$ contains only the data set mapped into d -dimensional space ($d \ll h$), and the columns of V are the basis of the eigenvectors associated with the d smallest eigenvalues of \bar{M} . More details about ONPP can be found in (Kokiopoulou and Saad 2007).

Figure 13.5 plots the visualizations of the squirm in three-dimensional space using ONPP. We can see that the image sequence of the living tongue exhibits random trajectories in three-dimensional subspace. More importantly, images of an artifact tongue that does not squirm will not exhibit any such random trajectories in the subspace. (For a description of our related experiments, see Sect. 13.5.4). This is supported by Eq. (13.1). That is, if the matrix X is composed of the same column vectors and V is the nonsingular matrix, then the result of mapping Y will be composed of the same vectors. Hence, only one point is obtained in the subspace. We propose to use this circumstance as a qualitative proof in our liveness detection procedure.

The novel framework proposed here is based on the analysis of the low dimensional manifold, as mentioned above. Figure 13.6 is a flowchart for our liveness detection procedure. The first step is to determine whether the input image sequences are valid against both the front and profile views of the tongue. The second step is to calculate the dynamic descriptor of the image sequence. The dynamic descriptor comes from the statistical variant and can be defined as in the following equation:

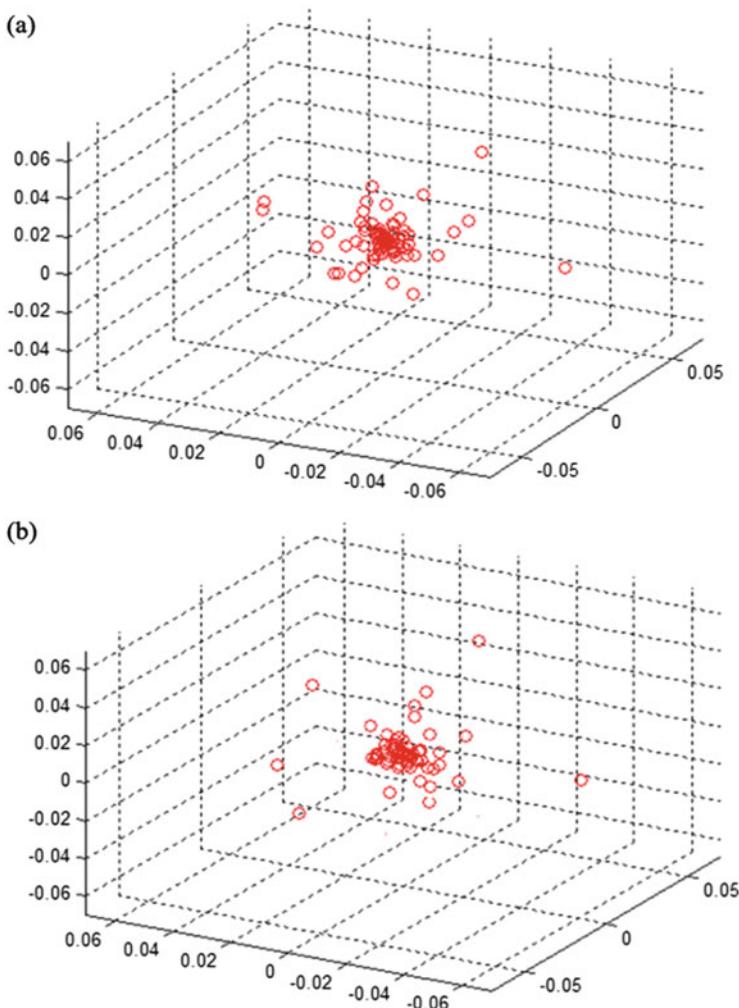


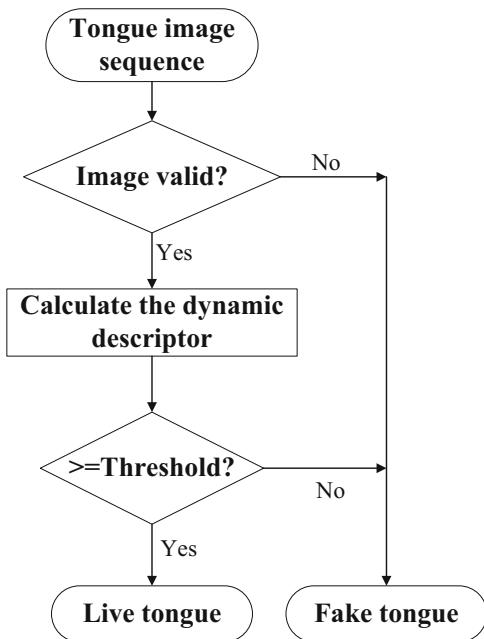
Fig. 13.5 Mapping the tongue squirm image sequences into 3D subspace: (a) is from the front view image sequence and (b) is from the profile view image sequence

$$D = \left\| \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu)(x_i - \mu)^T \right\| \quad (13.2)$$

where N is the number of samples assigned to the subject and μ is the mean value of the samples x_i .

Then, the value of the dynamic descriptor is compared with the threshold and a decision is made based on the result of this comparison.

Fig. 13.6 Flowchart for tongue liveness detection



13.3.2 Squirm Features Extraction

Embedded tongue squirm features can be extracted and visualized using manifold learning technologies (Kokiopoulou and Saad 2007), with the different trajectories of the squirms of different subjects being visualized in the three-dimensional subspace and clustered in different areas of the subspace (illustrated in Fig. 13.7). The task of enrollment thus requires the collection of feature vectors from a given subject and the learning of the parameters, i.e. the mean value μ and covariance matrix Σ , which can be calculated by Eq. 13.3:

$$\left\{ \begin{array}{l} \mu_j = \frac{\sum_{i=1}^{N_j} x_i}{N_j} \\ \Sigma_j = \frac{1}{N_j - 1} \sum_{i=1}^{N_j} (x_i - \mu_j)(x_i - \mu_j)^T \end{array} \right. \quad (13.3)$$

where N_j is the number of samples assigned to subject j . Once the parameters for each tongue squirm have been found, determining a probe class is straightforward by minimum distance method.

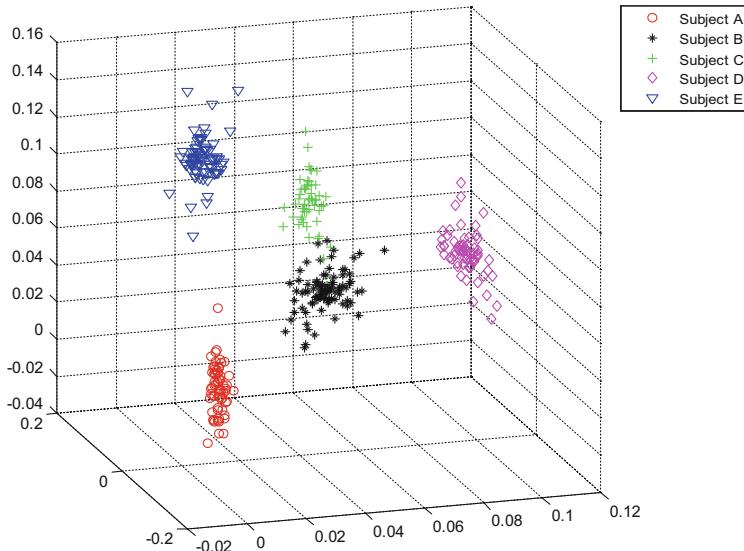


Fig. 13.7 Different tongue squirm trajectories in 3D subspace from five different subjects

13.4 Extraction of Static Physiological Features

In this section, we focus on introducing the ways of extracting physiological static features of an individual tongue, i.e. its geometric features, crack features and texture features.

Before feature extraction, it is necessary to obtain the contours of the tongue from the captured tongue image. The accuracy and robustness of the tongue contour extraction method are crucial for this recognition system. Tongue contour extraction (sometimes known as tongue segmentation) is made difficult by the fact that the surface color of the tongue is very similar to that of the ambient tissue. The literature contains many techniques for solving this problem (Pang et al. 2005). In this work, we carry out segmentation in the red channel (shown in Fig. 13.8b) of the original RGB images, since the tongue is basically red and tongue images contrast more strongly in the red channel. Figure 13.8c shows the contours of the tongue in profile and front views.

13.4.1 Geometric Features

The geometric features of the tongue can be obtained by a set of control points: $P_1, P_2, \dots, P_{11}, P_{tip}$ and P_m (shown in Fig. 13.9). Note P_1 and P_2 are obtained from the corners of the mouth. We define the part of the tongue that is of interest; the part below segment L_{P_1, P_2} in Fig. 13.9a, from the corners of the mouth and the tip of the



Fig. 13.8 Tongue contour extraction. (a) The original image; (b) The red channel of image (a); (c) Extracted tongue contours

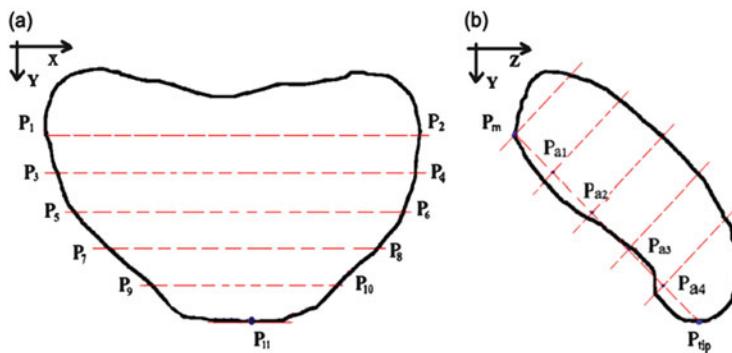


Fig. 13.9 The tongue feature model, front (a) and profile (b) views

tongue. The following describes the ways we measure (1) the width of the tongue, (2) the thickness of the tongue, and (3) the curvature of the contour of the tongue to serve as our measurement vectors:

- (1) Width: The width vector \bar{W} is constructed from the lengths of five segments. Four segments ($L_{P_3P_4}, L_{P_5P_6}, L_{P_7P_8}, L_{P_9P_{10}}$) are parallel to the segment $L_{P_1P_2}$. These four segments are equidistant, according to the following formula:

$$d(L_{P_1P_2}, L_{P_3P_4}) = d(L_{P_3P_4}, L_{P_5P_6}) = d(L_{P_5P_6}, L_{P_7P_8}) = d(L_{P_7P_8}, L_{P_9P_{10}}) \quad (13.4)$$

where $d(\cdot)$ represents the distance between two segments.

- (2) Thickness: The thickness of the tongue is defined as follows. Take a line between P_{tip} and P_m (Fig. 13.9b). The equidistant points on $L_{P_mP_{tip}}$ are labeled $P_{a1}, P_{a2}, P_{a3}, P_{a4}$. Crossing these points, we can get the orthogonal lines of the segment $L_{P_mP_{tip}}$, whose lengths within the profile view contour are used for the thickness vector \bar{T} .
- (3) Curvature of the contour of the tongue: We measure curvature of the contour of the tongue by using total curvature functions (TCF) (Pikaz and Dinstein 1995).

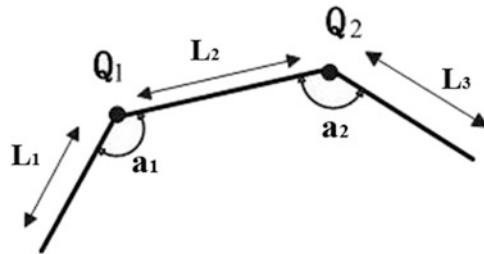


Fig. 13.10 Total curvature measures. L_i : length of the segment between Q_i and its predecessor point; L_2 : length of the segment between Q_1 and Q_2 ; L_3 : length of the segment between Q_2 and its successor point; a_i : interior angle at Q_i ; a_2 : interior angle at Q_2

The total curvature function is an approximate estimation method and defined for one segment between the two points Q_1 and Q_2 , as illustrated in Fig. 13.10, where the curvature at Q_1 can be formulated as:

$$C_1 = a_1(L_1 + L_2) \quad (13.5)$$

and the curvature at Q_2 is formulated as:

$$C_2 = a_2(L_2 + L_3) \quad (13.6)$$

The total curvature value (TC) of the segment L_2 between Q_1 and Q_2 is formulated as:

$$TC = L_2 * (C_1 - C_2) \quad (13.7)$$

We then use these TC to build the vector \overline{Cur} with the curvature values at the control points $P_3, P_4, \dots, P_9, P_{10}$ in Fig. 13.9a.

Since the components of these vectors have different dimensions, they have a large dynamic range. Thus, it is necessary to normalize these components into a common range. The final step is to combine the three measurement vectors, forming the geometric feature vector.

13.4.2 Crack Features

Cracks are obvious line features which are primarily found on the central part of tongue surface. To extract this information, we had set up a sub-image of the segmented tongue image as a region of interest (ROI). This region was selected using the coordinates system $P_{corner}OP_{tip}$ and had an area of 256*256 pixels, corresponding to the rectangular area enclosed by the white line in Fig. 13.11a.

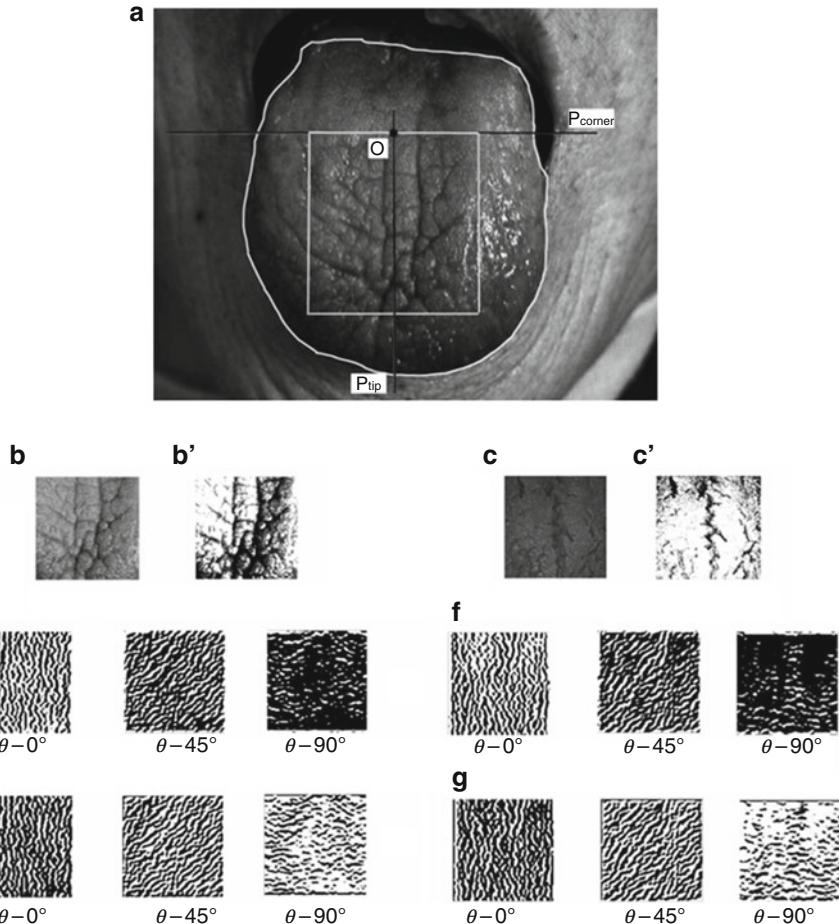


Fig. 13.11 (a) Illustrates the obtaining of the ROI; (b) and (c) are two samples of the cracks in the ROI; and (b') and (c') give the results of the histogram equalization from (b) and (c) respectively. (d) and (f) are respectively the real parts of the Gabor filtered results of (b') and (c'). (e) and (g) are the imaginary parts of the Gabor filtered results of (b') and (c')

To extract the crack features, we applied two-dimensional Gabor filter (Kong et al. 2003; Jain et al. 1998), which has the following general form (Kong et al. 2003):

$$G(x, y, \theta, u, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left\{\frac{x^2 + y^2}{2\sigma^2}\right\} \exp\{2\pi i(ux \cos \theta + uy \sin \theta)\} \quad (13.7)$$

where $i = \sqrt{-1}$; u is the frequency of the sinusoidal wave, θ controls the orientation of the function, and σ is the standard deviation of the Gaussian envelope. In order to make the Gabor filter more robust against brightness, it is set to zero DC (direct current) with the application of the following formula (Kong et al. 2003):

$$G'(x, y, \theta, \mu, \sigma) = G(x, y, \theta, \mu, \sigma) - \frac{\sum_{i=-n}^n \sum_{j=-n}^n G(i, j, \theta, \mu, \sigma)}{(2n + 1)^2} \quad (13.8)$$

where $(2n + 1)^2$ is the size of the filter. Figure 13.11d–g show some results of this procedure filtered using the Gable filter.

An input tongue sub-image $I(x, y)$ is convolved with G' . Then, the sample point in the filtered image can be coded as two bits, (b_r, b_i) using the following rules:

$$\begin{cases} b_r = 1 & \text{if } \operatorname{Re}[I \otimes G'] \geq 0 \\ b_r = 0 & \text{if } \operatorname{Re}[I \otimes G'] < 0 \\ b_i = 1 & \text{if } \operatorname{Im}[I \otimes G'] \geq 0 \\ b_i = 0 & \text{if } \operatorname{Im}[I \otimes G'] < 0 \end{cases} \quad (13.9)$$

Using this coding method means that only the phase information in the sub-images is stored in the crack vector. This feature extraction method was introduced by Daugman (1993) for use in iris recognition.

13.4.3 Texture Features

The crack features describe the overall attributes of a tongueprint and texture features provide even more detailed information of a tongueprint. In general, some tongueprints are very different and can be discriminated by their crack features. However, some are very similar with cracks or have no cracks. Hence, additional texture features are required for recognition.

The texture discrimination method used here is based on frequency domain analysis. Figure 13.12 shows four typical tongue surface textures and their corresponding frequency domain images.

Texture feature representation is used to describe the features in a concise and easy-to-compare way. We can represent the frequency domain images using the polar coordination system (r, θ) . The frequency domain image is divided into small parts by a series of circles which have the same center, as shown in Fig. 13.13a. The energy in each ring area R_i can be defined as:

$$R_i = \sum_{\theta=0}^{\pi} \sum_{r=K(i-1)}^{Ki} I_f(r, \theta) \quad i = 1, 2, \dots, K \quad (13.10)$$

where I_f is the frequency domain images in the polar coordination system, and K is the number of the pixels in the rings for integration. We call R_i the R features.

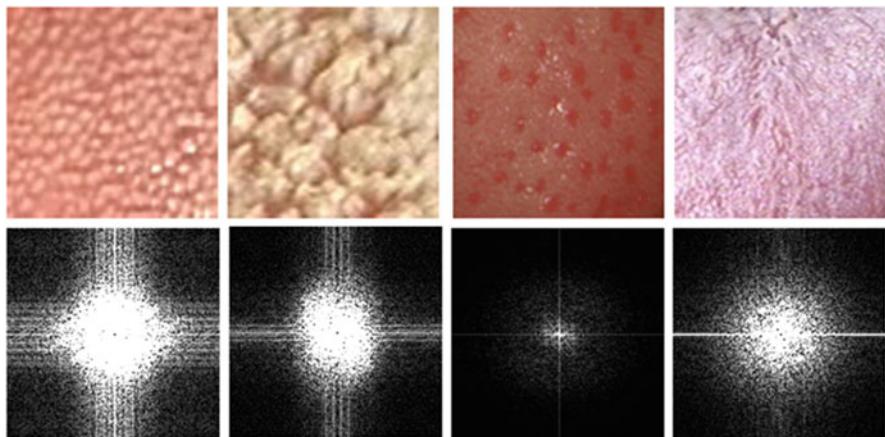


Fig. 13.12 Textures from different tongues and their frequency domain images. The upper row is the original texture image. Below each image is its corresponding frequency domain image

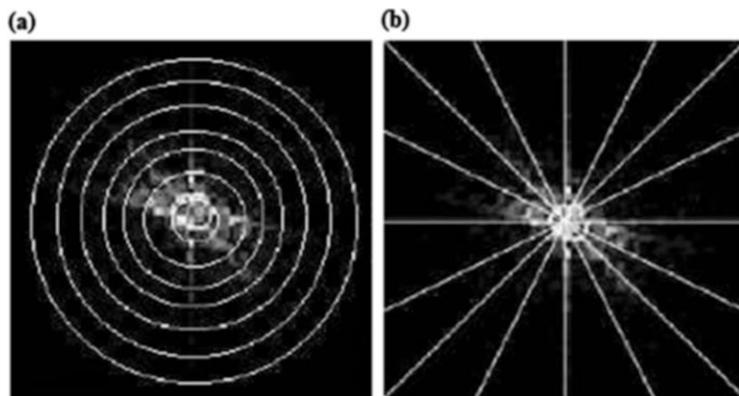


Fig. 13.13 Segmentation of frequency domain images (Wenxin et al. 2002). (a) is used for R features and (b) is for Θ features

Meanwhile, the frequency domain image can be divided by a series of lines that go through the center of the image, as shown in Fig. 13.13b. The energy in each fan-shaped part Θ_i can be defined as:

$$\Theta_i = \sum_{\theta=\pi(i-1)/M}^{i\pi/M} \sum_{r=0}^L I_f(r, \theta) \quad i = 1, 2, \dots, M \quad (13.11)$$

where L is the length of the line used for integration and M is the number of fan-shaped parts, Θ_i are referred to as Θ features.

13.5 Experimental Results

In this section, we validate the feasibility and effectiveness of a biometric identifier based on dynamic tongueprint by experiments.

13.5.1 Database

Research on the human tongue as a biometric has been hampered by the lack of a suitable tongue image database. In this chapter, we present a newly developed tongue image database. This is the first attempt at making a tongue image database that will be available to the biometric research community. The database collected tongue images from 174 subjects, 66.4% male and 33.6% female from a variety of ethnic/racial ancestries. Most of the samples came from the young and middle-aged. The reasons for this restriction are as follows: first, the tongue shapes of the young and middle-aged are more stable than those of children. Second, the old people are more difficult for keeping samplings. Table 13.1 shows the composition of the database.

To ensure that the database is useful for assessing and comparing algorithm recognition techniques, the tongue images were captured at high-resolution (1280*1024) in a studio environment to obtain both the textures and shapes of the tongues. Each subject was required to stick out his or her tongue with the tongue spread. The front and profile views were captured simultaneously using our special device. Figure 13.14 shows some examples from three different subjects in the database.

To evaluate the performance of the novel biometric identifier, we conducted experiments on the five types of tongue representations: geometric features, crack features, texture features, squirm features, and the fusion of physiological static features and squirm features. To compare the results, we first calculated the templates from the gallery set. We then calculated the matching score of the new tongue image set (containing front and profile views) against the stored template. In the experiment with the fusion approach, the overall match score is computed according to the matching score level fusion method (Kittler et al. 1998).

The geometric features, texture features and squirm features from the gallery samples were averaged to generate the stored template in each case and the

Table 13.1 Composition of the tongue image database

	Sex		Age		
	Male	Female	20–29	30–39	40–49
Number of samples	115	59	103	49	22
Percentage (%)	66.4	33.6	59.2	28.2	12.6



Fig. 13.14 Some examples of three subjects in the tongue image database

Euclidean distance was used for discrimination. In contrast, the stored feature template for the crack features from the gallery samples were directly compared with the probe features by using the minimum Hamming distance.

13.5.2 Identification Experiments

In our identification experiments, we used a closed universe model as described by Phillips et al. (2000). In this model, every subject in the probe set is also in the gallery set. All templates in the gallery set are assumed to have an identity known at the enrollment phase. Ninety-two subjects were used for both the probe and gallery sets. We carried out each experiment three times by randomly choosing the images for the gallery and probe sets three times. Figure 13.15 shows the average recognition rate for this set of experiments. Each curve represents a different type of tongue representation. The average rank one recognition rates of the fusion rule is 95%. The average rank one recognition rates of the other four inputs are 89.3%, 79.4%, 72.5% and 71% respectively. The recognition performance obtained in these experiments was encouraging.

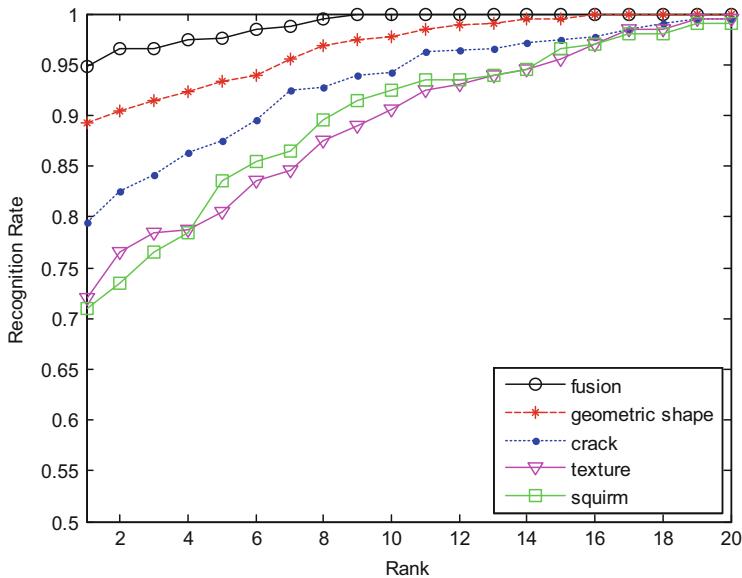


Fig. 13.15 Recognition rates of the five types of tongue representation input

13.5.3 Verification Experiments

Our verification experiments used an open universe model as described by Phillips et al. (2000). In this model, a subject in the probe set may or may not be present in the gallery set. The experiments used a probe set of 136 subjects and a gallery of 142 subjects with 102 subjects being present in both the probe and gallery sets. In other words, templates from 136 probe subjects were compared with templates of 142 gallery subjects resulting in a total of 19,312 verification attempts. Of these attempts, 102 were genuine and 19,210 were impostors. Again, we carried out each experiment three times by randomly choosing the images for the gallery and probe sets three times.

The following gives the results for using each of the five types of tongue recognition approaches for verification as a receiver operating characteristic (ROC), which shows the false acceptance rate (FAR) and false rejection rate (FRR) at each threshold value. An effective method will produce the lowest possible figure, but they are actually antagonists, meaning they must be traded off against each other. For this reason, we quantify the verification performance here by using the equal error rate (EER). Then the average of the experimental results is illustrated in Fig. 13.16. From Fig. 13.16, we can see that the lowest value of EER is 4.1%, which is obtained by the fusion approach. Moreover, the EER of the other four inputs are 9.6%, 15.2%, 19.1% and 20.4%.

In summary, in this set of experiments, the fusion approach performed better than the other approaches utilizing solely geometric features, crack features,

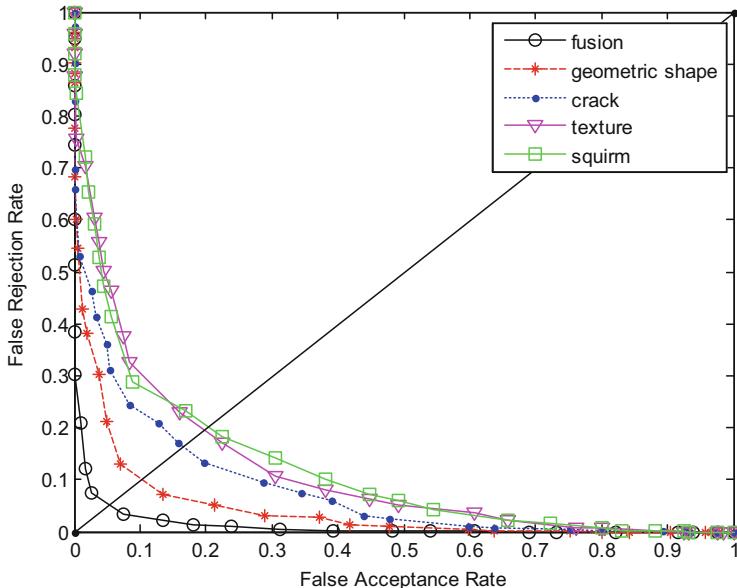


Fig. 13.16 ROC of the five types of tongue representation input

textures features, or squirm features. At the same time, geometric features always perform better than using crack features, texture features, or squirm features. The poorer performance of crack features is inherent as cracks on the tongue surface are obvious in some cases, but not in others. The same applies to texture features, which furthermore, can also be affected by the image quality. As to squirm features, its performance as a sole feature is affected by the fact that the squirm data sets can sometimes overlap in the subspace.

13.5.4 Liveness Detection Experiments

It may be difficult to depict all the kinds of attack methods because it is very difficult to enumerate all attempts for attacking the system by forge. In this chapter, the liveness detection method is exploited to resist the main fake approach, i.e., using a tongue model to spoof the tongueprint recognition system. In this case, we assumed the same illumination as the whole process takes place in a studio environment and we used an apparatus to fix the location of the head. Thus, the fake does not have a dynamic characteristic, such as a living human tongue. Even if the model is in motion, due to its 2-D planar structure and static characteristic, the appearances of images are nearly unchanged. Accordingly, these facts can be used for liveness detection.

Table 13.2 The values of dynamics descriptor

Image sequence	Dynamics descriptor		
	Mean value	Minimum	Maximum
Fake tongue	0.03	0	0.16
Live tongue	1.04	0.67	1.63

Table 13.3 Experimental results of the proposed liveness detection method

Threshold	FR
0	0.5
0.005	0.03
0.01	0.02
0.15	0.01
0.2	0
0.4	0
0.7	0.005
0.8	0.05
0.9	0.14
1.1	0.27
2.0	0.5

Take the image sequence of the front view for example. The dynamics descriptors are tabulated in Table 13.2. Table 13.2 shows the dynamics descriptors for the image sequence of the front view. Clearly, the values for the fake and live tongues are very different. As well, we can easily discriminate them by setting a threshold value.

Table 13.3 provides the results for a total of 100 live and 100 fake tongue image sequences used in evaluating our liveness detection system. The False Rate (FR) in the table is defined as $FR = \frac{FAT + FRT}{total\ times} \times 100\%$ where the FAT is the number of the counterfeit tongueprints that were accepted as tongueprints from living tongues. Similarly, the FRT shows the number of times that the live tongueprints were rejected as counterfeit.

Looking at Table 13.3, it is immediately observable that the False Rate (FR) can be reduced to zero only if we select the correct threshold. In theory, if the threshold is greater than zero, then FR should be zero. In practice, the results are affected by the noise so that we cannot obtain an ideal performance.

13.6 Summary

The tongueprint is a promising identifier because the tongue is a well-protected, but accessible internal organ that is difficult to counterfeit or spoof and its visual features cannot be reverse engineered. It becomes even more attractive as an

identifier when we make use of the tongue's involuntary squirm which not only provides liveness proof, but also cues for use in recognition.

The contribution of this chapter is not to use some new techniques, but for the first time, make claim that the tongue can be a new member in the biometrics family. The approach proposed in this chapter makes use of both static and dynamic features of the human tongue for personal authentication. Furthermore, liveness detection based on involuntary squirm is another highlight for tongue biometrics. In experiments based on our unique tongue image database, in every case, the fusion approach outperformed each of the single tongue representation approaches using geometric features, crack features, texture features, and squirm features alone. The dynamic tongueprint identifier using the fusion rule can achieve a 95% recognition rate in rank one in the identification mode and 4.1% for EER in the verification mode.

To sum up the above arguments, we can say that tongueprint is a qualified member of the biometrics family. The fact that the dynamic tongueprint is a feasible biometric identifier leaves a way open for its adoption in a variety of settings, such as law enforcement or where a system must be fool proof against counterfeiting with liveness detection, and in conjunction with other biometric modalities (Li and Juwei 1999; Lin and Anil 1998). Perhaps most obviously, tongue information can be combined into a feature template with face information, producing more accurate recognition performance. Ultimately, as with some other biometrics, the most difficult problem for the use of the tongueprint as a biometric identifier may be the need for the user's cooperation during the capture procedure.

References

- Abate A, Nappi M, Riccio D, Sabatino G (2007) 2D and 3D face recognition: a survey. *Pattern Recogn Lett* 28:1885–1906
- Bowyer K, Hollingsworth K, Flynn P (2007) Image understanding for iris biometrics: a survey. *Comput Vis Image Underst* 110(2):281–307
- Brand R, Isselhard D (1998) Anatomy of orofacial structures, 6th edn. St. Louis, Missouri
- Daugman J (1993) High confidence visual recognition of persons by a test of statistical independence. *IEEE Trans Pattern Anal Mach Intell* 15:1148–1161
- Daugman J (2004) How iris recognition works. *IEEE Trans Circuits Syst Video Technol* 14:21–30
- Haxby J, Hoffman E, Gobbini M (2000) The distributed human neural system for face perception. *Trends Cogn Sci* 4:223–233
- Jain A, Healey G (1998) A multiscale representation including opponent color features for texture recognition. *IEEE Trans Image Process* 7:124–128
- Jain A, Bolle R, Pankanti S (1998) Biometrics: personal identification in networked society. Kluwer Academic, Boston
- Jain A, Pankanti S, Prabhakar S, Lin H, Ross A (2004) Biometrics: a grand challenge. In: Proceedings of the 17th international conference on pattern recognition (ICPR), vol 2, pp 935–942
- Kittler J, Hatef M, Duin R, Matas J (1998) On combining classifiers. *IEEE Trans Pattern Anal Mach Intell* 20:226–239
- Knight B, Johnston A (1997) The role of movement in face recognition. *Vis Cogn* 4:265–274

- Kokiopoulou E, Saad Y (2007) Orthogonal neighborhood preserving projections: a projection-based dimensionality reduction technique. *IEEE Trans Pattern Anal Mach Intell* 29:2143–2156
- Kollreider K, Fronthaler H, Bigun J (2005) Evaluating liveness by face images and the structure tensor. In: Proceedings of the fourth IEEE workshop on automatic identification advanced technologies, pp 75–80
- Kollreider K, Fronthaler H, Bigun J (2009) Non-intrusive liveness detection by face images. *Image Vis Comput* 27(3):233–244
- Kong W, Zhang D, Li W (2003) Palmprint feature extraction using 2-D Gabor filters. *Pattern Recogn* 36:2339–2347
- Lee L, Berger T, Aviczer E (1996) Reliable online human signature verification systems. *IEEE Trans Pattern Anal Mach Intell* 18:643–647
- Li S, Juwei L (1999) Face recognition using the nearest feature line method. *IEEE Trans Neural Netw* 10:439–443
- Lin H, Anil J (1998) Integrating faces and fingerprints for personal identification. *IEEE Trans Pattern Anal Mach Intell* 20:1295–1307
- Moon Y, Chen J, Chan K, So K, Woo K (2005) Wavelet based fingerprint liveness detection. *Electron Lett* 41:1112–1113
- O’Gorman L (2003) Comparing passwords, tokens, and biometrics for user authentication. *Proc IEEE* 91:2021–2040
- OToole A, Roark D, Abdi H (2002) Recognizing moving faces: a psychological and neural synthesis. *Trends Cogn Sci* 6:261–266
- Pang B, Zhang D, Wang K (2005) The bi-elliptical deformable contour and its application to automated tongue segmentation in chinese medicine. *IEEE Trans Med Imaging* 24:946–956
- Phillips P, Hyeonjoon M, Rizvi S, Rauss P (2000) The FERET evaluation methodology for face-recognition algorithms. *IEEE Trans Pattern Anal Mach Intell* 22:1090–1104
- Pikaz A, Dinstein I (1995) Matching of partially occluded planar curves. *Pattern Recogn* 28:199–209
- Ratha N, Karu K, Shaoyun C, Jain A (1996) A real-time matching system for large fingerprint databases. *IEEE Trans Pattern Anal Mach Intell* 18:799–813
- Roweis S, Saul L (2000) Nonlinear dimensionality reduction by locally linear embedding. *Science* 290:2323–2326
- Sanchez-Reillo R, Sanchez-Avila C, Gonzalez-Marcos A (2000) Biometric identification through hand geometry measurements. *IEEE Trans Pattern Anal Mach Intell* 22:1168–1171
- Sandström M (2004) Liveness detection in fingerprint recognition systems. Department of Electrical Engineering, Linköping University
- Tistarelli M, Bicego M, Grossi E (2009) Dynamic face recognition: from human to machine vision. *Image Vis Comput* 27(3):222–232
- Toth B (2005) Biometrics liveness detection. *Inf Secur Bull* 10:291–297
- Turk M, Pentland A (1991) Eigenfaces for recognition. *J Cogn Neurosci* 3:71–86
- Wan V, Renals S (2005) Speaker verification using sequence discriminant support vector machines. *IEEE Trans Speech Audio Process* 13:203–210
- Wang L, Tan T, Ning H, Hu W (2003) Silhouette analysis-based gait recognition for human identification. *IEEE Trans Pattern Anal Mach Intell* 25:1505–1518
- Wenxin L, David Z, Zuoqun X (2002) Palmprint identification by Fourier transform. *Int J Pattern Recogn Artif Intell* 16:417–432
- Zhang D (2000) Automated biometrics – technologies and systems. Kluwer Academic, Boston
- Zhang D, Wai-Kin K, You J, Wong M (2000) Online palmprint identification. *IEEE Trans Pattern Anal Mach Intell* 25:1041–1050

Chapter 14

Online 3D Ear Recognition

Abstract The three-dimensional shape of the ear has been proven to be a stable candidate for biometric authentication because of its desirable properties such as universality, uniqueness, and permanence. In this chapter, a special laser scanner designed for online three-dimensional ear acquisition was described. Based on the dataset collected by our scanner, two novel feature classes were defined from a three-dimensional ear image: the global feature class (empty centers and angles) and local feature class (points, lines, and areas). These features are extracted and combined in an optimal way for three-dimensional ear recognition. Using a large dataset consisting of 2000 samples, the experimental results illustrate the effectiveness of fusing global and local features, obtaining an equal error rate of 2.2%.

Keywords 3D ear recognition • Feature extraction • Global and local feature classes

14.1 Introduction

Biometric authentication is of great importance for applications in public security (Pang et al. 2011; Zhang 2000; Jain et al. 1999). Nowadays, several novel biometrics, including palmprints (Zhang et al. 2003), veins (Zhang et al. 2007), and ears (Abaza et al. 2013; Abaza and Ross 2010; Hurley et al. 2005), have been developed to meet the needs of different security requirements.

With advances in three-dimensional (3D) imaging technology, 3D biometric authentication has drawn increasing attention from researchers. Examples include 3D face (Kakadiaris et al. 2007; Samir et al. 2006), palmprint (Zhang et al. 2009, 2010; Li et al. 2012), and ear recognition (Yan and Bowyer 2007; Chen and Bhanu 2007, 2009; Islam et al. 2011, 2012; Zhou et al. 2012). A 3D ear image is robust to imaging conditions, and contains surface shape information that is related to anatomical structure. In addition, it is insensitive to environmental illuminations. Yan (Yan and Bowyer 2007) utilized both color and depth images to determine the ear pit for automated 3D ear segmentation. Furthermore, they proposed an improved Iterative Closest Point (ICP) algorithm for 3D ear point cloud matching.

Chen (Chen and Bhanu 2007) gave a 3D ear recognition method founded on a Local Surface Patch (LSP) and ICP algorithm. Moreover, they proposed an indexing approach (Chen and Bhanu 2009) that combines feature embedding and a support vector machine-based learning technique for ranking their hypotheses. Islam et al. presented a local 3D features extraction method based on the key point detection (Islam et al. 2011, 2012). Zhou et al. presented a 3D ear recognition system combining local and holistic features (Zhou et al. 2012). Zhang et al. introduced a sparse representation framework into the field of 3D ear identification (Zhang et al. 2014). Chen and Mu proposed a hybrid multi-keypoint descriptor sparse representation-based classification (MKD-SRC) method to solve one sample per person problem in ear recognition (Liu et al. 2015).

Even though good results were achieved in these studies, there is no overall system for online 3D ear recognition. First, most of the current methods use commercial laser scanners to acquire the 3D range image, for example, the widely used Minolta VIVID Series (Zhang et al. 2014; Chen et al. 2015). Although these scanners are general-purpose and high-performance, they are expensive and cumbersome. Second, previous 3D ear recognition methods focused on a single aspect, that is, mostly local features, while global features such as the ear-parotic area angle, and the ear hole shape have not been discussed or used. Given these considerations, a laser scanner specifically designed for 3D ear acquisition and recognition was first developed using the laser-triangulation principle. The scanner provides 2D intensity images and 3D point-cloud data for subsequent recognition, and the total scanning and transmission time is less than 2 s. Based on the 3D ear images collected by our laser scanning device, two feature classes consisting of five features were defined. The empty center shape and the angle feature represent the depth and orientation of a 3D ear, and are treated as global features. The point, line, and area features describe key points, shapes, and the local area of the 3D ears. They are treated as local features. By combining these global features with local features, a hierarchical structure was introduced for 3D ear recognition. The 3D ears are pre-classified using global features and then recognized using local features. Thus, much time can be saved and accuracy can be improved in 3D ear recognition. Therefore, the 3D ear recognition system achieves both a high efficiency and accuracy.

The purpose of this study was to create a 3D ear recognition system using equipment that is practical for real applications. The main contents of this Chapter can be summarized as follows. Firstly, the global and local features categories in 3D ear are proposed. Secondly, multi-forms of features in 3D ears have been defined and extracted. Thirdly, multi-features fusion and hierarchical recognition of 3D ears have been discussed. Finally, a complete solution for 3D ear authentication has been achieved. The results on the collected 3D ear data show that the system is efficient and accurate.

14.2 Scanner Design for Online 3D Ear Acquisition

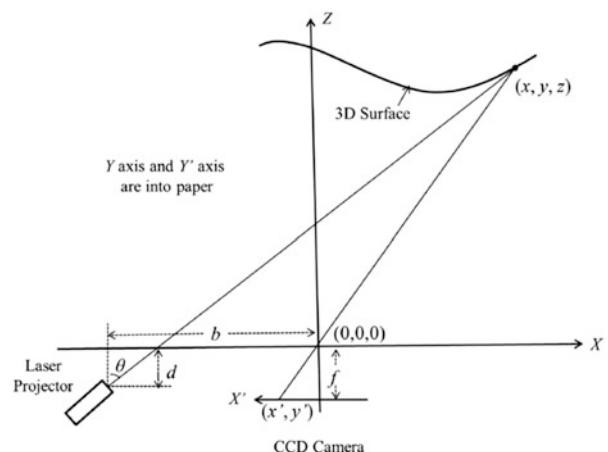
The 3D ear scanner we developed is based on the laser triangulation principle (Liu et al. 2015). Figure 14.1 illustrates the imaging principle of laser triangulation. In the reference X-Y-Z coordinates, the 3D coordinates (x , y , z) can be calculated according to Eq. (14.1).

$$\vec{p} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \frac{x'(b - d \tan \theta)}{x' + f \tan \theta} \\ \frac{y'(b - d \tan \theta)}{x' + f \tan \theta} \\ \frac{f'(b - d \tan \theta)}{x' + f \tan \theta} \end{pmatrix} \quad (14.1)$$

Figure 14.2 illustrates the framework of the 3D ear recognition system. The system consists of two main parts: hardware and software. To meet the requirements of online recognition, the hardware and software should be optimized for speed and accuracy. At the same time, its portability and cost for real applications should be considered. The laser scanner developed for 3D ear acquisition is shown in Fig. 14.3a. Figure 14.3b shows two groups of typical 3D ear samples captured by our device, where each row is the 3D point cloud from one ear viewed at different angles.

Table 14.1 provide a performance comparison of our proposed device with the Minolta Vivid 910 range scanner that is a widely used commercial scanner and has been used to acquire 3D ear data for UND data set. The acquisition time refers to the total scanning and transmission time, accuracy refers to the depth precision of the measurement, dimensions refer to the width, height and length of the scanner, in addition, the weight and price are also listed. Although the measurement accuracy of our acquisition system is inferior to that of Vivid 910, it has a higher speed,

Fig. 14.1 Imaging principle of laser-triangulation imaging



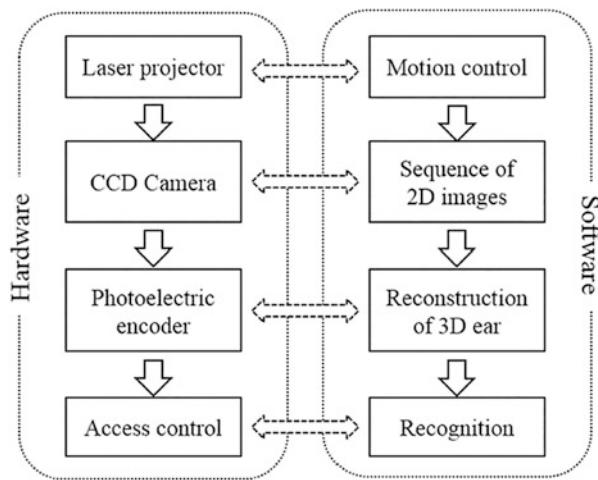


Fig. 14.2 Framework of the 3D ear recognition system

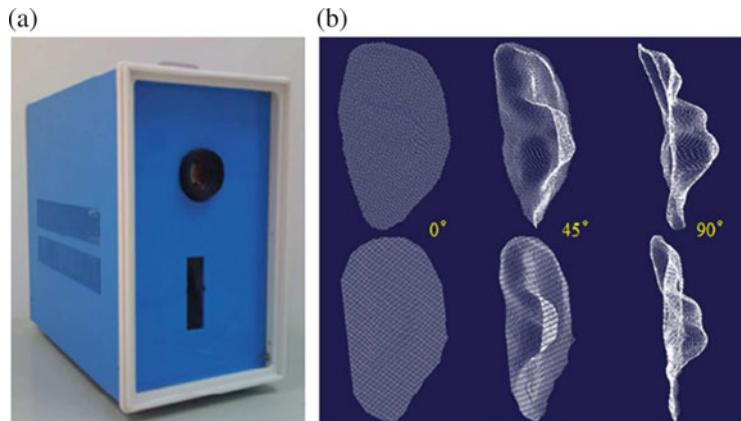


Fig. 14.3 Proposed 3D ear acquisition system: (a) 3D ear acquisition device and (b) 3D ear samples viewed at different angles (each row is collected from a single ear)

Table 14.1 Comparison of the scanning device

	Acquisition time (s)	Accuracy (mm)	Dimensions (mm)	Weight (kg)	Price (USD)
Vivid 910	4	±0.1	213 × 413 × 271	11	>20,000
Our scanner	2	±0.5	140 × 200 × 200	3	<1000

smaller size, and much lower cost. Moreover, the device could provide original frames of laser lines that describe the fundamental structure of 3D features. All these traits make the specially designed device suitable for 3D ear acquisition in practical biometrics applications.

A 3D ear database was established using the developed 3D ear acquisition device by collecting 3D ears on two separate occasions separated by an interval of around 1 month. On each occasion, the subject was asked to provide two samples. The database contains 2000 samples from 500 volunteers consisting of 341 males and 159 females. The volunteers were students and staff of the Shenzhen Graduate School of Harbin Institute of Technology. The written consents were obtained from the participants prior to the study. The study was approved by the Academic Committee of the Department of Computing of Harbin Institute of Technology, Shenzhen Graduate School, which ensures that research programs are consistent with academic ethics. The 3D ear acquisition study was discussed in a meeting of the committee, and written approval was subsequently granted by the Department Head. Because our research work does not involve patients or privacy, and all the participants have given written consent to the use of their ear images for academic purposes, all the data and figures are fully available from the figshare.com and Biometrics Research Center of The Hong Kong Polytechnic University.

14.3 3D Ear Global and Local Feature Classes

Prior to feature extraction, the 3D ears were normalized using a projection density method (Huang et al. 2009). After that, a 3D image of the ear is formed as a normalized posture in unified X-Y-Z coordinates, where all features are extracted from the 3D point cloud of the ear.

14.3.1 *Global Feature Class*

Two global features, empty center and angle, are defined in the proposed system.

Empty Center Feature

In the normalized X-Y coordinates, the boundary points of the ear were first detected (Fig. 14.4a), then the connected areas were labeled (Fig. 14.4b). The connected areas that are less than a threshold was removed then (Fig. 14.4c). Lastly, the connected pixels inside the ear were selected as the empty center feature (Fig. 14.4d).

The template matching technique is used to calculate the distance between two empty center features. The distance is defined as: $D = E_1 \oplus E_2 / E_1 \cup E_2$, where E_1 and E_2 are the empty center features of different samples. To avoid displacement interference, the test image was shifted by ± 40 pixels left-right and up-down, where the minimum distance is taken to be the difference of the two empty center areas (Fig. 14.5).

Figure 14.6 shows the empty center feature vectors extracted from Sample A, Sample B, and Sample C. Sample A and Sample B are from the same ear, and

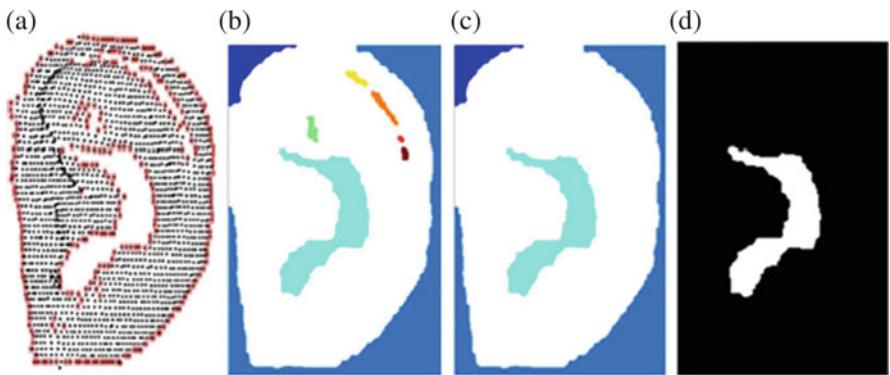


Fig. 14.4 Empty center feature extraction

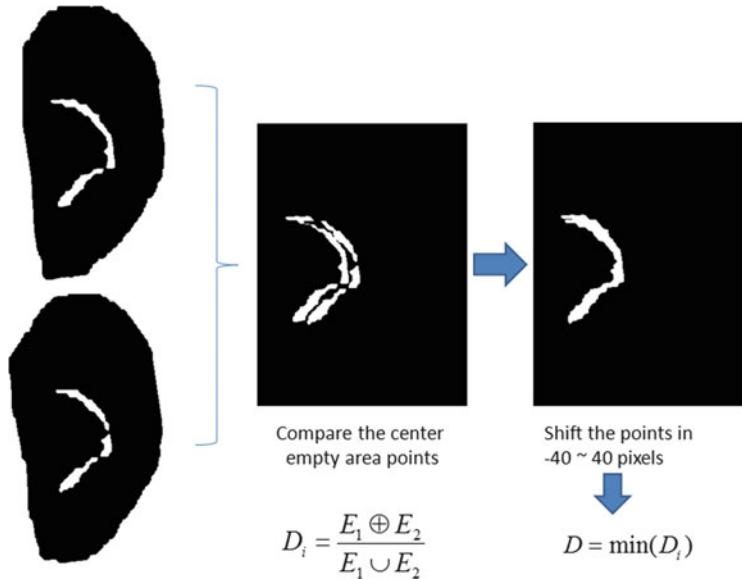


Fig. 14.5 Matching empty center features

Sample C is from a different ear. The distance between Sample A and Sample B is 0.23, and the distance between Sample B and Sample C is 0.56, which indicates that the empty center feature vectors from the same ear are alike and those from different ears are dissimilar.

Angle Feature

In Fig. 14.7, there is an angle between the ear and parotic area of a person (Liu et al. 2014). It can be assumed that there is a plane, $A_fx + B_fy + C_fz + D_f = 0$, which represents the 3D points on the parotic region (green circle shown in Fig. 14.8).

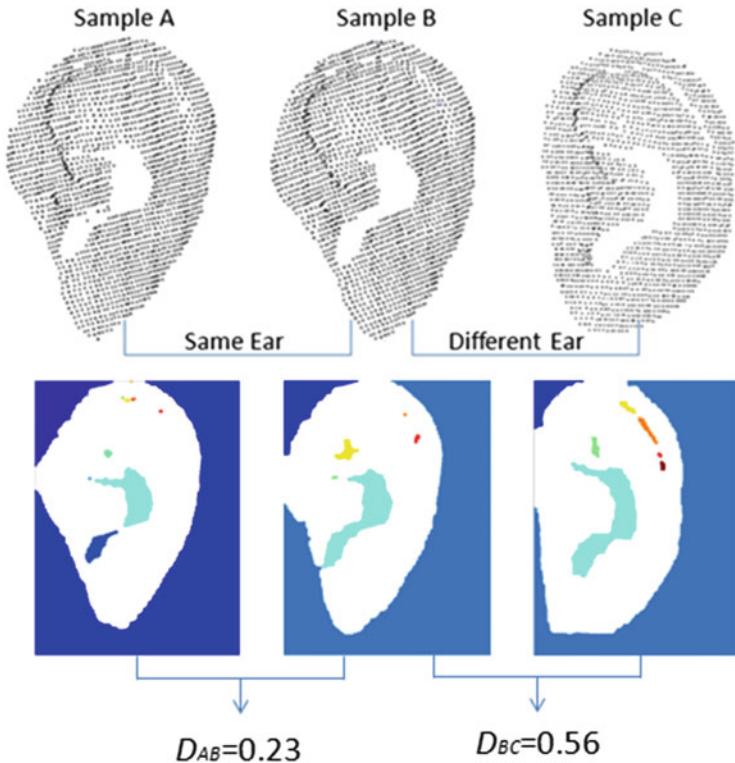


Fig. 14.6 Discriminating the same ear and different ears using the empty center feature vector

And there is another plane, $A_e x + B_e y + C_e z + D_e = 0$, represents the 3D points on the ear edge. Thus, the normal vector of the parotic plane can be obtained as $n_f = (A_f, B_f, C_f)$, and the normal vector of the ear plane is $n_e = (A_e, B_e, C_e)$. The angle θ between the parotic and ear planes can be defined as follows:

$$\theta_1 = \arccos \left(\frac{\langle n_f, n_e \rangle}{\|n_f\|_2 \|n_e\|_2} \right) \quad (14.2)$$

Where $\langle n_f, n_e \rangle$ is the inner product of normal vectors n_f and n_e . The $\|n_f\|_2$ and $\|n_e\|_2$ are L₂-norms of n_f and n_e respectively.

Hence,

$$\theta = \begin{cases} \theta_1 & \text{if } \theta_1 < 90^\circ \\ 180^\circ - \theta_1 & \text{otherwise} \end{cases} \quad (14.3)$$

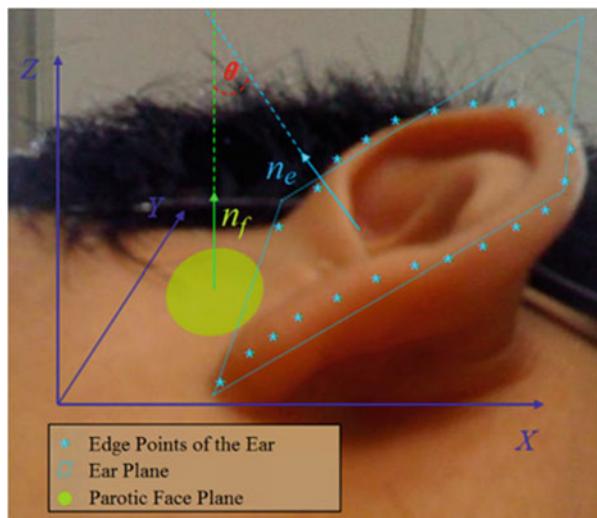


Fig. 14.7 Angle feature extraction

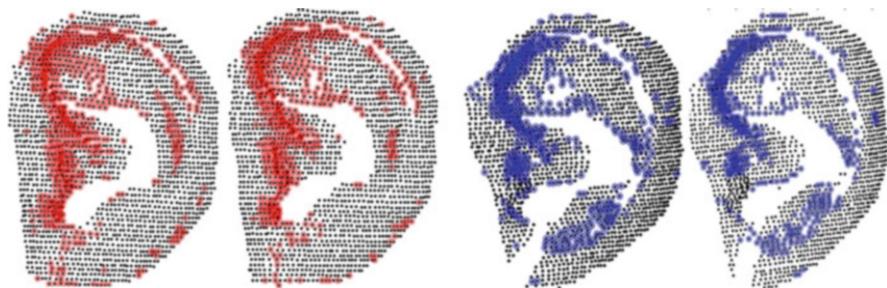


Fig. 14.8 Detected key-points: The *left* two samples are from one ear and the other two samples are from a different ear. The detected key-points are marked by *red* and *blue* respectively (Color figure online)

14.3.2 Local Feature Class

Three categories of local features in the 3D ear image were defined: point, line, and area features.

Point Feature

The 3D ear model consists of a number of points in 3D coordinates. Therefore, if the key points that are stable for the same ear and distinguishable for different ears could be found, then the 3D ear models would be recognized using these key points.

The aim of key-point detection is to select points on the 3D ear surface that can be identified with high repeatability in different models of the same surface. Islam and Mian proposed a key-point detection and feature extraction method that is effective on 3D ears (Islam et al. 2011) and faces (Mian et al. 2008). Although the

core of our key-point detection technique is similar to theirs, the technique is modified to make it suitable for the 3D ears data captured by our proposed device. In addition, the point feature is defined differently.

The input to the algorithm is a point cloud of the ear $E = \{P_1, \dots, P_n\}$. For each point $P(x_i y_i z_i)^T$, where $i = 1, \dots, n$, a local surface is cropped from the point cloud using a sphere of radius r centered at P and recorded as $\text{SetL} = \left\{ \left[x_j \ y_j \ z_j \right]^T \mid \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2 + (z_j - z_i)^2} < r \right\}$. The principle component analysis is then applied on the data points SetL . The difference between the eigenvalues along the first two principal axes of the local surface is computed as d . The value of d indicates the extent of asymmetry around the center point P , which is zero if SetL is planar or spherical. It is then compared to a threshold t , and if $d > t$, the point $P(x_i y_i z_i)^T$ is selected as a key-point. At the same time, the angular separation φ between the third principal axes and the original unified Z coordinate was calculated. Let $K_m = [x_m \ y_m \ z_m \ d_m \ \varphi_m]^T$ (where $m = 1, \dots, n_k$) record the key-point information. Set K_m is used at a later stage of feature extraction. Parameters r and t are empirically chosen as $r = 5$ mm and $t = 2$ mm.

Figure 14.8 shows examples of key-points detected on four different point clouds scanned from two individuals. It illustrates that key-points are stable in the ear data of the same individual, and distinguishable for the ear data of different individuals.

After key-point detection, features are extracted from set K_m (as shown in Fig. 14.9).

First, the normalized ear was divided into 12 average fan-shaped parts, where each sector is further divided into four equidistant parts. Each part is marked as F_l , where $l = 1, \dots, 48$. Thus, the x-y-z values of key-point set K_m fall within these 48 parts.

Second, for each F_l , the statistical histograms of d and φ were calculated. The histogram bins of d are set to 2, 3, 4, 5, 6, and 7, and the bins of φ are set to 0, 1, 2, and 3. Next, the number in each bin was counted to obtain a 10-dimensional vector. If there is no key-point in F_l , the vector was set to [0,0,0,0,0,0,0,0,0,0].

Finally, all 48 vectors were connected to obtain a 480-dimensional vector V_p as the final point feature vector. The difference between two ears is calculated using the Euclidean distance between their V_p vectors.

Figure 14.10 shows the point feature vectors extracted from different samples. Sample 1 (S1) and 2 (S2) are from the same ear, and Sample 3 (S3) is from a different ear. The red curve is the point feature vector of S1, the blue curve is the point feature vector of S2, and the black curve is the point feature vector of S3. The distance between S1 and S2 is 33.7, and the distance between S1 and S3 is 127.4. It can be seen that the point feature vectors from the same ear are very similar, and those from different ears are dissimilar.

Line Feature

To calculate the line feature, a rectangle was fitted on the normalized ear in the X-Y coordinates, define $(M+N)$ lines, V_1, \dots, V_m (which divides the rectangle equally in

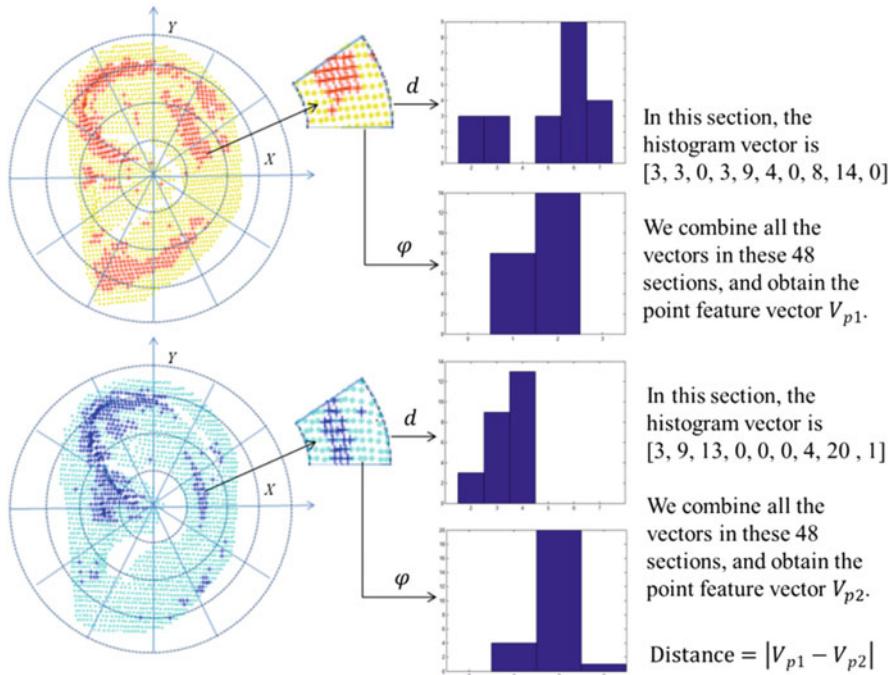


Fig. 14.9 Extraction of the point feature vector

the horizontal direction), and H_1, \dots, H_n (which divides the rectangle equally in the vertical direction), as shown in Fig. 14.11. Next, the 3D points on each line were obtained and their z values were recorded. Each line was then divided equally and the z crossing point values were marked as $z1, z2, \dots, z10$ (or $z1, z2, \dots, z20$ for V_1, \dots, V_m). These z values were used to form the line feature vector L ($V_1, \dots, V_m, H_1, \dots, H_n$), where the vector is of length $(20 \times m + 10 \times n)$.

Figure 14.12 shows the line feature vectors extracted from the same samples as those in Fig. 14.11. Parameters $m = 2$ and $n = 3$ were used in the experiment to test the discrimination of the line feature. The distance between S1 and S2 using the line feature is 7.02, and the distance between S1 and S3 is 41.12. It can be seen that the line feature vectors from the same ear are very close, and the line feature vectors from different ears are further apart.

Area Feature

In order to compute the area feature, the 3D ear was fitted into a fixed block and divided into $m \times n$ equal areas (see Fig. 14.13). All coordinate points in the area are defined as $(x_i, y_i, z_i) i = 1, \dots, N$, where N is the number of the points in the area. All the coordinates of these points constitute an $N \times 3$ matrix W as follows:

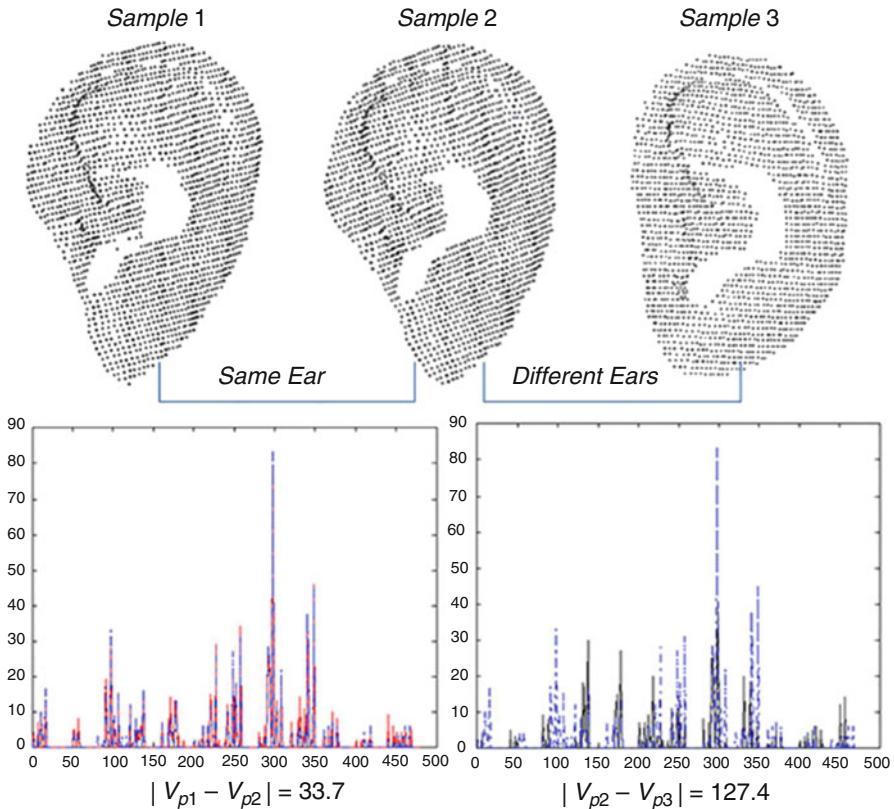


Fig. 14.10 Discriminating between the same and different ears using the point feature vector

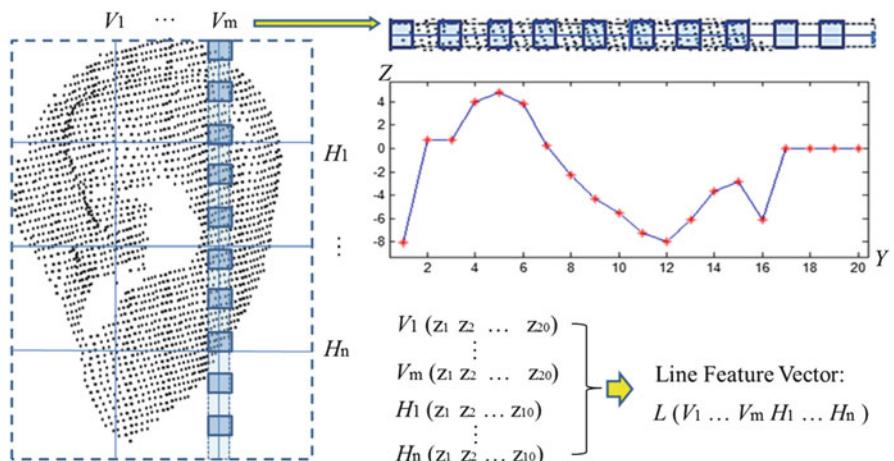


Fig. 14.11 Extraction of the line feature vector

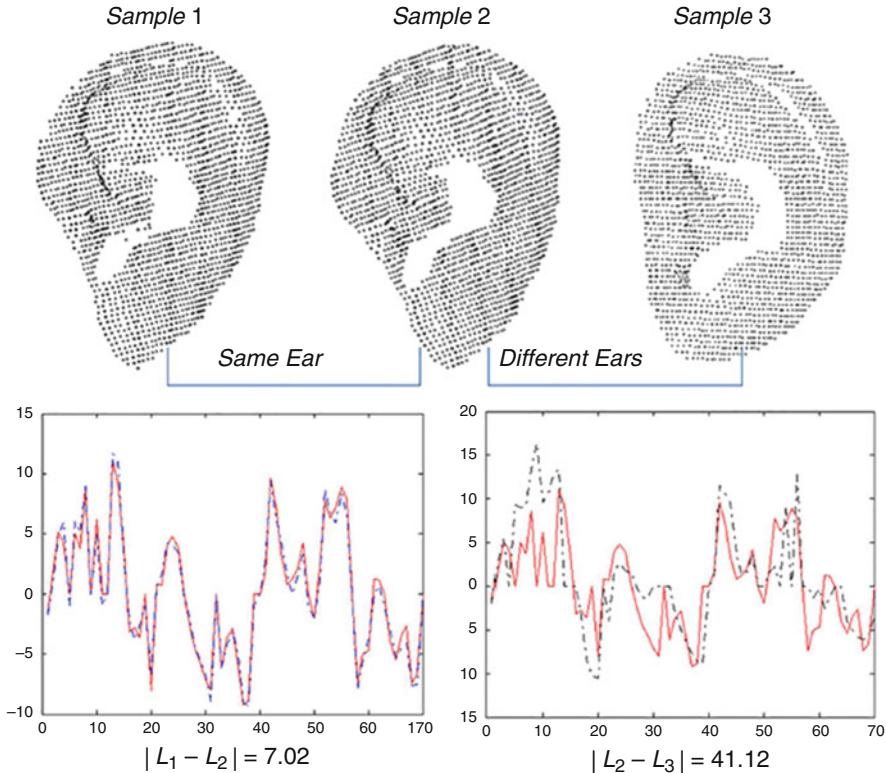


Fig. 14.12 Discriminating between the same and different ears using the line feature vector

$$W = \begin{bmatrix} x_1 & y_1 & z_1 \\ \dots \\ x_N & y_N & z_N \end{bmatrix} \quad (14.4)$$

Principle component analysis (Pang et al. 2010a, b) is performed on W and the resulting normal vector is represented as $V_N(i, j, k)$.

The average is calculated using

$$(\bar{x}, \bar{y}, \bar{z}) = \frac{1}{N} \sum_{i=1}^N (x_i, y_i, z_i) \quad (14.5)$$

The scatter matrix is given as $S = \sum_{i=1}^N (W_i - \bar{W}_i) \times (W_i - \bar{W}_i)^T$, the eigenvectors of S are Φ , and the first column of Φ is the normal vector $V_N(i, j, k)$. It is clear that $V_N(i, j, k)$ can be thought of as the direction of matrix W . In addition, the center of gravity of W can be represented as $V_C(\bar{x}, \bar{y}, \bar{z})$. As a result, the normal vector V_N , center of gravity VC , and min/max z values VZ are calculated and joined to form a

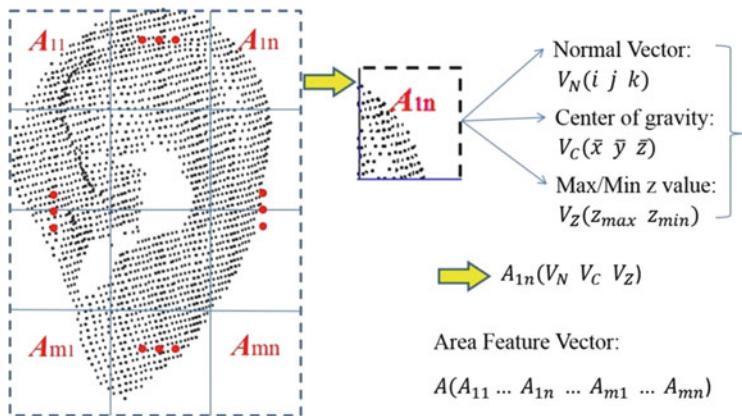


Fig. 14.13 Extraction of the area feature vector

vector AN for each area. The area feature subsequently becomes the vector consisting of all $m \times n$ vectors A , ($A_{11}, A_{12}, \dots, A_{mn}$). Figure 14.14 shows the area feature vectors extracted from S1, S2, and S3. The distance between S1 and S2 is 6.89, and the distance between S2 and S3 is 27.78, which indicates that the area feature vectors from the same ear are alike and those from different ears are not alike.

14.4 Experimental Results and Discussion

The experiments were divided into two parts: feature optimization and verification experiments. As mentioned above, our database contains a total of 2000 different samples from 500 individual ears. A PC with Intel Core 2 CPU @2.33 GHz and 2 GB memory was used in our experiments.

14.4.1 Feature Optimization

Because the parameters used in the definition of each local feature may influence the length of the feature vector as well as the equal error rate (EER) of the verification experiments, the feature optimization experiments were performed to determine the most effective values for these parameters.

In our point feature, the number and distribution of the key-points determines the point feature vector. Hence, threshold t is the parameter that needs to be optimized. Figure 14.15 shows the different key-points extracted using different thresholds, while Table 14.2 shows the EER for different thresholds. Considering the time consumed, the feature optimization experiments were performed on a sub-dataset

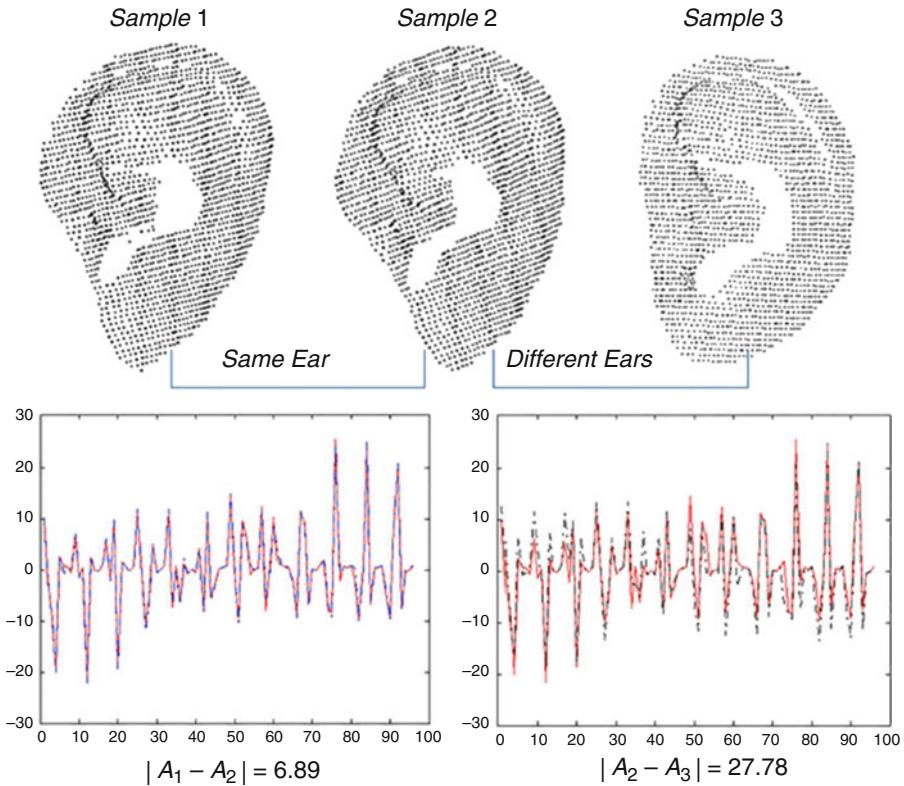


Fig. 14.14 Discriminating between the same and different ears using the area feature vector

that contains 100 different sample ears. From Table 14.2, it can be seen that the best result is achieved when $t=2$.

The line feature vector is determined by the number of horizontal and vertical lines. Therefore, the line number is the parameter that needs to be optimized here. Figure 14.16 shows the different lines across the ear. Table 14.3 shows the EER obtained using different line numbers, where 12 lines obtains the lowest EER.

Because the number of blocks determines the area feature vector, this parameter is the one that must be optimized. Figure 14.17 shows the different blocks on the ear and Table 14.4 shows the EER obtained using different block numbers. It can be seen that the best result is achieved when there are 48 blocks.

14.4.2 Matching Using Local Features

The matching experiments were carried on all 2000 samples, and performed using the local feature class (point, line, and area features) as well as their feature-level

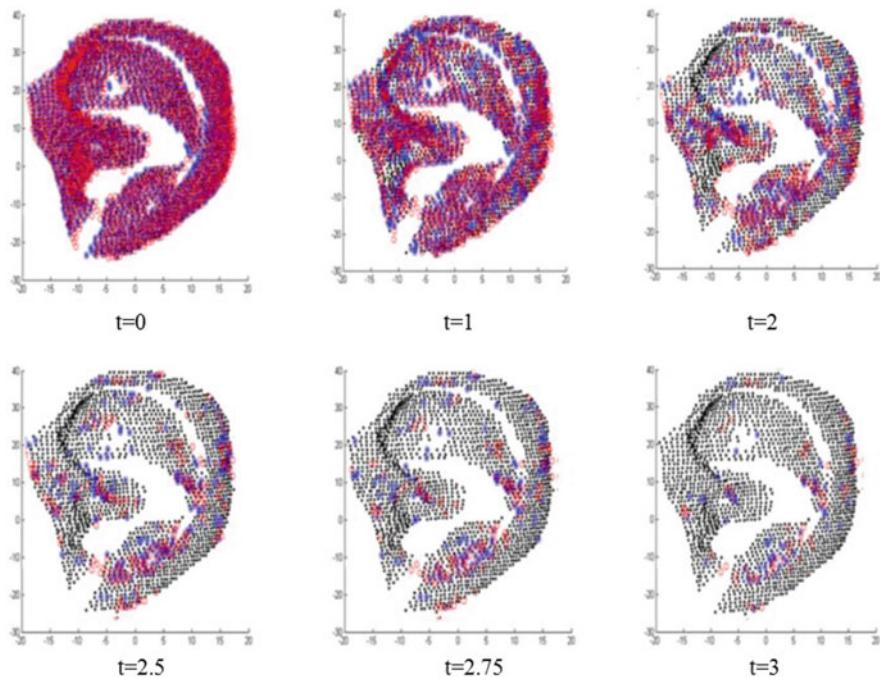


Fig. 14.15 Point feature optimization

Table 14.2 Point features with different t parameters

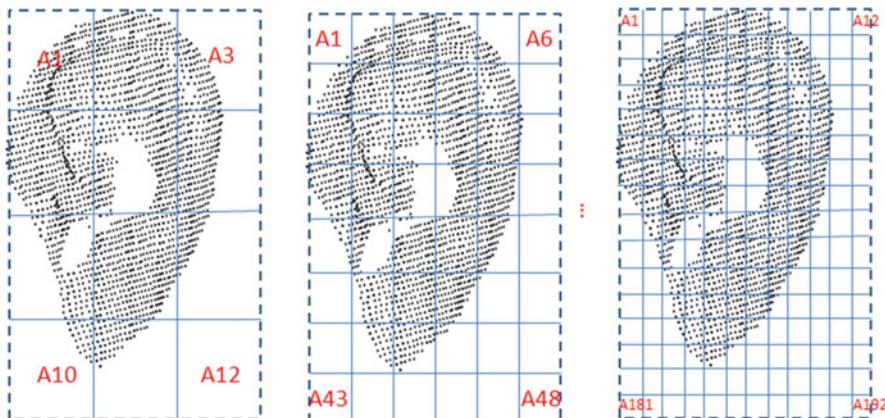
t	0	1	2	2.5	2.75	3
No. Points	2009	1320	462	207	136	82
EER (%)	2.5	2.2	2.2	5.3	7.8	17.2



Fig. 14.16 Line feature optimization

Table 14.3 Line features for different line numbers

Line No.	5	12	26
Vector length	70	170	370
EER (%)	3.6	3.0	3.2

**Fig. 14.17** Area feature optimization

fusion. Since all the local features (point, line, and area features) are defined in form of vectors (VP , VL , VA), the most direct strategy for feature-level fusion is to joint different vectors into one fusion feature vector. Therefore, the fusion feature vectors can be described as follows:

$$\left\{ \begin{array}{l} V_{P+L} = \text{normalization}(\text{joint}(\text{normalization}(V_P), \text{normalization}(V_L))) \\ V_{P+A} = \text{normalization}(\text{joint}(\text{normalization}(V_P), \text{normalization}(V_A))) \\ V_{L+A} = \text{normalization}(\text{joint}(\text{normalization}(V_L), \text{normalization}(V_A))) \\ V_{local} = \text{normalization}(\text{joint}(\text{normalization}(V_{P+L}), \text{normalization}(V_A))) \end{array} \right. \quad (14.6)$$

The function `normalization` normalizes the feature vector into unit vector. The function `joint` combines two feature vectors into one fusion feature vector. Table 14.5 shows the EER results of different local features and their combinations. It can be seen that the optimal result is achieved when all local features are fused together.

14.4.3 Recognition with Global Feature Indexing

Different from the weighted fusion method, the global and local features fusion is implemented in a hierarchical procedure. The 3D ears are pre-classified using global features and then recognized using local features. Thus, much time can be

Table 14.4 Area features with different block numbers

No. Blocks	12	48	192
Vector length	108	432	1728
EER (%)	5.0	4.3	4.3

Table 14.5 Matching results for different local features

Features	Point	Line	Area	Point + Line	Point + Area	Line + Area	All features
EER (%)	4.7	4.2	5.1	3.3	4.5	3.6	2.8

saved and accuracy can be improved in 3D ear recognition. The flowchart of the overall recognition with global feature indexing is shown in Fig. 14.18a. For a given ear sample, the procedure is as follows:

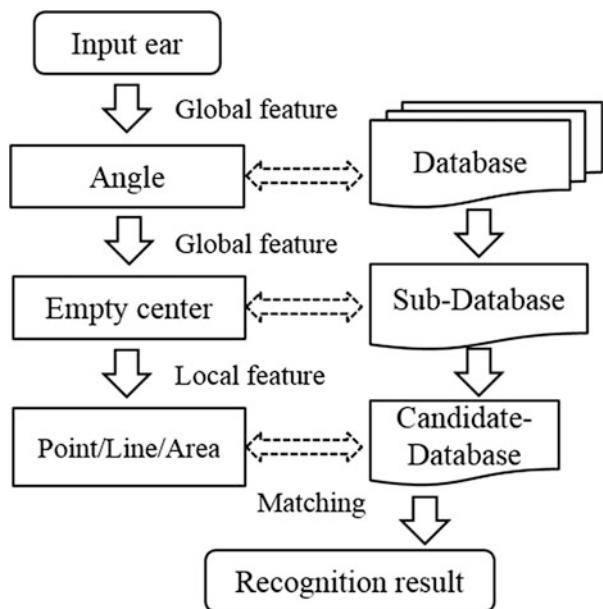
1. Extract the global features of the test sample Angle (G_t), Center (G_t).
2. Compare Angle (G_t) with global features Angle (G_i) $i = 1, \dots, N$ of all ear models (in our experiments, $N = 500$) to obtain the matching distance Dist [Angle (G_t), Angle (G_i)].
3. If Dist [Angle (G_t), Angle(G_i)] is smaller than threshold $T(\beta)$, the ear model is treated as a matched candidate and place it into a sub-database.
4. Match test ear G_t with the sub-database ears using the empty center feature and adjust the candidate sub-database G_i accordingly.
5. Extract the local features of $VLocal_t$ and the local features of the ear models in the candidate sub-database $VLocal_i$ ($i = 1, \dots, k$), where k is the total number of ears it contains.
6. Match local features between $VLocal_t$ and $VLocal_i$ to measure the differences between the test ear and candidate ears (in our experiments, the Euclidean distance was used).
7. The candidate ear that is closest to the test ear is the recognition result.

Figure 14.18b shows the receiver operating characteristic curve of the results obtained by combining both global and local features together, where the EER is 2.2%. It can be seen that the fusion of global and local features achieves the smallest EER of all schemes, and is even better than single feature matching. This is reasonable, because more information usually leads to more accurate recognition.

14.4.4 Performance Analysis

To better measure the performance of the proposed method, six criteria (database, acquisition device, feature extraction method, average matching time, EER, and online properties) were used to compare the proposed method with other 3D ear recognition methods. The results are shown in Table 14.6.

Fig. 14.18 Fusion of global and local features: (a) flowchart of recognition with global feature indexing and (b) receiver operating characteristic curve of the global and local feature fusion



From Tables 14.1 and 14.6, it can be seen that our 3D ear scanner has a lower price (approximately 5% that of the Vivid 910), and a smaller size (approximately 25% that of the vivid 910). Meanwhile, the overall recognition time (including acquisition and recognition time) is less than 2.5 s, and the EER on a database with 2000 samples is 2.2%. So far, our 3D ear recognition system is the only system offering an overall solution for both 3D ear data acquisition and optimized recognition. Its performance is sufficient to meet the online system requirements for a real-time application.

14.5 Summary

In this chapter, two novel feature classes, global and local features, were defined and extracted from 3D ear point clouds. The global feature class includes the empty center and ear-parotic area angle, whereas the local feature class consists of point, line, and area features. The experimental results show that all features are stable for the same ear and distinguishable between different ears. Furthermore, global features can be used for indexing, while the combination of both global and local features produces matching results with an EER of 2.2% on our 3D ear database of 2000 samples. Using our own developed scanner and the optimized recognition method, a real-time 3D ear recognition system is achieved.

Table 14.6 Comparison with existing 3D ear recognition methods

Reference	Dataset size (ID/subjects/images)	Device used	Method applied	Matching time (s)	Reported EER	Online
Yan and Bowyer	UND/415/1386	Vivid 910	ICP	1.5	1.2%	N/A
Chen and Bhanu	UND-F/302/942 UCR/155/902	Vivid910 Vivid300	LSP	3.7	2.3%–4.2%	N/A
Islam et al.	UND-F/302/942 UND-J/415/830	Vivid 910	L3DF + ICP	0.06	2.3%–4.1%	N/A
Proposed method	HIT/500/2000	Our scanner	Global + Local	0.5	2.2%	YES

References

- Abaza A, Ross A (2010) Towards understanding the symmetry of human ears: a biometric perspective. In: Fourth IEEE international conference on biometrics: theory applications and systems (BTAS), pp. 1–7
- Abaza A, Ross A, Hebert C, Harrison MAF, Nixon MS (2013) A survey on ear biometrics. ACM Comput Surv (CSUR) 45(2):22
- Burge M, Burger W (2000) Ear biometrics in computer vision. In: 15th International conference on pattern recognition, pp 822–826
- Chen H, Bhanu B (2007) Human ear recognition in 3D. IEEE Trans Pattern Anal Mach Intell 29 (4):718–737
- Chen H, Bhanu B (2009) Efficient recognition of highly similar 3D objects in range images. IEEE Trans Pattern Anal Mach Intell 31(1):172–179
- Chen L, Mu Z, Zhang B, Zhang Y (2015) Ear recognition from one sample per person. PLoS One 10(5):e0129505
- Choraś M (2005) Ear biometrics based on geometrical feature extraction. Electro Lett Comput Vis Image Anal 5(3):84–95
- Huang C, Lu G, Liu Y (2009) Coordinate direction normalization using point cloud projection density for 3D ear. In: Fourth international conference on computer sciences and convergence information technology, pp 511–515
- Hurley DJ, Nixon MS, Carter JN (2005) Force field feature extraction for ear biometrics. Comput Vis Image Underst 98(3):491–512
- Iannarelli AV (1989) Ear identification. Paramount Publishing, Freemont, CA
- Islam SM, Davies R, Bennamoun M, Mian AS (2011) Efficient detection and recognition of 3D ears. Int J Comput Vis 95(1):52–73
- Islam S, Bennamoun M, Owens RA, Davies R (2012) A review of recent advances in 3D ear-and expression-invariant face biometrics. ACM Comput Surv (CSUR) 44(3):14
- Jain A, Bolle R, Pankanti S (1999) Biometrics: personal identification in networked society. Springer Science & Business Media, New York
- Kakadiaris IA, Passalis G, Toderici G, Murtuza MN, Lu Y, Karampatziakis N et al (2007) Three-dimensional face recognition in the presence of facial expressions: an annotated deformable model approach. IEEE Trans Pattern Anal Mach Intell 29(4):640–649
- Li W, Zhang D, Lu G, Luo N (2012) A novel 3-D palmprint acquisition system. IEEE Trans Syst Man Cybernet A Syst Hum 42(2):443–452
- Liu Y, Zhang B, Zhang D (2014) Ear-parotic face angle: a unique feature for 3D ear recognition. Pattern Recogn Lett 53:9–15
- Liu Y, Lu G, Zhang D (2015) An effective 3D ear acquisition system. PLoS One 10(6):e0129439

- Mian AS, Bennamoun M, Owens R (2008) Keypoint detection and local feature matching for textured 3D face recognition. *Int J Comput Vis* 79(1):1–12
- Pang Y, Wang L, Yuan Y (2010a) Generalized KPCA by adaptive rules in feature space. *Int J Comput Math* 87(5):956–968
- Pang Y, Li X, Yuan Y (2010b) Robust tensor analysis with 1-norm. *IEEE Trans Circuits Syst Video Technol* 20(2):172–178
- Pang Y, Yuan Y, Li X, Pan J (2011) Efficient HOG human detection. *Signal Process* 91 (4):773–781
- Purkait R, Singh P (2008) A test of individuality of human external ear pattern: its application in the field of personal identification. *Forensic Sci Int* 178(2):112–118
- Samir C, Srivastava A, Daoudi M (2006) Three-dimensional face recognition using shapes of facial curves. *IEEE Trans Pattern Anal Mach Intell* 28(11):1858–1863
- Yan P, Bowyer KW (2007) Biometric recognition using 3D ear shape. *IEEE Trans Pattern Anal Mach Intell* 29(8):1297–1308
- Zhang D (2000) Automated biometrics: technologies and systems. Springer Science & Business Media, Berlin
- Zhang D, Kong W-K, You J, Wong M (2003) Online palmprint identification. *IEEE Trans Pattern Anal Mach Intell* 25(9):1041–1050
- Zhang YB, Li Q, You J, Bhattacharya P (2007) Palm vein extraction and matching for personal authentication. In: International conference on advances in visual information systems. Springer, Berlin, pp 154–164
- Zhang D, Lu G, Li W, Zhang D, Luo N (2009) Palmprint recognition using 3-D information. *IEEE Trans Syst Man Cybernet C Appl Rev* 39(5):505–519
- Zhang D, Kanhangad V, Luo N, Kumar A (2010) Robust palmprint verification using 2D and 3D features. *Pattern Recogn* 43(1):358–368
- Zhang L, Ding Z, Li H, Shen Y (2014) 3D ear identification based on sparse representation. *PLoS One* 9(4):e95506
- Zhou J, Cadavid S, Abdel-Mottaleb M (2012) An efficient 3-D ear recognition system employing local and holistic features. *IEEE Trans Inf Forensics Secur* 7(3):978–991

Chapter 15

Book Review and Future Work

Abstract Traditional biometric technologies, such as fingerprint, face, iris, and palmprint, have been well studied and introduced in many research books. However, these technologies have their own advantages and disadvantages, and there is not one kind of biometric technology can be fit for different applications. Many new biometric technologies have been developed in recent years, especially for some new applications. This book describes some new biometric technologies, such as High resolution fingerprint, Finger-Knuckle-Print, Multi-Spectral Backhand, 3D fingerprint, Tongueprint, and 3D ear. Many efficient feature extraction, matching and fusion algorithms are introduced and some potential systems have been developed in this book. It may serve as a handbook of biometrics authentication and be of use to researchers and students who wish to understand, participate, and/or develop a biometrics authentication system. It would also be useful as a reference book for a graduate course on biometrics.

Keywords Biometrics application • Spoofing attack

Traditional biometric technologies, such as fingerprint, face, iris, and palmprint, have been well studied and introduced in many research books. However, these technologies have their own advantages and disadvantages, and there is not one kind of biometric technology can be fit for different applications. Many new biometric technologies have been developed in recent years, especially for some new applications.

This book describes some new biometric technologies, such as High resolution fingerprint, Finger-Knuckle-Print, Hand back skin texture, 3D fingerprint, Tongueprint, and 3D ear. Many efficient feature extraction, matching and fusion algorithms are introduced and some potential systems have been developed in this book. It may serve as a handbook of biometrics authentication and be of use to researchers and students who wish to understand, participate, and/or develop a biometrics authentication system. It would also be useful as a reference book for a graduate course on biometrics.

In this chapter, we first recapitulate the contents of this book in Sect. 15.1. Then, Sect. 15.2 discusses the future of biometrics research.

15.1 Book Recapitulation

This book includes 15 chapters, and the contents are represented into four main parts: High Resolution Fingerprint Recognition, Finger-Knuckle-Print Verification, Other Hand-Based Biometrics, and Some New Head-Based Biometrics. These different techniques have been implemented into different biometric systems, respectively. The experimental results under different challenging conditions have shown the superiority of these techniques.

Chapter 1 introduces recent developments in biometrics technologies, some key concepts in biometrics, and the importance of developing new biometrics systems.

Chapter 2 proposes a new approach to aligning high resolution partial fingerprints based on pores, a type of fingerprint fine ridge features, which are abundant on even small fingerprint areas. Comparing with representative minutia based and orientation field based methods, the experimental results show that the proposed method can more accurately locate corresponding feature points, estimate the alignment transformations, and hence significantly improve the accuracy of high resolution partial fingerprint recognition.

Chapter 3 presents a dynamic anisotropic pore model to describe pores more accurately by using orientation and scale parameters. The fingerprint image is first partitioned into well-defined, ill-posed, and background blocks. According to the dominant ridge orientation and frequency on each foreground block, a local instantiation of appropriate pore model is obtained. Finally, the pores are extracted by filtering the block with the adaptively generated pore model. Extensive experiments are performed on the high resolution fingerprint databases. The results demonstrate that the proposed method can detect pores more accurately and robustly, and consequently improve the fingerprint recognition accuracy of pore-based AFRS.

Chapter 4 discusses the optimal resolution for an AFRS using the two most representative fingerprint features: minutiae and pores. We first designed a multi-resolution fingerprint acquisition device to collect fingerprint images at multiple resolutions and captured fingerprints at various resolutions but at a fixed image size. Then, we carried out a theoretical analysis to identify the minimum required resolution for fingerprint recognition using minutiae and pores.

Chapter 5 presents a new biometric authentication system using finger-knuckle-print (FKP) imaging. A specific data acquisition device is constructed to capture the FKP images, and then an efficient FKP recognition algorithm is presented to process the acquired data in real time. The local convex direction map of the FKP image is extracted, based on which a local coordinate system is established to align the images and a region of interest is cropped for feature extraction. For matching two FKPs, a feature extraction scheme which combines orientation and magnitude information extracted by Gabor filtering is proposed.

Chapter 6 studies image local features induced by the phase congruency model, which is supported by strong psychophysical and neurophysiological evidences, for FKP recognition. In the computation of phase congruency, the local orientation and the local phase can also be defined and extracted from a local image patch. These

three local features are independent of each other and reflect different aspects of the image local information. We compute efficiently the three local features under the computation framework of phase congruency using a set of quadrature pair filters. We then propose to integrate these three local features by score-level fusion to improve the FKP recognition accuracy. Such kinds of local features can also be naturally combined with Fourier transform coefficients, which are global features.

Chapter 7 explores an effective FKP recognition scheme by extracting and assembling local and global features of FKP images. Specifically, the orientation information extracted by the Gabor filters is coded as the local feature. By increasing the scale of Gabor filters to infinite, actually we can get the Fourier transform of the image, and hence the Fourier transform coefficients of the image can be taken as the global features. Such kinds of local and global features are naturally linked via the framework of time–frequency analysis. The proposed scheme exploits both local and global information for the FKP verification, where global information is also utilized to refine the alignment of FKP images in matching. The final matching distance of two FKP is a weighted average of local and global matching distances.

Chapter 8 proposes a method to deal with pose variation problem in FKP system by reconstructing the query sample with a dictionary learned from the template samples in the gallery set. The reconstructed FKP image can reduce much the enlarged matching distance caused by finger pose variations; however, both the intra-class and inter-class distances will be reduced. We then propose a score level adaptive binary fusion rule to adaptively fuse the matching distances before and after reconstruction, aiming to reduce the false rejections without increasing much the false acceptances.

Chapter 9 introduces several popular features in 3D fingerprint recognition, such as scale invariant feature transformation (SIFT) feature, ridge feature and minutiae. To extract these fingerprint features accurately, an improved fingerprint enhancement method has been proposed by polishing orientation and ridge frequency maps according to the characteristics of 2D touchless fingerprint images. Therefore, correspondences can be established by adopting hierarchical fingerprint matching approaches. Through an analysis of 440 3D point cloud finger data (220 fingers, 2 pictures each) collected by a 3D scanning technique, i.e., the structured light illumination (SLI) method, the finger shape model is estimated. It is found that the binary quadratic function is more suitable for the finger shape model than the other mixed model.

Chapter 10 studies the hand back skin texture (HBST) pattern recognition problem with applications to personal identification and gender classification. A specially designed system is developed to capture HBST images, and an HBST image database was established, which consists of 1920 images from 80 persons (160 hands). An efficient text on learning based method is then presented to classify the HBST patterns. First, textons are learned in the space of filter bank responses from a set of training images using the l_1 -minimization based sparse representation (SR) technique. Then, under the SR framework, we represent the feature vector at each pixel over the learned dictionary to construct a representation coefficient

histogram. Finally, the coefficient histogram is used as skin texture feature for classification.

In Chap. 11, we propose a novel palmprint acquisition system based on the line-scan image sensor. The proposed system consists of a customized and highly integrated line-scan sensor, a self-adaptive synchronizing unit, and a field-programmable gate array controller with a cross-platform interface. The volume of the proposed system is over 94% smaller than the existing palmprint systems, and the verification performance of the proposed system is comparable to the best area camera-based systems.

Chapter 12 explores a novel hand-based biometric system, door knob hand recognition system (DKHRS), is proposed. DKHRS has the identical appearance of a conventional door knob, which is an optimum solution in both physiological factors and psychological factors. In this system, a hand image is captured by door knob imaging scheme, which is a tailored omnivision imaging structure and is optimized for this predetermined door knob appearance. Then features are extracted by local Gabor binary pattern histogram sequence method and classified by projective dictionary pair learning.

Chapter 13 presents a new number for the biometrics family, i.e. tongueprint, which uses particularly interesting properties of the human tongue to base a technology for noninvasive biometric assessment. The tongue is a unique organ which can be stuck out of the mouth for inspection, whose appearance is amenable to examination with the aid of a machine vision system. Yet it is otherwise well protected in the mouth and difficult to be forged. Furthermore, the involuntary squirm of the tongue is not only a convincing proof that the subject is alive, but also a feature for recognition. That is to say, the tongue can present both static features and dynamic features for authentication. However, little work has hitherto been done on the tongue as a biometric identifier. In this work, we make use of a database of tongue images obtained over a long period to examine the performance of the tongueprint as a biometric identifier. Our research shows that tongueprint is a promising candidate for biometric identification and worthy of further research.

In Chap. 14, in order to expand the use of palmprint biometrics, we propose a novel palmprint acquisition system based on the line-scan image sensor. The proposed system consists of a customized and highly integrated line-scan sensor, a self-adaptive synchronizing unit, and a field-programmable gate array controller with a cross-platform interface. The volume of the proposed system is over 94% smaller than the existing palmprint systems, and the verification performance of the proposed system is comparable to the best area camera-based systems.

15.2 Future Work

Biometrics systems have been explored and developed in last decades and also made great achievements. Large numbers of manufacturing and engineering enterprises, even education industry, government and military departments, have

implemented biometrics systems as a form of identification and access control for the purpose of security. The world biometric techniques and applications are undergoing revolutionarily changing, and we expect a great future for biometrical systems. Our vision for the future of biometrics systems and applications are as follows:

1. Biometrics systems can completely replace the conventional identification systems. The conventional identification systems, knowledge-based and token-based authentication, nowadays are still the mainstream for personal identification. When it comes to their drawbacks and flaws, biometrics can overtake these traditional systems by its overwhelming beneficial properties. Biometrics systems can secure your text messages, phonebook contacts and electronic images on your devices. Since data thieves can be implemented in many ways by obtaining conventional identification patterns, passwords or number codes, these patterns can only protect a device to a certain degree. However, biometrics systems can prevent these frauds from occurring in the first place. Thus the flourish of biometrics systems is unpredictable and inevitable. Meanwhile, significant efforts are still required.
2. Multi-biometrics systems can be widely used in identification systems. Hybrid multi-modal biometrics systems can provide more accurate identification results. Multi-biometrics systems can simultaneously capture and utilize more than one biologic attributes or characteristics, such as both fingerprint and finger vein images with the single touch of a finger for more precise identification. The other application example is to combine fingerprint scanners and voice recognition as forms of biometric security. It will become popular multi-biometrics techniques that biometrical hardware systems incorporate state-of-the-art technologies integrating more than one separate sensor into a single unit. However, fundamental research is still required in this area.
3. Low cost, efficient and effective identification performance. Practical identification systems should satisfy the requirements which are lightweight, flexible, secure, efficient, accurate and durable hardware requiring low power requirements. Researchers are trying to find the way that how to effectively and efficiently represent and recognition biometrics patterns. Future biometrics systems can quickly and accurately capture biology characteristics regardless of the conditions or hash physical environment. Moreover, more accurate biometrics systems can be developed for serving for end users, such as recognizing a person with 99.999% accuracy. It is a tendency that consumers and businesses will progress more comfortable with using them as well. More research efforts are needed on how to tactfully design biometrics systems.
4. The thriving use biometrics for healthcare systems. Multi-biometric patient safety systems can be developed for patient safety. There is a rise in the use of iris biometrics for patient identification in healthcare. This application is to identify the patients by their biometrics characteristics, and then medical institutions can promptly retrieve the patients' medical records. To this end, healthcare institutions can introduce this technology as a means to help prevent

the creation of duplicate medical records and overlays, predigest the registration process, eliminate medical identity theft and fraud at the point of service. Moreover, it also can protect patient privacy and heighten patient safety. Thus, the biometrics system can perform an irreplaceable role in healthcare service. However, much more research efforts still should be paid in innovative mechanisms of biometric systems used in healthcare system.

5. Overcoming external limitations and influences in biometrics authentication. Varieties of limitations are still unsolved when biometrics systems perform the operations of identification exploiting any single biometric feature or traits. External negative influences on identification accuracy are produced by these limitations such as nonuniversality, noisy sensor data, lack of individuality, intra-class variations and spoofing attacks. Some other factors also should be pointed out for study in depth. For example, capabilities of fingerprints identification systems should adjust to identify these images captured with moist or dry fingerprints in any type of weather conditions, even extreme in nature. Facial identification systems suffer the negative effects from the variations of poses, facial expression, disguises, illumination conditions and environments. All these limitations and influences can remarkably influence the identification accuracy. Biometric research and development institutions should pay much close attention to explore the approaches to overcome these limitations. Although abundant competing algorithms and techniques have been proposed and applied in biometrics recently, significant efforts are still required.

6. High-Resolution and the 3D Biometrics applications. High-resolution biometrics and 3D biometrics are two significant and recently-developed biometric techniques. Typical high-resolution biometric systems include high-resolution finger identification system and signature authentication system, and so on. A high-resolution fingerprint identification system allows fingerprints to be authenticated at three different levels i.e. pattern level, minutia point level, as well as the pore and ridge contour level. The algorithm is another key point of this kind of system. The algorithm developed for the high-resolution biometrics should be able to adequately exploit the much information provided by the system. The high-resolution system cooperating with a fine algorithm dedicated to the high-resolution data is able to produce higher authentication accuracy than the low-resolution system.

3D biometric techniques acquire 3D-image data of biometric traits and the corresponding systems are therefore called 3D biometric systems. Examples of 3D biometrics include 3D face and 3D fingerprint. While the high-resolution and the 3D biometric systems promise higher accuracy, they also suffer from some problems. First, they usually involve a high device cost. Second, since the system retains a large quantity of information from the subject, it is necessary for the system to have a large memory and a high computation performance. This will continue, particularly with more practical techniques and applications.

7. Exploring new areas of biometric security. Innovation and creativity are the impetuses of evolution of human civilization. We have already witnessed a remarkable growth in various biometrics systems and applications over the

past years. Recently, thermal imaging technique, electrocardiogram (or called heart rhythm) based technique and electroencephalogram signals from imagined activities are exploited as biometric characteristics for personal identification. The significance of an uptick in biometrics applications also reflects the requirement for more essential types of biometric characteristics. Biometrics researcher should try their best to find other more fraud resistant biological characteristics and traits for more robust, effective and efficient identification systems.